Crafting digital doubles

Enhancing shape acquisition and material representation

Dissertation

zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Sebastian Thiel

geb. Werner
aus
Bergisch-Gladbach

Bonn, Juli, 2020

Angefertigt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Matthias Hullin

2. Gutachter: Prof. Dr. Reinhard Klein

Tag der Promotion: 25.06.2020
Erscheinungsjahr: 2020

Abstract

The demand for digital applications follows the unhindered growth of the digitization of tasks across many fields, including the entertainment industry, education and science but also work environments. Common tasks thereby involve the generation of photorealistic renderings for movie scenes, illustrations of objects that are not directly accessible or placing virtual objects in real scenarios as required for augmented reality setups. Independent of the specific field, visualization applications are often based on computer-generated imagery (CGI). A key component of CGI are digital doubles, the digital recreation and combination of real-world objects, encoding their geometrical and optical properties.
In this thesis, we will present three approaches to widen the gamut and enhance the quality of digital doubles. We improve shape acquisition techniques for low-cost off-the-shelf hardware solutions and introduce a new digital material that unlocks the recreation of worn surfaces with microscopic scratches. In Chapter 2 we present a structured-light-based range finding approach which utilizes a conventional projector and camera. We employ a standard acquisition technique with minimal changes to the illumination patterns to reconstruct reflection profiles using a closed-form solution. This allows not only for a significantly more reliable detection of invalid measurements but also enables data-driven editing and error correction. This addresses one of the key challenges of structured-light approaches, one the go-to solutions for high-quality 3D scanning. As their performance relies on the choice of illumination patterns, processing methodology and acquisition effort, it is of great importance to decide if a measurement is valid or not, where the latter would introduce severe artifacts into a reconstructed shape and, if invalid measurements exist, to correct these if possible.
In Chapter 3 we further enhance available range finding methods, again utilizing off-the-shelf hardware. Here we instead focus on correlation time-of-flight setups that use a time-dependent amplitude-modulated illumination for an encoding/decoding scheme. Such systems are usually employed for non-static scenes or mobile measurement setups and handheld devices. Our approach allows to focus on a specific range of interest, much like a zoom lens on conventional cameras. Thereby we effectively reduce the measurement range drastically but, in turn, achieve a trade-off due to a significantly improved depth resolution. We show that a sensible choice of signals for the coding scheme leads to strong improvements for these range finding setups especially in low light scenarios.
In Chapters 4 and 5 we turn our attention to digital materials, the second major building block for digital doubles. We are especially interested in effects that are commonly found on everyday life items but are often overlooked. This is especially true for materials that are affected by erosion, exhibit a patina or show defects due to wear: These are very subtle effects, but digital doubles created without these characteristics often appear too sterile,

revealing their artificial nature. We focus on microscopic surface defects, tiny scratches that cannot be resolved with the naked eye but show iridescent colors under strong directional illumination, greatly influencing the perceived appearance. We first perform a qualitative analysis of the underlying surface structure. Based on the gained insights, we then develop our surface model. With this, for the first time macroscopic variation can be combined with microscopic surface features in a physically-based illumination model, which we show to result in a faithful creation of digital doubles that show iridescent scratches.

In conclusion, this thesis provides approaches that can be applied to significantly enhance the quality of digital doubles with respect to shape and digital material alike. For the former, we focus on the improvement of acquisition techniques that cover both, range finding for static and moving scenes due to the choice of different hardware our methods built upon. For the latter we provide an illumination model that is capable of recreating finely detailed surface defects that cause diffraction, an often overlooked but important effect for realistic appearance. These approaches combined lead to new possibilities in the context of computer-generated imagery.

Acknowledgements

Now as the journey of writing my PhD-thesis comes to an end, I want to thank everyone who accompanied me on this way. Without the continuous support of many people during that time this work would not have been possible.

First of all, I would like to thank Prof. Dr. Matthias Hullin for giving me the opportunity to write my PhD-thesis in his working group. It would have been difficult to complete this thesis without his support and feedback over the last few years. I have enjoyed to work in his group and I am very grateful for all the opportunities he gave me to travel to conferences and connect with partners in industry.

I would also like to thank Prof. Dr. Reinhard Klein, who kindly agreed to be the second referee of my thesis. Furthermore, thanks to Prof. Dr. Thomas Schultz and Prof. Dr. Klaus Lehnertz for joining the PhD examination committee.

I am also very grateful to Dr. Michael Weinmann, Clara Callenberg and Julian Iseringhausen for proof-reading my thesis and giving me constructive feedback.

I would like to say thank you to my colleagues in the institute for their support, their encouragement and the nice working atmosphere during the last few years. A special thanks goes to Clara Callenberg for our morning coffee meetings and a lot of motivating discussions.

I also want to thank my family and all my other friends, who always believed in me and supported me all the time during my PhD.

I want to give a very special thank you to my wife, Vivien Thiel. Always believing in me, her support and encouragements have given me the motivation and strengths needed to finish my PhD project. She definitely is my rock in a stormy sea, my refuge.

List of publications

This work forms a cumulative disseration whose core contributions are the two publications listed below, sorted according to their occurrence throughout this thesis.

1. Sebastian Werner, Julian Iseringhausen, Clara Callenberg, and Matthias Hullin. Trigonometric Moments for Editable Structured Light Range Finding. In Hans-Jörg Schulz, Matthias Teschner, and Michael Wimmer, editors, *Vision, Modeling and Visualization*. The Eurographics Association, 2019. ISBN: 978-3-03868-098-7. DOI: 10.2312/vmv.20191315

2. Sebastian Werner, Zdravko Velinov, Wenzel Jakob, and Matthias B Hullin. Scratch iridescence: wave-optical rendering of diffractive surface structure. *ACM Transactions on Graphics (TOG) (Proc. SIGGRAPH Asia)*, 36(6):207, 2017

Contents

CHAPTER 1

Introduction

Computer generated imagery (CGI) makes up a significant part of the images we perceive every day, which occurs in the context of movies and games but also in advertisements, education and work environments. The purpose of CGI in movies and games is to generate appealing imagery by creating virtual worlds and scenarios via adding visual effects or landscapes into movie scenes for example. Advertisement use-cases often include the digital recreation of a product, such as a piece of furniture, within a virtual booth under the illumination of the customer's choice [XRi19]. In particular, customers want the appearance of the digital version of the product they consider to buy to match the real object. Education can rely on CGI for many advantages, ranging from simple illustrations and movies that visualize complex learning content to complex systems that implement interactive learning applications across many fields [Lee12]. In work-related contexts, CGI is often employed in training programs for specialized staff. For example, aircraft maintenance and operation workers are taught using assembly- and flight simulators. On the other hand, medical personnel can be trained using *augmented reality* systems [DFP+10; SOB+04]: Nowadays surgeons can rely on virtual-reality guiding systems that fuse and visualize data from many diagnostic systems for more precise, minimally invasive techniques that lead to better medical results [FLR+98; SNT+98; Shu04; HWR+10].

At the core of these techniques lies the digitization and digital reproduction of real-word scenarios and datasets. Of special interest for a faithful digital recreation within a computer-graphics context are the geometry as well as the optical properties of the objects to be digitized. This includes simple everyday life objects such as chairs, cutlery or other household items for advertisements or the entertainment industry, cockpit interiors and digital planes for flight simulators but also "objects" whose digital recreation is difficult to achieve directly, such as organs or biological tissue. We call such digital representations *digital doubles*, which resemble a corresponding real object. Computer graphics thereby focuses on the digital version of an object's shape and material, which encode the optical properties as well as the geometry respectively. Hence, a digital double's quality is governed directly by the quality of its associated components, which in the end drives its visualization capabilities and governs the possible use-cases.

The computer graphics community strives to achieve a photorealistic recreation of ob-

jects, which imposes high needs in terms of accuracy and quality onto a digital double's components. In general the properties of a real object can be digitized by two approaches:

Measurement of object properties On the one hand, the acquisition of object and material properties such as shape and reflectance behavior by measurements can lead to a precise digital double of the object under investigation. This is especially relevant for fields like medical imaging or cultural heritage that rely on the faithful and exact reconstruction of individual objects [BRB+19a; SK12]. To achieve high-quality digital doubles, highly sophisticated measurement setups have been developed to capture both, shape and material of real objects [WK15; GGG+16]. Depending on the surface material, a wide variety of different acquisition techniques exists, starting with methods for the digitization of simple opaque surfaces up to setups to measure human skin, teeth or even eyes [MWL+99; KRP+15; BBN+14]. In addition to improvements on the hardware side of measurement devices, recent advances in the field of material acquisition strive to provide new parametrization techniques to reduce the acquisition effort as well as the resulting amount of data [LKY+12; BBP15; GGG+16].

Still, such data-driven approaches are in general limited by the resolution of the measurement setup as well as the amount of data that has to be processed. If the detail, either in terms of geometry or reflectance properties or both, is too fine, acquisition either misses important information or requires high sampling rates, which leads to large datasets that are difficult to process.

Modeling object properties On the other hand, the shape and material can be modeled. The former, for example, can be created using 3D modeling frameworks [Ble18] in a purely manual fashion, utilizing machine-learning based approaches [WWL+19; THP+19] or domain-specific solutions [BRB+19b] that infer shape of very specific objects. For simple shapes such approaches often yield satisfactory results that can be achieved with moderate effort. In contrast, objects with highly detailed geometry such as busts or toys for example have been proven difficult to model or immensely increase the effort required for a faithful digital recreation of its geometry .

Illumination models that describe a digital material's interaction with light to recreate an observed visual behavior either are developed in a heuristic manner or are derived from physical principles. Heuristic illumination models [Pho75; Bli77] most often are driven by parameters that are chosen to match a required result, for example the color of a surface. In contrast, physically-based illumination models [CT82; ON94; WML+07; JHY+14] rely on theoretical concepts that aim to describe the scattering process process which forms the material's appearance.

Both, heuristic and physically-based illumination models allow for an abstraction in a plausible way whilst maintaining faithful visual renditions of real materials. They are based on parameters that describe the material which can be freely chosen or directly measured, which circumvents the need to measure the full optical behavior. While parameters of heuristic models are chosen by an artist to match an expected visual result, parameters of physically-based illumination models directly connect to a mathematical description of the scattering process, allowing for a physically meaningful choice of parameter values.

Currently available illumination models are often tailored to specific types of- or even specific materials directly, to make use of simplifications, for example in terms of reduced dimensionality to allow for highly performant solutions. CGI heavily depends on the efficient evaluation of these models for visualization purposes. Such material-specific models however only "sample" but do not cover the full spectrum of real materials.

In this thesis, we focus on solutions to enhance available methodologies that widen the gamut of available digital doubles. In particular, in the following sections we will provide an overview about three new approaches that build upon existing methodologies with the aim to increase the quality of digital doubles. To this end, Chapter 2 presents a range-finding technique that can be used to obtain highly reliable shape information of even complex stationary scenes or objects based on commonly available hardware. In Chapter 3, we focus on different types of depth sensors that are primarily used for scenes in motion, due to their much faster acquisition times. Here, available solutions often suffer from low illumination intensity. We mitigate this problem by proposing a depth-focusing operation mode that works with standard hardware but significantly increases depth resolution locally. Finally, Chapters 4 and 5 present our solution to a problem which remained unsolved up to this point: An illumination model that combines spatial detail with wave optics. In particular, we present an illumination model that simulates worn surfaces with microscopic scratches that diffract incident light, leading to iridescent colors.

The following sections provide a more detailed motivation for each of these approaches. In addition, we present different relevant aspects and considerations that our solutions build upon and present relevant related work.

1.1 Digital doubles

Digital doubles form a key component of any visualization application in the computer graphics context. Both the visualization of a material without an underlying shape as well as the rendering of an object without a material applied are bound to fail the initial goal to create appealing, often even photorealistic imagery. Thereby, the *completeness* of the digital double, i.e the existence of all its components as well as their individual quality are of great importance to achieve sufficient CGI. Shapes that are distorted by artifacts are as great a problem as illumination models that create physically implausible renderings of digital materials. Thereby, digital doubles cannot only be the captured or modeled recreation of an object existing in reality, but can also be a combination of materials and shapes obtained from different sources. Figure 1.1 shows the components of a digital double consisting of data from different sources using the example of our methods as presented in Chapters 2, 3, and 5. For simplicity, we split the discussion of this example into the two components that are required to form a digital double, its shape and digital material:

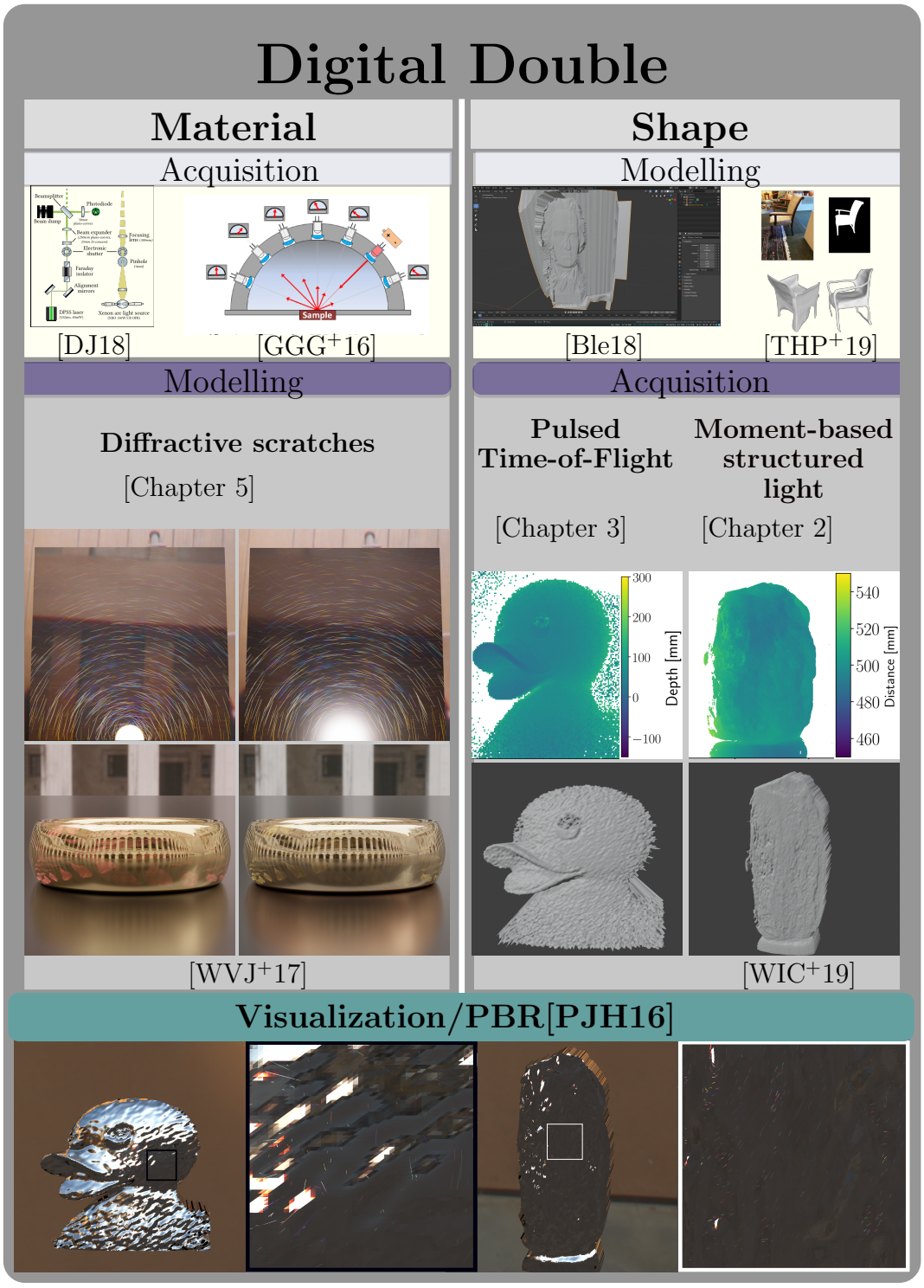

Figure 1.1: Digital doubles consist of two components, the shape and the digital material. Both can either be acquired via modeling or direct acquisition using specific measurement setups. In this work, we present a material model to simulate worn surfaces with microscopic scratches Chapter 5. In addition, we develop two approaches that enhance available shape acquisition techniques based on two different methodologies: Time-of-Flight range finding Chapter 3 and structured light shape acquisition Chapter 2. Digital doubles consist of a combination of the shape and material components, which can be visualized using physically-based rendering (PBR) techniques. This can lead to both a photorealistic representation of an existing object or a recreation of virtual new objects by merging material and shape of different origin.

Shape

A digital double's shape (Fig. 1.1, right panel) represents the object' boundaries within 3D space. This can either be achieved by explicitly modeling the geometry by hand using dedicated software [Ble18] or can be inferred, for example using machine-learning approaches (such as [THP+19]) on images. On the other hand, shape acquisition methods can be employed which commonly are divided into active and passive sensing techniques. A passive technique, such as stereo vision based approaches, does not interfere with the object or scene observed and the corresponding sensor hardware only records light that is coming from the scene. In contrast, an active sensing technique optically modifies the scene by actively illuminating it. At the core of such technique lies a combination of sensors and light sources (for active solutions). Each of the sensors thereby acquires a single depth value per sensor pixel, encoding the distance of the object seen in the pixel's intensity value. Such a *depth map* as displayed in Fig. 1.1 (Shape - Acquisition) yields information of the 3D points visible from the current sensor view point. Such ensembles of 3D points are called *point clouds*, which are commonly triangulated for visualization purposes, leading to a *triangle mesh* [BPK+07]. This comprises a list of *vertices* (the 3D points) and faces or triangles, formed by three vertices respectively, here depicted directly below the corresponding depth maps.

On a side note, it is usually not sufficient to use the point cloud obtained from a single view point to craft a full shape reconstruction of the original object due to self-shadowing effects or other difficulties. For example, the "backside" of the object is not visible from only a single view. Instead, the object is rotated in front of the range finding setup to generate a single depth map per view. From the depth maps, 3D points can be extracted, yielding a dense point cloud which then can be merged and triangulated.

We here focus on two active range finding techniques that are further introduced in Section 1.2.1 and Section 1.2.2. Both build upon existent hardware setups that allow for accurate range finding but strongly suffer from measurement noise and artifacts that arise due to complex light propagation within the measurement area.

The digital material

The material of a digital double defines its interaction with light and hence governs its visual appearance. A mathematical representation of the underlying scattering events is commonly performed by defining illumination models that describe the optical properties of the material. A common approach to such illumination models is to describe the light-material interaction only at the surface, which can be achieved by the *bidirectional scattering distribution function* (BSDF) [NRH+92]. The BSDF encodes the ratio between the incident and outgoing light energy for every surface point and direction possible, which requires an eight-dimensional function at least (four for incident and outgoing angles, four for incident and outgoing surface points). In practice, however, the BSDF is often decomposed into the *bidirectional reflectance distribution function* (BRDF) and the *bidirectional transmittance distribution function* (BTDF). These resemble the reflected and transmitted light components respectively [BDW81] and, in turn, allow for illumination models tailored to the specific

component.

Following this argument, in this work we focus on the development of an illumination model that closes the gap between spatial variation of macroscopic surface defects, such as scratches on worn surfaces, and the implementation of wave optics due to microscopic surface roughness: If the scratches exhibit cross sections in the order of the optical wavelength (below $1\,\mu$m), the incident light is diffracted. In such cases, individual scratch cannot be observed with the naked eye, however their distribution gives rise to a characteristic appearance. We introduce this approach in more detail in Section 1.2.3.

1.2 Contributions

The field of computer graphics is strongly driven by the desire to provide digital doubles of high quality with respect to both shape and material representation. The former thereby is often acquired using measurements for two reasons. First, the capture of the shape of a real-world object delivers an exact (up to the resolution of the measurement device) digital copy of its geometrical properties and second, it is often less time-consuming than manual modeling. On the contrary, the optical properties govering the scattering process are often represented using physically-based illumination models due to the fact that measurements would have to sample a high-dimensional space to create a faithful digital material. Figure 1.1 gives a coarse overview of our contributions with respect to the components required to craft digital doubles. In particular, in this section we focus on the motivation, related work and insights that lead to our methods for enhanced range finding and new illumination models for worn surface materials.

1.2.1 Trigonometric Moments for Editable Structured Light Range Finding

Structured light methods remain one of go-to technologies in high-quality 3D scanning, specifically for the acquisition of single objects and simple scenes in the context of embedded systems, for industrial automation or virtual reality. Especially of use are *phase-shifting-based* approaches that project sinusoidal patterns of different frequencies and phase shifts onto the scene, thereby achieving high depth resolution with spatial sub-pixel accuracy [SFP+10; Zha18]. Most common structured light approaches can be implemented using an off-the-shelf projector for active illumination and a conventional camera for data acquisition. The working principle thereby relies on establishing the correspondence between camera pixels and the sub-pixel locations of the illuminating projector pixels which works well especially for simple scenes. Reconstruction of the depth then follows via triangulation of the rays formed by camera-projector pixel correspondence pairs, yielding an independent depth value per (camera) pixel. On the other hand, machine learning techniques have become more and more popular also for shape acquisition [RRT+16; RVR+17], where the triangulation step is replaced by a 'learning-based' depth estimation.

Due to the ascent of digital image sensors in recent years, the (spatial) resolution of structured light methods is primarily bound by the resolution of the projector, as CCD and CMOS

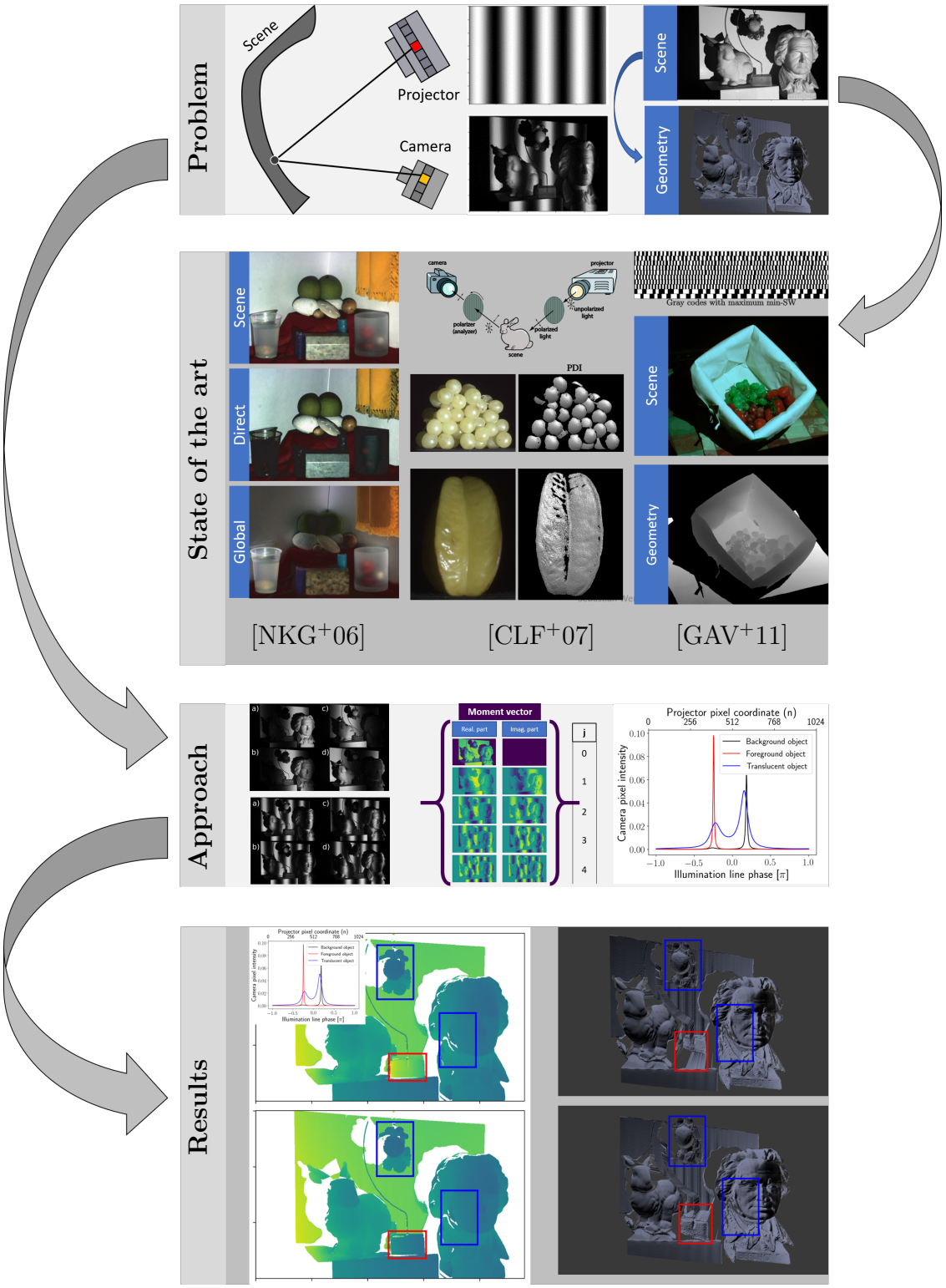

Figure 1.2: Phase-shifting based structured-light range finding techniques achieve high depth resolution with spatial sub-pixel accuracy. A common problem in such methods are specular or translucent materials, concave geometry or even both. To mitigate measurement artifacts that stem from such effects, several approaches exist that aim to separate the global component due to complex light transport, either by specialized illumination patterns or additional hardware. Instead of removing data, we use the full information content of standard phase-shifting measurements. We utilize specific frequencies for our sinusoidal illumination patterns, which allows for a closed-form reconstruction that reveals a reflectance profile per camera pixel. This enables us to extract more reliable depth measurements or even reconstruct fore- and background of translucent objects.

sensors reach extreme pixel densities with up to 102 megapixels resolution for high-end DSLR and mirrorless cameras. Figure 1.2 gives an overview about our active range finding technique based on a standard phase-shifting structured light approach. Such approaches commonly suffer from more complex scene geometries or materials, such as translucent objects. These involve non-local light transport (e.g. interreflections, sub-surface scattering), adding a *global component* that leads to severe artifacts in the reconstructed depth maps [GBR+01]. A number of approaches aim to reduce or completely remove the global component, either by introducing additional hardware [XZJ+19], by carefully choosing suitable illumination patterns [NKG+06; TAH+07; CSL08; GAV+11], or both [CLF+07; MHP+07; HL11].

These methods share the goal to separate the global component from the measurement entirely, either before the data is acquired or in a first processing step. We instead draw inspiration from the field of *light transport analysis* which deals with the acquisition of the full reflectance field of a scene. Often, such approaches rely on structured light measurement modalities. A different data acquisition and processing scheme however allows to virtually project patterns onto the scene that were never actually measured or to exchange the viewpoint of camera and projector [SCG+05; SD09]. In addition O'Toole et al. [OMK14] showed that the analysis of light transport can be used to perform a much more accurate removal of the global component, however at the cost of highly specialized hardware.

Instead of removing data, our approach [WIC+19] uses the full information content of standard phase-shifting measurements by exploiting a mathematical re-presentation of the data to form a complex-valued vector containing all measurements. We thereby do not alter the acquisition pipeline except for specifically choosing which frequencies the sinusoidal patterns exhibit. With this, we effectively sub-sample the spectrum of the underlying reflectance field per camera pixel. We do not recover the full spectrum but a subset of its *trigonometric moments*. We draw upon findings from Peters et al. [PKH+15] to perform a spectral estimation post-processing step which allows to recover a virtual reflectance profile per camera pixel using an analytic expression. This allows for an almost instantaneous reconstruction, which resembles an illumination with a line, swept over the scene. However, instead of shifting this line pixel by pixel on the projector and taking hundreds of images (1280 images for a 720p projector), we recover the corresponding information content with only 20 images using sinusoidal illumination patterns with 5 frequencies and 4 phase shifts each.

We in addition exploit the acquired closed-form reflectance profiles to gain insights about the underlying light transport, allowing for a significantly more robust identification of pixels with valid (i.e physically plausible) measurements, effectively reducing the number of artifacts in the reconstruction. Furthermore, we are able to perform a data-driven semi-automatic editing which can be used to correct errors, simply by picking the most likely reflectance profile characteristics. For example, this editable range finding allows us to acquire the shape of translucent materials as well as to remove them from the measurement.

All in all, this gives the opportunity to use standard phase shifting techniques for structured light range finding and, with only minor adjustments, obtain more reliable and accurate shapes to craft digital doubles.

1.2.2 A new operation mode for depth-focused high-sensitivity ToF range finding

Time-of-Flight (ToF) range finding setups are deployed in many applications and can be found in robotics, closely tied to exploration or automated manufacturing, in motion capture and 3D mapping as well as in biometrics [KBK+09; KHD+09; PVB+11; VWG+12; WOV+12]. Counting towards active range finding techniques, they consist of a light source-sensor pair, where light is sent into a scene to encode the distance between light source, object and sensor in the time a ray of light has traveled. In contrast to such *pulse-based* (P-ToF) approaches [GD67; Koe68; IY01; YIM07], most off-the-shelf systems nowadays use an illumination with light that is amplitude-modulated periodically in the time domain (usually several 10 MHz), usually employing a sinusoidal or rectangular modulation. This encodes the time a ray has traveled in a phase measurement [SXH+97; LSB+00], which is reconstructed by "comparison" with a demodulation signal at the sensor. By comparison of the original emitted signal and the received version, these *correlation ToF* (C-ToF) methods are able to reconstruct the time-of-flight and hence distance. Using specialized sensors, this decoding step is performed on a per-pixel level in hardware, allowing for sufficiently fast acquisition of depth maps for motion capture for example. We also rely on such off-the-shelf C-ToF sensors with the aim to increase depth accuracy in low-light scenarios, Figure 1.3 shows an overview of our method.

The specialized hardware comes at the cost of reduced spatial resolution, where state-of-the-art modules deliver up to 640x480 pixels [Luc19], which reveals a major drawback in comparison to structured light methods. Despite of their low spatial resolution, ToF systems remain an important pillar of depth sensing methodologies.

Thanks to advances in sensor technology, the amount of pixels increases between subsequent sensor generations. In contrast, the depth resolution is driven by the working principle and heavily depends on design choices such as the time- and power budget, which is especially limited for mobile applications. In contrast to P-ToF systems, correlation-based ToF setups are more prone to errors from measurement noise [BS08], resulting in a considerably lower depth resolution.

To circumvent this problem, available solutions aim to optimize the modulation-demodulation signal pairs, either by a combination of high- and low-frequency signals [JCP+10; JBP+11], higher harmonic suppression [PDC+08; PDC+10] or by developing different modulation signals that are specially tailored to a certain task [GVN+18].

These approaches thereby aim to achieve a high depth resolution over the full, frequency-limited measurement range. In contrast, we propose PC-ToF Chapter 3, a two-step depth acquisition scheme which allows for depth focusing and increased accuracy. To achieve this goal, we draw from an observation by Gupta et al. [GVN+18], who show that the accuracy of a C-ToF setup relies on the gradient of the participating modulation and demodulation signals. Instead of optimizing this characteristic for the full measurement range, we focus on locally maximal gradients, which can be achieved by combining a pulsed illumination signal with a rectangular demodulation. This, however, effectively limits the range of valid measurements to a fraction of the full measurement range: Our prototype as presented in Chapter 3 achieves a measurement range of about 80 cm around a specific depth, whereas

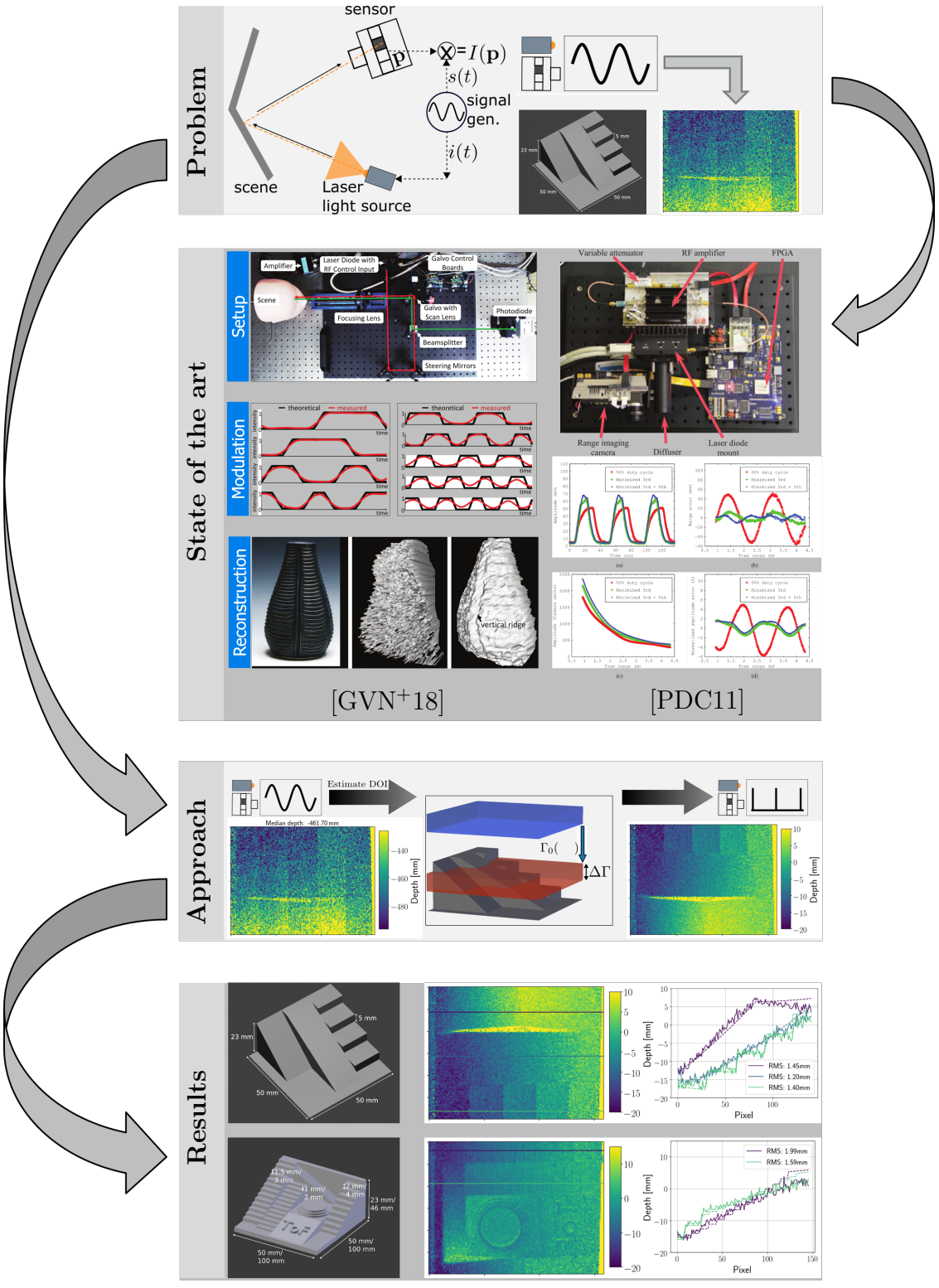

Figure 1.3: Overview of our pulsed-correlation Time-of-Flight (PC-ToF) range finding approach. Available off-the-shelf correlation Time-of-Flight imaging solutions suffer from strongly reduced accuracy in low signal-to-noise-ratio scenarios. To circumvent this problem, available solutions aim to optimize the modulation-demodulation signal pairs to support as high depth resolution as possible over the full measurement range. We propose to use PC-ToF in a two-step depth acquisition scheme, where a first, rough estimate is obtained using standard correlation Time-of-Flight with sinusoidal modulation-demodulation, followed by a depth-focused PC-ToF measurement that exploits a trade-of between measurement range and depth accuracy to locally obtain high depth resolution, especially for low SNR.

the standard C-ToF approach yields valid measurements over a range of 15 m. Still, our method allows to sample the full measurement range, effectively increasing the acquisition time significantly but allowing to "focus" on a certain depth, which especially useful for example for scanning small objects.

To circumvent the need for such a costly sampling strategy, we perform a two-step acquisition by a first, rough depth estimate, followed by a high-sensitivity depth measurement. To this end, we first employ standard C-ToF with sinusoidal modulation-demodulation to achieve a notion of the depth we want to focus on. After the first measurement, the rough depth estimate allows to shift the sensitive measurement range of PC-ToF towards the depth of interest. This is followed by a focused measurement which employs a sub-nanosecond laser pulse illumination signal in combination with a rectangular demodulation.

With this method, we trade the global sensitivity of a standard C-ToF setup for measurements with strongly localized high sensitivity, especially for low-light scenarios. Using real-world experiments, we show that our technique is capable of achieving depth resolutions up to 2 mm under optimal conditions using a modulation frequency as low as 10 MHz and an optical power as low as 0.6 mW. In general, we are able to halve the depth RMS error in comparison to standard correlation ToF measurements at equal optical power. This makes PC-ToF especially viable for low-power applications, for example for mobile devices. We implement our approach as an additional mode of operation on standard correlation Time-of-Flight hardware which in the end allows to "switch" to a more accurate range finding technique for improved quality of a digital double's shape.

1.2.3 Scratch iridescence: Wave-optical rendering of diffractive surface structure

Detailed modeling and rendering of surface defects can dramatically improve the realism of digital doubles, hence the pursuit of such models has been a topic of great interest to the rendering community at large [DRS10]. Measuring the scattering and reflectance distribution functions of materials is often performed to explicitly capture data of special materials in movie productions or industry applications. It is usually not desirable for materials that exhibit a very detailed scattering distribution with high angular dependency, especially in combination with spectral dependency, as the required high sampling rates lead to large amounts of data. This is especially true for surface materials with irregularities and imperfections, which remains one of the greatest challenges in computer graphics with respect to both, measurements and modeling. We focus on a gap in the spectrum of digital materials for digital doubles, that the computer graphics community so far was not able to close. Under bright, directional lighting, which occurs for example in kitchen environments due to spotlights, surfaces of materials such as glass, metal or plastic objects exhibit iridescent colors. The reason for this is diffraction, a wave-optical phenomenon that becomes visible as the light is reflected off microscopic surface features that are in the order of the optical wavelength. Diffraction, however, poses a challenging problem for measurements, as the strong dependency of the diffraction patterns on viewing and lighting conditions [Goo96; LLL10] requires spectral sampling with very high angular resolution, rendering the data acquisition unfeasible.

Still, the mathematical description of such patterns ultimately relies on the height varia-

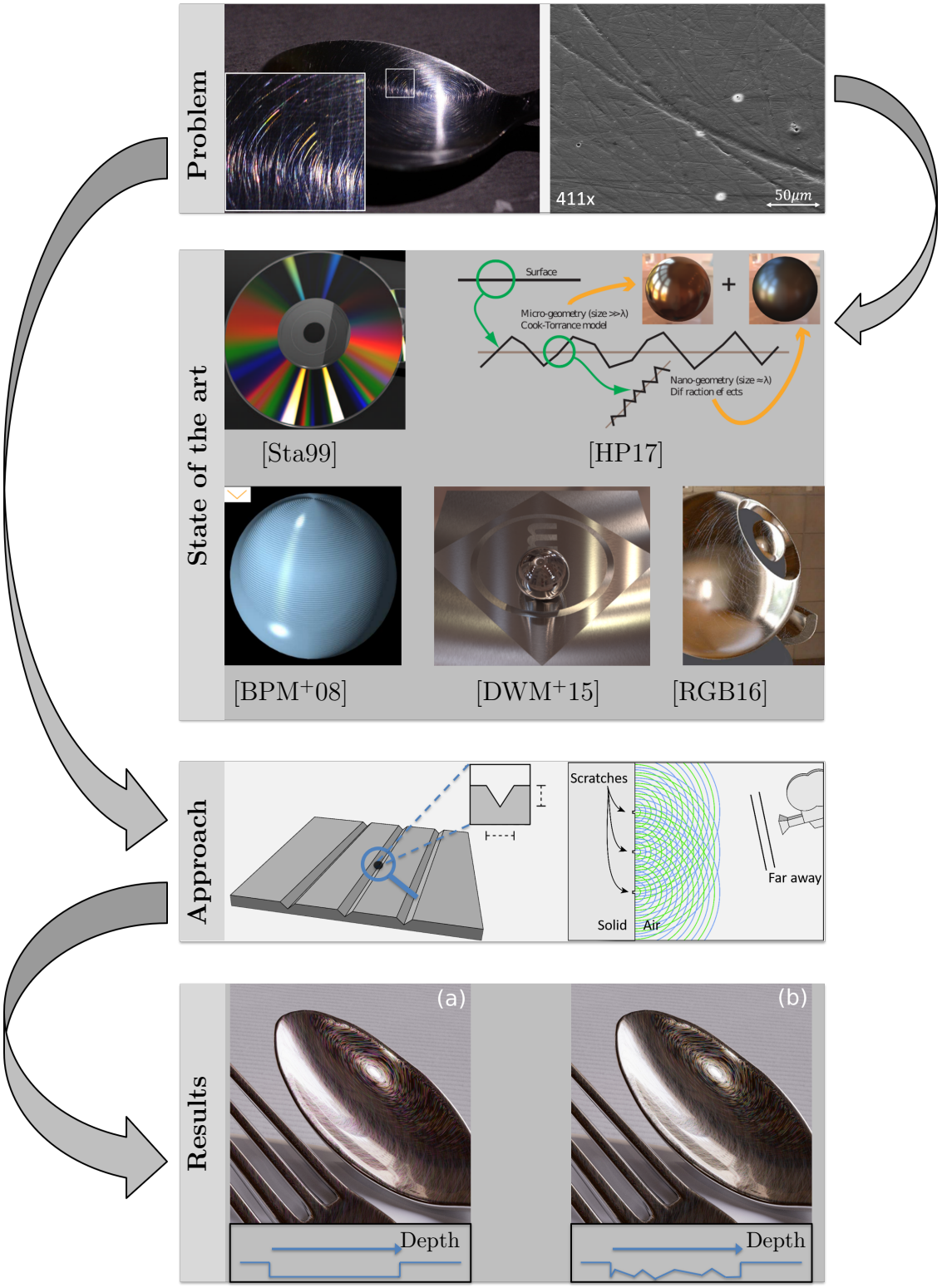

Figure 1.4: Overview of our method to simulate iridescent microscopic scratches on worn surfaces. Their cross-section is in the order of the optical wavelength, causing diffraction. Available models either focus only on wave-optical effects of surfaces without macroscopic variation or rely on geometrical optics to simulate scratches on surface materials. We approach this problem by merging both methodologies. The scratch geometry is described by linear segments with constant cross-section. For these, we are able to analytically compute wave-optical reflectance distributions. This allows us to create photorealistic renderings of worn objects such as spoons. The realism of such visualization strongly depends on the variations of model parameters: Varying the depth along a scratch results in a more natural look whilst a constant depth can be seen as a manufactured sample.

tions encountered on the surface material, such as dents or scratches, which could be used as input for data-driven approaches [CHB+12; DTS+14; MMR+13] able to recreate diffraction effects. However, a measurement as well as a simulation of the surface irregularities in this case would have to be performed on a sub-micrometer resolution, creating large amounts of data for even small patches of material.

Other available illumination models either support surface defects such as scratches, patination or erosion on surface materials [DRS10; MDG01; BPM+04; YHM+16; DWM+15] with respect to only geometrical optics, where light can be described as a ray. Illumination models that allow to simulate wave-optical effects due to microscopic surface structure on the other hand are mostly limited to random surface height variations [HTS+91; SFD+00; BB17].

To approach this problem, we draw inspiration from works that explicitly model the scratch geometry. Figure 1.4 shows an overview of our approach. Similar to prior work [RGB16; BPM+08], we separate spatial and optical information by describing the scratch layout as a curve and its reflectance behavior using a profile at each position along the it. Like Holzschuch and Pacanowski [HP17] we use non-paraxial scalar diffraction theory [HVK+00] to express the diffracted reflectance as a superposition of reflections from individual scratches. We thereby implicitly apply the far-field approximation [LLL10], which is a reasonable assumption given that the scratched surface is observed from at least a few cm away. Similar to Sun et al. [SFD+00] we derive the BRDF from the explicit calculation of the scattered complex wavefront, maintaining as much generality as possible.

To this end, we derive a closed-form solution to the diffraction integral for linear scratch segments. On the mathematical side, this imposes a problem with respect to the separability of multiple scratches, as the diffraction integral requires the Fourier transform of the full surface heightfield. To solve this problem, we draw inspiration directly from physics. Extended light sources, which we have to consider for any realistic lighting scenario, exhibit a limited *spatial coherence* [LLL10], which acts as a spatial filter, effectively selecting scratches that contribute for a given region on the surface. This allows us to reproduce not only diffraction effects but also the mutual interference created by dense scratch ensembles, providing natural transitions from localized glint-like iridescence to smooth surface reflectance representing the superposition of many reflections at large viewing distances.

We demonstrate that our model is capable of recreating the overall appearance as well as characteristic detail effects observed on real-world examples, such as metal plates or spoons. The realism of such visualization strongly depends on the variations of model parameters: Varying the depth along a scratch results in a more natural look whilst a constant depth can be understood as a manufactured sample. With this illumination model, we widen the gamut of available digital materials for the creation of digital doubles.

1.3 Outline of this work

This work evolves around the idea of digital doubles, where we aim to improve and enhance available methods or methodologies. In particular, we directly focus on the building blocks of digital doubles, the shape and material of such a digital recreation. This cumulative thesis is organized as follows: In Chapter 2 we present our first project, *Moment-based structured*

light based on our paper "Trigonometric Moments for Editable Structured Light Range Finding" [WIC+19], which was presented at VMV 2019. We here introduce an improved structured light range finding technique that relies on commonly available off-the-shelf structured light phase-shifting solutions. These commonly require a single projector and conventional camera to project sinusoidal patterns (in the spatial domain) onto an unknown scene. We represent the data acquired using such a standard approach in a different mathematical basis. With minimal assumptions to be fulfilled, we are then able to perform a closed-form density estimation that enables the data-driven refinement of measured depth maps and significantly enhances the reliability of acquired depth estimates. This is particularly useful for the reconstruction of scenes and objects that exhibit complex light transport, be it via interreflections between objects within the scene or "difficult" materials such as translucent surfaces.

In Chapter 3 we further enhance available range finding methods, again utilizing off-the-shelf hardware. Here however we focus on the use of correlation time-of-flight setups that use a time-dependent amplitude-modulated illumination for an encoding/decoding scheme. Our approach allows to focus on a specific range of interest, thereby effectively reducing the measurement range drastically but trading this for a significantly improved depth resolution, especially useful to scan small objects or specific areas of the scene with increased accuracy. We show that a combination of a pulsed illumination and rectangular demodulation signals for the coding scheme leads to strong improvements for range finding setups especially in low light scenarios.

In Chapters 4 and 5 we focus on the second major building block for digital doubles: With two new approaches that enhance the available shape acquisition techniques at hand we turn our attention to digital materials. Of particular interest to us are effects that are commonly found on everyday life items but are often overlooked. This is especially interesting for materials that are affected by erosion, exhibit a patina or show defects due to wear: These are very subtle effects, however, whenever a digital double is created without these characteristics, it appears too sterile in almost any environment.

We focus on microscopic surface defects, tiny scratches that cannot be resolved with the naked eye but show iridescent colors under strong directional illumination, greatly influencing the perceived appearance. To this end, we perform a qualitative analysis of the underlying surface structure and properties such microscopic scratches exhibit as discussed in Chapter 4. Building on these findings, we present our surface model in Chapter 5, which was presented at SIGGRAPH ASIA 2017, corresponding to our paper "Scratch iridescence: Wave-optical rendering of diffractive surface structure" [WVJ+17]. With this, for the first time macroscopic variation can be combined with microscopic surface features in a physically-based way, which we show to result in a faithful creation of digital doubles that exhibit such phenomena.

Each chapter is preceded by a preface, summarizing the relations of each project to the superordinate creation of digital doubles and stating the challenges we face as well as the contributions we make to the respective field. Due to the different underlying concepts the projects rely on, the notations given in the respective chapters are decoupled from each other. Finally, in Chapter 6 we present a conclusion followed by a discussion about possible

directions for future work.

Preface – Moment-based structured light

Range finding techniques comprise a crucial part in the field of digital material appearance: The perceived realism of any digital double is only as good as its geometric model, either finely crafted by an artist or accurately measured. Structured light methods make up a huge part of off-the-shelf solutions for 3D shape acquisition, ranging from table-top systems for hobbyists to industry-grade setups for commercial asset generation used for a variety of products. Originally a stereo-vision-based approach, the 3D information is acquired by finding correspondences between camera and projector pixels. The shape of the object that is to be measured thereby distorts an a priori known illumination pattern projected onto it, encoding the underlying geometry. In combination with phase-shifting approaches that are aimed at sub-pixel accuracy on the projector side, the spatial resolution of these approaches is nowadays still mainly limited by the projector resolution due to high pixel densities on conventional DSLR or mirrorless cameras.

The procedure, being based on visual information, in general exhibits two directly apparent bottlenecks in terms of performance and quality: First, parts of the object that cannot be directly illuminated or viewed, for example due to self-shadowing, cannot be measured. Second, the quality of the digital double increases with the number of views and the number of pixels per view, the acquisition time however increases drastically when aiming for high angular coverage. The former is readily fixed by rotating the object in front of the acquisition setup: The implicit higher angular coverage often allows to acquire data from parts that are shadowed in different views.

In addition, a crucial part that is critical for all available approaches is the identification of unreliable measurements: Either originating from aforementioned problems that entirely rely on the measurement setup and shape of the object or from the reconstruction procedure, the user needs to make sure that the shape is not distorted by false positives, for example caused by shadowing, too strong specular reflections or caustics.

To enhance shape acquisition techniques that rely on structured light methods, we present a procedure that relies on standard phase-shifting approaches but, by introducing a different representation of the acquired data, allows for a heavily improved pixel-wise reliability measure and in addition enables a data-driven editing step. We employed this editing capabilities in a semi-automatic fashion to refine uncertain measurement points, which reduces the amount of measurements required for the digitization and allows for improved quality of the resulting digital double. The chapter

presents our work as published in [WIC+19], the supplemental videos are available at `https://osf.io/etdgb/?view_only=a60cc22ccdf54afbb9be279495d75508`.

CHAPTER 2

Moment-based structured light

Abstract Structured-light methods remain one of the leading technologies in high quality 3D scanning, specifically for the acquisition of single objects and simple scenes. For more complex scene geometries, however, non-local light transport (e.g. interreflections, subsurface scattering) comes into play, which leads to errors in the depth estimation. Probing the light transport tensor, which describes the global mapping between illumination and observed intensity under the influence of the scene can help to understand and correct these errors, but requires extensive scanning. We aim to recover a 3D subset of the full 4D light transport tensor, which represents the scene as illuminated by line patterns, rendering the approach especially useful for triangulation methods. To this end we propose a frequency-domain approach based on spectral estimation to reduce the number of required input images. Our method can be applied independently on each pixel of the observing camera, making it perfectly parallelizable with respect to the camera pixels. The result is a closed-form representation of the scene reflection recorded under line illumination, which, if necessary, masks pixels with complex global light transport contributions and, if possible, enables the correction of such measurements via data-driven semi-automatic editing.

2.1 Introduction

Structured light methods are one of the go-to solutions in current 3D scanning setups, ranging from use cases in standard shape acquisition for digitization to embedded systems for industrial automation and virtual reality. The underlying working principle relies on establishing the correspondence between camera pixels and the sub-pixel locations of the illuminating projector pixels. We acquire a 3D representation of an unknown scene by projecting illumination patterns onto the scene and acquire the reflectance with a camera. The scene geometry thereby distorts the illumination pattern and hence encodes the underlying structure. As active illumination is used in structured light setups, a controlled generation of correspondences is possible, allowing to extract the encoded 3D information via triangulation of pixel-pixel or pixel-line pairs. To this end, a large variety of patterns have been developed, ranging from simple linesweep illumination, where one captures the projection of a single line of projector pixels, to binary codes to reduce the number of required acquisitions, to phase shifting methods [SFP+10] that rely on the use of fringe patterns with continuous intensity

variation for subpixel accuracy. These methods all share the desire to reduce the acquisition time and number of recorded images as well as to increase the depth resolution [Zha18]. These avaliable systems often rely on the assumption that an observed scene point reflection uniquely maps to direct illumination from the light source, implicitly requiring diffuse materials to dominate the scene. Real-world scenes often times violate this assumption: The incoming illumination of a 3D scene point originates not only directly from a certain projector pixel but is perturbed by complex light transport effects such as subsurface scattering, interreflections or volumetric scattering, adding a so-called *global component* to each camera pixel intensity. Considering the required camera-projector correspondences for triangulation this leads to severe problems as the mapping no longer is unique, which can lead to large errors in the recovered shape. To mitigate this undesired effect, a number of approaches has been developed which aim at reducing or completely removing the global component, either by introducing additional hardware [XZJ+19], by carefully choosing suitable illumination patterns [CSL08; GAV+11], or both [CLF+07]. Instead of trying to remove the global component, we aim to explicitly capture all available information in a way that is most suited for 3D reconstruction via triangulation. In theory, this is given by pixel-pixel correspondences via measuring the camera-pixel reflectance per projector pixel, yielding the full *4D light transport tensor*. Alternatively a standard linescan procedure can be performed, for which a lower number of acquisitions is needed. Still, both methods require extensive measurement and/or computational effort. Our method on the other hand relies on a phase-shifting approach based on a small number of sinusoidal illumination patterns with specifically chosen frequencies. This allows us to employ a closed-form reconstruction scheme to estimate a 3D subset of the light transport tensor, which resembles a dense linesweep illumination yielding a *pixel response* per camera pixel. With this we obtain a functional dependence of the camera pixel intensity based on the linesweep position. By explicitly reconstructing the linesweep light transport tensor from the sparse measurements, we are able to perform an in-detail analysis of global illumination effects and to choose the most likely correspondence as well as perform a data-driven semi-automatic refinement procedure.

2.2 Related work

In this related work section we focus on structured light methods for complex scenes, such as those containing large amounts of global illuminations or translucent materials. For a more concise review about the state-of-the art in structured light techniques we refer the reader to the reports of Salvi et al. [SFP+10] and Zhang [Zha18]. We also consider time-of-flight (TOF) techniques related to our approach. In particular, there exists a subclass of TOF approaches that aim to correct for non-direct illumination [WOV+12; HHG+13; KWB+13; FSK+14; PKH+15] by estimating the time-resolved reflectance profile per pixel. These techniques then allow to separate the direct and global component in the time domain, whereas our approach focusses on the spatial domain to do so.

2.2.1 Phase shifting

Continuous coding by phase-shifting based on sinusoidal illumination patterns is a well known technique used in structured light geometry acquisition [SFP+10]. A common setup consists of a digital projector, which projects a series of phase-shifted sinusoidal patterns onto the scene, and a camera that captures the reflected light. Ideally, phase information is extracted for each camera pixel, encoding the camera-projector pixel correspondence. For higher accuracy usually multi-frequency measurements are performed, for which phase unwrapping is required. Our method inherently relies on a multi-frequency phase-shifting approach with sinusoidal illumination patterns to acquire the *trigonometric moments* of the scene per camera pixel.

2.2.2 Shape estimation under global illumination

3D shape acquisition based on structured light setups relies on the detection of direct reflections of the illumination. The global component is caused by subsurface scattering, interreflections, and ambient light and can strongly affect and perturb the acquired correspondences, which can lead to severe errors in the shape estimation [GBR+01]. Key to many applications is to remove or mitigate the global component, either by careful choice of illumination patterns or by hardware modifications.

One of the first to acquire shape in the presence of global illumination were Nayar et al. [NIK91], who presented an iterative approach for Lambertian objects. Chandraker et al. [CKK05] estimate 3D data using a shape-from-shading technique. Shape-from-shading is also used by Chen et al. [CGS06], who employ an interactive photometric method to obtain specular reflections. Modifications of the measurement procedure were proposed by Park et al. [PK08], who move the camera for global illumination mitigation. More recently, Fanello et al. [RRT+16; RVR+17] used learning techniques to estimate disparity maps from infrared structured light data and augmented stereo vision systems respectively. Both approaches effectively *learn* disparity maps and yield very efficient high accuracy depth reconstruction for standard (not translucent) materials.

A milestone with respect to phase-shifting-based structured light is the work of Nayar et al. [NKG+06], who showed that high-frequency illumination patterns can be used to effectively separate the direct and global component. Talvala et al. [TAH+07] remove glare from high dynamic range images using Nayar's separation approach by selectively masking light that generates the glare. On the other hand, Nayar's direct-global separation technique was introduced to structured light systems by Chen et al. [CLF+07], who combine it with a polarization difference imaging (PDI) acquisition step. As multiple scattering events depolarize the reflected light [HvdH81; SNN03; SK05], this provides an additional filter to reduce the global component. Afterwards, Chen et al. [CSL08] proposed a modulated illumination signal for improved reduction of the global component. Similarly, Holroyd et al. [HL11] proposed an active multi-view stereo technique with high-frequency illumination that is invariant to global illumination. Ma et al. [MHP+07] extended the idea of PDI in conjunction with a shape-from-shading approach to use circularly polarized spherical gradient illumination for the recovery of translucent objects.

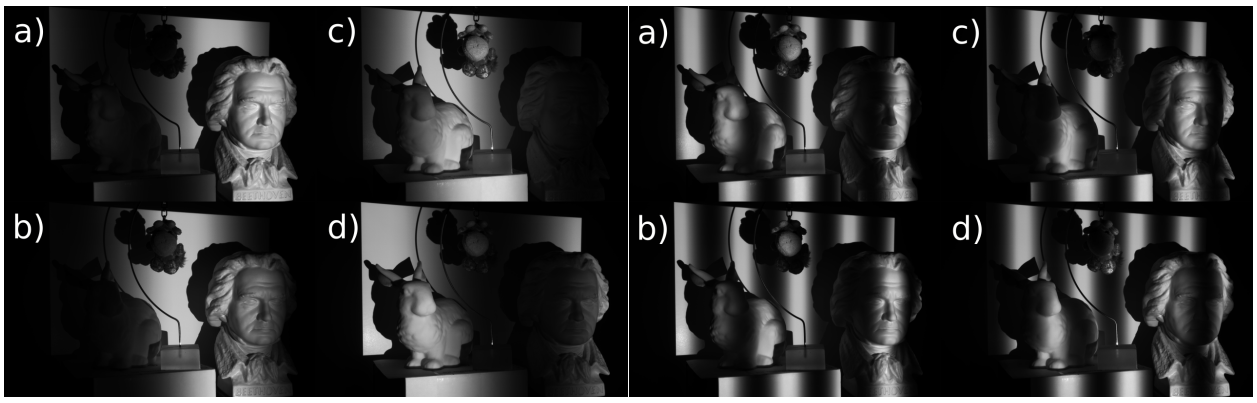

Figure 2.1: Example images acquired with our illumination procedure. We use sinusoidal illumination patterns with different frequencies ranging from $\nu = 1$(**left**) to $\nu = 4$ (**right**) and acquire images for four equi-spaced phases in the range $[0, 2\pi]$ (**a**, **b**, **c**, **d**).

2.2.3 Pattern optimization and light transport acquisition

The aforementioned techniques rely on the elimination of the global component, combined with mostly off-the-shelve shape acquisition techniques acting on the remaining, direct component. In contrast, Gupta et al. [GAV+11] combine the complementing properties of Gray codes and logical codes for depth measurement. We consider this approach closely related to the idea of moment-based structured light, although our method reconstructs a continuous pixel response that contains more information than a binary coded pattern. Also connected to our approach are methods that try to acquire the full reflectance field of a scene, such as [SCG+05; SD09] who use this concept for dual photography. In the context of shape acquisition, light transport analysis has been successfully used to estimate shape and surface normals under the assumption of Lambertian surfaces [LNM10]. Reddy et al. [RRC12] perform a direct-global separation based on frequency domain considerations. More recently, O'Toole et al. [OMK14] showed that 3D shape acquisition greatly benefits from light transport analysis. Despite being very efficient, their approach relies on special hardware that is capable of high-speed acquisition and modulation.

Our approach extends the capabilities of existing phase-shifting methods without the overhead of additional hardware components. Instead, we rely on a frequency-domain approach to reduce the number of required input images and perform spectral estimation to recover a scene response that encodes the pixels' reflected intensity in the presence of a linesweep illumination. This approach has three major advantages: First, a linesweep illumination renders our approach specifically suited for triangulation as we directly enforce camera-projector pixel correspondences. Second, we require only a small fraction of the number of acquisitions compared to a conventional linesweep. Furthermore, we show that by explicitly recovering the 3D (linesweep) light transport tensor we are able to reliably separate direct from global components and refine our results by applying a data-driven editing technique, which substantially improves the resulting depth map quality.

2.3 Phase shifting for structured light

Projecting a sinusoidal pattern onto a scene allows a unique mapping between camera pixel and projector pixel: Each point along a line across the sinusoid can be assigned a specific phase value. Non-flat geometry within the scene will deform this pattern, yielding a phase deviation in a recorded image which encodes the underlying 3D shape. A decoding step that matches the deformed and originally projected patterns then allows a phase retrieval. To this end, the *N-step phase shifting* is used where subsequent captures of shifted versions of the projected pattern are performed, usually with K equally spaced phase shifts $\delta\varphi = 2\pi/K$. The illumination signal of a projector pixel (m, n) can then be described as

$$I_k(m,n) = \frac{1}{2}\left[\cos(j \cdot \nu + \delta\varphi_k) + 1\right];\ \delta\varphi_k = k \cdot 2\pi/K, \tag{2.1}$$

where j denotes an integer multiple of the frequency ν and k the k-th phase shift. To express the light transport within a scene, it is common to consider the 4D light transport tensor which can be understood as a density $\rho(x, y, m, n)$, where (x, y) are coordinates in the camera system denoting the respective camera pixels and (m, n) with $m \in [0, \ldots, M]$, $n \in [0, \ldots, N]$ the corresponding projector pixels. In other words, this tensor maps the illumination intensity present at pixel (m, n) (or any superposition of multiple pixels) of the projector to the received intensity value at the camera pixel (x, y), thus encoding the full light transport within the scene. For a single phase shift $\delta\varphi_k$ and fixed j the measurement procedure then acquires the camera pixel intensity

$$\begin{aligned} I_k(x,y) &= \int_0^M \int_0^N I_k(m,n)\rho(x,y,m,n)\,dm\,dn \\ &= \frac{1}{2}\int_0^M \int_0^N \left[\cos(j \cdot 2\pi\frac{n}{N} + \delta\varphi_k) + 1\right]\rho(x,y,m,n)\,dm\,dn \end{aligned} \tag{2.2}$$

which is the convolution of the light transport tensor and the active illumination, containing all the global light transport effects and we assumed the base frequency $\nu = 1$ along the (horizontal) n-axis of the projector. In the structured light literature, the density ρ is usually split into a global and direct part ρ_G and ρ_D where the direct part maps exactly one camera- to one projector-pixel and the result of Eq. 2.2 for a single frequency j is often written as [CLF+07]

$$\begin{aligned} I_k(x,y) = &\frac{1}{2}\left[L_d(x,y) \cdot \cos(\Phi(x,y) + \delta\varphi_k)\right] \\ &+ \frac{1}{2}\left[L_d(x,y) + L_g(x,y)\right] \end{aligned} \tag{2.3}$$

where L_d denotes the direct reflection obtained from ρ_D, depending on the phase $\Phi(x, y)$ of the surface point observed in camera pixel (x, y), encoding the surface point's local geometry. L_g on the other hand is the global component, which is independent on the phase and originates from ρ_G. With at least three different, equally spaced phase shifts $\delta\varphi_k$ the global and direct components can then be separated and the phase $\Phi(x, y)$ can be estimated, yielding point correspondences for triangulation, see [SFP+10; CLF+07] for a more detailed explanation.

It is clear that the full knowledge of the 4D light transport tensor would yield insight into the global illumination effects present within the scene, regardless of their physical origin, as L_g could be computed and corrected for. The measurement of this tensor, however, is either costly in terms of acquisition time [SCG+05] and data storage or requires extensive numerical reconstruction [SD09], rendering it unfeasible for shape acquisition techniques based on structured light.

In contrast, we do not aim to reconstruct the full 4D light transport tensor but a 3D subset of it, which is inherently related to the scene as if being illuminated by a linesweep, rendering it highly suited for projector-camera pixel correspondence finding. We rely on the standard phase shifting approach and re-interpret the measurements to obtain a vector of *trigonometric moments* per camera pixel.

2.4 Using trigonometric moments for structured light

Peters et al. [PKH+15] utilized a spectral estimation technique based on the maximal Burg entropy [Bur79] to reconstruct time-resolved scene responses from data captured with a specialized camera. Their technique relies on a sinusoidal illumination modulation (in the time domain) and presents a closed-form reconstruction scheme for the scene response per camera pixel. Key to this technique is the acquisition of multiple images with modulation frequencies that are the integer multiple of a chosen base frequency, which results in measurements of the so-called *trigonometric moments*. We transfer this technique to the concept of structured light shape acquisition based on phase-shifting, where a scene is illuminated with sinusoidal patterns of choosable frequency in spatial domain.

In general, the trigonometric moments b_j of a 1D density $h(\varphi)$ are described as

$$b_j = \int_0^{2\pi} h(\varphi) e^{ij\varphi} \, d\varphi \tag{2.4}$$

where as before j denotes the j-th multiple of the base frequency (assumed to be 1) and $h(\varphi)$ is a 2π-periodic density distribution function. The reconstruction of the underlying periodic function $h(\varphi)$ is then performed using a maximum entropy formalism that maximizes the Burg entropy H (see Peters et al. [PKH+15]) given as

$$H[h(\varphi)] = \int_0^{2\pi} -\log(h(\varphi)) \, \mathrm{d}\varphi. \tag{2.5}$$

The closed form solution then reads

$$h(\varphi) = \frac{1}{2\pi} \frac{e_0^T \cdot \mathbf{B}^{-1} \cdot e_0}{|e_0^T \cdot \mathbf{B}^{-1} \cdot \mathbf{s}(\varphi)|^2}; \; \boldsymbol{s}_j = e^{ij\varphi} \tag{2.6}$$

where $e_0^T = (1, 0, \ldots, 0)^T$ and the measurement matrix $\mathbf{B}$ is given as

$$\mathbf{B} = \begin{pmatrix} b_0 & b_{-1} & \cdots & b_{-m} \\ b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \cdots \\ b_m & \cdots & b_1 & b_0 \end{pmatrix}$$

where $b_{-j} = \overline{b_j}$ denotes the complex conjugate of the measured trigonometric moment for frequency j. In the case of sinusoidal patterns projected onto a scene and observed with a camera, the phase φ corresponds to a position in the direction of the linesweep along the n-axis. Peters et al. showed, that a rough estimate for the trigonometric moment measurements b_j can be obtained by performing four measurements with equidistant phase shifts, to equally sample the real and imaginary part of the complex exponential in Eq. 2.4. In particular, we can thus describe a trigonometric moment measurement of a camera pixel (x, y) as

$$b_j(x, y) = \sum_k I_k(x, y) e^{ij\varphi_k};\ \varphi_k \in \left\{0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi\right\} \tag{2.7}$$

which is a superposition of the acquired pixel intensities $I_k(x, y)$ for each phase shift k for a specific frequency $\nu \cdot j$, forming the vector of trigonometric moments $\boldsymbol{b}$ per camera pixel.

Our main result is the estimation of a 3D tensor (a subset of ρ) using the aforementioned maximum entropy estimation technique: Per camera pixel we reconstruct a pixel response in dependence on the phase φ, the 3D tensor hence resembles a linesweep illumination and is decribed via

$$h(x, y, \varphi) = \frac{1}{2\pi} \frac{e_0^T \cdot \mathbf{B}^{-1}(x, y) \cdot e_0}{|e_0^T \cdot \mathbf{B}^{-1}(x, y) \cdot \mathbf{s}(\varphi)|^2};\ \varphi = 2\pi \cdot n/N \tag{2.8}$$

The reconstruction is computed independently per camera pixel and relies on the closed form Eq. 2.8, allowing for fast and parallelizable computation. In addition, our method does not require any posterior phase unwrapping step: Taking a closer look onto the denominator in Eq. 2.8 reveals that the phase dependence of the expression can be understood as a Fourier series with frequencies j. As long as the lowest frequency spans the full scene, which we take care of, this expression reconstructs the full 2π-periodic density with respect to $j = 1$. We have to be aware that the ability to decompose the light transport components with respect to the phase does not resolve the full global light transport: The projector pixels forming the line of illumination along the m-axis at each position n still all contribute, which cannot be separated with this technique, a problem common to most phase shifting techniques.

2.5 Measurement setup and procedure

Our setup consists of a 14-bit 1920 x 1200-pixel FLIR Grasshopper 3 camera and a Casio XJ-A142 projector with a native resolution of 1024 x 768. Prior to our measurements, we performed a radiometric and geometric calibration of the stereo setup. In particular we have to take care of the linearity of the projector output. To this end, we first calibrate the camera and then project a medium frequency sinusoidal pattern ($j = 4$) onto a flat diffuse target. We acquire images for 64 equally spaced phase shifts of this pattern and compute the Fourier transform of the center camera pixel along the phase axis. With proper calibration, we acquire a contrast of 2000:1 for the first to second frequency component of the signal. We chose an exposure time of 16 ms to match the refresh rate of the projector of about 60 Hz and average 20 frames for noise reduction, establishing measurements at about 3 FPS.

Our measurement procedure closely follows the standard phase shifting technique with sinusoidal illumination patterns: Per frequency $j \cdot \nu$ we project four shifted patterns with phase shifts $\delta\varphi_k = 2\pi\frac{k}{K}$ with $k = [0 \dots K-1]$, $K = 4$ onto the scene. We choose the base

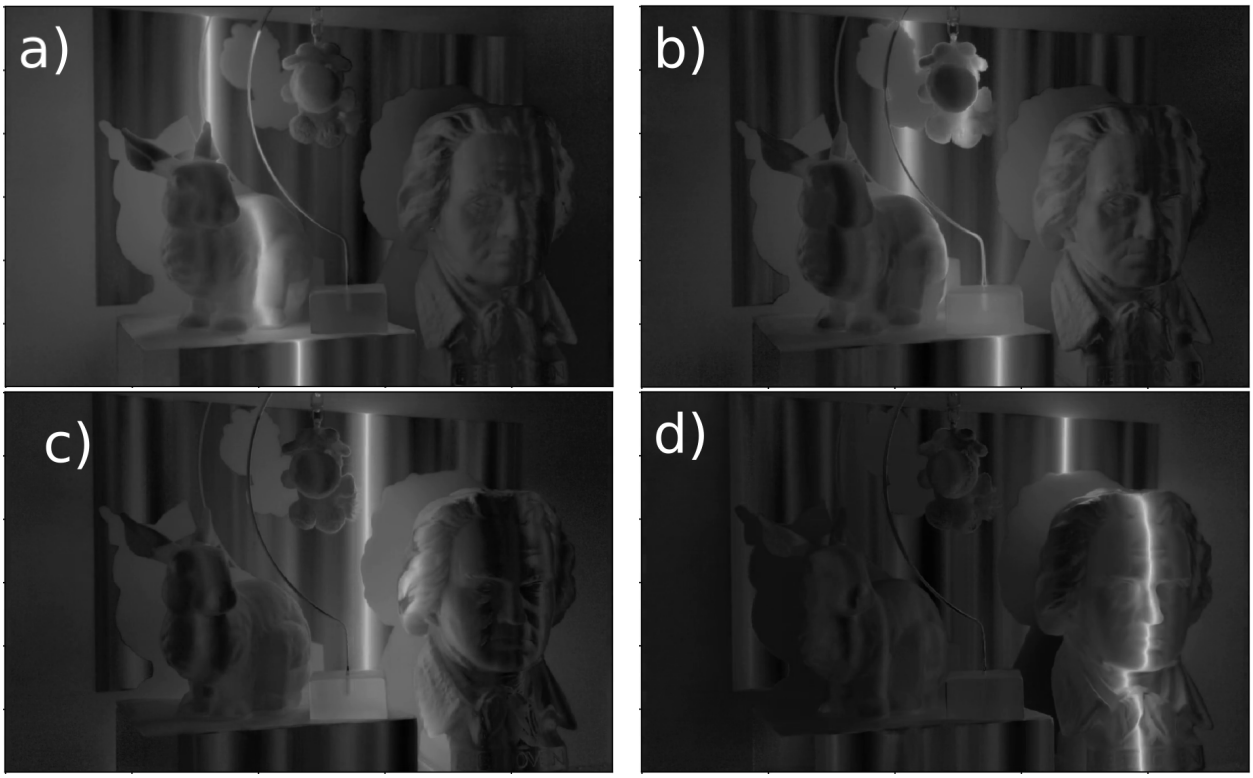

Figure 2.2: Reconstructed linesweep illumination based on the density estimation described in Sec. 2.3, computed from a total of 20 images, measuring 5 frequencies with 4 phases each. Note that for visibility we chose a logscale representation.

frequency $\nu = 1$, so that the period of the lowest frequency exactly spans the projector image.

Establishing point correspondences With our closed-form reconstruction technique at hand we are now able to estimate the pixel response of each camera pixel with respect to a linesweep illumination. In Fig. 2.2 we show four reconstructed example images for different positions of the (virtual) linesweep, videos of the reconstructed sweep are available in the supplemental material. As we inherently know the projector pixel corresponding to each line position, we in principle could start with a standard triangulation procedure at this point. Fig. 2.3 (right) shows an exemplary phase map where we assumed that the line position corresponds to the global maximum of the pixel responses. Here, two problems common to structured light phase shifting become apparent: First, shadows cannot be directly lit and therefore introduce errors and second, objects that exhibit either complex geometric or material properties do not necessarily have a dominant direct reflection. Taking a closer look at the reconstructed responses in Fig. 2.3 (left) reveals that diffuse materials, such as the fore- and background objects in this case indeed have a dominant direct reflection. The translucent object however shows a more complex response where the global maximum matches with the background reflection, introducing a wrong depth estimation at all such pixels.

To avoid such problems we utilize the benefit of having the full pixel response information at hand: We assume that direct reflections are not necessarily the global maximum of the

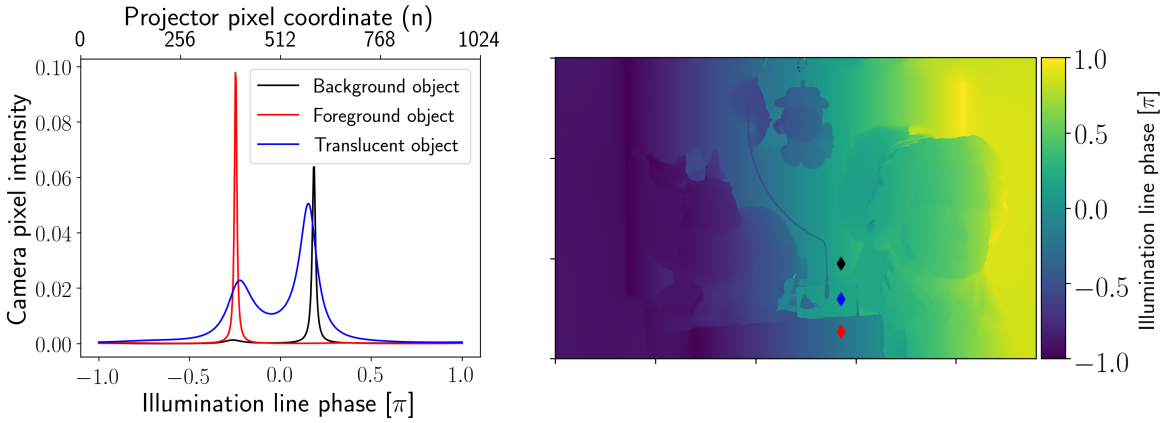

Figure 2.3: **Left**: Pixel responses of pixels observing scene points with different objects/materials, corresponding to the color-coded pixels in the right panel. **Right**: Phase map of the scene depicted in Fig. 2.1. The estimated phase per pixel corresponds to the global maximum of each camera pixel response.

pixel response but instead a direct reflection is only present when the global maximum is much stronger than the next strongest local maximum. The reasoning behind this is that a direct reflection can only deliver the highest intensity if the light is not distributed within the scene by additional scattering effects, which would result in strong secondary local maxima. It is then easy to provide a confidence map, which we compute as the ratio between the two highest maxima found per pixel. Critical points of the density $h(\varphi)$ have to fulfill the polynomial equation [PKH+15]

$$\sum_{j=0}^{J}\sum_{l=0}^{J}\overline{(\mathbf{B}^{-1}\cdot e_0)_j}\cdot(\mathbf{B}^{-1}\cdot e_0)_l\cdot(j-l)\cdot z^{J+j-l}=0 \tag{2.9}$$

where $z=\exp(i\varphi)$ and the polynomial is of degree $2\cdot J$. It therefore can have up to $2\cdot J$ roots than can be computed directly from the measured trigonometric moments, without the need to actually calculate the pixel response. For our measurement setup we thus obtain eight maxima, their respective position and strength, and sort these in descending order. We then compute our confidence as

$$C(x,y)=\frac{\hat{I}_0(x,y)}{\hat{I}_1(x,y)} \tag{2.10}$$

where $\hat{I}_0$ denotes the strongest and $\hat{I}_1$ the next strongest maximum. The confidence then tells us how likely it is that a pixel contains only a direct reflection. Note that this scale is not linear and is only bounded towards lower values; a perfect direct-reflection-only pixel would show a confidence of infinity whereas the lowest confidence is reached when the two peaks have equal strength, yielding $C=1$. We found that a confidence of $C>5$ is a safe albeit conservative threshold to identify camera pixels that contain a dominant direct reflection. Fig. 2.5 (top left) shows a mask created from such *direct pixels*.

The zeroth moment The zeroth moment is defined by

$$b_0 = \int_0^{2\pi} h(\varphi)\, d\varphi \tag{2.11}$$

and hence captures the absolute brightness the scene achieves due to the active illumination without any modulation of the pattern present ($j = 0$). Related work in the structured light community only utilizes measurements with frequencies larger than zero to not capture more images than necessary. In fact however, the zeroth moment contains important information in general, but especially for our reconstruction technique: The zeroth moment controls the sparsity of the reconstruction [PKH+15]. In addition, we use the zeroth moment to mask unreliable data that is produced by shadows during the capture. Assuming that a pixel that is shadowed does not receive direct light for any frequency, we average the absolute values of all moment measurements per pixel and apply a standard thresholding scheme to find shadows. Considering that indirect illumination can still reach a shadowed scene point we found that a reliable threshold to identify shadows is 2% of the maximum signal available in the so-formed image. The result is a mask that neglects all such *shadow pixels*, as for example in Fig. 2.5 (bottom left). Note that especially translucent materials are not detected in the shadow mask but are found to contain complex global light transport and hence are denoted as non-direct pixels. Vice versa, direct pixels may be found in shadow regions due to interreflections, but are removed by the shadow mask.

Correcting errors Camera pixels that are neither direct nor shadow pixels exhibit multiple (mostly two in our scenes) strong maxima in the pixel response (see Fig. 2.3, left). Instead of directly neglecting this data, our method allows for a data-driven semi-automatic editing. In particular, a standard processing procedure of the acquired data works like the following:

1. We reconstruct the per-pixel scene response according to Eq. 2.8.

2. We compute the 3D position of each scene point as seen by a camera pixel using a standard triangulation technique [HZ03].

3. The pixels with high enough confidence are marked as *direct pixels*.

4. The pixels with low enough intensity are marked as *shadow pixels*.

5. For direct pixels, the global maximum encodes the pixel correspondence and we choose the triangulated 3D position accordingly.

6. A semi-automatic editing step can be used to refine the reconstruction based on three available procedures and the available scene response. We manually select a region of interest containing the pixels we want to refine. We then perform one of the following actions:

   Background selection Choose the local maximum that corresponds to the further away 3D position.

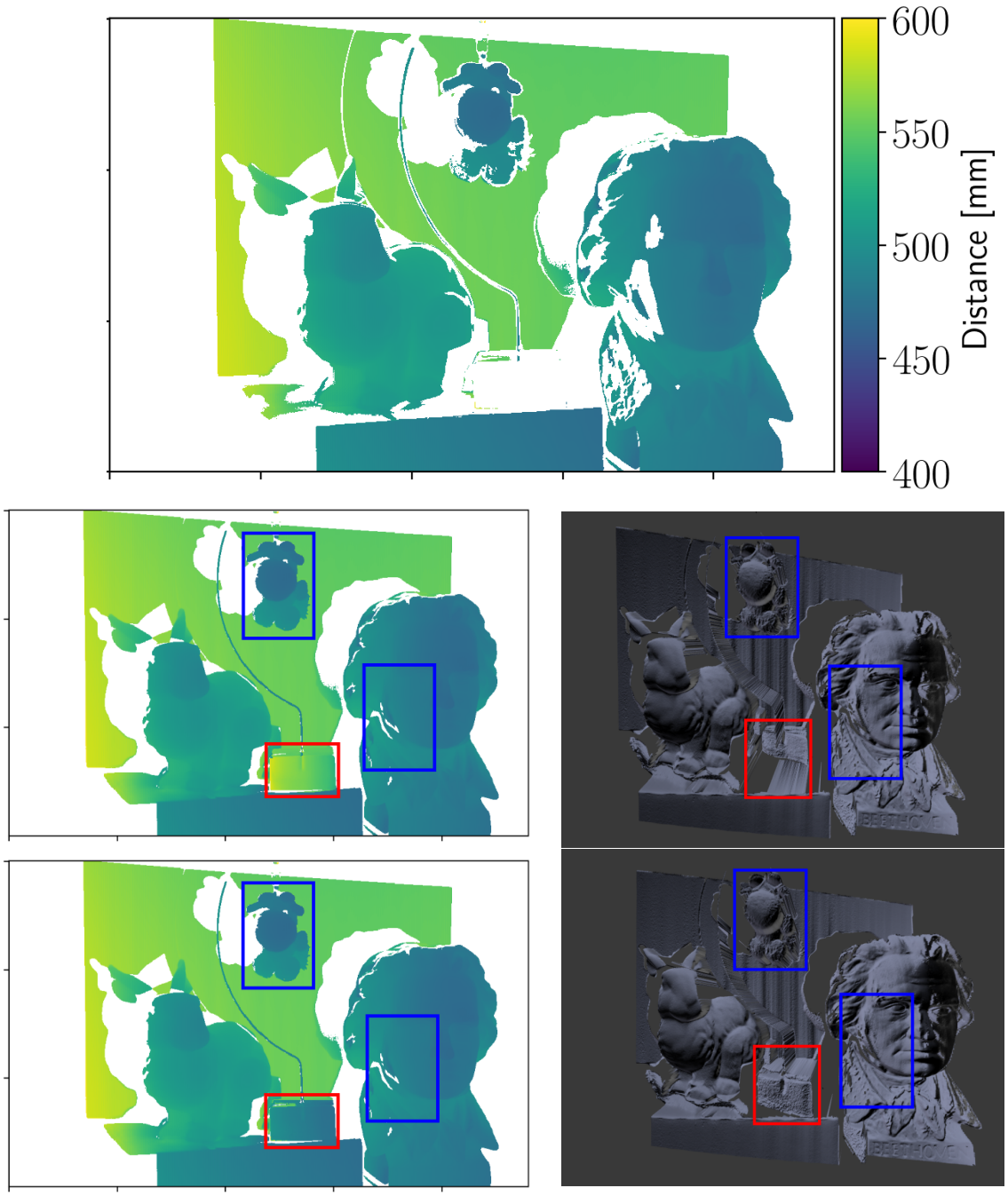

Figure 2.4: Depth maps and renderings of the 3D point clouds obtained from our data-driven semi-automatic editing step. Starting from our direct pixels with reliable depth information (**top**), we can apply our correction scheme and adjust uncertain data with respect to adjacent direct pixels. The red boxes denote a background (**bottom left**) and foreground ((**bottom right**)) selection for the correction process. The pixels within the blue boxes result from our smoothing approach. Note that the foreground correction works better due to the geometry of the problem, the background correction suffers from the dependency on interreflections.

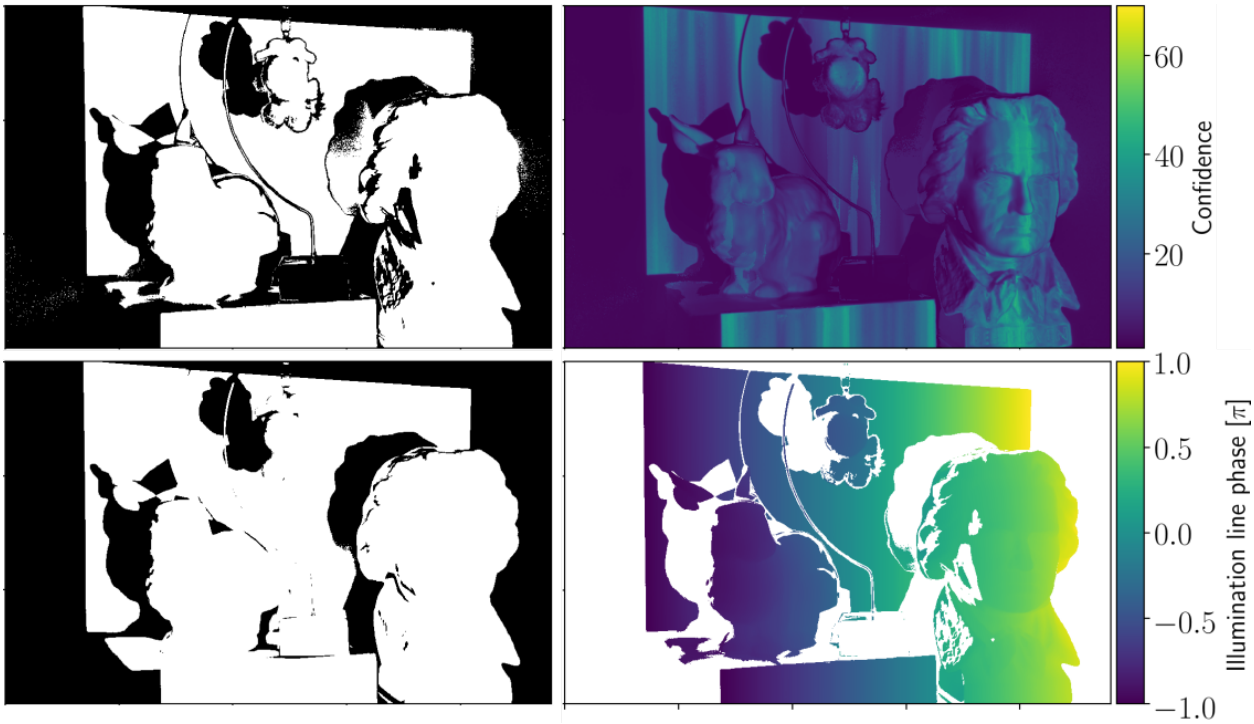

Figure 2.5: **Left**: The direct mask only containing pixels with $C > 5$ (top) and the shadow mask which removes pixels with too little information present (bottom). In both cases, white denotes a retained pixel. **Right**: The confidence map (top) denotes the ratio between the two highest maxima found in the pixel responses. Based on this information we remove non-reliable pixels from the phase map and only obtain direct pixels (bottom).

Foreground selection Choose the local maximum corresponding to the closer 3D position.

Smooth flood fill Within a chosen region, iteratively find pixels that are adjacent to direct pixels. We then choose the maximum that is closest to the (global) maximum of the direct pixel and set the now edited pixel to be recognized as direct as well for the next iteration. This performs similar to a flood fill procedure [Tor16].

The fore- and background selection share the benefit that they are independent on the absolute height of the maxima: The driving parameter is the distance obtained via triangulation with respect to each single maximum, the selection only uses the location of the maxima for correspondences camera-projector pixel correspondences.

Note that this refinement procedure is only possible because we estimated the pixel response in the first place and only works in the semi-automatic fashion presented. A fully automated procedure would require a more elaborate classification of pixel responses. Additionally, a well calibrated camera-projector setup is required to obtain reliable distance information for each maximum. A showcase video of the edit process can be found in the supplemental material.

2.6 Results

In this section we present results based on measurements using 5 frequencies ($j \in [0 \dots 4]$) with 4 phases per frequency, yielding a total of 20 images. The reconstruction itself takes under one minute for images with a resolution of 600x960 pixels on an Intel Core i7 processor. Due to the semi-automatic interactive refinement procedure, our method is not only able to acquire range information with a low number of acquisitions, but can also be used to acquire the shape of translucent materials, see Figure 2.4. Since the complete line sweep response is reconstructed at each camera pixel, we can selectively choose to estimate the shape of translucent objects by choosing the local intensity maximum corresponding to the depth value closest to the camera instead of the global intensity maximum. Light that passes through the translucent object and hits another scene point generates a second response (cf. Fig. 2.3). By choosing this response we are able to remove the translucent object from the depth estimation., although the reconstructed depth value is biased due to unaccounted refraction events at the translucent surfaces and its accuracy is limited because of interreflections. To fix this it would be necessary to ray trace through the translucent object by sampling its full light transport tensor, which is unknown. Also, because they originate from specular reflection, some points on the translucent object's surface have a very high confidence value and are not touched by our depth map editing operator (see the direct pixel mask in Fig. 2.5). The foreground correction tool on the other hand (cf. Fig. 2.4, bottom row) is able to reconstruct the translucent object, as the geometry of the scattering process is much simpler. In addition, we perform a smoothing correction on parts of the scene, denoted by the areas encoded in blue. In areas with a low confidence value, the refinement step estimates the most plausible candidate peaks by comparison with neighboring pixels of high confidence, see Fig. 2.5 (lower right). By applying this smoothness prior, we are able to fill in previously unreliable depth measurements.

Comparison We compare our method to the polarization difference imaging approach by Chen et al. [CLF+07]. Similar to [CSL08], their method is designed to reconstruct translucent materials, however without limiting its capabilities with respect to diffuse materials and scenes without global light transport, which is a limitation of the modulation-based method. Their method requires a total of 120 images and delivers accurate reconstructions of translucent materials as shown in Fig. 2.6. For scenes with little global light transport (bottom), both methods perform equally well, yielding reliable depth estimates for the crystal lamp. Minor differences occur due to the different masking of direct pixels. Strong differences arise for more complex scenes. To understand this effect, we have to consider the confidence map and corresponding direct mask of the complex scene in Fig. 2.7. Chen et al compute the confidence as the ratio between high-frequency and low-frequency pixel intensity, which is a common approach since low-frequency illumination patterns contribute more to the global component. This makes it difficult to fine tune the threshold needed to separate direct and global components. Here, we used a threshold of 0.5 to achieve a good balance between retaining the translucent object and masking out incorrect measurements. Still, we obtain false positive results due to interreflections and problematic scene geometry. We have found our confidence map generation more reliable with less false positives in these cases, since the ratio of largest local maximum intensity values yields a direct measurement for the complexity of the light transport at a given camera pixel. The data-driven correction approach then can be used to refine unreliable range estimates without having to hallucinate new data, filling in data where possible and otherwise removing pixels strongly influenced by global light transport.

Quantitative results To obtain a measure of accuracy for our method in comparison to PDI, we chose to estimate distances for both, large- and small-scale situations. The former is realized by placing a planar target in front of the system and measuring the corresponding distance, whereas the latter relies on the acquisition of a sphere in front of this plane, as shown in Fig. 2.8 (left panel). For the planar target, we perform a pose estimation to obtain reference data (tangent normal and position). For the sphere, we measure the physical object's diameter and fit its position using a least squares optimization routine. We then compute the difference between the reference data and the distance estimation obtained with both methods. Note that we base our computations only on the direct pixels (cf. Sec. 2.5) as we consider the others to only hold invalid information. Figure 2.8 reveals that for large scales such as the planar target, our method suffers from the low frequencies used: Over the depicted range of roughly 30 cm we obtain a RMSE of about 18 mm. The error obtained for PDI is only slightly less, however our method also shows a low-frequency ripple-like structure that is not found in the results of the method by Chen et al., which will be discussed in the next paragraph. In contrast, our method provides comparable measurements and masks out invalid (non-direct) pixels very reliably, leading to much less artifacts. These artifacts are also the reason that the obtained RMSE for PDI is much worse than for our method; Removing these artifacts yields very similar errors.

Ringing artifacts and higher frequencies For the results presented in this section, we are working at the lower limit of phase shifts needed to correctly sample the modulated

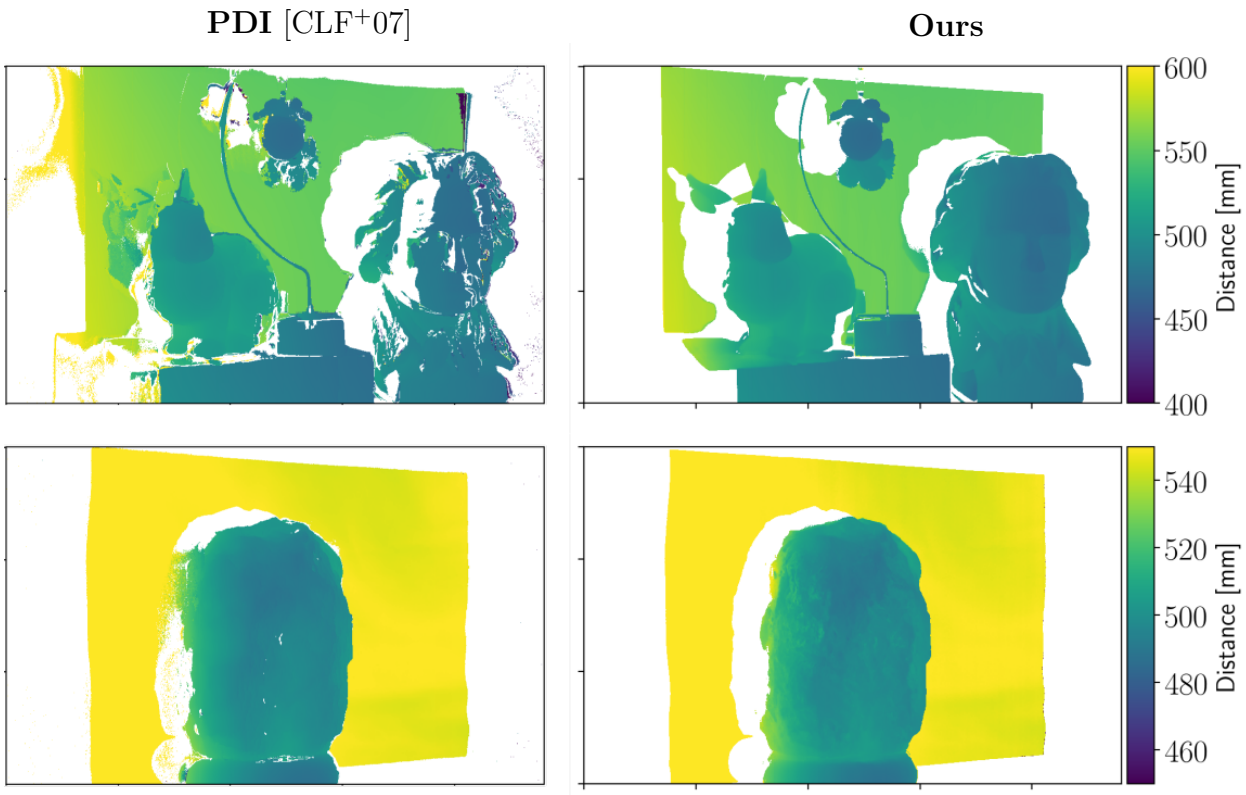

Figure 2.6: Comparison with the PDI approach by Chen et al. [CLF+07] for our most complex scene (top) and a crystal lamp (bottom). Both, the PDI method and ours truthfully reconstruct the foreground translucent object but show severe differences for scenes with strong global light transport. Additional comparisons can be found in the supplemental material.

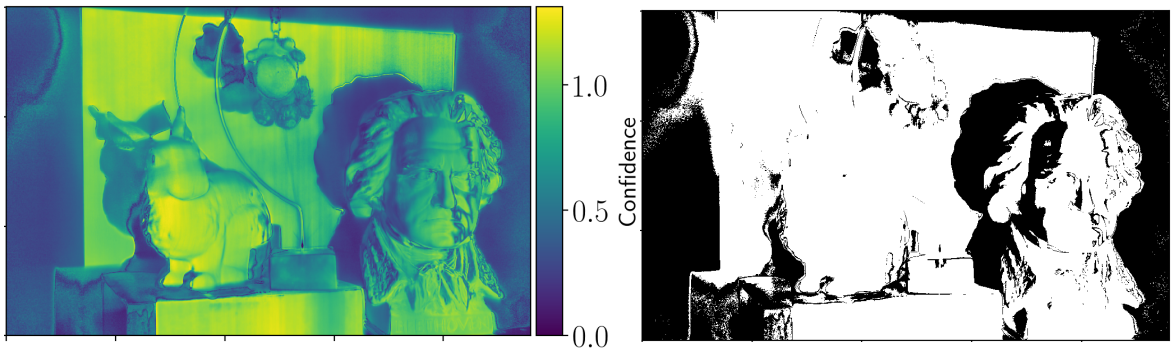

Figure 2.7: Confidence map (**left**) and direct pixel mask (**right**) for the PDI measurement. The mask is obtained by thresholding the confidence at 0.5. A lower threshold would yield more reliable results but would exclude the translucent object.

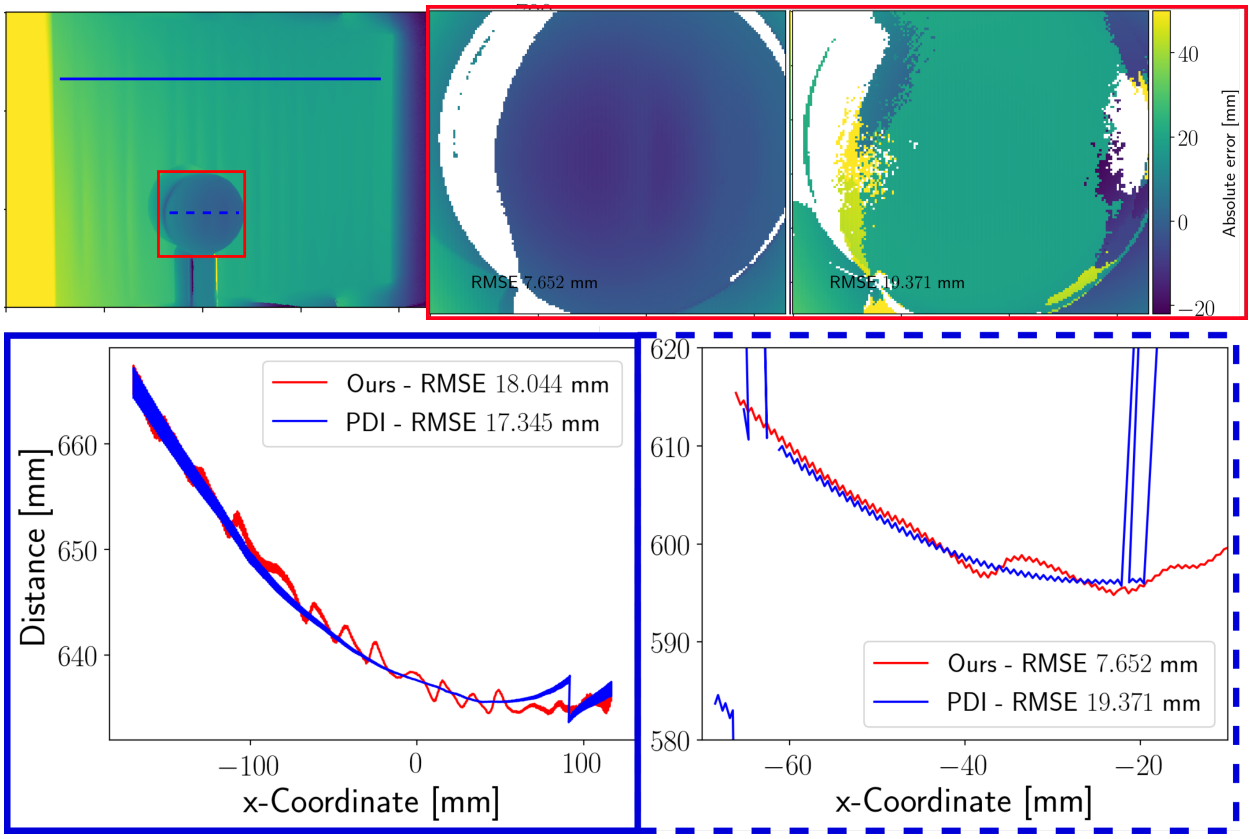

Figure 2.8: Quantitative comparison between PDI [CLF+07] and our approach. We reconstruct distance on two scales, a planar target covering the full field-of-view and a sphere. The reconstructions are performed based on the direct pixels only. **Top left**: Depth map obtained with our approach. The red rectangle contains the sphere used for the local analysis (middle panel), the blue lines correspond to the slices shown in the right panel. **Top right**: Absolute error obtained for the sphere target. Our method performs better on small scales, which can be attributed to the more robust choice of direct pixels (19 mm vs 8 mm). **Bottom**: Reconstructed distances corresponding to the blue lines. In contrast to the PDI, our method suffers from the low frequencies used, resulting in a higher RMSE for large scales (18 mm vs 17 mm). On the other hand, our method more reliably masks out invalid pixels, resulting in less artifacts. The planar target here appears curved due to the distance measurement with respect to the camera.

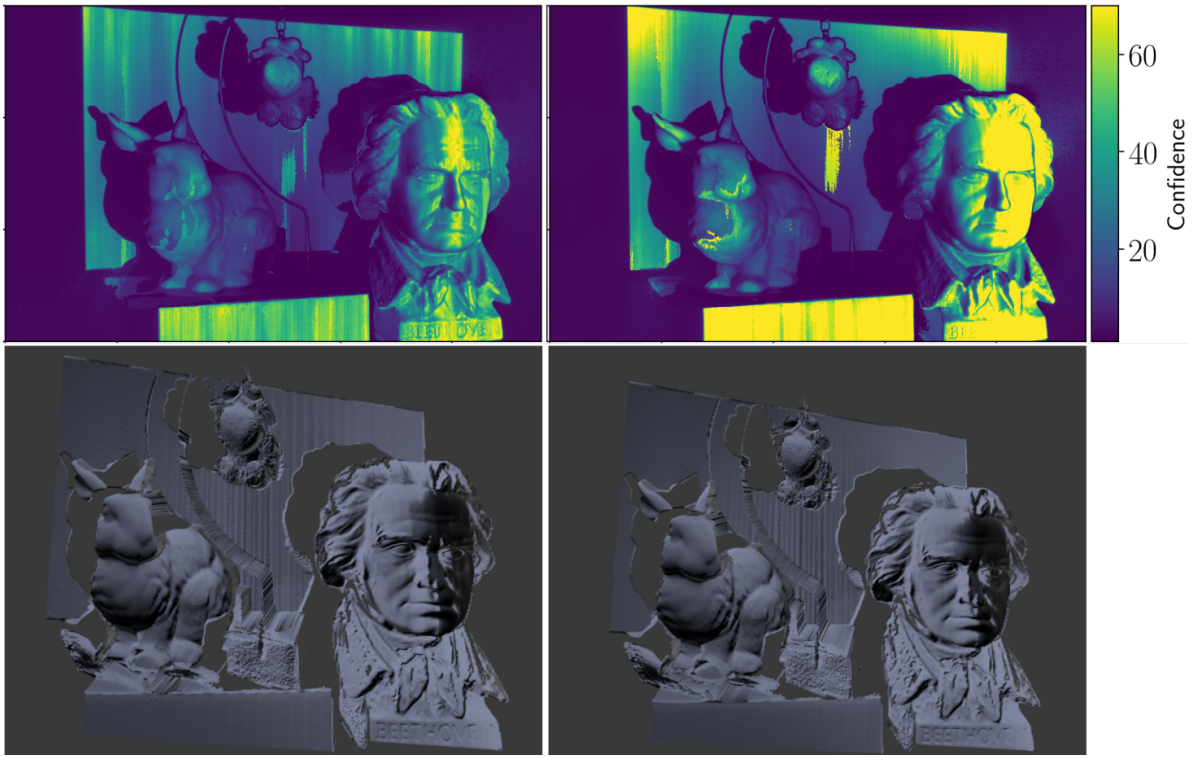

Figure 2.9: Increasing the number of phases while fixing the number of frequencies results in higher confidence (top) and less artifacts in the reconstructions (bottom). **Left**: Measurement with 8 equally distributed phase shifts (40 images). **Right**: Measurement with 16 phaseshifts (80 images).

illumination. This comes at the cost of ringing artifacts. Taking a closer look at the point clouds in Figures 2.10 and 2.4 such artifacts are visible especially on the far wall, and are also indicated in the corresponding confidence maps and the log-scale linesweep images and videos, where they manifest as local maxima and minima alongside the scanline. To correct this, measurements with higher frequencies and/or more phases per frequency can be acquired. Both cases linearly increase the amount of acquisitions, as our reconstruction relies on subsequent integer multiples of the base frequency and for each frequency an equal amount of phase shifts is performed. Figure 2.9 shows the confidence maps and reconstructed 3D point clouds for 8 and 16 phases. For a measurement with 8 phase shifts, the confidence map quality increases significantly, making a direct-global separation more clear. Ringing artifacts are reduced compared to the 4-phase measurement but still present, which is well visible within the confidence map and reconstruction. With 16 phase shifts, the reconstruction does not differ noticeably from the previous measurement, ringing artifacts are only slightly reduced, hinting at the need for higher frequencies to mitigate this effect.

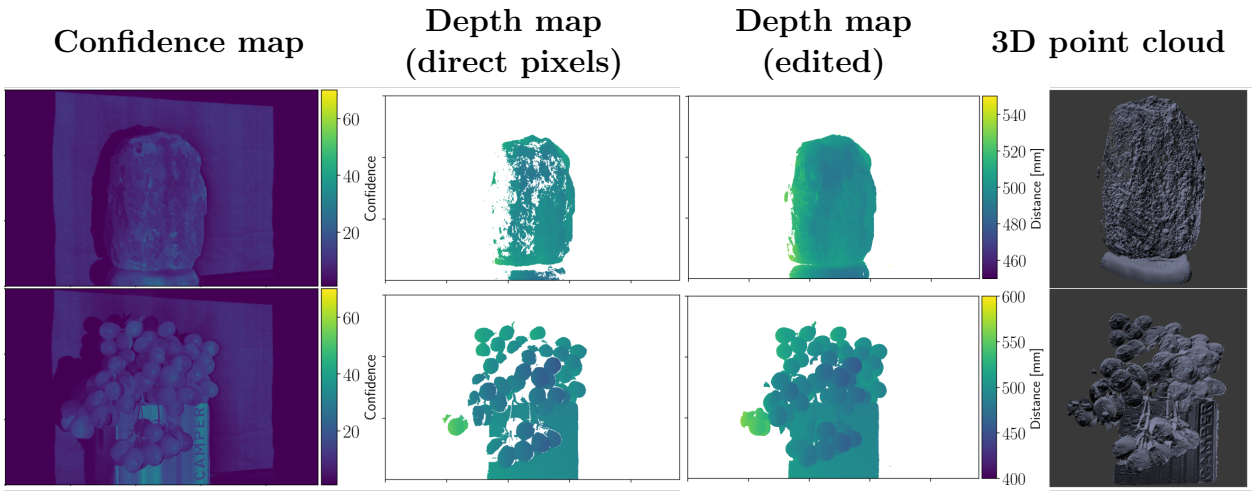

Figure 2.10: After data acquisition, we reconstruct the scene response of each camera pixel and extract the local extrema. The ratio of the two highest local maxima gives the confidence map (**first column**), which is a lower bounded measure of how likely the pixel only contains a direct reflection. We then compute the 3D positions of the direct pixels via triangulation (**second column**) which yields a reliable depth estimate containing holes. Using our correction steps, we can "fill in" missing values at uncertain non-direct pixels (as long as they are not identified as shadow pixels) and provide a more complete depth estimation without hallucinating new data or interpolation (**third column**). We thereby only rely on the actually measured scene responses. With these datasets we are then able to provide a 3D point cloud for e.g. meshing purposes (**fourth column**). Additional results can be found in the supplemental material.

2.7 Conclusions

We presented a frequency-domain approach that utilizes a closed-form spectral estimation to reconstruct the reflectance field per camera pixel as if illuminated by a linesweep. Linesweep illumination has the advantage of separating contributions along the sweep direction, rendering it useful for the reduction of global-illumination contributions. Naïvely, such a measurement scheme would require an amount of acquisitions equal to the number of projector pixels along the sweep direction. In contrast, we can reduce the number of acquisitions to only 20 for a reliable reconstruction and confidence estimation and, due to the maximum entropy constraint of the reconstruction, reduce noise drastically.

The method performs well for scenes with and without global light transport contributions by translucent objects or interreflections. The latter can efficiently be masked out using our confidence measure whereas the former can be analyzed and edited. Based on the pixel response, we can then choose to either render a translucent object virtually invisible or to reconstruct its shape.

Currently, our method relies on acquisitions with integer multiples of a base frequency and an equal number of phase shifts per frequency, which introduces ringing artifacts due to undersampling. To mitigate this effect, measurements with higher frequencies are required, which for our method would drastically increase the amount of acquisitions. Hence, one major point for future work would be the incorporation of a non-equal number of phase shifts per frequency as well as multiple base frequencies. Still, even with as little as 5 frequencies and 4 phase shifts, corresponding to a total of 20 images captured, our method reliably separates direct from global components, which is especially useful for scenes with global light transport.

Acknowledgements This work was supported by the Computational Imaging Group from the Stuttgart Technology Center of Sony Europe B.V., the German Research Foundation (HU 2273/2-1) and by the ERC Starting Grant ECHO. We thank Christoph Peters for many helpful discussions and suggestions.

Appendix

In the following we present additional results that accompany our paper and were left out due to space constraints or to improve readability. In addition to the images provided in this document, we supply videos with the following content:

1. One video per scene showing the reconstructed linesweep with a slightly upscaled brightness for better visibility.
2. One video per scene showing the reconstructed linesweep in logscale.
3. A video showcasing the editing of the complex scene, containing multiple objects including the translucent plastic block.

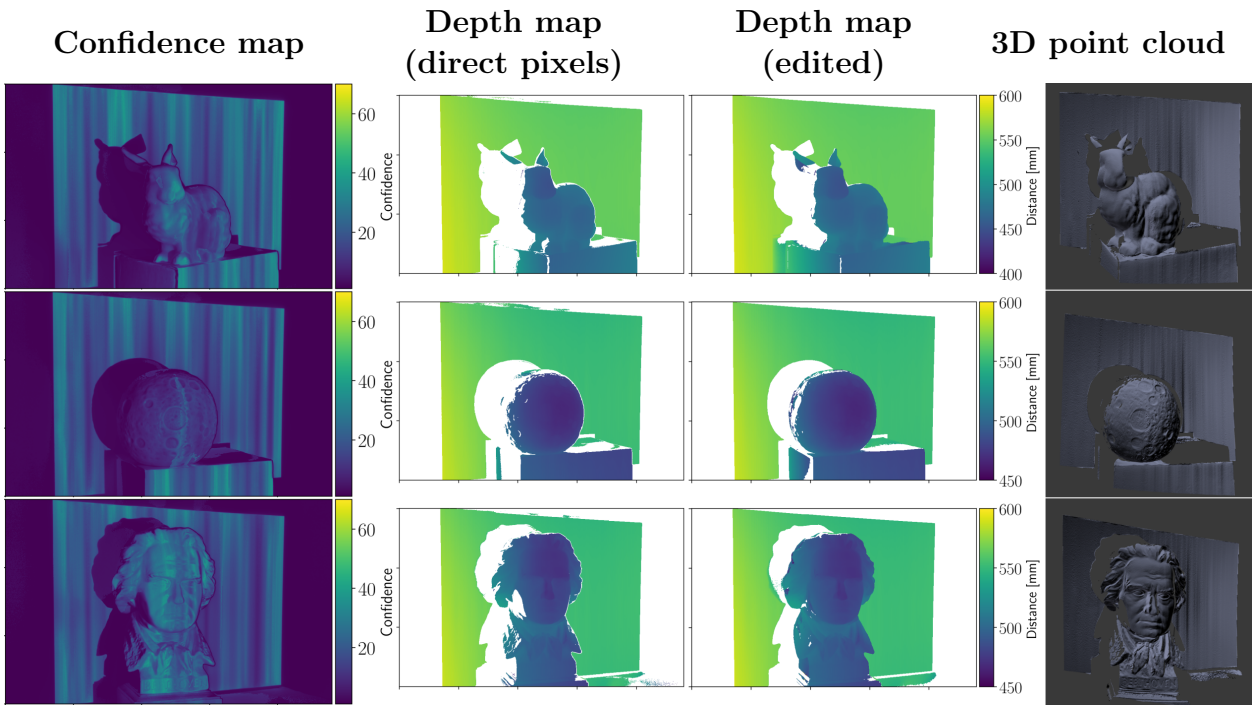

Figure 2.11: Confidence maps, depthmaps and final reconstructions obtained for three additional scenes acquired with our measurement setup. The measurements were performed with 5 frequencies ($j = [0 \dots 4]$ and 4 equally spaced phase shifts per frequency. This completes the results shown in Figure 2.10 in the paper.)

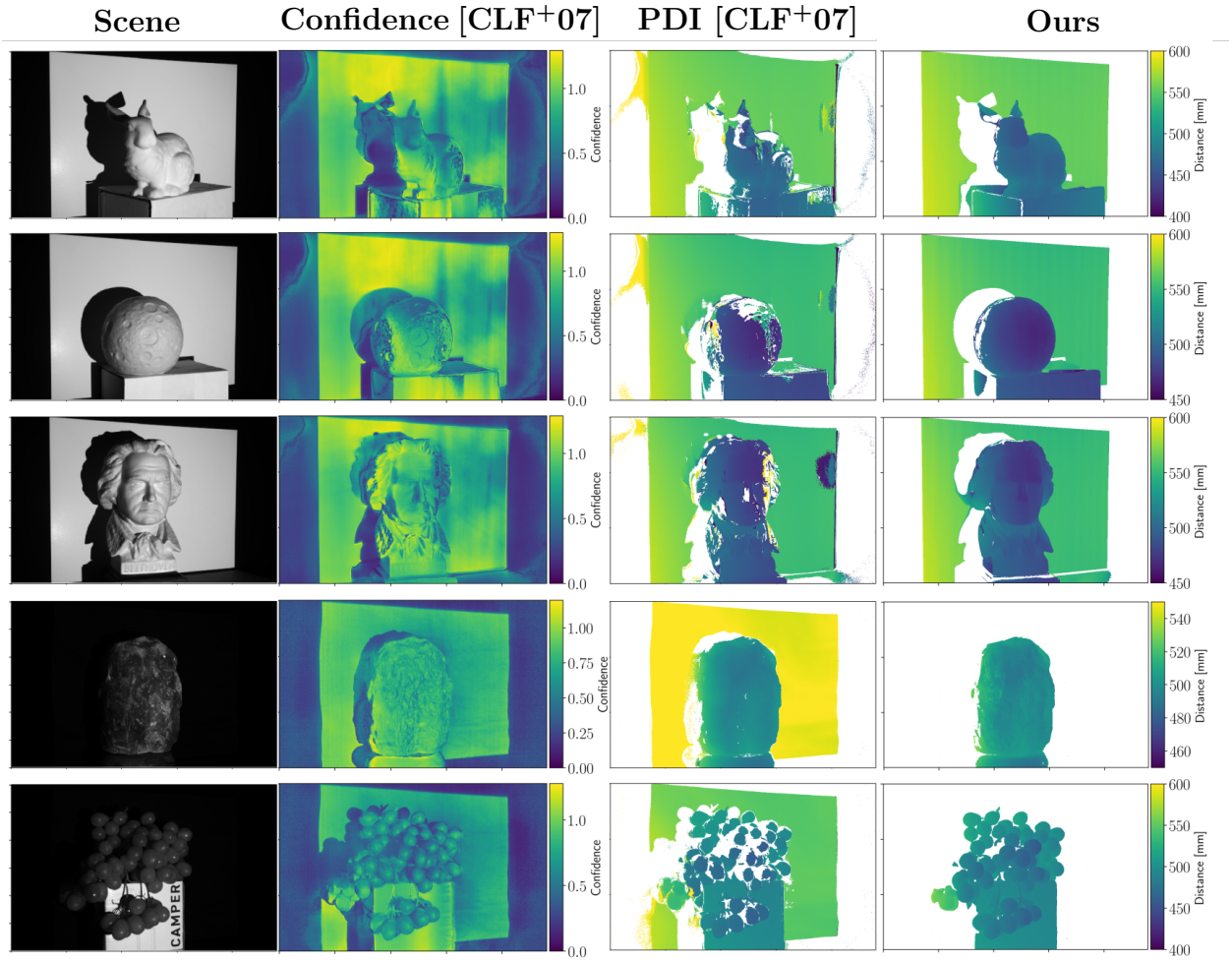

Figure 2.12: Comparison between the polarization difference imaging (PDI) approach by Chen et al [CLF+07] with our results for the scenes not presented in the paper. From left to right, we present the original scene image, the confidence computed following Eq.5 in [CLF+07], the final result obtained with the PDI method (and masking out "unreliable" pixels) and our final result after masking and editing.

Preface – Pulsed Correlation Time-of-Flight

Time-of-Flight (ToF) range finding setups are commonly employed for 3D shape acquisition. In contrast to structured light methods, the acquisition times of such devices is fast enough to not require perfectly static scenes which makes them especially useful for mobile applications and handheld devices, where the user has to manually "scan" the object of interest from multiple view points. To avoid the need for costly high-speed electronics, current off-the-shelf ToF solutions rely on indirect measurements: Amplitude-modulated light is sent into the scene, usually with sinusoidal or rectangular modulation. Upon scene traversal, the signal undergoes a phase shift that is related to the distance traveled. Correlation with a demodulation signal at the sensor is used to reconstruct this phase, and and hence distance, at the sensor, directly in hardware. Such sensors employ a specialized type of pixels, which limits the available spatial resolution. The depth resolution on the other hand is primarily defined by the choice of modulation/demodulation signal pairs and design choices such as time- and power budget. Available methods often suffer strongly from measurement noise, especially in scenes with low illumination intensity. This renders the results prone to artifacts and errors in the resulting depth estimations, ultimately leading to low-quality shapes that impact the quality of an associated digital double.

We approach this problem by enhancing available off-the-shelf solutions with a new mode of operation, pulsed correlation Time-of-Flight imaging. Here, we combine two low-power measurements: First, we perform a rough depth estimation with the standard C-ToF approach already implemented in the hardware by default. With this coarse notion of depths associated with visible objects within the observed scene, we employ a depth-focusing step in combination with a high-sensitivity acquisition that is achieved by replacing the modulation/demodulation signal pair by a pulsed illumination and a rectangular demodulation.

With this, we are able to significantly reduce the noise in extracted depth maps and even allow the acquisition of shapes under circumstances that would lead to gaps and strong errors in standard operation mode. Our method hence mimics a feature long used for conventional cameras: The ability to "zoom in and out", thereby selecting a depth of interest. We employ this method without the need for additional hardware, which makes it especially useful for mobile applications on a tight power budget, allowing to acquire shape for crafting digital doubles with increased precision.

CHAPTER 3

Pulsed Correlation Time-of-Flight

Abstract We introduce pulsed correlation Time-of-Flight (PC-ToF) sensing, a new operation mode for correlation time-of-flight range sensors that combines a sub-nanosecond laser pulse source with a rectangular demodulation at the sensor side. In contrast to previous work, our proposed measurement scheme attempts not to optimize depth accuracy over the full measurement range: With PC-ToF we trade the global sensitivity of a standard C-ToF setup for measurements with strongly localized high sensitivity – we greatly enhance the depth resolution for the acquisition of scene features around a desired depth of interest. Using real-world experiments, we show that our technique is capable of achieving depth resolutions down to 2 mm under optimal conditions using a modulation frequency as low as 10 MHz and an optical power as low as 0.6 mW. In general, we are able to halve the depth RMS error in comparison to standard correlation ToF measurements at equal optical power. This makes PC-ToF especially viable for low-power applications.

3.1 Introduction

Time-of-flight (ToF) range finding setups support a vast amount of applications, ranging from robotics closely tied with exploration and automated manufacturing to motion capture and 3D mapping, as well as biometrics [KBK+09]. They are all connected by the common need for truthful representations of the three-dimensional environment. As a consequence, all applications share the desire for both – high spatial resolution as well as precise depth estimation. Thanks to advances in sensor technology, the former rises with every generation of sensors, whereas the depth resolution depends on design choices, such as the time and power budget of any such sensor and hence is fundamentally limited by noise. Especially for low *signal-to-noise ratio* (SNR) measurements, accurate detection of distances becomes a challenge that received a lot of attention from the scientific community. There exists a number of range finding approaches based on ToF measurements, which can be divided into two classes that differ in both hardware requirements and reconstruction techniques.

Direct pulsed time-of-flight range finding. Direct pulse-based ToF systems (P-ToF) [GD67; Koe68] were the first ToF systems to be employed for range finding purposes. These setups emit a single short (pico-/nanosecond) laser light pulse into the scene. The sensor then

receives a delayed pulse after a certain travel time. The time delay between emission and acquisition is directly proportional to the distance travelled and subsequently the depth of the scene. Recent pulse-based systems rely on the determination of the pulse-shape, altered by scene traversal [IY01; YIM07] and implementation of a fast image shutter in front of the sensor chip. The simplicity of the underlying concept comes at the cost of elevated hardware requirements, enabling the measurement of time delays in the order of picoseconds in low SNR scenarios. Due to these limitations, such systems often consist of a single-pixel sensor only and require time-consuming pixel-wise scanning of the scene. Despite those shortcomings, the strength of P-ToF systems lies in their high depth resolution.

Correlation time-of-flight range finding. To alleviate the need for fast and costly hardware, *amplitude-modulated continuous-wave* (AMCW) ToF systems have been developed that consist of temporally modulated light sources and sensors [SXH+97; LSB+00]. At the core of these *correlation time-of-flight* (C-ToF) setups lie the modulation (at the light source) and demodulation (at the sensor) functions, that are used to code and decode the illumination signal. Current C-ToF setups utilize sinusoidal or square coding functions. Upon scene traversal, the amplitude modulated illumination undergoes a phase shift with respect to the original signal emitted by the light source. This phase shift is proportional to the traveled distance and is acquired using a correlation measurement between the emitted and received signal. As modulation and demodulation functions are periodic, these measurements implicitly are limited to the so-called *unambiguity range*, which depends on the frequency of the modulation signal.

Depth resolution enhancements for C-ToF systems. In comparison to P-ToF systems, correlation-based ToF systems are more prone to errors from measurement noise [BS08], resulting in a considerably lower depth resolution. In recent years a great amount of research has been done to mitigate noise artefacts such as effects from higher harmonics of modulation-demodulation signal pairs [PDC+08; PDC+10] for C-ToF systems. Dual-frequency setups for example try to enhance depth resolution without the loss of unambiguous measurement range by combining high- and low-frequency measurements [JCP+10; JBP+11]. More recently, Gupta et al. [GVN+18] presented a framework for general C-ToF range finding, which allows for the simulation and computation of the depth resolution performance for arbitrary modulation-demodulation signal pairs. In addition, they also used their framework to develop an optimized Hamiltonian coding function, which achieves depth resolutions below 1 cm over the full ambiguity range. Closely tied to the work presented in this paper, Payne et al. [PDC11] discuss the optimal choice of the duty cycle for the chosen illumination signal for the special case of sinusoidal and square illumination modulation. They point out that a reduction of duty cycle results in an increased peak power and thus better SNR for the illumination signal. All these approaches share the desire for improvements of the depth sensitivity over the full unambiguity range, which are inherently deemed to result in a tradeoff due to their respective relations to the modulation frequency. This is due to the fact, that the limited bandwidth of the illumination signal(s) directly relates to the depth variations the procedure can truthfully distinguish. With *pulsed correlation time-of-flight sensing* (PC-ToF), we propose a dual-measurement scheme that explicitly makes use of this

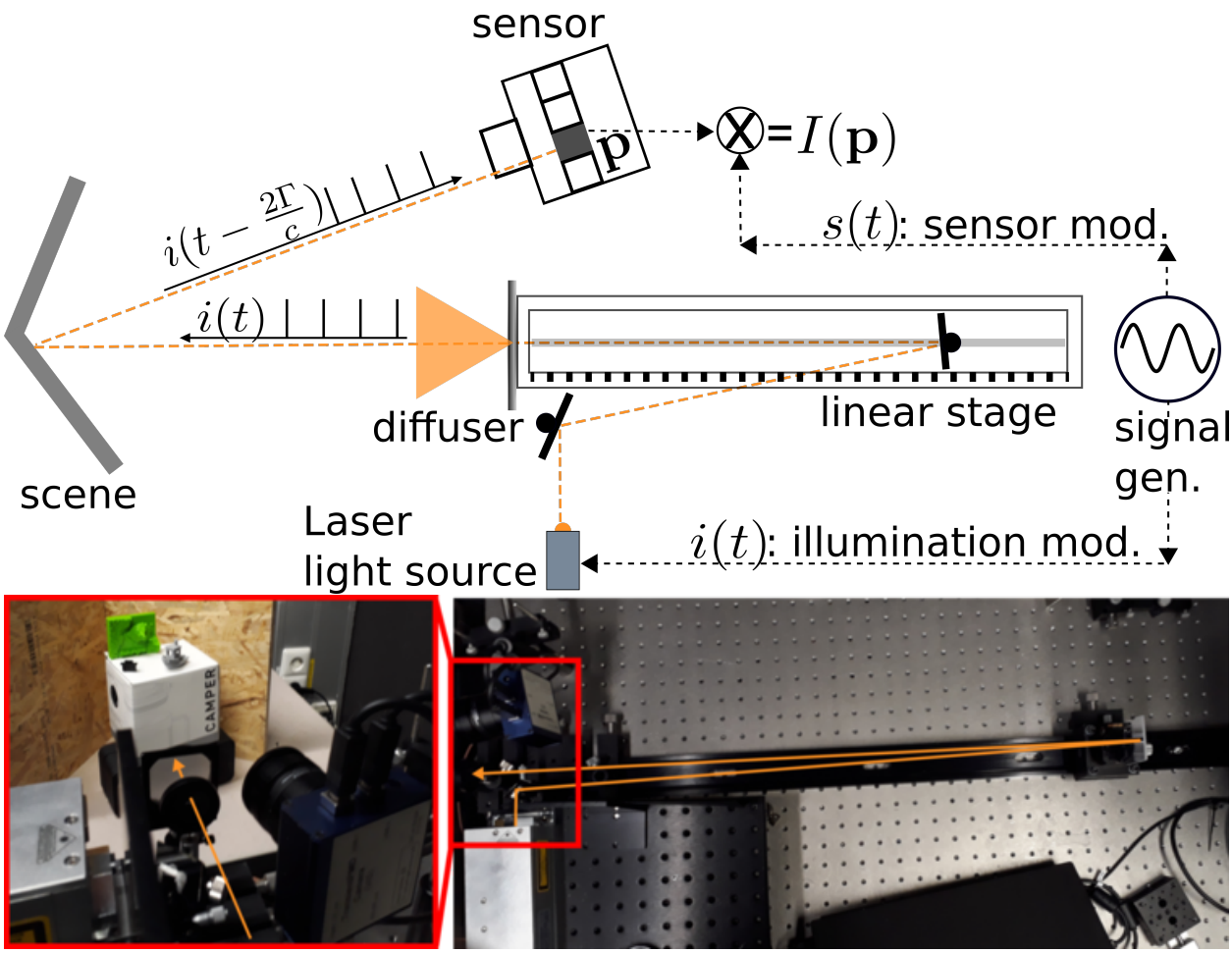

Figure 3.1: **Upper panel**: Schematic visualization of our correlation ToF setup. A signal generator drives both – modulation of the illumination and the sensor gain for the two modes of operation necessary for our PC-ToF acquisition. **Lower panel**: Pictures of our lab setup, as indicated in the upper panel.

information content. This novel hybrid approach combines the high depth resolution and noise resilience of P-ToF with the low-cost hardware of C-ToF setups at the cost of ambiguity: We replace the (continuous) modulation function with pulsed illumination but maintain a continuous demodulation signal. The two steps of a PC-ToF acquisition are:

1. Obtain a rough depth estimate with standard C-ToF range finding methods and select a *depth of interest* (DOI) for close inspection.

2. precisely measure the depth around a user-specified DOI with depth resolutions down to 2 mm, utilizing our hybrid approach.

The power consumption of C-ToF systems is dominated by optical output power and sensor modulation. Its light efficiency and the use of slow modulation frequencies make PC-ToF especially suitable for low power applications, for example in mobile hardware.

3.2 Correlation time-of-flight image formation

We will briefly revisit the image formation model for correlation ToF as described in [GNH+15; GVN+18] and, for simplicity, adapt their notation: Correlation time-of-flight setups (see Fig. 3.1) consist of an amplitude-modulated light source and a gain-modulated sensor, with modulation functions $i(t)$ and $s(t)$ respectively. Like [GVN+18; FSK+14; HHG+13; KWB+13], we assume the absence of any indirect or multi-bounce light, which allows us to describe the scene response as a single scattering event at the precise depth Γ. This results in a shift of the modulation function $i(t)$. In general, the *irradiance* $E(\mathbf{p}, t)$ that arrives at pixel $\mathbf{p}$ can then be described as

$$E(\mathbf{p}, t) = E_a(\mathbf{p}) + E_c(\mathbf{p})\, i(t - 2\Gamma/c) \tag{3.1}$$

where c is the speed of light, $E_a(\mathbf{p})$ denotes the ambient light component and $E_c(\mathbf{p})$ is the mean pixel irradiance due to the modulated light, encoding the optical properties of the scene. The shifted normalized illumination modulation is described by $i(t - 2\Gamma/c)$. The sensor then records the *pixel intensity*

$$\begin{aligned} I(\mathbf{p}) &= \int_0^\tau E(\mathbf{p}, t) s(t) \mathrm{d}t \\ &= I_a(\mathbf{p}) + E_c(\mathbf{p}) \int_0^\tau i\left(t - 2\Gamma(\mathbf{p})/c\right) s(t)\, \mathrm{d}t, \end{aligned} \tag{3.2}$$

where τ is the exposure time and $I_a(\mathbf{p}) = \int_0^\tau E_a(\mathbf{p}) s(t) \mathrm{d}t$ is the incident ambient light. Eq. 3.2 can be understood as a cross-correlation function. From this, we define the *normalized correlation function*

$$C(\Gamma) = \int_0^\tau i\left(t - 2\Gamma/c\right) s(t)\, \mathrm{d}t. \tag{3.3}$$

This way, we are able to express the full image formation process of correlation time-of-flight imaging via the *image formation equation*

$$I(\mathbf{p}) = E_c(\mathbf{p}) C(\Gamma) + I_a(\mathbf{p}). \tag{3.4}$$

This equation reveals three unknowns $E_c(\mathbf{p}), \Gamma, I_a(\mathbf{p})$ which have to be determined pixel-wise. We require $K \geq 3$ measurements or samples of the correlation function $C_k(\Gamma)$ for $k \in \{0, \ldots, K\}$. These measurements are commonly realized by inserting an additional phase shift θ_k into the demodulation function, such that

$$s(t) \to s_k\left(t + \theta_k/\omega\right);\ \theta_k \in [0, 2\pi) \tag{3.5}$$

and data acquisition is performed for K equally spaced phases. Gupta et al. [GVN+18] continue to develop a *depth precision measure* $\bar{\chi}_c$, which encodes the average depth accuracy depending on the average optical properties encoded in $E_{c,\text{mean}}$ as well as the noise standard deviation $\Omega = \sqrt{\sum_{i=1}^K \sigma_k^2}$ (assumed to be constant) for K-tap correlation ToF measurements as

$$\bar{\chi} = \frac{E_c}{\Omega \Gamma_{\text{range}}} \int_\Gamma \sqrt{\sum_k \left(\partial C_i(\Gamma)/\partial \Gamma\right)^2} d\Gamma, \tag{3.6}$$

where Γ_{range} is the unambiguous depth range.

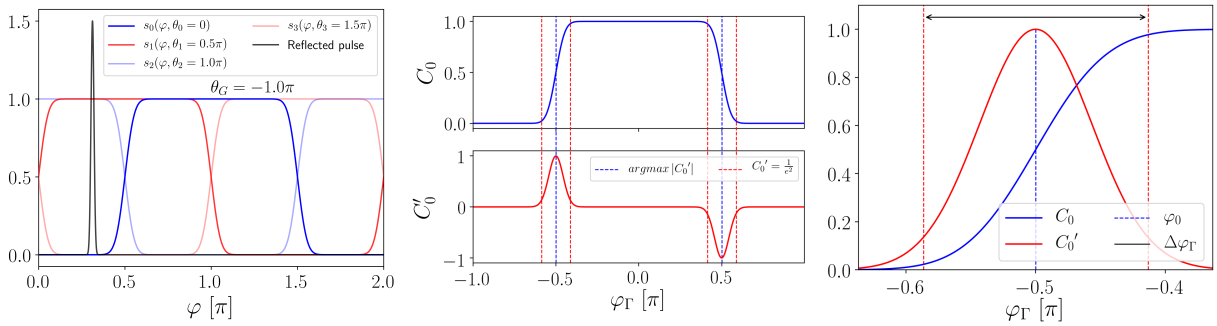

Figure 3.2: **Left**: Demodulation- and measured signals for our PC-ToF approach. Note that the modulation signal $i(\varphi)$ is rescaled for visualization. **Middle**: Correlation function $C_0(\varphi_\Gamma)$ and its first derivative $\partial C_0/\partial\varphi_\Gamma$ in dependence on the phase (and hence, depth). The *phase of maximum sensitivity* φ_0 is denoted by the dashed blue line as defined for $\theta_k = 0$. The *sensitive range*, here indicated as the region between two dashed red lines is related to the rise time of C_0. **Right**: Close-up of $C_0(\varphi_\Gamma)$ and C_0' around the first extremum φ_0.

3.3 A new operation mode for time-of-flight range finding

Our foremost aim is to increase the depth sensitivity not on a global scale (over the full ambiguity range) but locally. From Eq. 3.6 we directly see that depth sensitivity depends on the gradient of the (normalized) correlation signal. Ideally, $\partial C/\partial\Gamma \to \infty$ which would result in a vanishing rise time $T_{\text{Rise}} = t(\max(C)) - t(\min(C)) \to 0$. We base our considerations on the idealized case of a combination of pulse trains (Dirac comb) for our modulation signal and using a rectangular demodulation signal on the sensor side, both with frequency ν. To unify considerations and clarify, that we are limited to exactly one period of the modulation and demodulation signals before ambiguities arise, we switch the integration variable to phase φ via $\varphi = \omega t$; $\omega = 2\pi\nu$; $\varphi_\Gamma = \frac{2\omega\Gamma}{c}$.
We describe our (real) modulation i and demodulation s signals as a chain of Gaussian pulses and a smoothed rectangular signal chain. This is achieved by the convolution $\circledast$ of a Dirac comb and a square signal with a Gaussian $\mathcal{G}$ with standard deviations σ_M and σ_D respectively.

$$
\begin{aligned}
i(\varphi) &= \sum_n \left[\delta(\varphi - n\,2\pi - \varphi_\Gamma) \circledast \mathcal{G}(\varphi, \sigma_M)\right], \\
s_k(\varphi) &= \sum_n \left[\text{rect}\left(\frac{\varphi - n\,2\pi + \theta_k}{\pi}\right) \circledast \mathcal{G}(\varphi, \sigma_D)\right],
\end{aligned}
\tag{3.7}
$$

where the *pulse width* is assumed to equal the FWHM. To compute the correlation function (Eq. 3.3), we utilize the fact that the convolution of two Gaussians yields another Gaussian function with $\sigma = \sqrt{\sigma_D^2 + \sigma_M^2}$. Assuming the pulse width being smaller than the period of

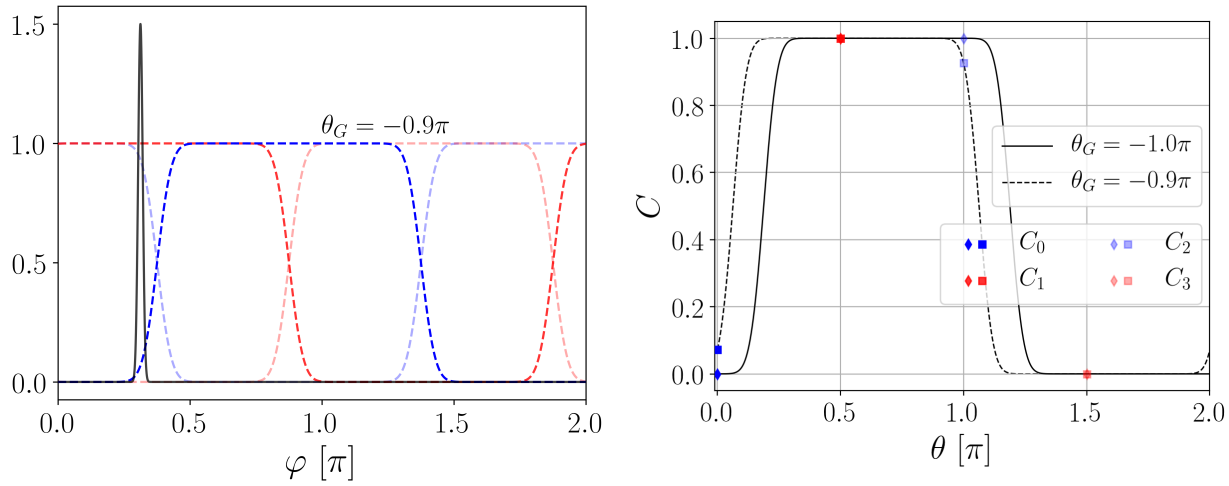

Figure 3.3: Signals corresponding to a single pixel measurement. The demodulation signals are color coded according to Fig. 3.2. **Left**: In contrast to Fig. 3.2 (left), we adjust the DOI by applying θ_G such that the reflected pulse lies within the *sensitive range*. This results in a coincidence of reflected pulse and rising signal edge of $s_0(\varphi)$. The 4 measured samples C_k now refer to a unique phase (see **right** panel, dashed line) and a small change in measurement value will result in a large change of the phase estimate.

the modulation (FWHM $<< \omega T$), we obtain

$$\begin{aligned} C_k(\varphi_\Gamma) =& \frac{1}{\omega}\int_0^{2\pi} \text{rect}\left(\frac{\varphi + \theta_k - \varphi_\Gamma}{\pi}\right) \circledast \mathcal{G}(\varphi, \sigma)\, d\varphi \\ =& \frac{1}{\omega}\frac{\sqrt{\pi}}{2\sqrt{a}}\{\text{erf}(\sqrt{a}\varphi_2) - \text{erf}(\sqrt{a}\varphi_1) \\ \varphi_{2,1} =& \pm\frac{\pi}{2} + \varphi_\Gamma - \theta_k;\; a = \frac{1}{2\sigma^2}. \end{aligned} \tag{3.8}$$

Depth sensitivity. The depth sensitivity (Eq. 3.6) is driven by the gradient of the normalized correlation function

$$\begin{aligned} \frac{\partial C_k(\varphi_\Gamma)}{\partial \varphi_\Gamma} =& \frac{1}{\omega}\Big\{\exp(-a(\varphi_\Gamma - \theta_k + \frac{\pi}{2})^2) \\ & - \exp(-a(\varphi_\Gamma - \theta_k - \frac{\pi}{2})^2)\Big\}, \end{aligned} \tag{3.9}$$

which essentially are two Gaussians located at $\varphi_\Gamma = \theta_k \pm \pi/2$, one with negative, the other with positive amplitude (see Fig. 3.2, middle). These Gaussians indicate that the maximum (absolute) gradient of the correlation function $C(\Gamma)$ is achieved at this local maximum and minimum respectively, which depend on the value of φ_Γ and hence the distance towards an observed object. This means that our PC-ToF approach exhibits strong sensitivity in a narrow range around a specific phase, the *phase of maximum sensitivity* φ_0.

The depth of interest (DOI). From Fig. 3.2 (middle, right) shows that only measurements with a specific depth achieve maximum sensitivity for a given θ_k. Here, we choose $\theta_k = \theta_0 = 0$ for reference and focus on the rising signal edge (first extremum of C_0'). The

phase of maximum sensitivity φ_0 depends on two parameters, the depth Γ at which the observed scene point is located and the phase shift θ_0. This means that φ_0 can be shifted towards another φ_Γ: We introduce an additional phase shift θ_G that acts on all θ_k equally (to still utilize the 4-tap hardware):

$$\varphi_0 = \varphi_\Gamma \mp \frac{\pi}{2} - \theta_0 = \frac{2\omega\Gamma_0}{c} - \frac{\pi}{2};\ \theta_k \to \theta_k + \theta_G \tag{3.10}$$

Physically, this allows to shift the rising edge of the demodulation function towards the reflected pulse (compare Fig. 3.3; Fig. 3.2 left). The depth corresponding to φ_0 we call the *depth of interest* (DOI) Γ_0.

Sensitive range. The phase of maximum sensitivity is restricted to one particular value and corresponding depth, whereat measurements within a surrounding phase interval also benefit from increased depth sensitivity. We call this interval the *sensitive range* $\Delta\Gamma$ and corresponding phase $\Delta\varphi_\Gamma$. Mathematically it can be described as the interval with nonzero first derivative $C_0' \neq 0$. For our approach, we hence define the sensitive range as the interval bounded by the points where the Gaussian reaches $1/e^2$ of its peak, often also called the *beam width*.

$$\Delta\varphi_\Gamma = 4\sigma;\ \Delta\Gamma = \frac{\Delta\varphi_\Gamma\, c}{2\omega} \tag{3.11}$$

For illustration, consider a PC-ToF measurement which has a true depth of Γ. To achieve maximum sensitivity, the ideal solution would be to set the DOI to the exact depth and acquire the necessary phase shift θ_G. This is the case, when the reflected illumination pulse coincides with the rising edge of the demodulation signal, as in Fig. 3.3 (left). If the true depth Γ and the DOI Γ_0 differ by more than $0.5\Delta\Gamma$, the reflected pulse coincides with one of the plateaus of the demodulation signals (Fig. 3.2 left) and the measured C_k will not change at all, if there is a small variation of the depth.

A priori the exact depth value is unknown. To still achieve a high resolution depth measurement we will choose the DOI Γ_0 such that it is close to the depth Γ.

3.3.1 Hardware

Our approach describes an additional mode of operation for existing correlation ToF range finding setups, which builds upon the following hardware:
First, we require a correlation ToF sensor. These devices are either externally modulated by a high-frequency signal or employ their own signal generator for this purpose. We utilize a PMD CamBoard nano (based on their 19k-S3 sensor) with external DDS modulation source [HHG+13], which also triggers the laser source. **Second**, we require the light source to emit narrow pulses, synchronized with the given sensor modulation signal. To this end we utilize an Omicron QuixX laser with a pulse width of FWHM $\leq 500\,\text{ps}$.
Third, for calibration we require a phase shift to be applied to either the modulation or demodulation signal. This phase shift needs to be adjustable with as high an accuracy as possible, as this affects the final resolution of the range imaging system. The modulation source allows setting the phase with 14 bits precision, leading to phase steps as small as $\Delta\varphi = 2\pi/2^{14}$. The parameters for our measurements, as well as the specific hardware used, can be found in Table 3.1.

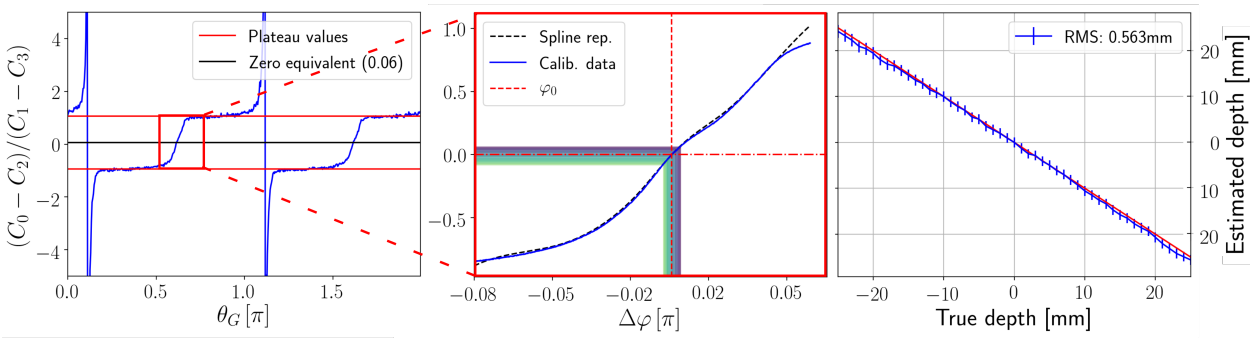

Figure 3.4: Calibration and validation procedure example for a single pixel. **Left**: Measured raw fraction $\Psi = \frac{C_0(\mathbf{p})-C_2(\mathbf{p})}{C_1(\mathbf{p})-C_3(\mathbf{p})}$ for 512 equally spaced phase shifts $\theta_G \in [0, 2\pi)$. We estimate plateau values and zero-crossing equivalent for the sensitive region (red rectangle). **Middle**: We acquire Ψ within the sensitive range enclosing the rising signal edge, which is performed at the highest possible accuracy in terms of θ_G (14 bit). This yields a θ_G-Ψ mapping that allows to estimate the offset $\delta\varphi$ of a measurement from a reference phase. To ensure a continuous lookup table, we fit a spline representation. **Right**: Results of our validation: We physically change the distance of the calibration target and estimate the corresponding phase using our calibration. The plot shows the depth averaged over all camera pixels, plotted against the ground truth depth.

Table 3.1: Parameters of the hardware used for our measurement setup (cf. Fig. 3.1) Frequency and corresponding opt. power are based upon PC-ToF operation.

Laser light source Omicron QuixX 852-150		**Sensor** PMD 19k-S3		**Lens** Fujinon HF35SA-1	
Wavelength	852 nm	Resolution	160x120	Focal length	35mm
Pulse width (FWHM)	<500 ps	Shutter time	1 ms	Aperture	f/2.0
Sensitive range $\Delta\Gamma$	$\lessapprox 0.75$ m			#Acquisitions	25
Measurement parameters					Notes
Frequency [MHz]	10	15	20	25	
Average opt. power [mW]	0.60	0.90	1.15	1.4	PC-TOF

3.3.2 Setup and measurement procedure

Fig. 3.1 shows a schematic illustration as well as pictures of our setup: The pulsed laser illumination is guided onto a mirror mounted on a linear stage before being reflected back onto a diffuser for uniform illumination of the scene. The linear stage allows to control the distance travelled which directly translates to a proportional phase shift. This is equivalent to adding a phase shift in hardware and is used for validation only, but could in principle also be utilized for calibration. The sensor observing the scene then retrieves a delayed version of the illumination signal, which is correlated with the demodulation signal on a per-pixel level.

Depth reconstruction. The CamBoard nano, as most available C-ToF systems, employs a four-tap measurement procedure that acquires four samples $\{C_0(\Gamma), C_1(\Gamma), C_2(\Gamma), C_3(\Gamma)\}$ of the correlation function per pixel $\mathbf{p}$ measured using demodulation functions shifted by $\theta_k \in \{0, \frac{1}{2}\pi, \pi, \frac{3}{2}\pi\}$. Instead of disclosing these four values, the CamBoard nano returns the differences of samples separated by $\Delta\theta = \pi$. For C-ToF systems utilizing sinusoidal modulation and demodulation signals it can be shown [SXH+97] that the unknown phase $\varphi_\Gamma(\mathbf{p})$ corresponding to the range $\Gamma(\mathbf{p})$ can be computed as

$$\varphi_\Gamma = \frac{2\omega\Gamma}{c} = \operatorname{atan}(\Psi); \quad \Psi = \frac{C_0(\mathbf{p}) - C_2(\mathbf{p})}{C_1(\mathbf{p}) - C_3(\mathbf{p})} \tag{3.12}$$

and we denote the argument Ψ as the *raw fraction*. This expression has two major benefits: First, the result of the differences is independent from ambient light $I_a(\mathbf{p})$ and second, the fraction of the two differences cancels out the scene dependent factor $E_c(\mathbf{p})$. Still, Eq. 3.12 is only valid for sinusoidal signals and results in strong systematic errors [PDC+08] for non-harmonic correlation functions such as ours. Instead, we will rely only on measurements of the raw fraction in dependence of a chosen depth of interest Γ_0 and phase shift θ_G.

Calibration. In contrast to the simple expression for sinusoidal correlation ToF (see Eq. 3.12), we cannot easily invert our correlation function for our pulsed approach (cf. Eq. 3.8). Instead we perform a calibration step in which we measure the raw fraction with a homogeneous calibration target (white diffuse plate) at a fixed distance to sensor and light source. Figure 3.4 (left and middle) visualizes the calibration process for a single pixel. **First**, we perform rough measurements of Ψ with 512 equally spaced phase shifts θ_G. This reveals the upper and lower plateaus of the (ideally) rectangular correlation function. In theory, these plateaus have equal absolute value and hence the phase of maximum sensitivity is at $\Psi = 0$. For real measurements, there is a small imbalance. The phase of maximum sensitivity is exactly in the middle of the two plateaus. The transition region between the plateaus is used to estimate the phase of maximum sensitivity on a per-pixel basis, which refers to the depth Γ at which the calibration target is placed (see supplemental material for details). **Second**, we measure Ψ within the sensitive region by stepping over the corresponding θ_G with as high precision as possible (14 bit). This yields lookup values that could be used directly for estimating phases from measured values. However, noisy measurements introduce ambiguities into a numerical inversion, which requires a smoothing step to obtain a monotonously rising function for unambiguous phase estimation. To this end, we fit a univariate spline

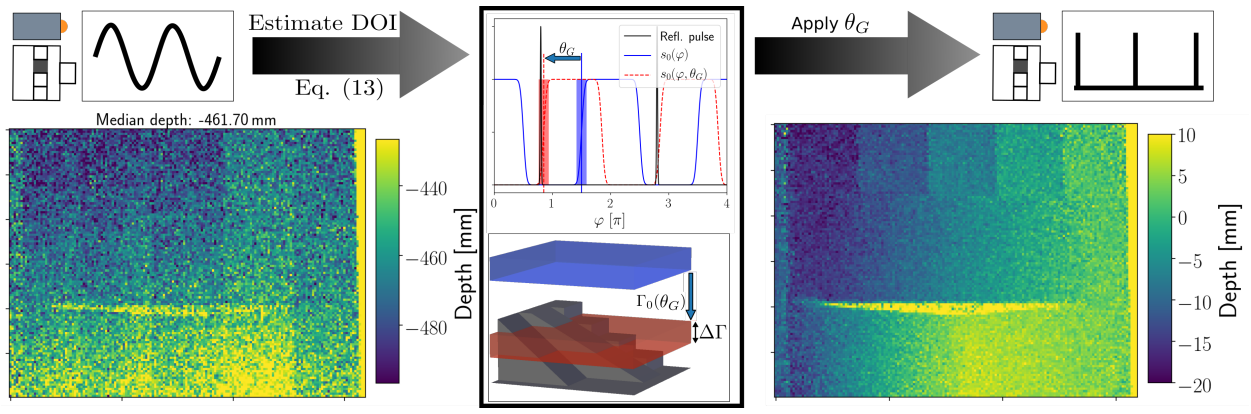

Figure 3.5: PC-ToF measurement procedure. **Left**: We perform a single ToF measurement with sinusoidal modulation and demodulation signals at 0.6 mW – The low SNR disallows to obtain a high resolution depth map but instead returns a rough depth estimate (about 50 cm) for features of the scene we are interested in, the depth of interest (DOI). **Middle**: Using Eq. 3.10 we compute the phase shift required to let the reflected pulse and the sensitive edge coincide, focussing our measurement onto the DOI Γ_0. The shaded rectangles here illustrate the sensitive range surrounding the DOI, both in phase space and for 3D depths. **Right**: We perform a single measurement at 0.6 mW after applying the shift and switching the mode of operation to our pulsed acqusition. The obtained depth reconstruction exhibits well improved depth resolution for all scene features that lie within the sensitive range.

representation to the measurements on a per-pixel level. The resulting lookup table can now be used to estimate the phase difference from φ_0. **Third**, we note that not necessarily all pixels of a C-ToF sensor exhibit the exact same behaviour, which in our case leads to a different spline representation and phase of maximum sensitivity per pixel. To account for this effect, we create a phase calibration mask. Details can be found in the supplemental material.

Validation. To validate our calibration, we need to assess how closely we can reconstruct changes of depth within a scene with the available calibration. As the calibration allows to measure an phase difference from φ_0, we perform measurements with a planar calibration target again but now we change the distance the light has to travel by moving the mirror on the linear rail (cf. Fig. 3.1). We perform a total of 51 measurements with offsets in the range of $[-2.5,2.5]$ cm at 1 mm accuracy with respect to the reference depth used for calibration. The results (cf. Fig. 3.4) indicate a good match between our depths obtained from a phase estimate using the spline representation for inversion and the ground truth depth values. Note that the validation is performed within a close range around the phase of maximum sensitivity, well within the sensitive range.

Performing a PC-ToF measurement. After validation, our measurement procedure (cf. Fig. 3.5) is straightforward: We first acquire a rough depth estimate using a low power C-ToF measurement with sinusoidal modulation and demodulation with our system. With this, we obtain a rough estimate of the depth the object of interest is located at and adjust θ_G such that the DOI is matched and interesting scene features lie within the sensitive

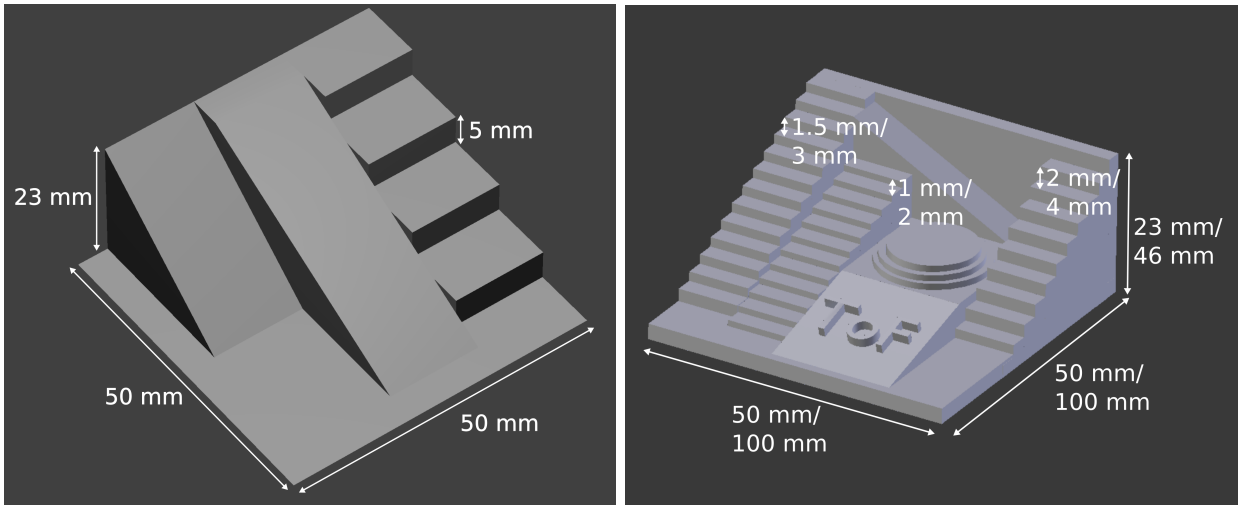

Figure 3.6: Pictures of the 3D models we printed and measured using our pulsed ToF approach. **Left**: Standard target box consisting of 2 ramps and stairs with a step height of 5 mm. This scene is used to validate the working principle of our approach. **Right**: Target box with detailed stairs. We generated 2 variants of this scene with the different measures separated as 1.5 mm / 3 mm. The most detailed stairs have a step height of 1 mm and are considered the limit test case for our method.

range. Another measurement, now in pulsed operation delivers a much better resolved depth estimate. All measurements, unless explicitly stated otherwise are obtained using a modulation frequency of 10 MHz at 0.6 mW optical power with the hardware stated in Table 3.1.

3.4 Results and conclusion

To assess the capabilities of our approach, we designed and 3D-printed three different targets, depicted in Fig. 3.6. The targets dimensions are chosen such that we cover a range of relative depth differences, starting from discrete steps of 5 mm down to 1 mm. Given the results from the validation measurement (cf. Sec. 3.3.2), we regard signal changes that originate from such small depth differences as the limit our approach can resolve.

Assessing the capabilities the method Fig. 3.8 shows the depth maps as well as cross-sections obtained from PC-ToF measurements in comparison to ground truth data for three different targets. Figure 3.7 shows the ability to focus on different depths, with very high noise on the distant object. A C-ToF measurement, as required to obtain a rough depth estimate (cf. Fig. 3.5) for our procedure reveals two objects at different depths, of which at least one is problematic to identify. Using the approximate depth, we set the respective phase shift θ_G to match the depths of interest to acquire a better resolved depth map of the rubber duck. In this case, parts of the foreground object cannot be reconstructed (white pixels) as the measured values for Ψ lie outside the sensitive region. We then shift the phase to include the foreground object, reducing the SNR in the background but recovering the full Mario figurine. In contrast to C-ToF, which suffers from the low SNR due to low optical

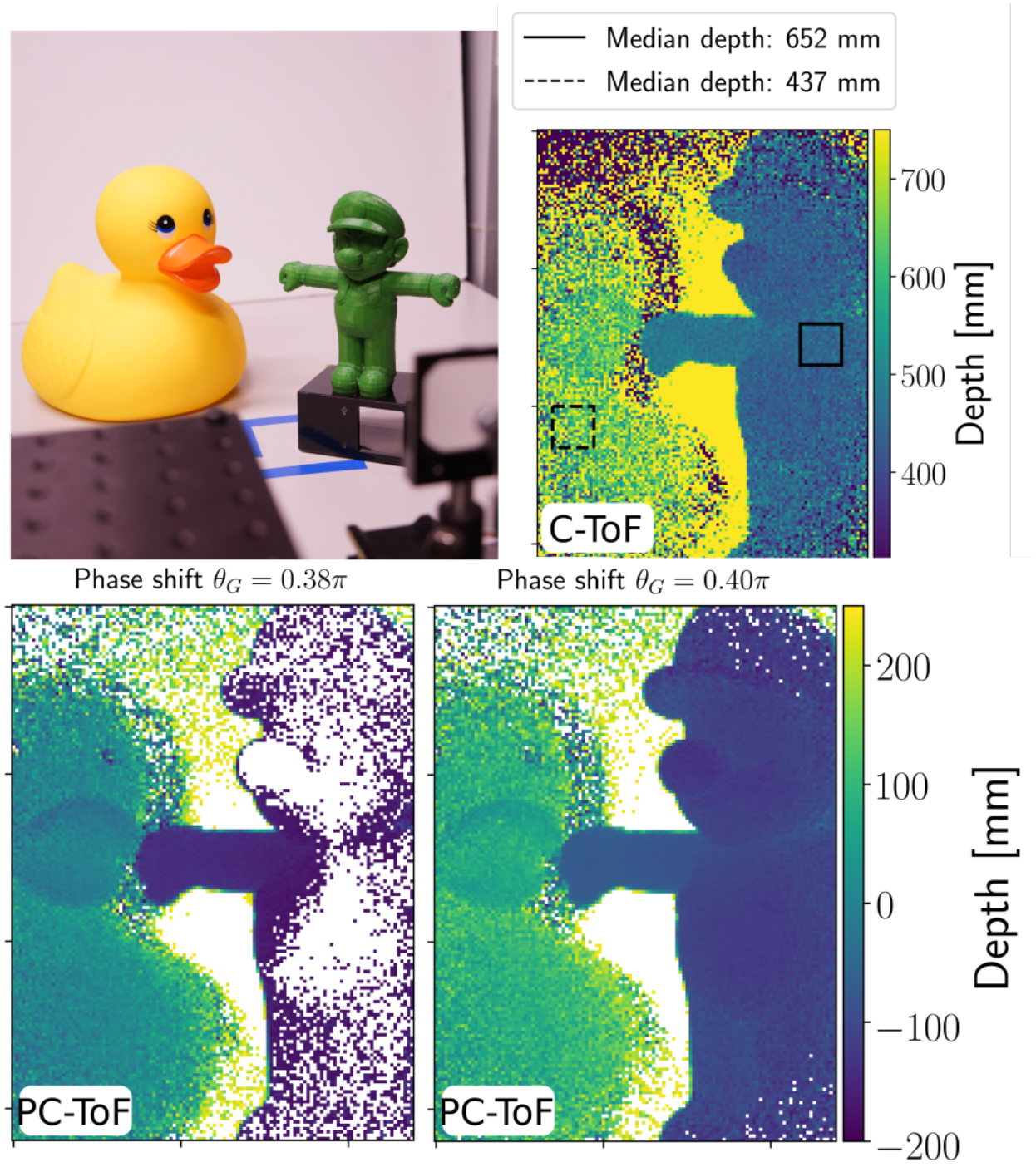

Figure 3.7: Depth focusing with our PC-ToF approach at 15 MHz, 0.9 mW. The scene consists of two objects at different depths, estimated by a low-power C-ToF measurement (**top right**). We use these depths to apply phase shifts θ_G to focus on the rubber duck and Mario figurine respectively (**lower row**). White pixels denote measurements outside the sensitive region.

power, our approach is able to achieve reconstructions with errors below 2 mm under optimal conditions. Included in these errors are also systematic problems that are equal for both, C-ToF and PC-ToF: Our ground truth is obtained via simulation, which relies on the correct pose estimation for the 3D models. This is a difficult task on its own, especially given the low resolution images.

Analysis of accuracy vs frequency vs power In order to assess the behaviour of PC-ToF with respect to modulation frequency and optical power, we performed measurements using the large target box variant (cf. Fig. 3.6 (right)) and computed the RMS error with respect to the ground truth. See Fig. 3.9 for the results. Our PC-ToF approach outperforms the standard C-ToF reference clearly, with RMS errors below 1 cm for optical power as low as 0.5 mW (see the inset for a close-up). As our method does depend on the shape of modulation and demodulation signal instead of frequency, we expect to see no strong gain (except for reduced measurement time) with higher frequencies, which can be established with our results.

Measurement of complex materials Common available ToF systems inherently suffers from the so-called multipath interference problem (MPI): Whenever multiple light paths end up in one sensor pixel, the resulting depth estimate for this pixel is shifted, usually to higher distance values. This is an effect often observed in concave corners or translucent objects. Our method relies on the working principles of available ToF hardware and hence does not circumvent this problem, which is well visible for example in Fig. 3.8 (bottom left). On the other hand, PC-ToF still outperforms C-ToF for example in the case of a translucent rubber duck target, best visible when considering the meshes extracted from the reconstructed depth maps as shown in Fig. 3.10. In particular, C-ToF achieves a comparable level of detail only for much higher optical power of 4.43 mW in contrast to 0.60 mW. The respective images can be found in the supplemental material.

PC-ToF performs especially well with optical power and modulation frequency as low as 0.6 mW and 10 MHz respectively, where it reveals depth differences as fine as 2 mm under optimal conditions. In general, our method outperforms standard C-ToF reference measurements, effectively halving the RMS error (cf. Fig. 3.9). These results also reveal that the approach is inherently dependent on the pulse shape of the modulation and rise time of the demodulation signal and hence the shape of those signals. The modulation frequency is only a secondary factor.
Thereby, our method relies on a two-step procedure to first obtain a rough depth estimate from low-power C-ToF measurements. We then extract a depth-of-interest onto which we focus by applying a corresponding phase shift, leading to a range finding measurement with much increased depth sensitivity within a certain sensitive region around the DOI.
In future work, we would like to test the limits of the approach for more contemporary ToF sensors, which operate at frequencies of 100 MHz or more. With higher frequencies, we expect shorter rise times of the sensor modulation and at the same time we can accumulate more laser pulses in the same exposure time, increasing the SNR (the supplemental material shows results for up to 25 MHz). This should also allow us to test our approach on scenes

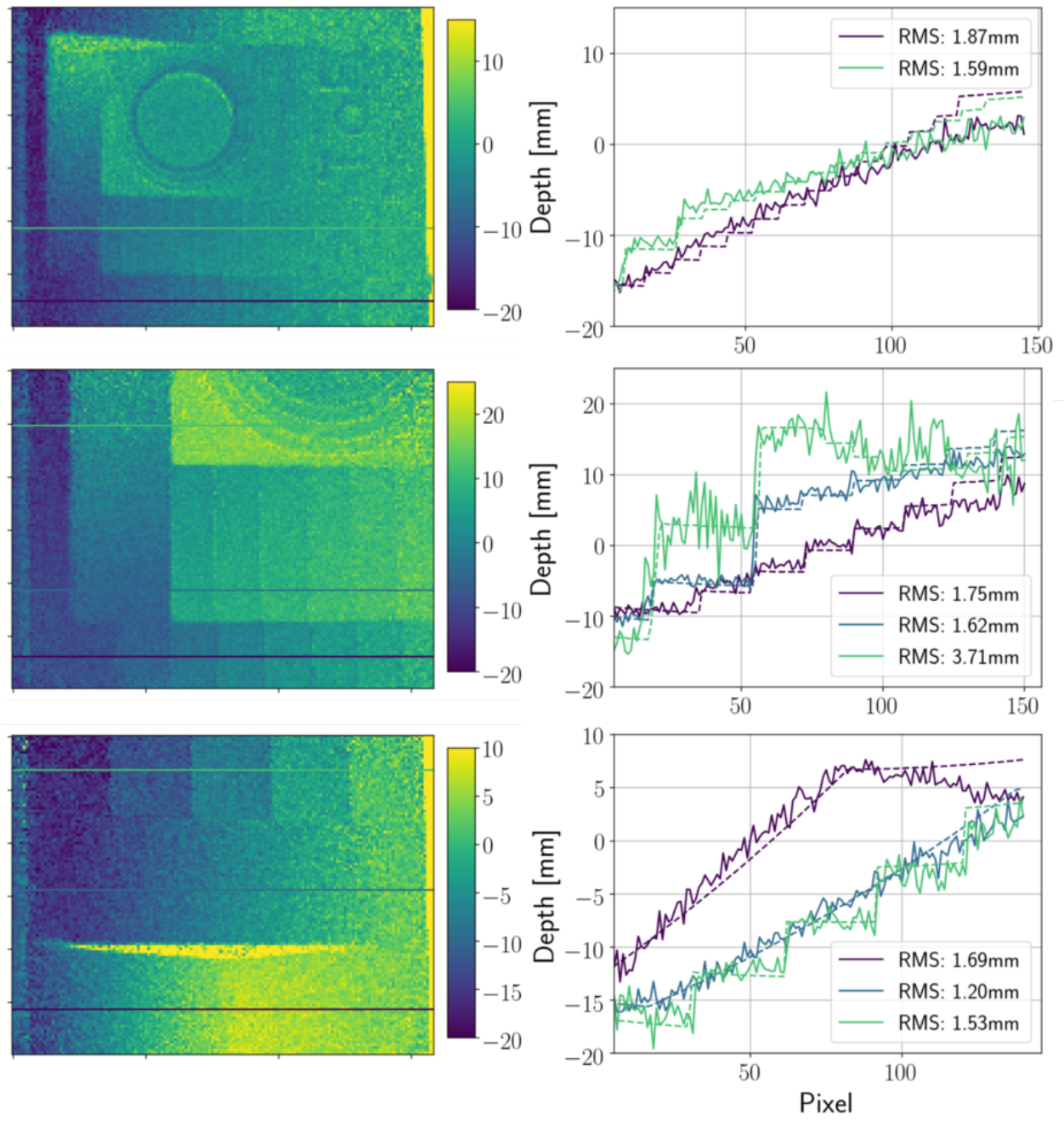

Figure 3.8: Depth maps (**left**) and comparison of depth slices (**right**) with our simulated ground truth data. All slices taken across stairs are averaged over 5 pixels in vertical direction. From top to bottom, the step heights of the stairs are 1 mm (green), 1.5 mm (violet); 2 mm (blue), 3 mm (violet); 5 mm (green). The remaining three slices (middle, green; bottom, blue and violet) are taken across a more complex inhomogeneous geometry and two slopes with same maximum height but different length respectively.

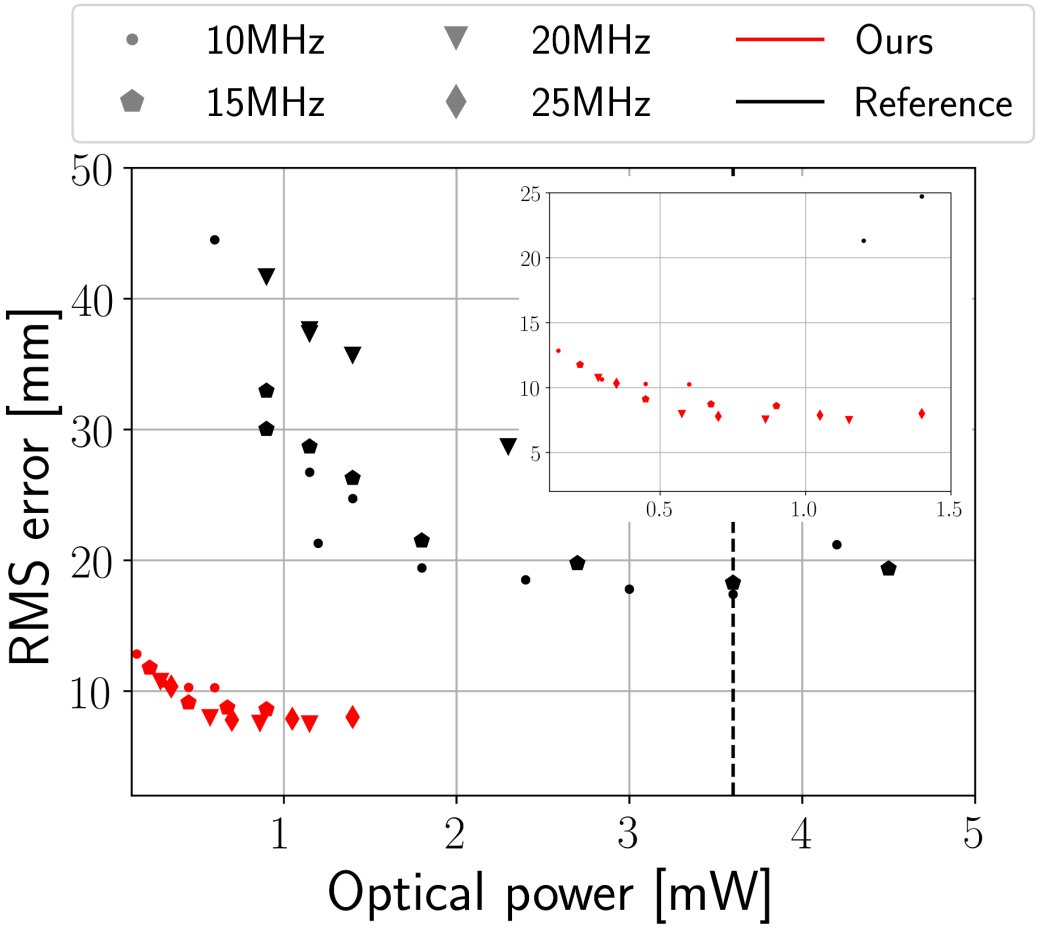

Figure 3.9: RMS error for varying frequency and optical power for both, C-ToF and PC-ToF for our reference target. The different methods are color-coded, the markers encode the respective frequencies. The inset shows a close-up of the low-power regime our method operates in. The dashed line marks the power at which we found some C-ToF measurements to be overexposed.

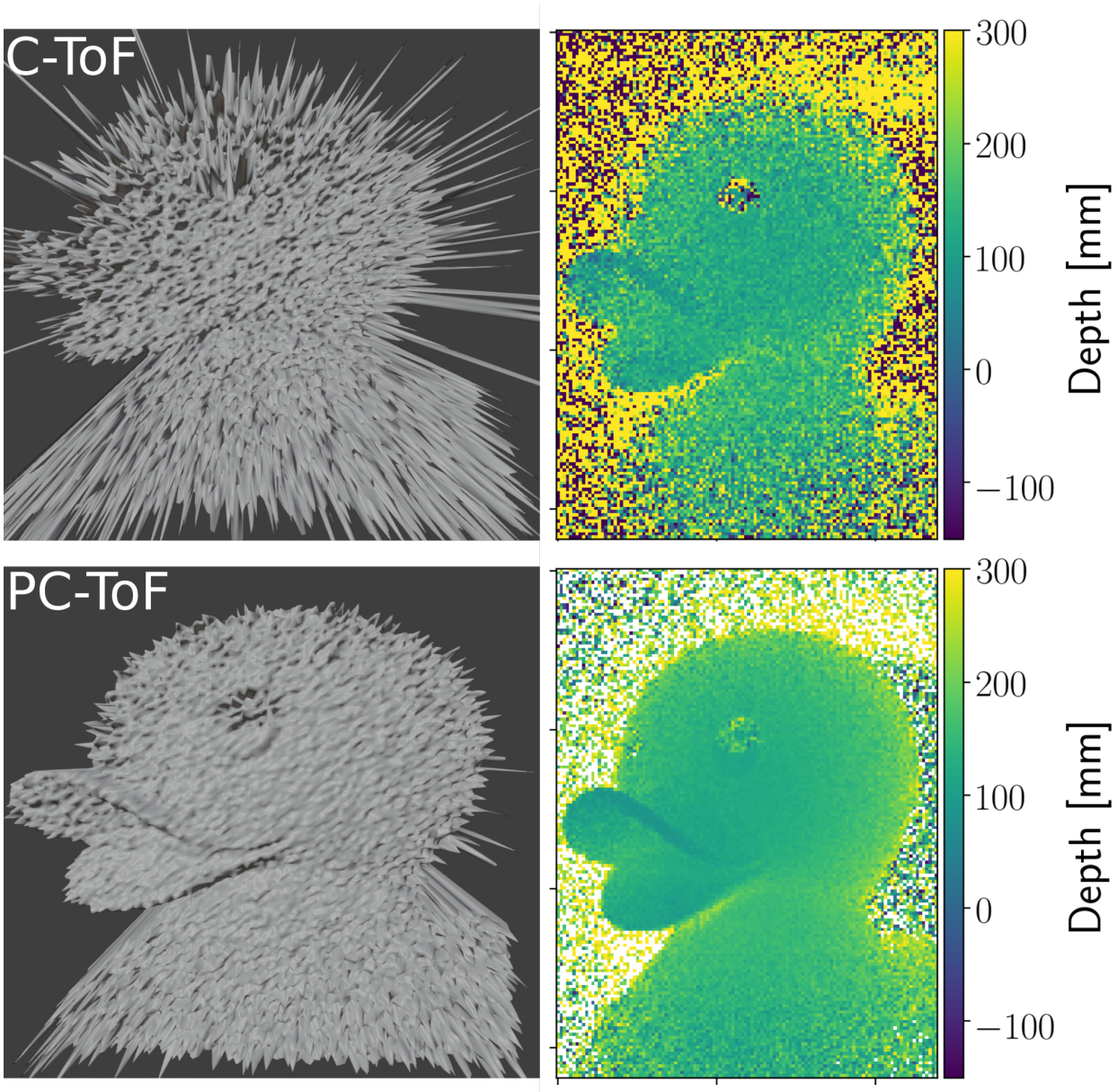

Figure 3.10: Equal power comparison for a complex material (rubber duck) at 10MHz and 0.6 mW. The left panel shows the renderings of the meshes reconstructed from the depth maps given in the right panel. **Top**: Results obtained with the reference C-ToF method. **Bottom**: Results obtained with our PC-ToF approach.

with larger depth ranges - focusing on different targets that ideally are allowed to be meters apart and reconstructing the depths within a few centimeters around the respective DOI with high accuracy. Future work should also include a way to circumvent the calibration procedure described in Sec. 3.3.2: One way to approach this problem could be to incorporate measurements of rise time and pulse shape of the modulation and demodulation signals, which can then be incorporated into the correlation signals theory. Either by approximation or numerical inversion we should then be able to acquire a phase-raw fraction mapping.
In the end, our approach is most suited for low power scenarios, such as mobile devices or for static measurement scenarios, where a detailed depth estimation for objects at a distinct range is required. With PC-ToF we trade the global sensitivity of a standard C-ToF setup (covering the unambiguity range) for highly increased sensitivity around a depth of interest: The broader the sensitive range, the less the maximum sensitivity and vice versa. In turn this allows for a task-specific tailoring of the pulse width and rise time, the two parameters that drive the depth resolution and sensitive range achievable with PC-ToF.

Appendix

In the following we present additional results that accompany our paper and were left out due to space constraints or to improve readability.

3.A Calibration mask extraction

Our calibration relies on a per-pixel measurement procedure which aims to detect plateau values of the raw fraction Ψ surrounding the sensitive region. This sensitive region is centered at the depth of interest (DOI) onto which our approach is able to focus. To enable the inversion of the Ψ-phase mapping, we record data for a lookup table in two steps: First, a rough sampling to obtain the approximate sensitive region and second, a fine sampling within the sensitive region. This process is depicted in Fig. 3.13. The result is a per-pixel phase φ_0 at which the DOI is matched.

Our approach relies on the setting of a single phase shift to focus on the DOI. To account for the difference in φ_0 between the different pixels, we choose the median over all φ_0 as a reference point and obtain a calibration mask, which employs a per-pixel phase correction with respect to the DOI.

3.B Accuracy analysis details

We perform measurements for various frequencies and optical power settings to assess the accuracy of our approach in terms of RMS error. To this end, we employ a measurement target whose geometry is known, and compare measurements for both, C-ToF and PC-ToF with simulated ground truth depth. A subset of the results acquired are shown in Fig. 3.11 for measurements obtained with 10 MHz at low optical power. These data show that PC-ToF outperforms C-ToF even for optical powers as low as 0.15 mW, however with strongly visible noise (top right image). Additionally it becomes visible that the setup inherently exhibits a lower SNR towards the right edge of the target, which is most likely due to a not perfectly uniform illumination by the diffuser.

To assess this effect, we acquire the RMS error not only for the full field of view as shown in the main paper, but distinguish between different positions on the target that relate to varying measurement conditions. We extract slices from the depth maps, as indicated by the color-coded horizontal and vertical lines in Fig. 3.12 (right) that correspond to a distinct target geometry (stairs or slope) and are close to the edges of the scene. For each of these slices, we compute the RMS error individually, the results are shown in Fig. 3.12 (left).

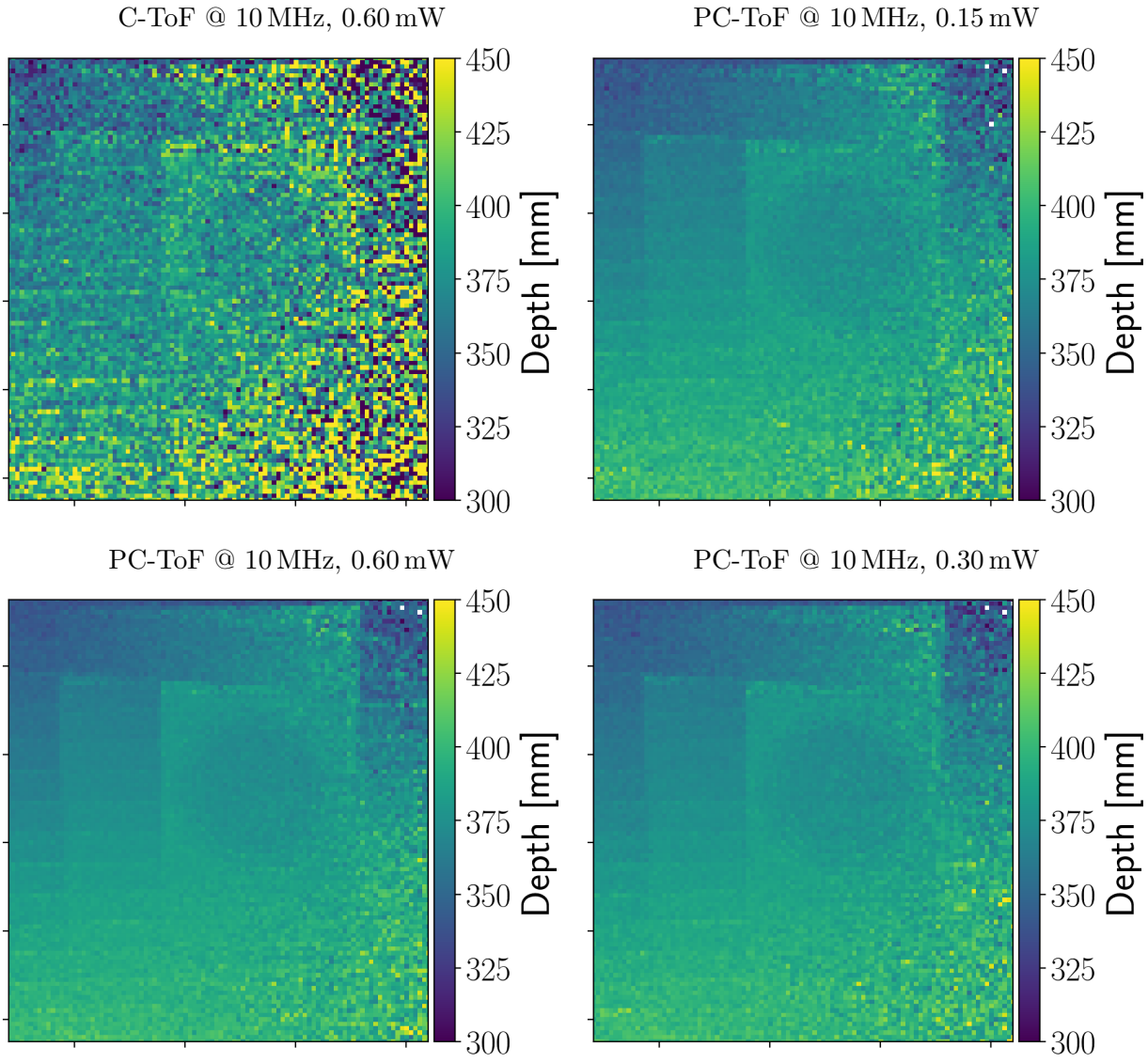

Figure 3.11: Depth maps for our reference target used for the RMS error analysis. We perform measurements with both C-ToF and PC-ToF for various frequencies and optical powers.

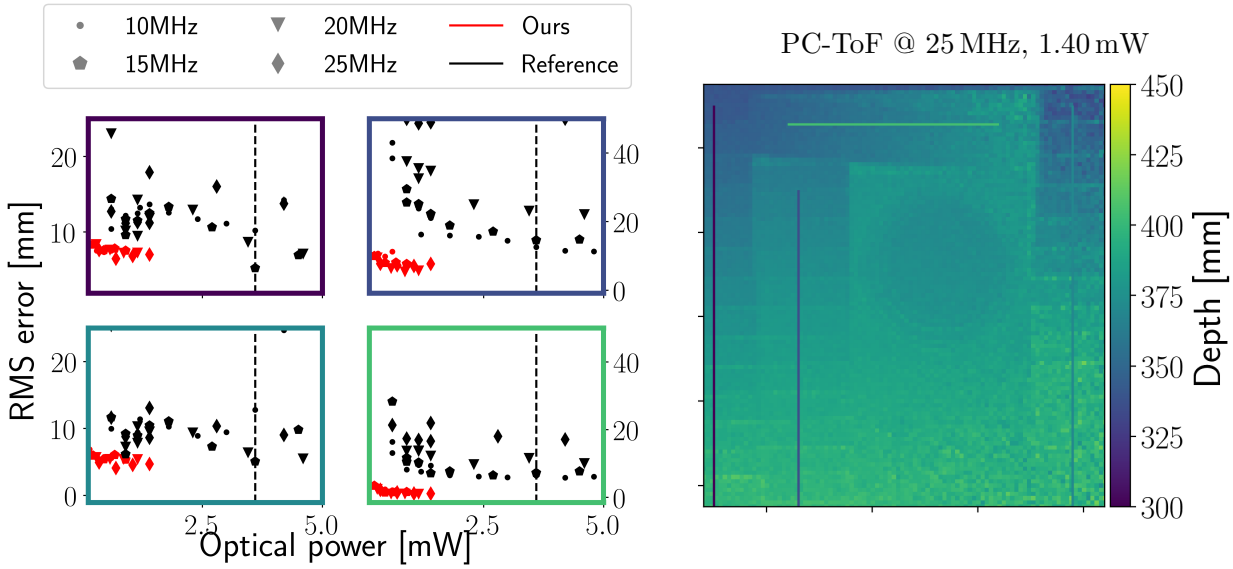

Figure 3.12: A more detailed analysis of the RMS error for C-ToF and PC-ToF for different frequencies and optical power settings. **Left**: Slice RMS errors obtained for the indicated settings. The border color corresponds to the colored lines in the right panel. **Right**: PC-ToF depth map with line indicators corresponding to the lines extracted to compute the RMS errors in the left panel. This figure complements Fig. 8 in the main paper.

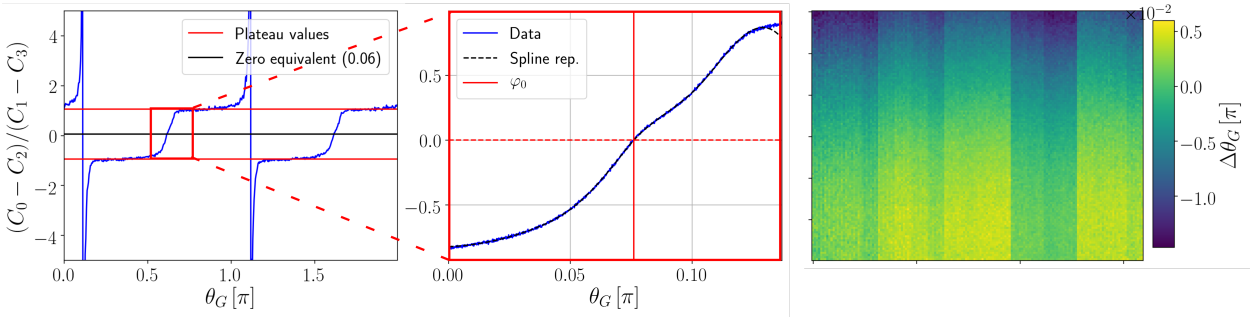

Figure 3.13: **Left**: Measured raw fraction Ψ for 512 equally spaced phase shifts $\theta_G \in [0, 2\pi)$. We estimate the values for the upper and lower plateaus of the correlation function (red horizontal lines). From the plateau values we compute a zero-crossing equivalent value, which is exactly half the difference between the plateaus. We further estimate the sensitive range, here denoted by the red rectangle. Middle: We perform an additional measurement of Ψ for phases θ_G within the sensitive range enclosing the rising signal edge, which is performed at the highest possible accuracy in terms of θ_G (14 bit). The data acquired exhibits noise, which leads to ambiguities when used for a lookup table. Instead, we fit a continuous spline representation to circumvent that issue. This way, we obtain a θ_G - Ψ mapping that allows to estimate the offset a measurement exhibits from a reference phase. **Right**: As we can only chose a single depth of interest via θ_G per measurement, we obtain the phase of maximum sensitivity for each single pixel and subtract it from the median over all pixels. This way,we obtain a calibration mask that employs a per-pixel phase correction with respect to θ_G . This figure complements Sec. 3.2 in the main paper.

3.C Complex material measurements for higher power/frequency

In section 4 we present the results of a depth reconstruction for a translucent (rubber duck) material in an equal power setup. The target is depicted in Fig. 3.14 under infrared light illumination, recorded with an IR camera. Here, the internal scattering within the material creates a glowing effect, indicating the complexity of the underlying light transport. Figure 3.15 contains additional results obtained with higher frequencies for both, C-ToF and PC-ToF for equal power as well as increased optical power in case of the reference C-ToF measurement. One of the major problems with translucent materials in ToF range finding is the existence of multipath interference, which introduces phase shifts towards larger distances, ultimately leading to stronger noise within the measurement. This especially affects concave geometry and fine surface details. In case of the rubber duck target, this becomes apparent for the reconstruction of the beak region: C-ToF is able to estimate the shape of the duck's beak with increased optical power (up to 4 times the power utilized for PC-ToF) but strongly suffers from noise. In contrast, the inherent increase in depth sensitivity for PC-ToF allows for a much smoother reconstruction of the beak even under these conditions.

Figure 3.14: The rubber duck target we use to showcase measurements of a translucent material. The image was taken under infrared illumination at 852 nm using our laser light source in combination with an IR camera. Internal scattering within the material is visible, visualizing the complex light transport involved.

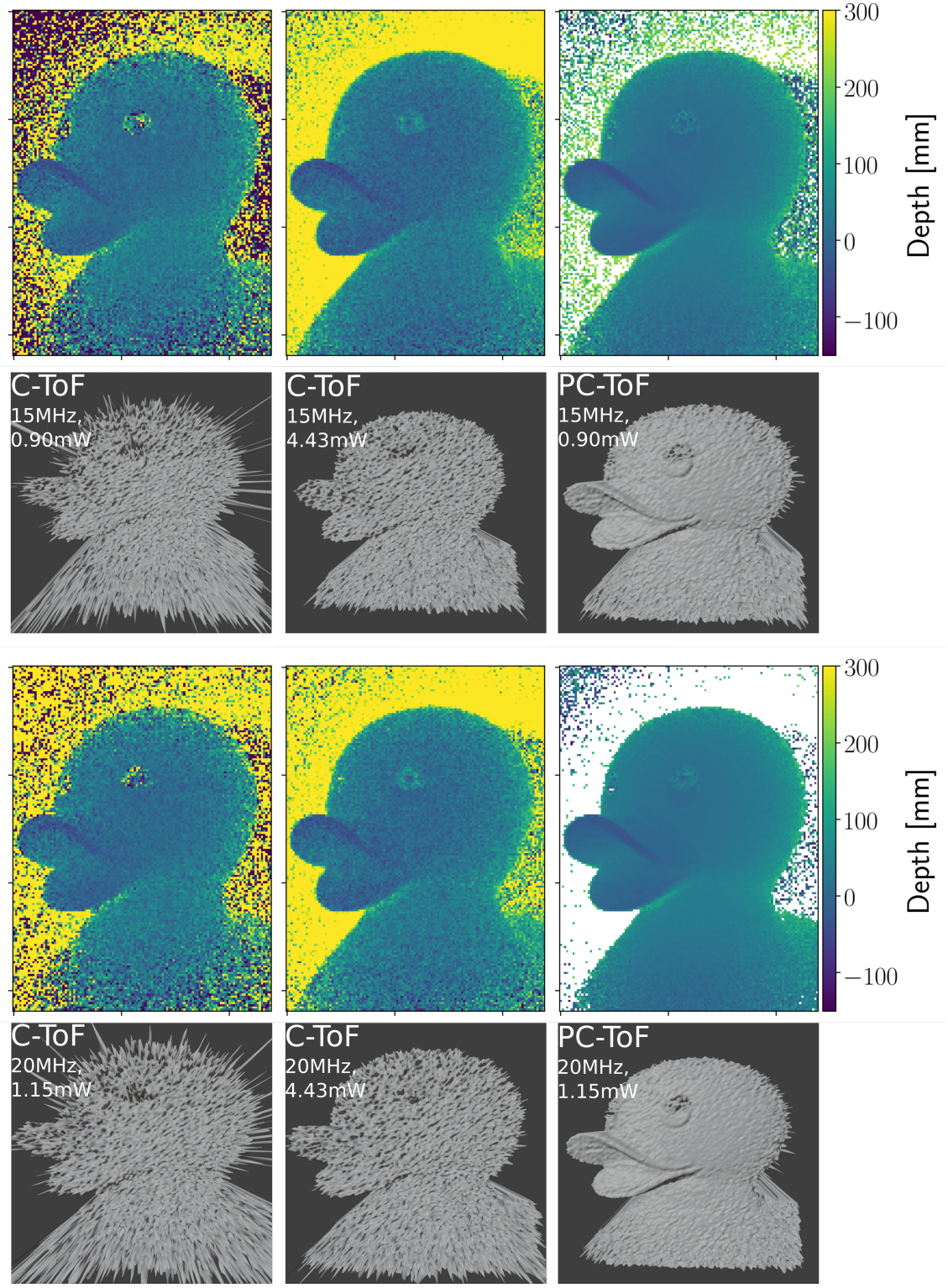

Figure 3.15: Additional results for the rubber duck target. We show measurements for the reference C-ToF as well as our PC-ToF approach as both, depth map and rendering of the extracted mesh. The first column shows the reference results for equal optical power with respect to our C-ToF measurements, which can be found in the third column. The middle column depicts C-ToF results for increased optical power, not yet reaching the quality of the PC-ToF acquisition. This figure complements Fig. 10 in the main body.

Preface – Scratch Iridescence

The computer graphics community is strongly tied to the field of digital material appearance, which is concerned with the appearance formation of digital doubles by defining the light-material interaction: Given a ray of light incident on an object within a physically-based rendering framework, the digital material associated with the intersected shape encodes the physical mechanisms that lead to the observed appearance. For dielectric materials, such as metal, ceramics or plastic this is often modeled using a bidirectional reflectance-distribution function (BRDF) that relates the incident and outgoing radiance at any surface point.

Driven by the desire for ever more realistic digital materials, the computer graphics community strives to develop BRDFs that closely recreate the properties of real-world materials. A particularly difficult task in this regard is the representation of hidden microscopic structures that, although not visible to the naked eye, strongly influence the observed appearance. In addition, such structures remain inaccessible with common hardware and require specialized measurement setups to generate an accurate digital representation. Here, we focus on the effect of diffraction: Microscopic structures with an extent in the order of the optical wavelength create colorful patterns visible in reflection. Due to these small surface features which can be found on almost any worn surface or manufactured items such as CDs and DVDs, a truthful acquisition of the underlying material would require a spatial resolution in the order of a few nanometers, rendering data-driven approaches unfeasible. We therefore cast the problem of creating a truthful representation to an approximate, physically-based digital material that is driven by, measurable characteristics. To this end, we consider the underlying wave-optical effects such as diffraction of single microscopic scratches as well as interference originating from scratch ensembles and provide a spatially-varying BRDF that enables the creation of digital doubles of worn surfaces with special interest in microscopic, iridescent scratches.

Before we develop the theory forming our BRDF model, we first aim to analyze the properties of the targeted effects in a qualitative fashion: In chapter 4 we start by a thorough observation of the direction-dependent optical properties of scratched surfaces of household items and deepen our knowledge about the underlying geometry of these surface features using optical microscopy as well as measurements obtained from a raster electron microscope. Our findings from this preliminary analysis form the framework for our later model and act as a sanity check for the qualitative appearance we recreate.

Based on the preliminary considerations we then introduce our spatially-varying reflection model in chapter 5 as published in [WVJ+17], the supplemental videos are available at `https://osf.io/etdgb/?view_only=a60cc22ccdf54afbb9be279495d75508`.

The presented model is capable of re-creating digital doubles of worn and manufactured materials alike, such as the presented household items and a CD. Due to the explicit computation of diffraction originating from each single participating scratch, the computation cost of the presented model is high, desiring a more efficient solution. This was later provided in [VWH18], which is a publication with equal contribution from Velinov et al. and the author of this thesis.

CHAPTER 4

Preliminary analysis of iridescent surfaces

Worn or scratched surfaces can be found on many everyday objects, being an omnipresent but still often overlooked part of our daily perception. As an example, we encounter the traces of wear on a daily basis, such as on clothes, working tools or cutlery but are seldom aware of them. Instead, when we are confronted with digital recreations of objects that lack the characteristic dents, scratches and damages which we would expect for example on tools in a workshop or cutlery in a kitchen, we strongly recognize the discrepancy as the materials look too "clean".

A wealth of physically-based models for materials has been developed alongside with methods to simulate or explicitly model surface detail, such as microfacet-based models [CT82; WML+07; JHY+14; DWM+15; HP16], bump- or displacement maps [Bli78] as well as normal maps [Mik08]. Unlike geometrical optics, on which such models are based, the particular interest of our research lies on the wave-optical effects of defects that have a size in the order of the optical wavelength, ranging from around half to several micrometers across at least a single but often all dimensions.

The materials exhibiting such defects range from glass to plastic or metallic objects that eventually, be it by wear or manufacturing in the first place, show such characteristic microscopic scratches. As an illustrative example consider a standard kitchen environment with its sinks, pots and cutlery - everyday use will without doubt carve fine grooves into the surface. Under strong illumination with white light, these scratches often reveal iridescent colors. In particular, these colors change and vary with a change in view point and lighting direction. This contrasts our expectations from geometrical optics, where light-surface interaction is decribed as an incident ray being deflected and possibly changing color due to the albedo of the material, but not introducing angle-dependent colors.

As an example consider the objects displayed in Fig. 4.1: A kitchen spoon (top) and a plastic tool cover (bottom, left) show the characteristic colorful scratches created by wear, whereas a CD (bottom, right) exhibits such "defects" similar in appearance but due to a manufactured surface structure. The materials are illuminated using a single spotlight with small angular extent.

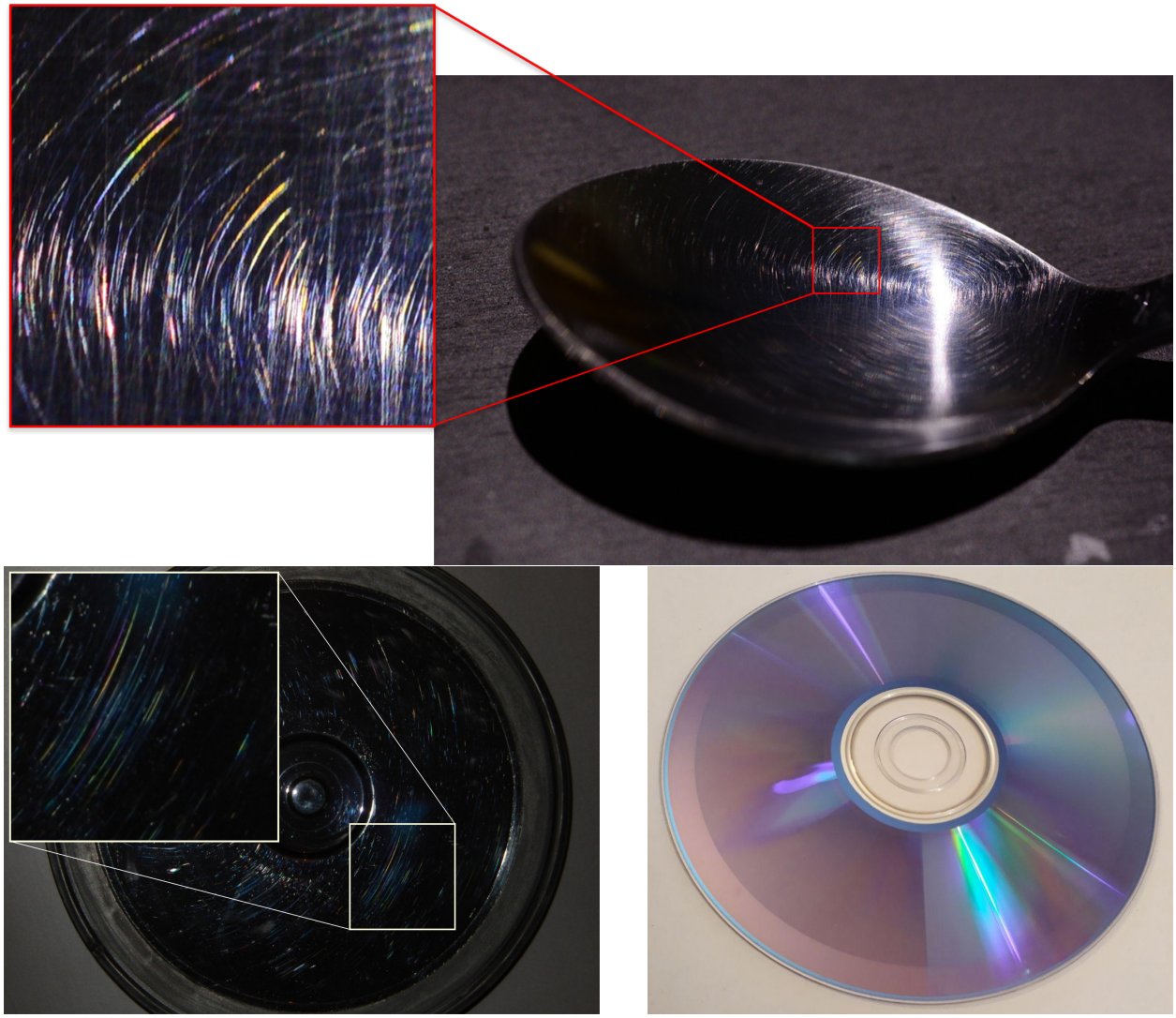

Figure 4.1: Examples for diffraction on everyday-use objects, caused by wear (left, middle) or manufacturing (right): **Top**: Used spoon under spotlight illumination. **Bottom, Left**: Plastic bottom of a grinding tool. **Bottom, Right**: Backside of an audio CD.

4.1 Diffraction

The phenomenon responsible for these iridescent colors is interference due to diffraction, which occurs when a wave interacts with an obstacle. Like water waves in a pond, light can be understood as a propagating wave, where multiple waves are able to enhance and diminish each other upon superposition, forming ripple patterns in the process. In the case of our pond scenario, water waves have a wavelength in the order of centimeters to decimeters. We can observe their interactions with each other as well as with objects in their path. The interaction of light with objects that we encounter every day is dominated by features much larger than the optical wavelength. In this case, light is often considered as a ray and its interaction with the surface can be sufficiently described by geometric considerations, comprising the concept of *geometrical optics*.

This concept is no longer sufficient when the object is made of a material which indeed exhibits for example defects in the order of the optical wavelength: Here light acts

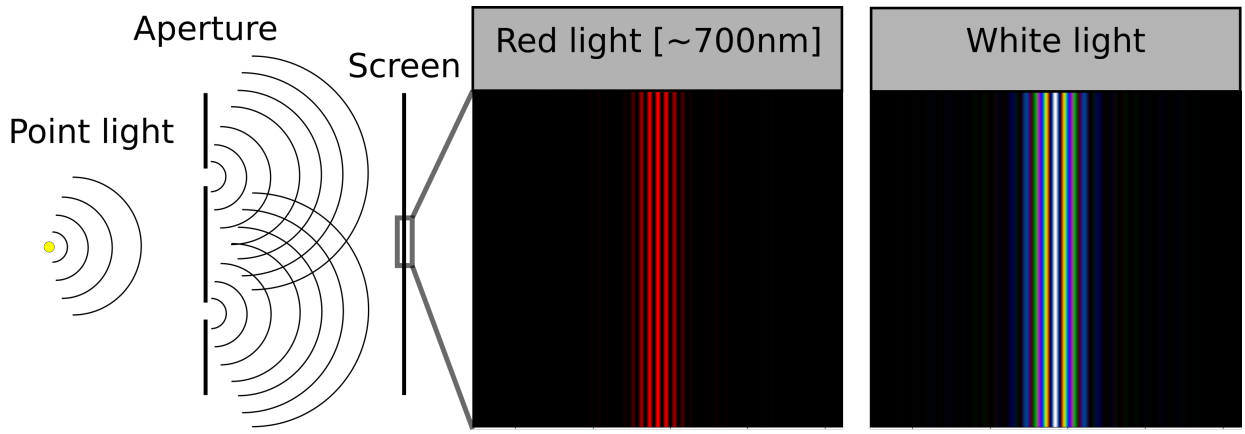

Figure 4.2: Simulated diffraction patterns obtained from a double slit aperture for red light (**left**) with a single wavelength and white light (**right**), consisting of a continuum of wavelengths. The aperture is illuminated by a point light source from the left, emitting *coherent light*. The slits of the aperture have widths in the order of the optical wavelength and can effectively be treated as two separate point light sources, generated upon incidence on the aperture. In the space on the left-hand side of the aperture the light emitted by these sources than interact, here indicated as the isolines forming the wavefronts of the spherical waves. The superposition of circular waves originating from the two slits creates a pattern of alternating high and low amplitudes per wavelength, showing a colorful pattern for white light.

similar to the water waves considered previously, which reveals the wave nature of light. A formal introduction and theoretical description is given in chapter 5.3. Here we restrict our considerations to an abstract explanation of the phenomenon: Each light source can be described by an ensemble of point light sources, each emitting a spherical wave with a spectrum corresponding to the wavelengths that describe the respective color of the light. The Huygens-Fresnel principle [KGS00] describes each point of a wavefront, usually visualized by the isolines of maximum amplitude (see Fig. 4.2 for example), as yet another point light source. The superposition of the single emitted spherical waves then forms the new wavefront. The same concept can be applied by considering each illuminated surface point as a new point source, emitting spherical waves with the same spectral properties. As an illustrative example, lets take a closer look onto Young's double slit experiment [LLL10]: A point light source is used to illuminate an aperture with two transmissive slits, whose width is in the order of the optical wavelength (a commonly used value is $1\,\mu$m). The transmitted light is observed on a screen behind the aperture, as illustrated in Fig. 4.2 (left). For simplification we assume that each of the two slits provides space for exactly a single point light source. This way, the wavefront incident from the left side spawns two sources at the slits' positions that now act as point light sources. These emit spherical waves into the space on the right side of the aperture. The corresponding waves interact with each other, forming a pattern that is visible on the screen. In Figure 4.2 (middle, right) we present two (illustrative) examples of such patterns. For an illumination with red light (wavelength of about 770 nm, middle) we obtain a characteristic diffraction pattern that shows alternating high- and low-amplitude stripes with the same wavelength. If we instead choose a white

light illumination (spectrum containing wavelengths in the range 400-800 nm, right), a single white stripe becomes visible, symmetrically enclosed by stripes of varying color, starting with red close to the white center, changing to green and then blue. This corresponds to the order of the respective wavelengths, starting with the highest (red) and ending with the lowest (blue), then repeating again. This can be explained by understanding white light as a superposition of all contained colors: The double slit creates a diffraction pattern similar to the one obtained for red light for every color, where the characteristic alteration between high and low amplitude is dependent on the respective wavelength. This way the observed separation of colors can take place.

In general, the interaction of light waves is called *interference*, where an enhancement of the amplitude due to interference is called *constructive interference*, the extinction *destructive interference* respectively. In addition to the specific wavelengths, the geometry of the scene is of importance: A lower distance between the two slits in the aperture for example leads to different points of destructive and constructive interference. The same holds for the position and orientation of both, screen and aperture.

4.2 Coherence

In most real world scenarios, global illumination plays a keyrole: Most shadows are not perfectly dark but rather consist of gradients due to interreflections within the scene. Similarly, point light sources are a means to test and validate lighting models. However, in more realistic scenarios, light sources do not necessarily illuminate the space uniformly and often have complex shapes.

Such *extended light sources* can be represented as an ensemble of point light sources, each emitting a spherical wave and their superposition forming the final wavefront of the extended light source. The single sources thereby do not necessarily emit their respective waves in phase, leading to additional effects that have to be considered. The degree of correlation between the phases of the point light sources contributing to a wavefront is qualified as *coherence*: A *coherent superposition* of waves takes place when all point sources have a constant phase difference. In contrast, if the phase difference is correlated but not constant, we speak of *partial coherence*. An *incoherent* light source on the other hand consists of point sources with perfectly random (uncorrelated) phase differences. As a consequence, destructive and constructive interference do not only depend on the geometry of the problem, but also on the degree of coherence of two or more light sources as well as their respective sizes, as larger light sources such as area lights are decomposed into point lights.

Figure 4.3 shows an illustration of the previously discussed double slit experiment (cf. Fig. 4.2) where we replaced the point- with an area-light source, which we consider to emit incoherent light. This forms a random wavefront that is incident on the aperture from the left hand side. Depending on the separation D of the two slits in the aperture, we can "select" points on the wavefront that are either close to each other or have significant distance and which now act as the new point light sources emitting light into space towards the screen. The closer the two slits are, the more correlated the points of the wavefront become, where the

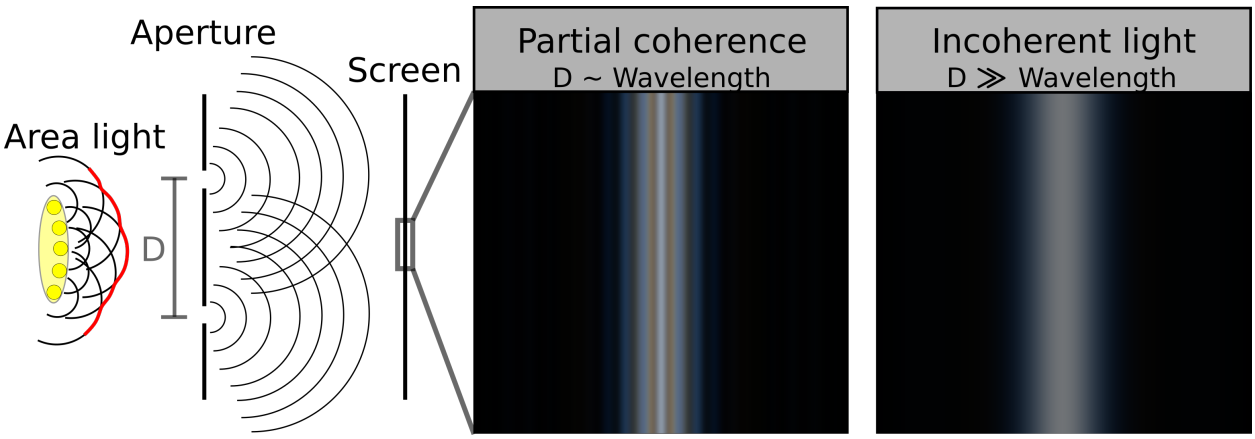

Figure 4.3: Simulated diffraction patterns for different distances of the two slits. In contrast to the point light scenario (Fig. 4.2), an area light source emits *incoherent light* due to its angular extent and can be represented as an ensemble of point light sources. The phase relation between parts of the wavefront, illustrated in red, are not constant. This results in *partially coherent* spherical waves originating from the aperture slits. Dependent on their separation D, the points of the incoherent wavefront forming the new point light sources at the slits are more or less correlated, resulting in blurred out (partially coherent) or even completely unordered (incoherent) superposition on the screen, removing the characteristic diffraction patterns.

exact same point, achievable with a single slit (the theoretical equivalent of both slits at the same position) has the maximum correlation. In this case, a perfectly coherent wave would impinge on the screen. For only small distances of the slits, the two spawned point light sources are still slightly correlated, leading to a partially coherent superposition on the screen. For illumination with white light, this results in a loss of contrast and color saturation, as constructive coherence is less dominant and the pattern is effectively "blurred". For even greater separation D between the two slits, color saturation is completely lost, as illustrated in Fig. 4.3 (middle, left), a phenomenon described as *spatial coherence*. As previously established, the microscopic scratches we consider form an ensemble of such light sources on the illuminated object's surface, obeying the concept of spatial coherence.

4.3 Properties of microscopic scratches

In contrast to the considerations related to Young's double slit experiment, the iridescent colors of scratches are observed due to the surface's reflection but not transmission. This places the observer (the screen) on the same side as the light source with respect to the aperture, which states a different problem than previously considered.

Following Babinet's theorem [BW13], the diffraction patterns of an aperture and its complement are equal. In particular this means that we can exchange two transmissive slits within an opaque aperture with only two opaque obstacles Fig. 4.4(a,b), effectively inverting the optical transmissivity, and obtain the exact same diffraction pattern in transmission. This in addition tells us, that to observe the diffraction pattern as a reflection on a screen, we can replace the transmissive slits with reflective elements within an otherwise opaque ab-

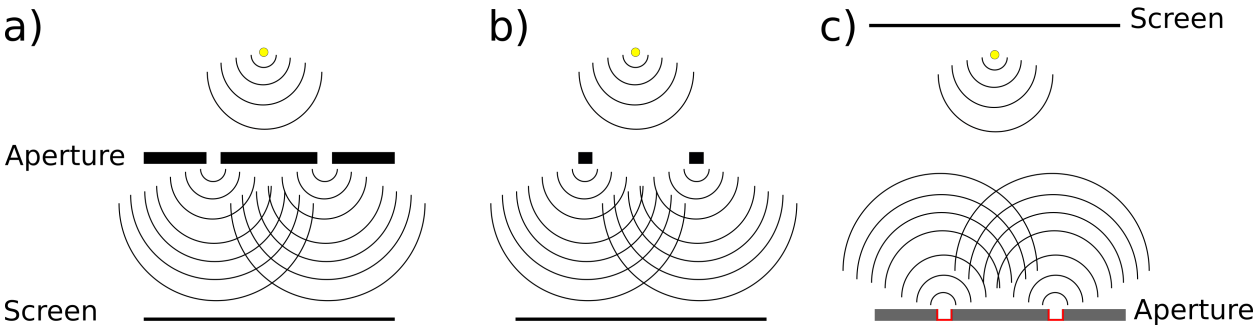

Figure 4.4: Babinet's theorem[BW13] states that the diffraction patterns of an aperture and its complement (inversion of optical transmissivity) equal each other. **a)**: In a standard transmission setup, an aperture containing two slits (c.f Fig. 4.2) forms a diffraction pattern that can be observed on a screen behind the aperture. **b)**: Following Babinet's theorm, the slits within the aperture can be replaced by only two opaque blocks of material, complementing the aperture of a). Because the light is diffracted around these obstacles, the same diffraction pattern is created. **c)**: Instead of transmission, scratches are observed in reflection. The aperture hence consists of an opaque, non-reflective material into which reflective scratches (red) are inserted. These now act as the slits in the transmission setup (see a)) and form the same diffraction pattern.

sorbing aperture Fig. 4.4(c). Being mostly of theoretical use, Babinet's theorem can hence be used to map the previously explained wave-optical principles, such as diffration and coherence, to the microscopic scratches: The cross section of a scratch can be seen as a reflective part within an otherwise non-reflecting opaque aperture(or rather, differently reflecting, as we will see in chapter 5.4). The digital recreation of a material's appearance by developing a reflectance model for surfaces that exhibit microscopic scratches is necessarily connected to the study of such materials using real-world examples. Unlike models in computer graphics that describe surface reflectance either by statistically representing surface characteristics or by explicitly calculating the light path a ray takes through features of the material, we have to be aware of two challenges our model has to solve: First, we aim to develop a physically-based model that explicitly describes microscopic scratches on both, worn and manufactured surfaces regarding their location and orientation. Available models that describe materials with diffracting surface features such as [Sta99; SFD$^+$00] rely on a statistical representation of the surface roughness, effectively creating random microscopic height variations according to a given distribution. These solutions are not feasible for our goal, as they are restricted to repeating patterns that are not found in scratches that cover multiple scales of extent in their lateral direction. On the other hand, available models that allow such detailed modeling of surface features [MDG01; BPM$^+$04; RGB16] do not incorporate wave optical effects, as they rely on geometrical optics considerations.

We approach these challenges by a preliminary analysis of scratch properties. Due to their size, these surface irregularities can not be observed with the naked eye, making the estimation of dimensions, such as length, width and depth a problematic task. For the very same reason, it is difficult to tell their shape. We start our investigation by the scratched metallic surface of a common kitchen pot. We utilize an optical microscope and capture the reflection of light under a strong directional illumination by a white LED.

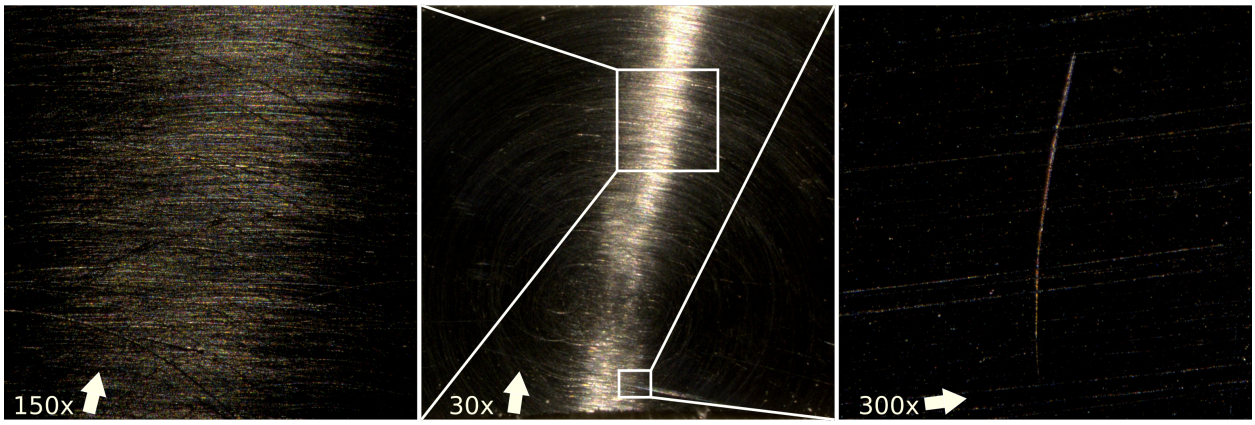

Figure 4.5: Images taken using an optical microscope with different magnification levels, indicated in the lower left of each panel. We also show the incident light direction projected onto the plane as the white arrow next to the zoom level. Note that while the left and middle panel share the same light direction, the right image shows a strongly magnified part at rotated incident light, revealing a single scratch that does not follow the dominant orientation.

Figure 4.5 shows an exemplary result we obtained from our microscopy analysis: The middle panel is a view of the scratched metal surface, observed at a magnification of 30 and under directional illumination according to the white arrow in the lower left. The surface is covered by scratches that follow a circular pattern around the center of the pot. Here one of the major properties of scratches is revealed: Scratches that are oriented perpendicular or close to perpendicular to the illumination direction show iridescent colors and reflect light, which we call the *specular condition*. Scratches whose orientation is rotated too far off the light direction on the other hand are barely visible. For further illustration, the left panel shows a close up part, where the reflecting scratch ensemble is interrupted by differently-oriented scratches, now visible as "black" scratches. Moving even closer, and rotating the direction of the incident light by almost 90° (right panel), we can identify single scratches that almost exactly cross the dominant direction, reflecting the light into the camera whereas the other scratches appear dark as they reflect the light according to their own orientation. The reflected light we perceive hence originates from the single scratch oriented perpendicular to the projected light direction, showing the characteristic colors that we expect due to diffraction.

Although being able to pinpoint single scratches, an optical microscope suffers from the fact that we actually observe the reflection of a scratch, which already encodes its local geometry: In analogy to the slits in Young's double slit experiment, we here expect the color to depend on the shape and size of the scratches' cross section. We are additionally confronted with another problem with respect to the visual analysis: The captured reflection covers a much larger amount of pixels, as the scratch itself would do, a limit we cannot surpass with optical methods.

To circumvent the reflection of the surface preventing analysis of the scratch geometry using a standard microscope, we utilize a raster electron microscope (REM), which allows us to visualize surfaces with extreme maginfication and without optical interactions. Due to the working procedure, the surface is always observed at an angle, making absolute mea-

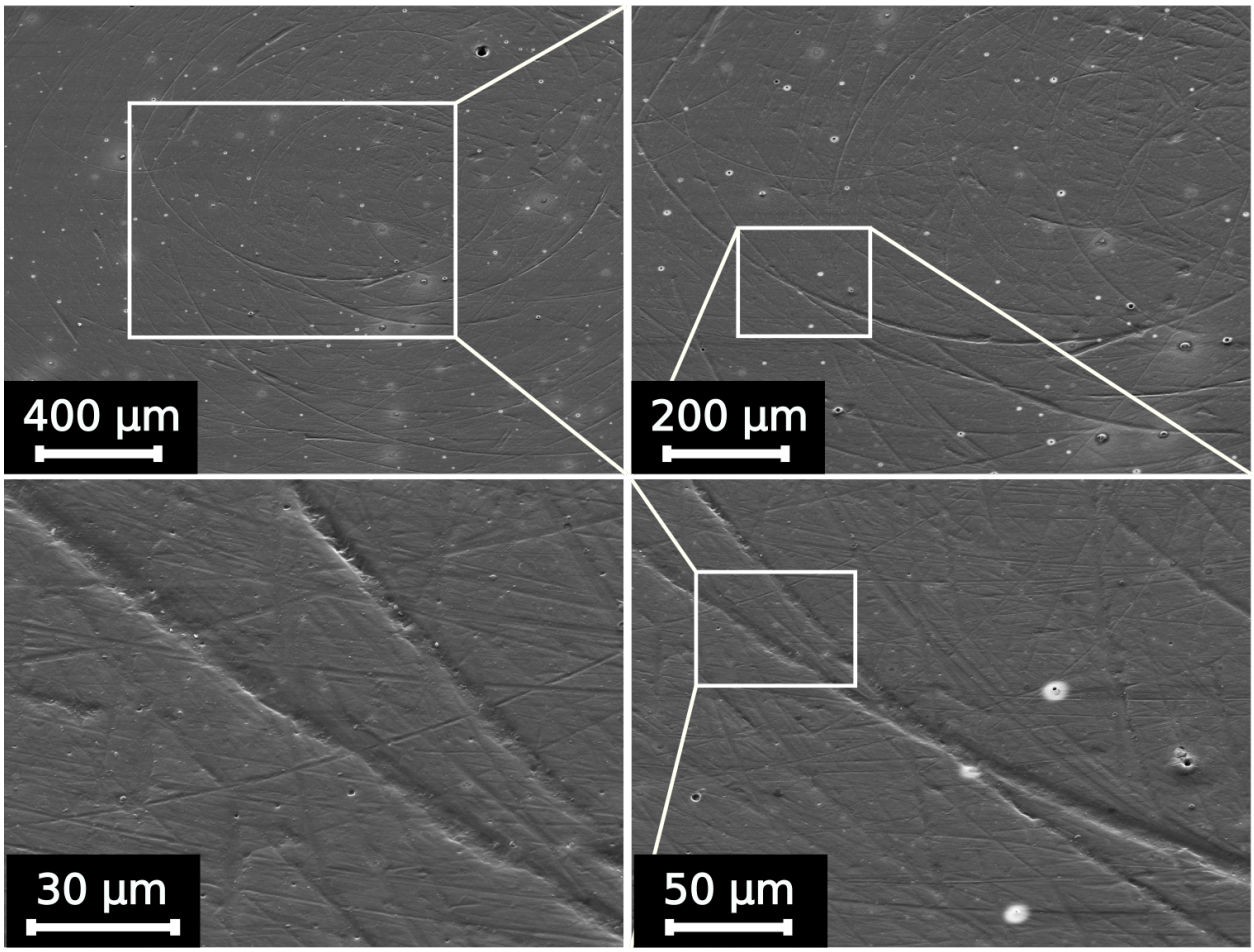

Figure 4.6: Raster electron microscope (REM) images obtained at 15 kV. In clockwise order, the images show the surface of a used cooking pot with increasing zoom level. With each increase in magnification, we obtain a clearer view of one of the largest scratches present on our sample, but also reveal that on each scale, there exist a multitude of even smaller scratches.

surements difficult, but allowing for good estimates of lateral and bi-lateral (cross section) extent of scratches. Figure 4.6 shows the REM images of a scratched metal surface at four different zoom levels, resolving features to a few micrometers. Here, larger scratches can now be imaged with sufficient detail, whereas for every magnification there appear to exist a plethora of yet smaller scratches, down to sub-micrometer extent. In principle, this indicates that surfaces such as these are fractal in nature, which poses an even greater challenge for the design of a digital reflection model. Scattering phenomena with surface features much smaller than the optical wavelength (around $0.4\,\mu$m in the lower limit) are described by Rayleigh scattering [You81; Bat84], which is negligible for our use case. Thus we restrict our considerations to scratches that, assuming their lateral extent to be much larger than the cross section, have a bi-lateral size of at least the lower limit of the optical wavelengths we consider. On the other hand, the largest scratches we observe in the REM images exhibit a cross section of up to at most $10\,\mu$m, forming the upper limit for later simulation and digital scratch generation. In addition, we can utilize the well resolved scratches in Fig. 4.6 to obtain an understanding of the scratch cross sections: Worn surfaces seldomly exhibit

scratches that are deeper than wide, which we attribute to the processes that create these scratches: Here, we observe surfaces that are exposed to wear, due to cleaning or use in combination with tools that exceed the scratches' dimension by many scales. The scratches are thereby generated by the tip of a fork for example or parts of a steel spoon during dish washing, carving only a minimum of material away, most likely in a v-shaped form or a tilted version thereof.

In contrast to material digitization, which is based on the measurement of materials, here we focus on the development of a physically-based diffraction model that can handle spatially varying microscopic scratches. Our model is aimed at high fidelity, enables the recreation of diffraction by surface irregularities in the order of the optical wavelength and can be driven with measured parameters if need be. As measurements itself are hard to achieve, our presentation relies on scratch properties we extract from optical observations and REM measurements, forming the framework within which we operate to deliver a physically plausible digital double of worn surfaces. In particular, our model relies on the following design choices extracted from our observations:

- A scratches' lateral extent is much longer than wavelength, not causing diffraction.
- The orientation of a scratch is given by its lateral direction on the surface plane.
- The bilateral extent is in the order of the optical wavelength, but not smaller than the lower limit of wavelengths considered (Rayleigh scattering).
- We neglect scratches with small (below Rayleigh limit) bilateral extent, circumventing the need to model every single scratch present on a surface, which would be problematic due to the fractal nature of the representation.

CHAPTER 5

Scratch Iridescence

Figure 5.1: Left: Inside of a cooking pot photographed under a halogen spotlight. Center and right: Renderings of scratched materials under environment lighting obtained using our model.

Abstract The surface of metal, glass and plastic objects is often characterized by microscopic scratches caused by manufacturing and/or wear. A closer look onto such scratches reveals iridescent colors with a complex dependency on viewing and lighting conditions. The physics behind this phenomenon is well understood; it is caused by diffraction of the incident light by surface features on the order of the optical wavelength. Existing analytic models are able to reproduce spatially unresolved microstructure such as the iridescent appearance of compact disks and similar materials. Spatially resolved scratches, on the other hand, have proven elusive due to the highly complex wave-optical light transport simulations needed to account for their appearance. In this paper, we propose a wave-optical shading model based on non-par-axial scalar diffraction theory to render this class of effects. Our model expresses surface roughness as a collection of line segments. To shade a point on the surface, the individual diffraction patterns for contributing scratch segments are computed analytically and superimposed coherently. This provides natural transitions from localized glint-like iridescence to smooth BRDFs representing the superposition of many reflections at large viewing distances. We demonstrate that our model is capable of recreating the overall appearance as well as characteristic detail effects observed on real-world examples.

5.1 Introduction

Modelling and rendering the real world with all its irregularities and imperfections remains one of the greatest challenges in computer graphics. A rich history of research on dirt, dust and fingerprints, weathering, patination, erosion [DRS10] and scratches [MDG01; BPM+04; DWM+15; YHM+16; RGB16] documents the massive amount of effort invested by our community to make computer graphics look less sterile and more realistic. In this paper, we focus on a subtle but very common effect observed on objects made of metal, glass or plastic. Under strongly directional lighting (like sunlight or a halogen spot), these surfaces exhibit colorful patterns that are caused by diffraction of light reflecting off microscopic surface details (Fig 5.1). Being fundamentally a wave-optical phenomenon, this effect cannot be reproduced by geometric optics models and requires a careful study of both diffraction by individual surface features at the microscopic scale as well as interference among multiple features. The most detailed wave-optical simulations conducted in computer graphics thus far have involved detailed finite-difference time-domain (FDTD) modeling of periodic microstructures on butterfly wings [MMR+13], an approach that is not feasible for structures beyond a few cubic micrometers.[1] In other works, the appearance of large-scale diffractive objects has been approximated by combining far-field scattering models for repetitive microscopic structures with traditional texturing approaches. Most objects we interact with in our everyday use, however, exhibit features across many scales, ranging from macroscopic ones that are resolvable with the naked eye to microscopic ones that are only indirectly visible due to their aggregate interaction with light. This complexity leads to an intricate variation of appearance along the spatial, angular and spectral dimensions that no model so far has been able to express.

We propose a new spatially varying bidirectional reflectance distribution function (SVBRDF) modeling surfaces covered with microscopic scratch particles, which enables us to simulate scratch iridescence and contributes a first step to providing a framework for rendering such phenomena. Simulating the appearance of scratches within modern physically based rendering systems involves a number of challenges; for instance, traditional analyses of the diffracted wave field rely on a paraxial (small-angle) assumption that would lead to grossly inaccurate results in a BRDF model that must support evaluation for any pair of incident and outgoing angles. Our model thus simulates the diffraction of light by microscale surface features using a *non-paraxial* scalar diffraction theory proposed by Harvey [HVK+00]. Our formulation leads to an SVBRDF with naturally coupled spatial, angular and spectral variation that exhibits multi-scale behaviour: at a distance, interference from a larger number of scratches causes it to resemble standard geometric optics models, while larger magnifications reveal iridescent reflections from small collections of scratches.

To model and simulate rough and scratched surfaces, we propose a vector graphics representation where an ensemble of fundamental primitives (linear scratch segments) resides on a base substrate. We show how the constituent scattering distributions can be expressed in closed form, and discuss automated and interactive techniques for placing large numbers of scratches on surfaces. Finally, we integrate our model into a modern physically-based ren-

[1] For instance, FDTD simulation of light waves in a $1\,\mathrm{mm}^3$ volume would require 2.4×10^{28} cell updates per second of simulated time ($\lambda = 500\,\mathrm{nm}$, grid resolution$\lambda/10$, time resolution $T/10$).

dering system and discuss efficient implementations of key operations, including importance sampling.

5.2 Related work

Detailed modeling and rendering of surface defects can dramatically improve the realism of renderings, hence the pursuit of such models has been a topic of great interest to the rendering community at large [DRS10]. Early work in this area includes methods by Buchanan and Lalonde [BL99] and Lu et al. [LKK00] who analyze general reflection properties of scratches. We group more recent related work into three main categories:

Explicit geometry Merillou et al. [MDG01] and Bosch et al. [BPM+04; BPM+08] derive BRDFs from the scratch cross-section (profile) at each shading point. Merillou et al. use a preset profile and texture mapping to position scratches, with scratch profiles split up into a number of tilted surfaces with associated procedural BRDFs. Bosch et al. enhance this model with generalized profile representations and curves on the surface of shaded objects to position scratches. Raymond et al. [RGB16] propose a multi-scale SVBRDF model based on a stack of coherently oriented scratch layers; their model relies on an accurate solution of interreflection within a scratch and supports multi-scale evaluation. All of these approaches separate spatial and optical information concerning the scratches, and their solutions only hold in the geometrical optics regime.

Microfacet models BRDF models based on microfacet theory are widely used in graphics and have proven effective in reproducing the appearance of real-world materials [NDM05]. Here, we only focus on models that specifically target rendering of scratched surfaces. Yan et al. [YHJ+14; YHM+16] numerically integrate the normal distribution function (NDF) of normal-mapped surfaces over the surface region observed within a single pixel, which yields an efficient multi-scale reflectance model capable of rendering high-resolution normal maps under directionally peaked illumination. Dong et al. [DWM+15] compare the use of microfacet and Kirchhoff scattering theory to predict surface appearance from the measured microgeometry of small surface patches. Also related to our work are image-based techniques that fit microfacet-based anisotropic reflectance models to measurements of finished wood [MWA+05] or brushed metal surfaces [WZT+08; DWT+10].

Diffraction Modelling diffraction by rough surfaces has been of great interest to the physics community and various models addressing the different characteristics of surfaces have been developed, ranging from Rayleigh-Rice vector perturbation theory (smooth surfaces) to more general ones such as Beckmann-Kirchhoff scattering theory (various roughness classes). A good overview as well as extensions of these scattering theories can be found in Krywonos [Kry06] .

In the computer graphics community, a variety of BRDF models has been developed to account for diffraction effects created by microscale surface features in ray-based frameworks. However, none of them targets the transition between texture and far-field diffraction that is needed for our purpose. One of the first to incorporate wave-optical scattering theory into

a BRDF model were Church and Takacs [CT09]. Löw et al. [LKY+12] later introduced their model to the graphics community and demonstrated its merits in numerical fits to measured reflectance data. He et al. [HTS+91] derived a BRDF model based on vector Kirchhoff theory for surfaces with roughness described by a Gaussian random process. Stam [Sta99] proposed a BRDF model based on scalar Kirchhoff theory, that is capable of rendering the diffraction effects of randomly distributed primitives (e.g. rectangular bumps) or Gaussian random surfaces. The method relies on the power spectrum of the autocorrelation function of the surface height variations. Sun et al. [SFD+00] derive an accurate far-field diffraction model to render compact discs modeled as a series of concentric tracks with a periodic arrangement of pits.

Cuypers et al. [CHB+12] propose a Wave Bidirectional Scattering Distribution Function based on statistical optics, computing the Wigner distribution function of microstructures to produce solutions that are valid in the near- and far-field. This approach heavily relies on analytical solutions for regular structures, and it does not generalize to complex non-periodic microstructures due to the exceedingly high memory requirements associated with the underlying four-dimensional representation. More recently, Dhillon et al. [DTS+14] developed a diffraction model based on heightfields acquired using an atomic force microscope; this data was used to generate look-up tables by truncating a Taylor-series expression of the BRDF. To better reproduce surface scattering effects based on statistical properties of heightfields, Holzschuch and Pacanowski [HP17] introduced the generalized Harvey-Shack theory to the computer graphics community. Lately, Belcour and Barla [BB17] extended microfacet-based models to recreate thin-film interference. To acquire and render diffraction effects from planar surfaces at small angles, Toisoul and Ghosh proposed an efficient framework [TG17]. Musbach et al. [MMR+13] proposed a reflectance model for iridescent biological structures based on detailed FDTD simulation of vectorial wave propagation. This approach is significantly more general than the previously discussed models, but the prohibitive cost of this type of simulation limits it to periodic structures. Levin et al. [LGX+13] use scalar Kirchhoff theory to predict BRDFs of a specific class of lithographically structured surfaces. While only remotely related to our work, their approach is one of the few to consider the effects of spatial coherence of the illumination source.

To overcome the discussed restrictions in the context of rendering scratched surfaces, our approach builds on an efficient representation tailored to this application. Similar to prior work [MDG01; BPM+04], we separate spatial and optical information by describing the scratch layout as a curve and its reflectance behavior using a profile at each position along the curve. We use non-paraxial scalar diffraction theory [HVK+00] to express the diffracted reflectance as a superposition of reflections from individual scratches. Similar to Sun et al. [SFD+00] we derive the BRDF from the explicit calculation of the scattered complex wavefront, maintaining as much generality as possible. This allows us to take into account spatial coherence to reproduce not only diffraction effects but also the mutual interference created by dense scratch ensembles.

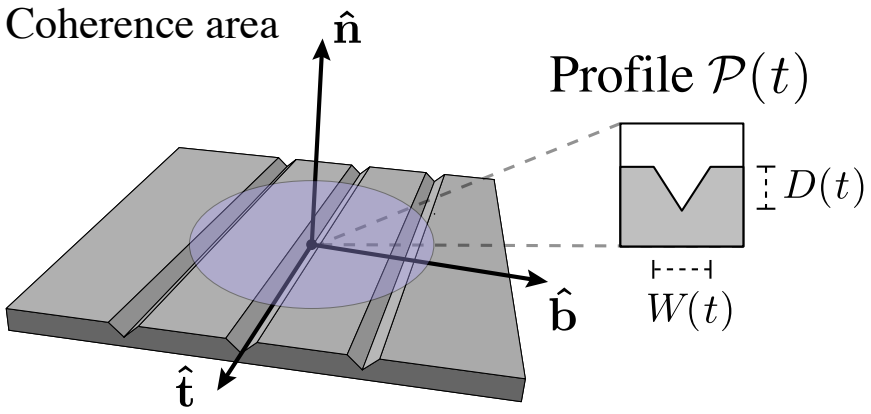

Figure 5.2: Local shading geometry: Scratches that lie within the coherence area contribute to the diffracted radiance. We represent each scratch by a parametric curve $\boldsymbol{r}(t)$, which brings its own local coordinate system with tangent and bitangent directions $\hat{\boldsymbol{t}}(t)$ and $\hat{\boldsymbol{b}}(t)$, respectively. The cross-section at any position t along the scratch is defined by a profile $\mathcal{P}(t, b)$. We use profiles that lead to analytical Fourier transforms and scale them in the $(\hat{\boldsymbol{b}},\hat{\boldsymbol{n}})$-plane using parameters W (width) and D (depth)

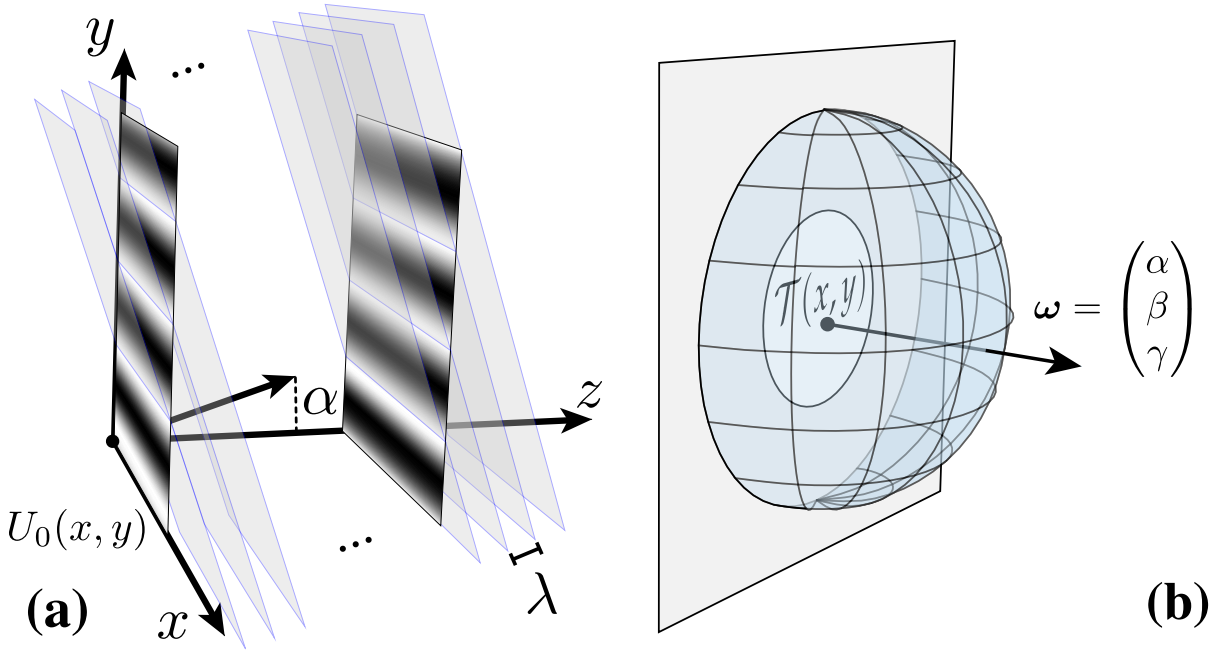

Figure 5.3: (a), Angular spectrum. Viewed along slices perpendicular to the z-axis, a monochromatic plane wave traveling in direction $(\alpha, 0, \sqrt{1-\alpha^2})$ causes vertical oscillations with frequency α/λ; translating the slicing plane incurs a corresponding phase shift. Using the Fourier transform, this relation can be used to express an arbitrary field $U_0(x, y)$ incident at $z = 0$ as a superposition of plane waves arriving from different directions. (b), We are interested in the far-field diffracted radiance $L(\boldsymbol{\omega})$, which is proportional to the squared amplitude of the plane wave traveling in the same direction.

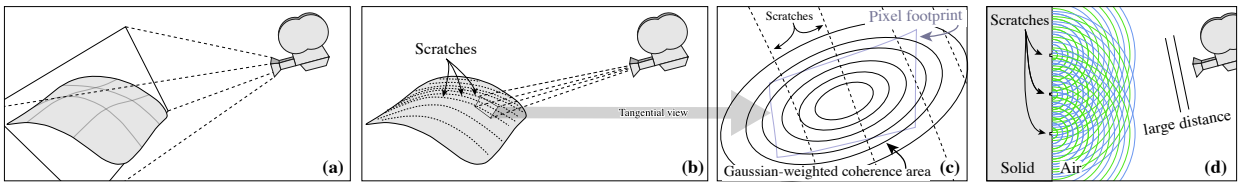

Figure 5.4: From scratches to diffracted radiance: **(a)** Macroscopic view of a surface containing scratch particles. **(b)** Ray differentials establish a mapping between the visible surfaces and pixels of the output image. **(c)** Our method considers scratches located inside a Gaussian-weighted coherence region around the surface region covered by a pixel. **(d)** Side view: each scratch in the coherence region scatters the incident light into wavelets that interfere with each other; although not shown here, our method also accounts for reflection from the non-scratched base surface. Since the distance to the camera is much greater than the wavelength of light, it is enough to consider diffracted radiance, a far-field approximation of the superposition of wavelets that remains accurate for all angles of observation.

5.3 Preliminaries

This section introduces the notation, geometric framework and relevant theory from prior work that serve as the foundation of our model. An overview of notation can be found in Table 5.1.

5.3.1 Scratch representation

We represent a scratch as a curve $\boldsymbol{r}(t)$ parameterized by the arc length t, with local tangent vector $\hat{\boldsymbol{t}}(t) = d\boldsymbol{r}(t)/dt$. At any location t along the scratch, the surface normal $\hat{\boldsymbol{n}}(t)$, tangent $\hat{\boldsymbol{t}}(t)$ and bitangent $\hat{\boldsymbol{b}}(t) = \hat{\boldsymbol{n}}(t) \times \hat{\boldsymbol{t}}(t)$ form an orthonormal coordinate frame. The geometric cross-section of the scratch at position t is defined by the *profile* $\mathcal{P}(t, b)$, which specifies a 1-dimensional height profile along the bitangential coordinate b. Our model relies on scratch profiles that have analytical 1D Fourier transforms, such as rectangle or triangle shapes. Their scale in the $(\hat{\boldsymbol{b}},\hat{\boldsymbol{n}})$-plane is given by the width and depth parameters $W(t)$ and $D(t)$, respectively (Fig. 5.2).

5.3.2 BRDF model

To quantify the interaction of light with a surface exhibiting microscale defects in physically-based rendering frameworks, we begin with the well-known definition of the bidirectional reflectance distribution function (BRDF)

$$f_r = \frac{dL_\mathrm{s}(\boldsymbol{x}, \hat{\boldsymbol{\omega}}_o)}{dE_i(\hat{\boldsymbol{\omega}}_i)}, \tag{5.1}$$

which relates differential irradiance to scattered radiance. Here, $\boldsymbol{x}$ represents a position on the surface, $\hat{\boldsymbol{\omega}}_i$ is the direction from which this surface is illuminated and $\hat{\boldsymbol{\omega}}_o$ is the direction from which it is observed. To compute the radiance scattered by a diffracting aperture, we rely on a far-field approximation for scalar diffraction theory, known as *diffracted radiance*, which was introduced by Harvey et al. [HVK+00]. It builds on tools from Fourier

Table 5.1: Overview of the notation used in this paper. The references point to the location where the quantity is introduced.

Symbol	Meaning	Reference
$\boldsymbol{x}, \hat{\boldsymbol{x}}$	Vector, unit vector	
$\hat{\boldsymbol{\omega}}_i, \hat{\boldsymbol{\omega}}_o$	Incident and outgoing light directions	
Spatial parameterization		
$\boldsymbol{x}=(x,y)$	Position on z-plane	
$U(\boldsymbol{x})$	Scalar field amplitude	Sec. 5.3.2
$U_0(x,y)$	Scalar field in plane $z=0$	Fig 5.3a, Eq 5.2
Plane-wave parameterization (spatial-frequency domain)		
(α,β,γ)	Vector of direction cosines	Fig 5.3b, Sec. 5.3.2
$\boldsymbol{\xi}$	$(\hat{\boldsymbol{\omega}}_o - \hat{\boldsymbol{\omega}}_i)/\lambda$	Sec. 5.3.2
$\boldsymbol{\xi}'$	Projection of $\boldsymbol{\xi}$ in scratch frame	Eq 5.11,5.13
Scratch representation		
t	Position along the scratch	
$\boldsymbol{r}(t)$	Point on scratch at position t	
$\{\hat{\boldsymbol{t}}, \hat{\boldsymbol{n}}, \hat{\boldsymbol{b}}\}(t)$	Local scratch coordinate frame (tangent, normal, bitangent) at t	Sec 5.3.1, Fig 5.2
b	Bitangential coordinate	
$\mathcal{P}(t,b)$	Scratch profile at t as function of b	
$W(t), D(t)$	Scratch width and depth at t	
η	Spatial phase integral	Eq 5.12, 5.23
$k, (k)$	Scratch index (for ensemble summation)	Eq 5.9
λ	Optical wavelength	
$\mathcal{G}$	Gaussian spatial filter	Fig. 5.4c, Sec. 5.4.1
δ_c	Coherence area diameter, $\delta_c = 60\mu m$	Sec. 5.4.1
$\mathcal{T}(\boldsymbol{x})$	Optical transfer function	Fig 5.3b, Sec. 5.3.2
A_s	Shading area	Eq 5.2

optics [Goo96], specifically the *angular spectrum*, and is accurate in the far field and for diffracting apertures that are significantly larger than the wavelength of the radiation (both, far-field and aperture condition are satisfied by our application). Being part of a *scalar* theory of light transport, these two tools assume that the electromagnetic field can be described by the (scalar) amplitude of oscillations, as opposed to the commonly used vectorial electric and magnetic fields. Let $U(x,y,z)$ denote the scalar amplitude of the electric field at position $(x,y,z)^T$, and let $U_0(x,y) \coloneqq U(x,y,0)$ denote a planar slice at position $z=0$ (here called the aperture plane, cf. Fig. 5.3). Harvey shows that the radiance diffracted by the aperture is obtained by computing $U(x,y,z)$ for $z>0$, which is given by a Fourier transform of the complex amplitude $U_0(\boldsymbol{x})$. Assuming a homogeneously illuminated diffraction aperture, the diffracted radiance can be written as

$$L(\hat{\boldsymbol{\omega}}_o, \alpha_i, \beta_i) = \gamma_i \frac{\lambda^2}{A_s} \left| \mathcal{F}\left\{ U_0(x,y)\, e^{2\pi i(\beta_i y + \alpha_i x)} \right\}_{\alpha,\beta} \right|^2, \tag{5.2}$$

where $\hat{\boldsymbol{\omega}}_o = (\alpha,\beta,\gamma)$, $\hat{\boldsymbol{\omega}}_i = (\alpha_i,\beta_i,\gamma_i)$ and A_s being the shading area. Importantly, this expression remains valid even for oblique angles of incidence. The variables of this param-

Table 5.2: References to assumptions used by our model

Assumption	**Reference**
Diffracted radiance	
Far-field scattering	Sec. 5.3.2; Eq. 5.2
Spatial coherence	
Equal coherence condition for the full scene	Sec. 5.4.1
Intensity drops off towards edges of a light source	Sec. 5.4.1
Surface representation	
Homogeneous base material	Sec. 5.4.2; Eq. 5.9
Scratches	
No intersections (Sum of transfer functions)	Eq. 5.9
No self-shadowing/masking	–
Linear segments with constant profile	Sec. 5.4.3; Eq. 5.11
Spatial–spectral separability	Sec. 5.4.3; Eq. 5.11

eterization are referred to as *direction cosines* with $\gamma = \sqrt{1-\alpha^2-\beta^2}$. Spatial coordinates in the above expression are expressed in units of the wavelength λ. The complex wavefront $U_0(\boldsymbol{x})$ can be described by the modulation of the wavefront of the incident light $U_i(\boldsymbol{x})$ with the so-called *transfer function* $\mathcal{T}(\boldsymbol{x})$ [LLL10; Goo96] of the diffracting plane:

$$U_0(\boldsymbol{x}) = U_i(\boldsymbol{x}) \cdot \mathcal{T}(\boldsymbol{x}). \tag{5.3}$$

Taking into account the assumption of homogeneous illumination ($U_i(\boldsymbol{x}) = U_i$), a change of variables enables us to rewrite the representation of the Fourier transform in a non-scaled coordinate system, which yields the BRDF

$$f_r(\boldsymbol{\xi}) = \gamma_i \frac{1}{A_s} \frac{1}{\lambda^2} \left| \mathcal{F}\left\{ \mathcal{T}(\boldsymbol{x}) \right\}_{\xi_{1,2}} \right|^2, \tag{5.4}$$

with

$$\boldsymbol{\xi} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \frac{1}{\lambda} \begin{pmatrix} \alpha - \alpha_i \\ \beta - \beta_i \\ \gamma - \gamma_i \end{pmatrix}. \tag{5.5}$$

A detailed explanation of these steps is provided in the supplemental material.

Equation 5.4 is the non-paraxial spectral BRDF for reflected light that is diffracted by a surface exhibiting microscale features represented by an optical transfer function $\mathcal{T}(\boldsymbol{x})$. Diffracted radiance shares similarities with the diffraction BRDF proposed by Stam [Sta99, Eq 7]. Stam's model explicitly utilizes Kirchoff theory (i.e. tangent-plane approximation and the Huygens principle) whereas diffracted radiance can be derived solely from the angular spectrum of plane waves [Kry06]. However, both approaches offer equivalent formulations for the far-field approximation at hand (i.e. homogeneous illumination). For our application to render spatially resolved scratches, we rely on diffracted radiance as it provides a convenient way to describe the amplitude and phase changes induced by light-surface interaction via the optical transfer function.

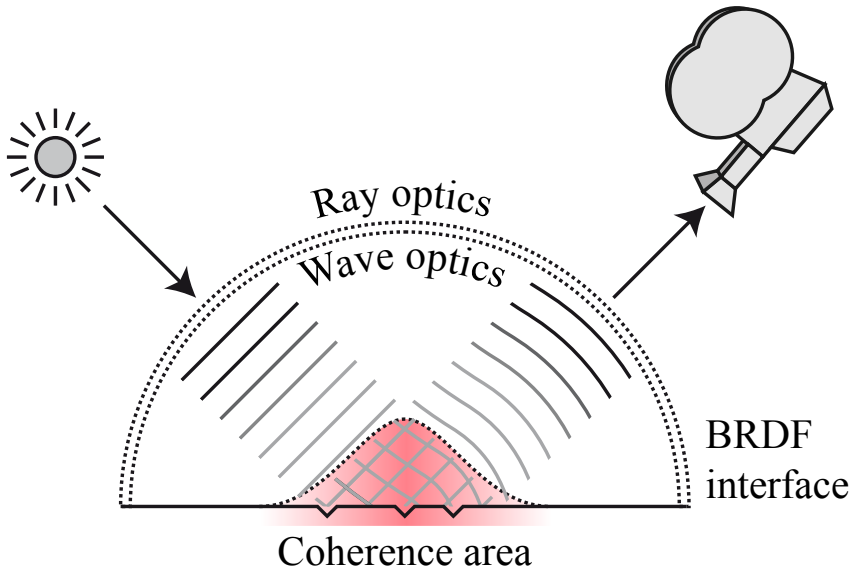

Figure 5.5: To be compatible with a standard global illumination renderer, our model constrains wave optics to the reflectance model. The coherence area, represented by the Gaussian filter $\mathcal{G}$, marks the portion of the surface where scattered light waves will superimpose coherently. The interface to the outside world is provided in terms of geometric optics and radiance units. The computation of the coherence area is explained in Sec. 5.4.1

5.4 A diffraction SVBRDF for scratched surfaces

The concept of ray tracing is fundamentally incompatible with the basic principles of wave optics. In fact, the wave-optical counterpart of a "ray" with sharply defined direction is a plane wave of infinite lateral extent and equal amplitude at each point on the sufficiently far away illuminated surface. Following Eq. 5.4, this leads to a Fourier transform of the whole surface regardless of its extent. Such an approach would be incapable of resolving localized surface features and would only be correct for point light sources or collimated beams. Since our goal is to create a model that permits resolving individual scratches affected by mutual interference, we draw on coherence theory to develop a physically justified interface between the long-range (ray-optical) light transport within a path tracing framework and the wave-optical scattering model. To this end, we will first introduce the concept of spatial coherence and use it to define a coherence window that will lead to a spatially varying BRDF (SVBRDF). We will then study the wave-optical contribution of a single scratch in isolation, and finally look at large ensembles of scratches.

5.4.1 Spatial coherence: resolving scratches

The van Cittert–Zernike theorem relates the angular extent of a light source illuminating a surface with a spatial filter on the surface via a Fourier transform [Goo96; LLL10]. For instance, a point-light source yields a constant infinite spectrum whereas a disk-shaped area light results in an Airy function: this *coherence function* defines the corresponding spatial weights. More intuitively, it defines which structures in the vicinity of the observed point, such as scratches, actually contribute to the wave-optical scattering that leads to diffraction

and interference. Usually, the distance between the first zero-crossings of the coherence function is used to define the *coherence area*, the extent of the filter. In the context of astronomy, a circular star of diameter $d = 0.07$ arcseconds e.g. has a coherence area of radius $r_c = 1.22\lambda/d \approx 1.9$m [LLL10]. After this first zero-crossing, the coherence function may predict further areas of coherence, which are suppressed if the light source is less intense around its edges. To approximate the coherence function, we make two simplifications: first, we assume that the same coherence condition holds throughout the scene, which allows us to define a global coherence function. Second, we will neglect coherence effects outside of the central peak, implicitly assuming a fall-off of the light-source intensity towards its corners. Similar to Dhillon et al. [DTS+14], we define the coherence function as a spatial kernel $\mathcal{G}(\boldsymbol{x} - \boldsymbol{x}_0)$ controlling the relative weight of points on the surface with respect to the intersection point $\boldsymbol{x}_0$ of the ray on the surface. In contrast to Dhillon et al., who primarily introduce $\mathcal{G}$ to reduce the cost of discrete Fourier transformations used by their technique, our model requires coherence for its ability to resolve spatial surface structure. We model objects covered with an irregular arrangement of scratches, of which only a subset contributes to the light scattered within a pixel. Similar to a short-time Fourier transform, the kernel $\mathcal{G}$ provides a natural windowing function that performs this selection. This is a substantial difference to prior work modeling regular structures [Sta99; DTS+14]. We use the spatial kernel

$$\mathcal{G}(\boldsymbol{x}) = e^{-\frac{1}{2}|\boldsymbol{x}-\boldsymbol{x}_0|^2/\sigma^2}, \tag{5.6}$$

which is an isotropic Gaussian in the tangent plane around $\boldsymbol{x}_0$ (see Fig. 5.5). The coherence area diameter containing the salient portion of the footprint is defined as $\delta_c = 6\sigma$. We now modify Eq. 5.4 to account for spatial coherence:

$$f_r(\boldsymbol{x}_0, \boldsymbol{\xi}) = \gamma_i \frac{1}{A_s}\frac{1}{\lambda^2}\left|\mathcal{F}\left\{\mathcal{T}(\boldsymbol{x}) \cdot \mathcal{G}(\boldsymbol{x} - \boldsymbol{x}_0)\right\}_{\xi_{1,2}}\right|^2. \tag{5.7}$$

A common literature value for δ_c for scenes under direct illumination by sunlight or a light bulb in a kitchen environment [MW95; DN15] is $\delta_c = 60\,\mu$m, which we choose to approximate such illumination situations. Note that Eq. 5.7 behaves as previously discussed: for a point-light source, $\delta_c \to \infty$ and $\mathcal{G}(\boldsymbol{x}) \to const.$, reducing the SVBRDF to the BRDF of Eq. 5.4, which does not resolve surface features. On the other hand, an infinitely extended light source yields $\delta_c \to 0$, i.e. a Dirac delta coherence function.

In a practical rendering system, the intensity of a pixel in the rendered image will generally be computed as part of a stochastic ray tracing process, which samples the pixel footprint with a spatial reconstruction filter centered at $\boldsymbol{x}_0$. Our implementation converts these pixel sub-samples to real-valued radiance before averaging them, which maps to spatial coherence as the coherent subsamples are first converted to radiance and then incoherently superposed [LGX+13].

5.4.2 Additive composition of transfer function

We now turn to the Fourier transform of the optical transfer function, a key part of our model that we utilize to describe amplitude and phase changes of incident light waves induced by height variations of the shaded surface. We define $\mathcal{T}(\boldsymbol{x})$ as

$$\mathcal{T}(\boldsymbol{x}) = A(\boldsymbol{x}) \cdot e^{i\phi(\boldsymbol{x})} = \sqrt{F(\boldsymbol{x})} \cdot e^{i2\pi(\gamma_i+\gamma)h(\boldsymbol{x})/\lambda} \tag{5.8}$$

where $A = \sqrt{F}$ is the amplitude factor of the surface material, F is the Fresnel factor and $\phi(\boldsymbol{x}) = 2\pi(\gamma_i + \gamma_o)h(\boldsymbol{x})/\lambda$ is the change of phase induced by the height variation of the microstructure. As the optical path traveled by rays depends on the incident and outgoing angle with respect to the surface normal, the optical path difference (and thus the phase change) exhibit a dependence on the view and light direction cosines γ_o and γ_i, respectively [Kry06; HP17]. We now apply this concept to scratched surfaces, where each scratch is described by its individual transfer function. Here, the transfer functions encode local surface height variations, which change the phase of incident light waves that interact with the scratches. We assume our materials to consist of a homogeneous *base material* that does not exhibit spatial variation. This base material exhibits defects (scratches) at distinct positions. We model this by defining our transfer function as a base transfer function, from which we first subtract *masks* covering the defects and then add the defects back at the same position. This implies that masks and scratches must cover the same area on the surface. More formally,

$$\mathcal{T}(\boldsymbol{x}) = \mathcal{T}_{\text{base}}(\boldsymbol{x}) - \sum_k \mathcal{T}^{(k)}_{\text{mask}}(\boldsymbol{x}) + \sum_k \mathcal{T}^{(k)}_{\text{scratch}}(\boldsymbol{x}), \tag{5.9}$$

where the superscript (k) denotes the k^{th} mask-scratch pair and $\mathcal{T}_{\text{base}}(\boldsymbol{x}) = A_{\text{base}} = \sqrt{F_{\text{base}}}$. This decomposition of surface structure into individual scratches, and the additive superposition of their contributions is the key to a practical implementation of our model, since it allows for an efficient analytical evaluation of the Fourier transform in Eq. 5.4. The analytic evaluation also depends on two further simplifications: first, we assume the base to be a perfectly flat mirror, hence surface roughness would need to be emulated by an intractably dense distribution of scratches. Sec. 5.5 introduces an alternative mask-based blending scheme that combines our model with existing rough BRDFs based on microfacet theory. Secondly, we neglect scratch-scratch intersections, and our model consequently slightly overestimates the surface area of scratched portions of a surface. In a region where two scratches overlap, the base contribution will be subtracted (masked) twice and replaced by the sum of two scratches. For a correct handling of intersections, the full transfer function would be generated by a sum of height fields in the complex phase instead, which would not allow for our analytic solution. As we show in Sec. 5.6, the effect of this approximative handling of intersections can lead to a re-distribution of energy from the specular reflection towards higher order diffraction lobes. Note that as long as scratches and base have the same amplitude factor, this error only affects the diffracted phase but does not violate energy conservation. We provide an overview of all assumptions and simplifications used by our model in Table 5.2.

5.4.3 Single-scratch transfer function

The linearity of the Fourier transform allows us to first consider a single scratch transfer function (or its transform) and later extend the concept to a full solution. The local geometry of each scratch is defined by the profile $\mathcal{P}(t, b)$ (Sec. 5.3.1, Fig. 5.2), which is a 1-dimensional transverse scratch height profile defined in terms of the bitangent coordinate b. The resulting scratch-space optical transfer function takes on the following form in local coordinates:

$$\mathcal{T}_{\text{scratch}}(b, t) = A_{\text{scratch}}(b, t) \cdot e^{i2\pi(\gamma_i+\gamma)\mathcal{P}(b,t)/\lambda}, \tag{5.10}$$

where $A_{\text{scratch}}(b,t) = \sqrt{F(b,t)}$. The integral of the transfer function along the scratch length yields a spatial representation of the amplitude and phase changes induced by the material's height variations.

Incorporating the scratch profile. Considering its shift and rotation properties, the Fourier transform of a scratch transfer function can be understood as the integral over the Fourier transform of the rotated and shifted 1D transfer functions along the scratch trajectory, whose intricacy is determined by profile variation and curvature of the scratch. For simplicity, we consider scratches consisting of linear scratch segments whose profiles do not change along the segment. There is no loss of generality, as arbitrary curves of varying profile can always be split into linear segments of constant profile (and thus having a constant transfer function).

In addition, we assume only the spatial phases to be affected by the Gaussian filter, i.e., the width of a scratch is negligible in comparison to its length with respect to the coherence area. This allows us to separate the spatial (position) and spectral (profile) components of each scratch and to express its transfer function in the scratches' own tangent space with axes $\hat{\boldsymbol{t}}$, $\hat{\boldsymbol{b}}$ and $\hat{\boldsymbol{z}}$. The Fourier transform of a single scratch transfer function then reads

$$\begin{aligned} &\mathcal{F}\left\{\mathcal{T}_{\text{scratch}}(b,t)\right\}_{\xi'_{1,2}} \approx \\ &\mathcal{F}\left\{\mathcal{T}_{\text{scratch}}(b)\right\}_{\xi'_2} \cdot \int dt\, \mathcal{G}(t)\, e^{-2\pi i(\boldsymbol{r}'(t)\cdot\boldsymbol{\xi}')} = \\ &\left[\int db\, \mathcal{T}_{\text{scratch}}(b) e^{-2\pi i b \hat{\boldsymbol{r}}'_2 \xi'_2}\right] \cdot \eta \end{aligned} \tag{5.11}$$

where

$$\begin{aligned} \eta &= \int dt\, \mathcal{G}(t)\, e^{-2\pi i(\boldsymbol{r}'(t)\cdot\boldsymbol{\xi}')} \\ &= \int dt\, e^{-|\boldsymbol{r}'(t)|^2/(2\sigma^2)}\, e^{-2\pi i(\boldsymbol{r}'(t)\cdot\boldsymbol{\xi}')} \end{aligned} \tag{5.12}$$

and the projection into tangent space is given by the inverse of the rotation matrix defining the orientation of the scratch so that

$$\boldsymbol{x}' = \boldsymbol{R}^{-1}\boldsymbol{x} = \begin{bmatrix} - & \hat{\boldsymbol{t}}^T(t) & - \\ - & \hat{\boldsymbol{b}}^T(t) & - \\ - & \hat{\boldsymbol{z}}^T & - \end{bmatrix} \boldsymbol{x}\,;\, \boldsymbol{r}' = \boldsymbol{R}^{-1}(\boldsymbol{r} - \boldsymbol{x}_0) \tag{5.13}$$

and $\boldsymbol{r}'$ is the relative scratch position. With Eq. 5.11 at hand, we are now able to express arbitrary scratch profiles. This profile is incorporated into an optical transfer function that can be used to express the corresponding diffraction effects.

5.4.4 Scratch ensemble solution

We will now derive a general solution for surfaces with arbitrarily large scratch ensembles. Recall that Eq. 5.11 provided the Fourier transform of the transfer function $\mathcal{T}^{(k)}_{\text{scratch}}$ for an individual scratch observed in isolation. Substituting this expression into Eq. 5.9 yields the superposition of a general scratch ensemble. Because we assume a homogeneous material,

the Fourier transform of the base transfer function $\mathcal{T}_{\text{base}}$ is given by the Fourier transform of the Gaussian filter kernel modulated by the square of the spatially uniform Fresnel coefficient (base amplitude). To compute the missing Fourier transform of the mask transfer function $\mathcal{T}^{(k)}_{\text{mask}}$, recall that scratches and masks share the same locations on the surface. Thus, the respective integral over the spatial phases $\eta^{(k)}$ is the same and mask and scratch only differ in their respective profile. The mask is simply a part of the base material that was cut out. Therefore, $\mathcal{T}^{(k)}_{\text{mask}}(b)$ is equal to the base transfer function with $A_{\text{mask}}(b) = A_{\text{base}}$ without any phase deviations because no height variations are present, and its spatial extent is restricted to the scratch width. This yields the Fourier transform of a rectangular function

$$\begin{aligned} \mathcal{F}\left\{\mathcal{T}^{(k)}_{\text{mask}}(b)\right\}_{\xi_2'} &= \mathcal{F}\left\{A_{\text{mask}}\,\text{rect}\left(\frac{b}{W^{(k)}}\right)\right\}_{\xi_2'} \\ &= A_{\text{base}}\,W^{(k)}\cdot\text{sinc}\left(\pi W^{(k)}\xi_2'\right). \end{aligned} \tag{5.14}$$

We are now able to combine Eq. 5.11, Eq. 5.14 and Eq. 5.9 with Eq. 5.7 to get the Fourier transform of the scratched surface. We split the result into more intuitive base- and scratch related terms to obtain

$$f_r(\boldsymbol{x}_0, \boldsymbol{\xi'}) = \gamma_i \frac{F}{\pi\sigma^2}\frac{1}{\lambda^2}\left|\mathcal{B}(\boldsymbol{\xi'}) - \mathcal{S}(\xi_2')\right|^2, \tag{5.15}$$

where $A_s = \pi\sigma^2$ is the area under the squared amplitude of the Gaussian (shading area) and F the Fresnel coefficient of the homogeneous material. We define

$$\mathcal{B}(\boldsymbol{\xi'}) = 2\pi\sigma^2 e^{-2\pi^2\sigma^2(\xi_1'^2+\xi_2'^2)}, \tag{5.16}$$

as the *base response* given by the Fourier transform of the filter kernel resembling the undisturbed reflectance of the material without scratches. On the other hand, the *scratch response*

$$\mathcal{S}(\boldsymbol{\xi'}) = \sum_k \left[\mathcal{F}\left\{\mathcal{T}^{(k)}_{\text{mask}}\right\}_{\xi_2'} - \mathcal{F}\left\{\mathcal{T}^{(k)}_{\text{scratch}}\right\}_{\xi_2'}\right]\eta^{(k)}(\boldsymbol{x}_0, \boldsymbol{\xi'}) \tag{5.17}$$

then defines the disturbance of the smooth heightfield by scratches. The scratch response thus relies on the scratches' profiles and their location on the surface with respect to the point of intersection encoded into the integral over the spatial phases $\eta^{(k)}$ (see App. 5.A for the full solution). In its simplest form using a rectangular scratch profile (see App. 5.B) and $A_{\text{mask}} = A_{\text{scratch}} = A_{\text{base}}$, the scratch response function is

$$\mathcal{S}(\boldsymbol{\xi'}) = \sum_k \mathcal{W}^{(k)}\mathcal{D}^{(k)}\eta^{(k)}(\boldsymbol{x}_0, \boldsymbol{\xi'}), \tag{5.18}$$

$$\mathcal{W}^{(k)} = W^{(k)}\text{sinc}\left(\pi\frac{W^{(k)}}{\lambda}\xi_2'\right), \tag{5.19}$$

$$\mathcal{D}^{(k)} = \left(1 - e^{2\pi i(\gamma_i+\gamma)D^{(k)}/\lambda}\right), \tag{5.20}$$

where we further separate the dependence of the individual diffraction patterns of the scratches on the respective width and depth via the width-term $\mathcal{W}^{(k)}$ and the corresponding depth term $\mathcal{D}^{(k)}$. This constitutes our reflectance function for rendering surfaces with microscale scratches. In Sec. 5.5, we explain how our rendering system efficiently looks up the scratches that are relevant for a given shading event.

Figure 5.6: We demonstrate our reflectance model on four example scenes. Top row from left to right: A golden door handle, a candle holder, both with randomized scratch distributions. A compact disc with circular scratches with constant separation. Bottom row: A spoon from different view points, the scratches are randomized. The top-down view reveals the scratches with iridescent colors that lie on a circle around the specular highlight. In the side-view the specular highlight only subtends a small fraction on the spoon and under incoherent illumination by the surrounding lights the scratches appear mostly white. At close-up again color as well as geometry of the scratches is revealed. We present complementary videos for these scenes in the supplemental material.

5.5 Usage in a rendering framework

In Sec. 5.4, we derived a BRDF for surfaces with micro-scale scratches that is compatible with standard ray tracing-based rendering systems. We use a standard backward path tracer that evaluates the BRDF at intersections found by tracing rays from the camera. The scratches are either applied to the surface by defining positions, directions, and other scratch parameters directly, e.g. by drawing them from a distribution, or by scratching an arbitrary mesh using an editing tool. In full-spectral rendering mode (all figures except Fig. 5.12), the renderer samples 16 wavelengths across the visible range. A reduced RGB version represents the primary colors by the wavelengths $\lambda_{\text{red}} = 700\,\text{nm}$; $\lambda_{\text{green}} = 520\,\text{nm}$ and $\lambda_{\text{blue}} = 440\,\text{nm}$.

Scratch data structure and lookup The scratch particles are represented by line segments, which we store in a bounding volume hierarchy (BVH) consisting of axis-aligned bounding boxes (AABB). The BVH is built using sorting on a space-filling curve (Morton code builder [LGS+09]) and efficient traversal is ensured by employing the skip-pointer structure proposed by Smits [Smi98]. To reduce spatial overlap between the elements of

this structure, we perform further splits. Since the shading cost of our model generally outweighs the intersection test cost, we do not directly split the scratch particles. We instead use a directed acyclic graph structure, where multiple AABBs can be associated with the same leaf element and spatially subdivide the scratches into a number of elements that can be adjusted for performance improvements. During intersection testing, we store only the unique intersections within the shaded area. When shading a point on the surface, we only consider scratches within the pixel footprint. To achieve the correct incoherent superposition of the coherent subsamples for a given pixel, we need to integrate over the pixel footprint (cf. Sec.5.4.1). This ensures the right multi-scale behaviour and allows for spatial variation at the cost of a larger number of required samples per pixel compared to standard BRDF models. Rays that strike the surface at an oblique angle might query many scratch particles, leading to poor performance due to a large pixel footprint. However, note that scratches that lie outside of a sphere whose diameter equals the Gaussian filter kernel ($\Delta_c = 6\sigma = 60\mu m$) only have a negligible contribution, hence we limit the BVH query to the intersection of this sphere and the pixel footprint.

Importance sampling the base response function To importance sample the base surface response Eq. 5.16, we generate two normally distributed samples and scale them by the standard deviation of the (Gaussian) target distribution in the angular spectrum, resulting in a sampled angular frequency $\boldsymbol{\xi}$. Specifically, we set

$$\boldsymbol{\xi}_{1,2} = [(\sqrt{8}\pi\sigma)^{-1}, (\sqrt{8}\pi\sigma)^{-1}]^T. \tag{5.21}$$

Next, the sampled frequency is used to map the incident direction onto a scattered direction $\hat{\boldsymbol{\omega}}_o = (\alpha_o, \beta_o, \gamma_o)^T$ by solving for α_o and β_o via $\alpha_i + \alpha_o = \xi_1$ and $\beta_i + \beta_o = \xi_2$, to obtain

$$\hat{\boldsymbol{\omega}}_o = \begin{pmatrix} \alpha_o \\ \beta_o \\ \gamma_o \end{pmatrix} = \begin{pmatrix} \xi_1 - \alpha_i \\ \xi_2 - \beta_i \\ \sqrt{1 - \alpha_o^2 - \beta_o^2} \end{pmatrix} \tag{5.22}$$

Two details must be noted regarding this step: occasionally, a sample satisfies $1 - \alpha_o^2 - \beta_o^2 < 0$, which does not lead to a valid scattered direction. These samples correspond to evanescent waves that do not propagate, and the associated sample is simply dropped. Secondly, sampling a position in the angular frequency domain and mapping it on the outgoing hemisphere corresponds to a change of variables that appears both in the sampling density and Monte Carlo weight of this sampling strategy. The mapping is simply the parallel projection from the unit disc to the unit hemisphere known as the Nusselt analog, and the Jacobian determinant factor for the Gaussian PDF associated with this mapping is the direction cosine γ_o.

Importance sampling of the scratch response function Importance sampling of the scratch response relies on a modification of a sampling technique that was originally developed by d'Eon et al. [dFH+11] in the context of hair rendering. Given an incident direction (ϕ_i, θ_i) expressed in the coordinate system of a hair fiber, this technique works by sampling a specular reflection from an ideally reflecting cylinder, producing a reflected direction on a

Dirac delta circle of azimuths with elevation angle $\theta_i = -\theta_o$. To account for roughness, the direction (ϕ_o, θ_o) is then perturbed by a random offset drawn from a spherical von Mises-Fisher distribution with concentration parameter κ. The spherical density of sampled directions has an explicit form in terms of a modified Bessel function of the first kind, specifically

$$p(\theta_o, \phi_o) = \frac{\kappa}{4\pi \sinh \kappa} e^{-\kappa \cos\theta_i \cos\theta_o} I_0 \left[\kappa \sin\theta_i \sin\theta_o\right]$$

Although disconnected from the explicit profiles of scratches, we found the resulting distribution to be an excellent match for the response function of individual scratches when interpreting the scratch tangent vector as a fiber direction and mirroring reflected directions that would enter the surface along the normal direction, doubling the density $p(\theta_o, \phi_o)$ for directions that lie in the upper hemisphere.

Our method applies multiple importance sampling via the balance heuristic [VG95] to combine sampling of the base surface and the weighted scratch profiles inside the coherence area into a single unified sampling strategy.

Combining other BRDFs with our model It is of great importance to be able to combine different BRDFs to achieve generality. To this end, we developed a simple modified alpha-blending step that marries correct wave-optical shading and mutual interference of scratch contributions to (in principle) arbitrary base BRDFs. In our implementation, we use microfacet models to enable a rough base appearance even in unscratched regions. To achieve this goal, we first retrieve all the scratches from our BVH that fall into the coherence area as before. Next, we calculate the weighted scratch area density, i.e. the normalized sum of all contained scratch areas weighted using the Gaussian spatial filter. This yields a spatially varying ratio between the base and scratch contribution. We use this ratio to blend between the chosen base BRDF and our scratch SVBRDF (Eq. 5.15), setting the base response $\mathcal{B} = 0$ in Eq. 5.15 and $A_{\text{mask}} = A_{\text{scratch}} = 1$ to ensure energy conservation.

5.6 Results

In this section we will first show example scenes rendered with our model to recreate the appearance of scratched surfaces. The corresponding render times and parameters can be found in the supplemental material. We will then discuss different aspects of our model in detail, including the impact of the coherent superposition of diffracted light, the possibilities to utilize and adapt our model to recreate realistic renderings and finally we will extend our model to incorporate not our specular base response function but an arbitrary microfacet model such as GGX [TR75].

Scratching arbitrary objects To facilitate authoring of assets, our editing tool allows the alteration of scratch particles in real time by drawing their spatial parameters from distributions or applying regular brush drawing techniques in 3D in combination with distribution based alterations of the optical (i.e. profile) parameters. We provide a detailed video that showcases this tool in the supplementary material of this paper. Additionally, we have implemented tools that allow conversion of 2D vector graphics images to scratch particles by projecting them from texture to object space. In the editor, an approximate real-time

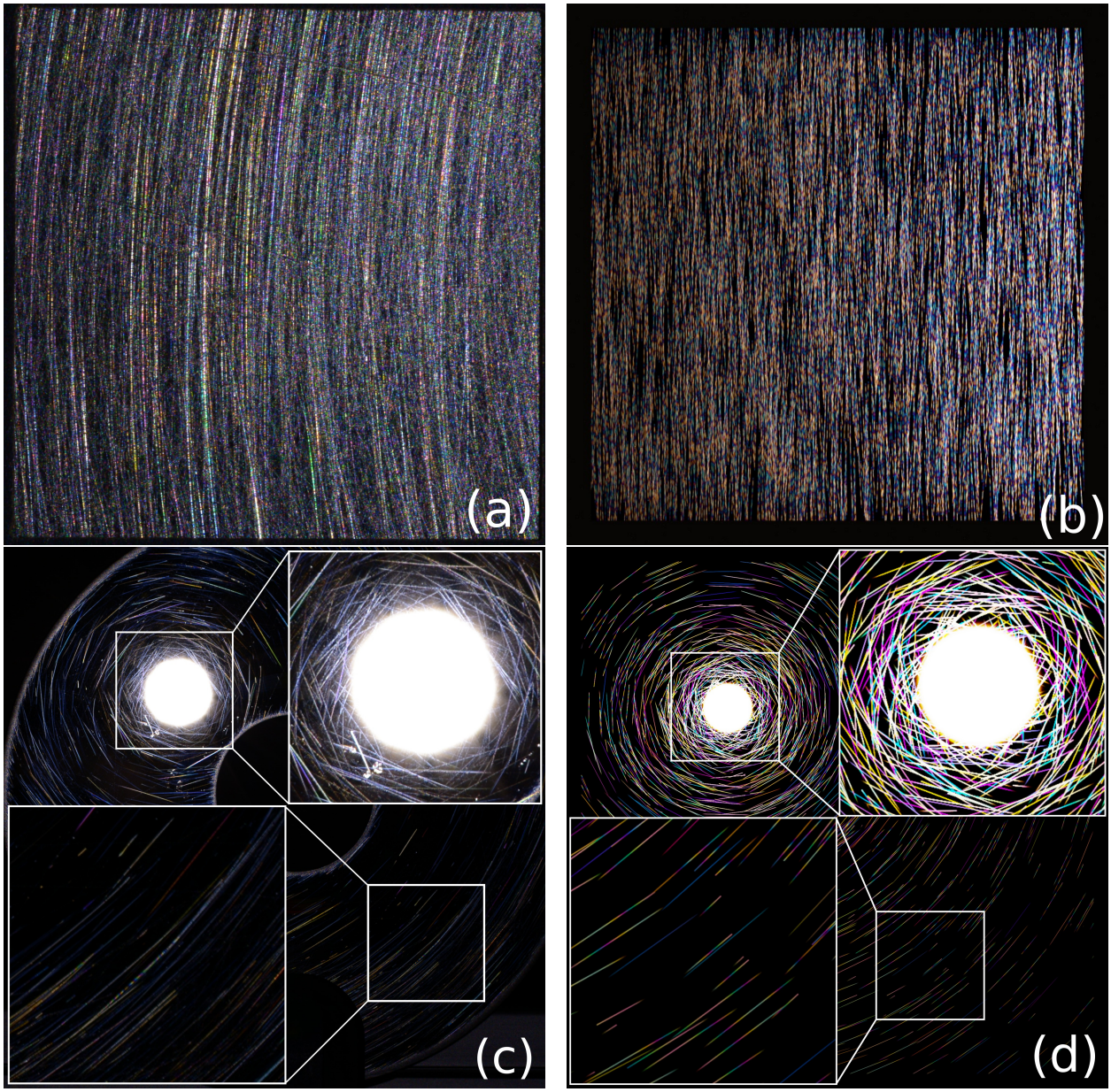

Figure 5.7: Qualitative comparison between photographs of real-world scenes (a,c), and renderings of their digital recreations (b,d). We show a $2\times2\,\mathrm{cm}^2$ part of a heavily scratched metallic plate (a,b) and a more sparsely scratched metallic disc with 12 cm diameter (c,d). All surfaces have scratches in all orientations. However, they only become visible under the specular condition, which leads to concentric structures around specular highlights. From these examples, it becomes clear that the problem of fully recreating real-world surfaces is mostly related to complexity. The reference examples show great variation, and their roughness spans many scales which are currently not covered by our model.

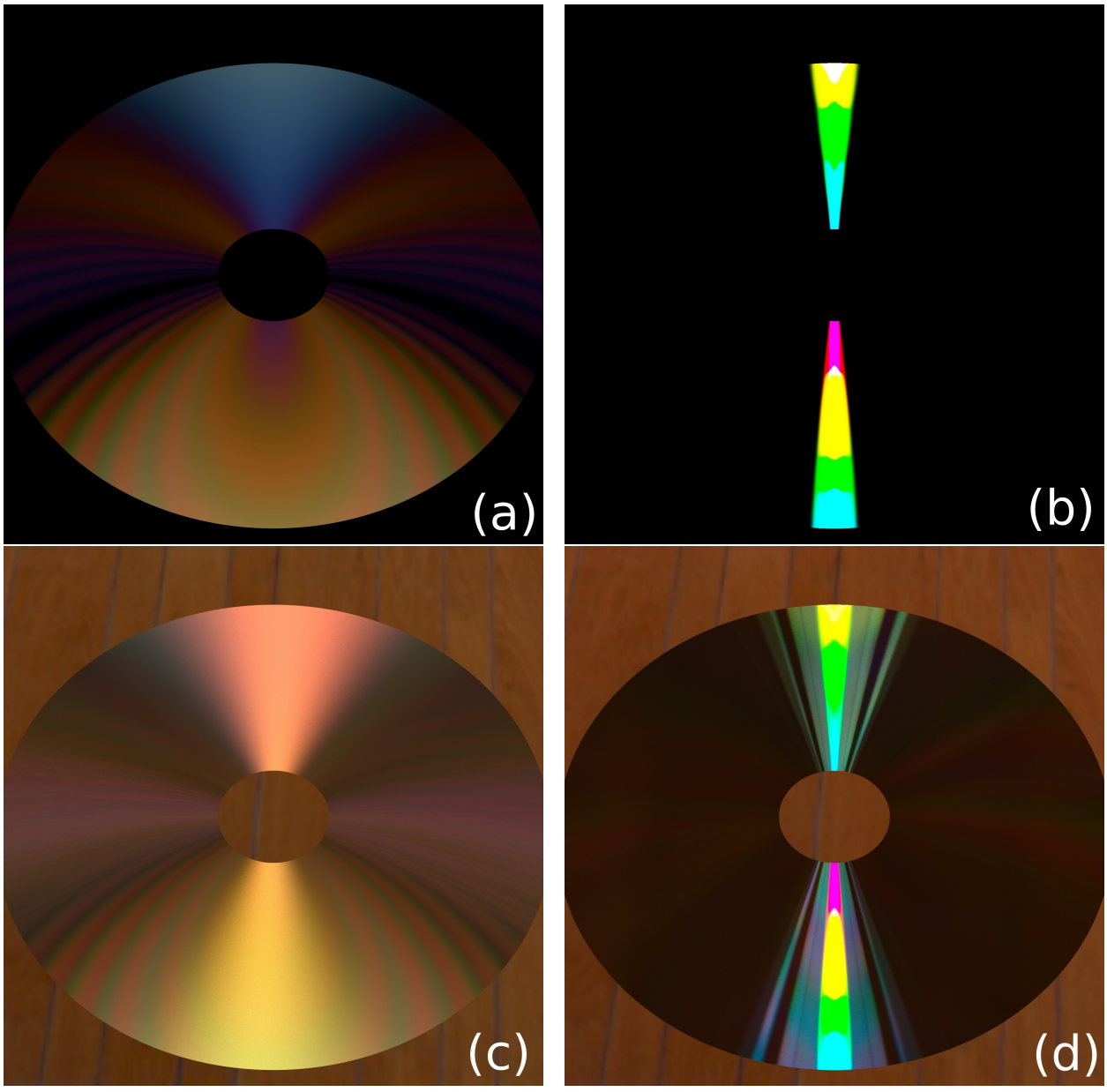

Figure 5.8: Qualitative comparison between Stam's model (a,c) and ours (b,d) without (top) and with (bottom) environment light. The different representation of the CD surface structure by the two approaches leads to diverging appearance. Stams model treats CD surfaces as a (random) collection of bumps which re-distributes energy to higher order diffraction lobes visible over the whole disc. Our model, on the other hand, models CD tracks as uniform circular scratches which suppresses diffraction in parts not fulfilling the specular condition.

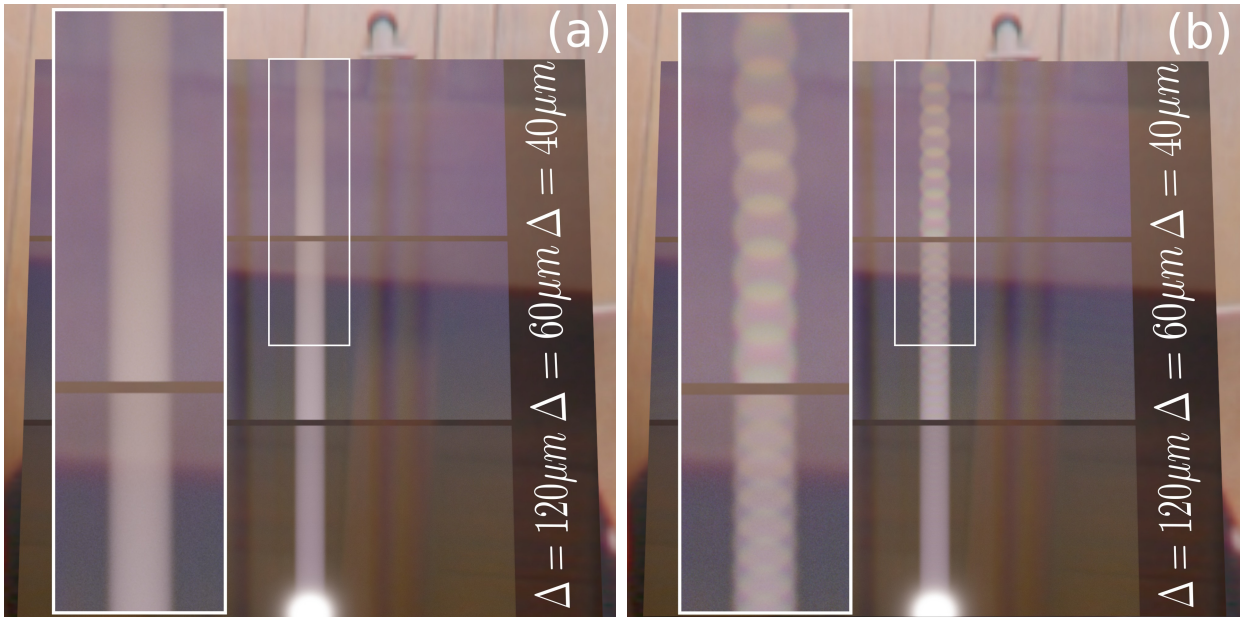

Figure 5.9: Comparison of coherent and incoherent scratch diffraction superposition. The scratches lie on a metallic plate (GGX microfacet BRDF) and are horizontally arranged as three gratings with different separations Δ between the uniform scratches. Thus, we expect the diffraction orders only to be visible in vertical direction (across the scratches). (a): An incoherent superposition of scratches within the coherence area leads to colored scratches due to single-scratch diffraction. (b): A coherent superposition of scratches (our model) not only accounts for single-scratch diffraction but is also able to recreate mutual interference effects such as higher diffraction orders of the underlying scratch grating which reveals the separation of colors especially in the area of high intensity (see zoom-ins).

BRDF model only shades the first intersected scratch with a single light sample; no coherent effects are taken into account.

Comparison to real-world data To provide an intuitive comparison of effects our model is able to recreate, Fig. 5.7 compares photographs of real-world scenes and renderings of their digital recreations. Since the underlying surface structure is unknown, we restrict ourselves to a purely phenomenological comparison and use heuristics to distribute scratches and set their parameters. To focus on microscopic features, we compare photograph (macro-lens close-up) and rendering of a densely scratched metal plate (a,b) whereas the comparison of a scratched disc (c,d) places emphasis on the macroscopic properties. A common phenomenon in real-world scratches, the change of color along scratches due to varying profile (parameters), is closely reproduced by our model, whereas high-intensity glints in low-intensity regions are not since unscratched regions are black (smooth-surface assumption). The discrepancies most likely results from two features currently not covered by our model: additional base-surface roughness and inclined scratch normals resulting from a non-uniform scratching process, which leads to more complex scratch profiles. However, our model qualitatively reproduces scratch visibility according to the specular condition (tangent not perpendicular to projected light direction) forming circular structures around the specular reflection of the light source. Both examples illustrate that our model is capable of recreating the phenomeno-

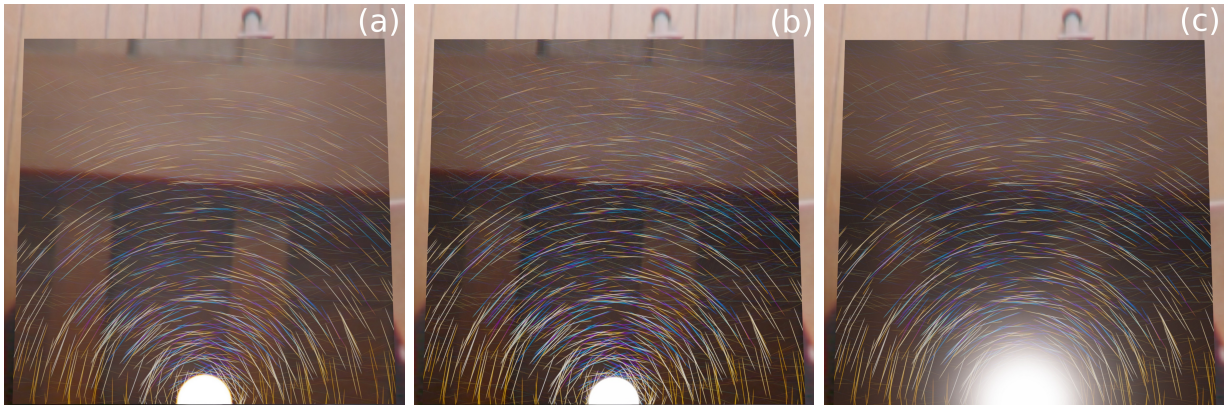

Figure 5.10: Comparison of scratches on a plate with our coherent (a) base response, a smooth/specular GGX (b) and a rough GGX (c) base reflectance. We observe a change in color saturation as interference between base and scratches is neglected and single scratch diffraction is overestimated

logical features of diffractive scratches, which could likely be improved by stronger emphasis on the modeling step. A convenient step to introduce surface roughness for example could be to utilize our base-blending approach, described in the following, in combination with the scattering model proposed by Holzschuch and Pacanowski [HP17].

Comparison to Stam's model Most available reflection models that incorporate diffraction are restricted to single heightfield inputs or analytic solutions for periodic structures. For a qualitative comparison we provide renderings of a CD in Fig. 5.8. The periodic surface structure can be used by both, our approach (b,d) and the model proposed by Stam [Sta99] (a,c), though using different representations. Stam defines the surface as a periodic ensemble of *bumps* and computes the spectral power density as input for the BRDF. We, on the other hand, create tracks that consist of a number of scratches. Track separation and scratch parameters are taken from Stam [Sta99] so that the main difference between the two representations is that ours does not account for gaps between scratches of one track. Both models diverge in appearance: Stam's model produces clearly visible higher order diffraction for tracks that do not fulfill the specular condition. The reason for this mainly lies in the surface representation: our model considers CD tracks as continuous scratches; the complex phasor defining the wave contribution (Eq. 5.15) is mainly driven by the scratch profile (and thus binormal direction). Stam's model, on the other hand, creates bumps of defined geometry which results in diffraction patterns in both, tangential and binormal direction. This leads to a (spectral) re-distribution of energy and therefore a change in color which is clearly visible in the upper row (area light only). The re-distribution of energy is emphasized in the bottom-row renderings (additional environment light) as it allows for the low-intensity higher order diffraction still to be visible. In contrast, our model is able to produce sparsely scratched surfaces such as shown in Fig. 5.7 where single scratches can be distinguished as well as the change of color along the scratch, a common phenomenon in non-manufactured real world scratches.

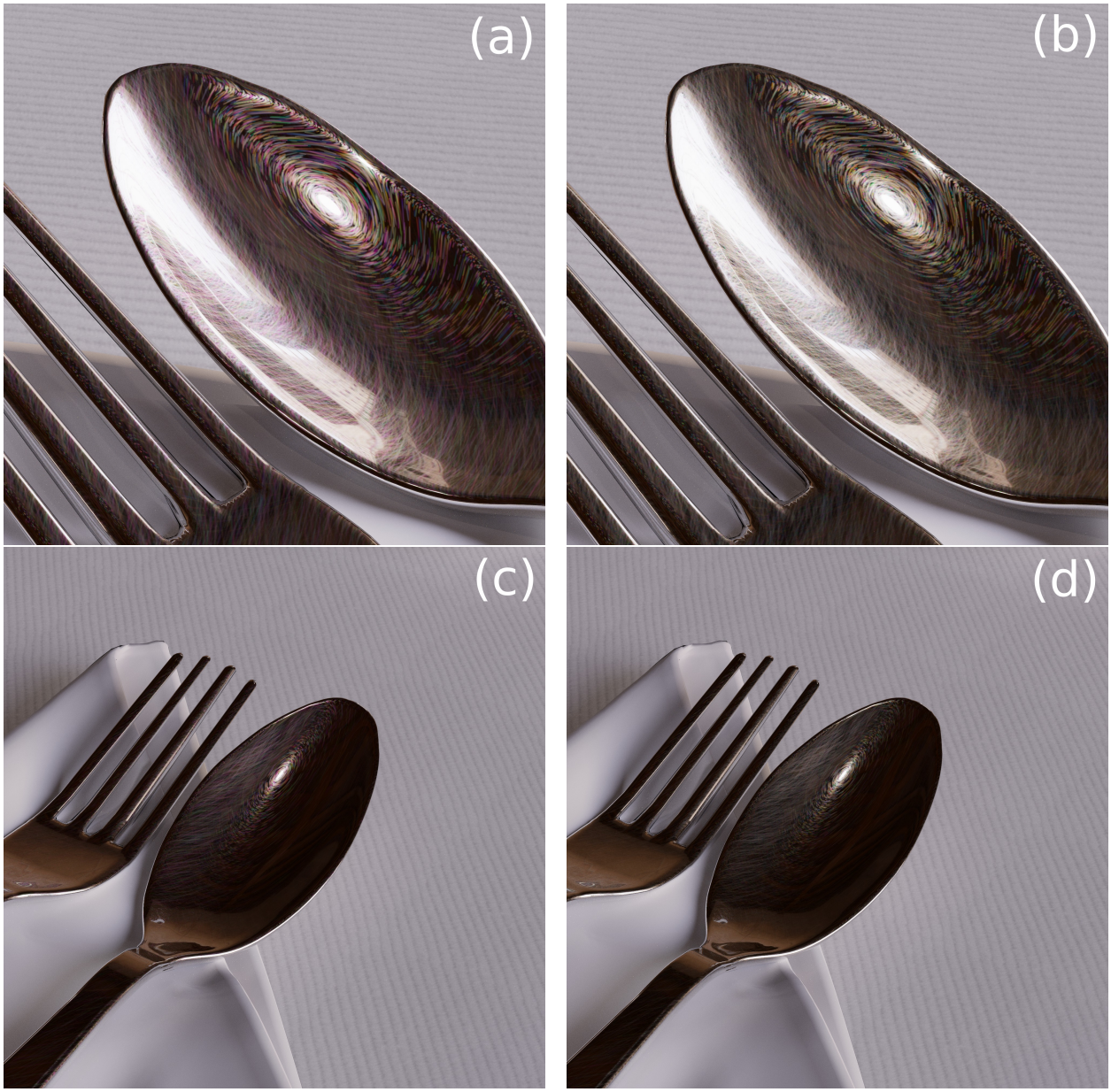

Figure 5.11: Variation of scratch parameters greatly enhances the realistic appearance of our renderings. Subimages (a),(c): without variation; (b),(d): with variation along the scratches.

Figure 5.12: Our editing tool allows us to map complex scratch patterns onto objects of our choice. Here we "engraved" a ring and a planar surface with an intricate vector pattern (inset). Depth and width were drawn from a Gaussian distribution.

Coherent vs. incoherent superposition The treatment of coherence is of great importance for effects such as mutual interference from structured surfaces, for example compact discs or holographic papers. Our model treats spatial coherence by applying a Gaussian weight to the contributions of scratches according to their position, as the complex diffracted amplitudes per scratch are superposed. Figure 5.9 reveals that without coherent superposition, effects such as diffraction orders generated by gratings are neglected and therefore the associated separation of colors cannot be reproduced. Our model on the other hand is able to closely reproduce such phenomena.

Profile variation Scratches on surfaces are created by multiple effects such as every-day wear or even manufacturing. Whereas manufactured scratches or structures mostly have a well defined geometry, scratches produced by wear do not. To account for this and more closely reproduce such surfaces, we vary width and depth of scratch profiles by sampling from a simplex-noise function [Per02]. The random number generator used to generate the noise is seeded by the scratch index k to ensure determinism. This feature increases realism with very modest impact on performance and memory footprint, since longer scratch segments do not need to be split up to incorporate such variations. Figure 5.11 compares the effect of this parameter variation to scratches of constant parameters.

Microfacet base blending In Fig. 5.10 we show results for our model with our simple blending approach in comparison to the coherent base response. Our fully coherent solution (a) shows good agreement with the specular GGX base (b) in terms of scratch colors and base reflectance. However, some changes are noticeable: first, the specular highlight of our model exhibits a red outline which is due to our separate treatment of wavelengths. Red light is scattered more strongly compared to smaller wavelengths, an effect which is not the

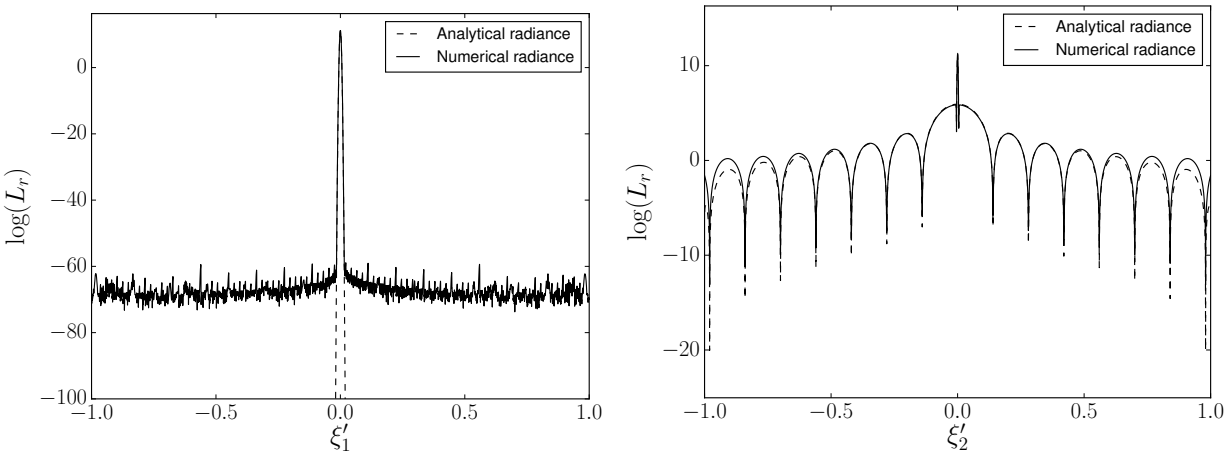

Figure 5.13: Comparison of numerical and analytical slices of the radiance distribution for a single scratch (Fig. 5.14 left). Left;$\boldsymbol{\xi}'_2 = 0$: The slice in tangential direction of the scratch reveals the impact of the Gaussian filter. Right;$\boldsymbol{\xi}'_1 = 0$: The radiance distribution in bitangential direction shows the characteristic diffraction pattern. Our model greatly enhances performances as the analytic solution exhibits a computational complexity of $\mathcal{O}(NM)$ for a single scratch opposed to $\mathcal{O}(NM \log(NM))$ for the FFT.

case for geometrical optics models. Second, we observe more saturated colors, which results from the lack of interference with the base. Due to the neglected scratch-base interaction, we do not correctly account for the re-distribution of energy and we overestimate the energy diffracted by the scratches. The use of different microfacet models allows us to incorporate surface roughness (c) which, by construction, does not affect the scratches but only the base response. In this way we are able to retain the iridescent effects of scratched surfaces with only minor differences and utilize the benefits of microfacet models.

Mapping complex scratch patterns Our editing tool also allows us to project (in principle) arbitrarily complex scratch patterns provided as vector graphics onto complex objects. Fig. 5.12 shows the mapping of a vectorized leaf-texture onto a ring and a plate.

Approximation evaluation To evaluate the impact of our assumptions regarding separability (we discard the Gaussian filter in bitangential direction) and scratch-scratch intersections, we compare the numerical radiance obtained via FFT against our model. To this end, we create surfaces which exhibit a number of scratches and rasterize these. The resulting heightfield (including scratch intersections) is used to create an optical transfer function via Eq. 5.8 and then input into the FFT. The radiance is obtained according to Eq. 5.7 with unit amplitude and the origin as the intersection point. We first show the radiance corresponding to a surface with a single scratch with tangent $(1, 0, 0)$ (Fig. 5.13 and Fig. 5.14(left)) to clarify single aspects of our approximation and problems that arise when using the FFT. A slice along the scratch (left, ξ'_1 is the direction cosine in tangential direction) reveals a paraboloid function that rapidly drops off as is expected for a Gaussian in logarithmic representation. We expect a Gaussian as the scratch is longer than the surface we consider, therefore extending the integration limits for $\eta^{(k)}$ to $(-\infty, \infty)$. This

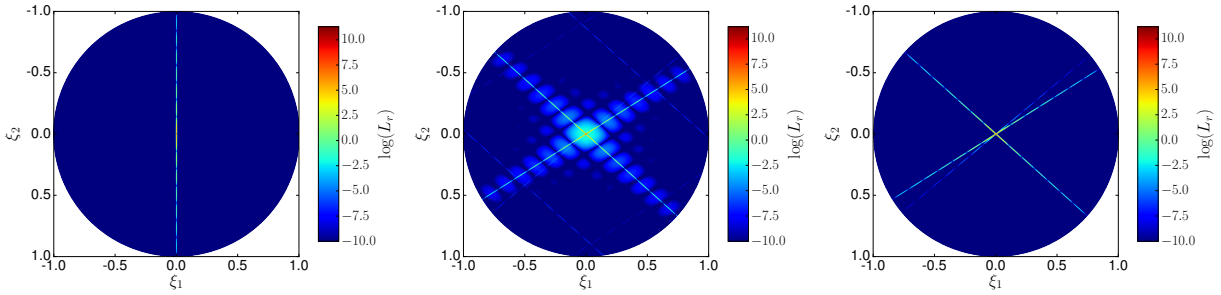

Figure 5.14: Numerical radiance obtained via FFT vs. our Model in direction cosine space. Left: Analytic single scratch radiance in direction cosine space. Middle; Right: A surface with 10 randomly distributed scratches. The numerical solution for the whole hemisphere (middle) shows ghosting artifacts as well as parallelogram-shaped structures of low intensity (PSNR=84.9 dB) which are not present in our analytic model (right). The former result from the discretization of the heightfield whereas the latter correspond to the neglected intersections of scratches within the coherence area. All plots are available in the supplemental material in high resolution. Additionally, plots for different discretization resolutions are available to emphasize the difference between ghosting and assumption impact.

behavior is reproduced by our model and in good agreement with the numerical solution. A slice across the scratch (right, ξ_2' corresponds to the bitangential direction) reveals the effects of an approximation in our model: since we neglect the convolution with the Gaussian filter in the bitangential direction (cf. Eq. 5.11), the radiance is underestimated for larger angles. Interpreting the convolution as a re-distribution of the diffracted energy, it should decrease the central peak while increasing the side lobe maxima. Taking these limitations into account, our model agrees well with the numerical results. For a surface with ten randomly distributed scratches (Fig. 5.14), the numerical solution for the whole hemisphere (middle) shows discretization-related ghosting artifacts, which are not present in our analytic model (right). Additionally, parallelogram-shaped structures are visible in the numerical solution that result from scratch-scratch intersection. Although of comparably low intensity (PSNR=84.9 dB) these correspond to a re-distribution of energy from the main- and side-lobes of the diffraction pattern to the parallelogram structures and thus, together with our approximations (see single scratch), can lead to an overestimation of higher order diffraction lobes resulting in overly saturated scratch colors. However, scales and primary (i.e. non-artifact) radiance distributions agree well, especially along the bitangential directions of the scratches, which reveals the characteristic diffraction distribution. Note that the analytical model delivers a significant performance increase as the computational complexity is only $\mathcal{O}(NM)$ for a single scratch in comparison to $\mathcal{O}(NM\log(NM))$. Precomputation would furthermore require a 4D lookup table with large memory footprint to sufficiently sample the spatial (Gaussian filter kernel) and spectral dimensions (FFT).

5.7 Discussion and future work

In this paper, we presented a wave-optical SVBRDF model for surfaces with iridescent microscale scratches. Our model encapsulates wave-optical computations in the shading

evaluation and accounts for diffraction and mutual interference involving multiple scratches under a arbitrary viewing and lighting conditions. By approximating the spatial coherence using a Gaussian filter, we are able to recreate both localized glint-like iridescence and higher orders of diffraction from grating-like structures. Our approach is flexible on both the model and the data side: by subdividing paths on the surface into line segments and analytically calculating the associated diffracted complex amplitudes, we can support arbitrary scratch profiles. Furthermore, the separation of the spatial (paths) and optical (profiles) components enables efficient control of parameters and even spatial variation of parameters along scratches. Additionally, our editing tool allows us to freely scratch arbitrary geometries and apply complex scratching patterns with ease. Our model has the following limitations that could be addressed in future work:

Footprint integration Our model subsamples the pixel footprint with a spatial filter resembling the coherence function to recreate local wave-optical diffraction. For full convergence, we rely on Monte-Carlo integration performed by the ray-tracer. In future work, an important step to improve performance would be to approximate the integral of the coherent SVBRDF over the pixel footprint to create a full multi-scale model.

Importance sampling Accurate sampling strategies for both, our base response functionas well as for microfacet models used for our base-blending readily exist. The sampling strategy for the scratch response function samples the azimuthal direction uniformly [dFH$^+$11], which is overly conservative considering for example the sinc-distribution (cf. Eq. 5.21 for the square profile. In future work, more efficient sampling schemes for typical diffraction patterns and, more challenging, mutual interference of scratch ensembles could be developed.

Other scratch profiles We only presented results for a rectangular profile function, showing that it is expressive enough for recreating such intricate effects as iridescent scratches. Additional renderings of rectangular and triangular profiles in comparison can be found in the supplemental material. We expect to obtain a wider variety of scattering distributions from different scratch profiles, leading to an increased degree of realism. Both, profiles with analytic solutions (like V-shaped or Gaussian grooves) as well as tabulated scattering distributions for arbitrary scratch profiles could be precomputed and used at moderate additional cost.

BSDF extension So far, we showed that our model is able to represent the reflection properties of scratched surfaces. On the other hand, the scattered radiance of our SVBRDF is only dependent on the difference between the incident and outgoing directional cosines. This implies that the model in principle generalizes to transmittance effects (thin materials without internal scattering, like foils or very light curtains) via a simple sign change. This would be an interesting avenue to explore, as many real world diffraction effects are observed in transmittance rather than reflectance.

Coherent base - Use NDF to generate base height variation Using our base-blending scheme we can in principle use arbitrary BRDFs as a base material. This approach

does not account for phase variations that are directly produced by the surface roughness of the unscratched material. A future direction would be to derive a modified *coherent* base response function that is driven by surface roughness, for instance in the form of microfacet models.

Comparisons against real-world material samples Given real-world data, both, experimental validation of our model and fitting of model parameters remains a major challenge due to the large number of unknowns involved. As of now, it is not yet clear what kind of input data would be needed to provide sufficient constraints, how to acquire such data, how to represent the surface (individual scratches or distribution parameters) and how to determine the model parameters in finite time. We look forward to studying these aspects in more detail in the future.

Appendix

5.A Gaussian weighted spatial phases

The integration of the spatial phases of the scratches (and masks) relies on the following assumptions: First, the width of a scratch is negligible compared to its length with respect to the coherence area. Second, the scratch segments we integrate are lines. Third, the profile $\mathcal{P}$ does not change over a segment. To regain spatial resolution we apply a spatial filter kernel on the surface which weights each phase according to its position relative to the point of intersection (the origin of the footprint, see Fig. 5.2) and approximates the coherence function. This provides us with a closed-form solution of the integral of Eq. 5.11 given by

$$\eta^{(k)} = \int dt\, \mathcal{G}(t)\, e^{-2\pi i(\boldsymbol{r}'(t)\cdot\boldsymbol{\xi}')}$$

We represent our scratches as the relative position $\boldsymbol{r}'(t) = \boldsymbol{r}'_0 + t\cdot\hat{\boldsymbol{t}}' - \boldsymbol{x}'_0$; $t \in [-L/2, L/2]$ and $\boldsymbol{r}'_0(t) = \boldsymbol{r}'_0 - \boldsymbol{x}'_0$ where L is the length of the scratch and the prime denotes the coordinates in tangent space. The integral to solve then reads

$$\begin{aligned}
\eta^{(k)} &= \int dt\, e^{-2\pi i(\boldsymbol{r}'(t))\cdot\boldsymbol{\xi}'}\, e^{-|\boldsymbol{r}'(t)|^2/(2\sigma^2)} \\
&= e^{-2\pi i \boldsymbol{r}'_0(t)\cdot\boldsymbol{\xi}'}\, e^{-\frac{|\boldsymbol{r}'_0|^2}{2\sigma^2}} \int dt\, e^{-2\pi i t\xi'_1}\, e^{-\frac{t^2+2t(\hat{\boldsymbol{t}}'\cdot\boldsymbol{r}'_{r,0})}{2\sigma^2}} \\
&= c_0 \cdot \left[\operatorname{erf}\left(\frac{a_0 + L/(2\sigma)}{\sqrt{2}}\right) - \operatorname{erf}\left(\frac{a_0 - L/(2\sigma)}{\sqrt{2}}\right)\right]
\end{aligned} \tag{5.23}$$

where L is the total length of the scratch and

$$\begin{aligned}
a_0 &= 2\pi i\sigma\xi'_1 + \frac{(\hat{\boldsymbol{t}}'\cdot\boldsymbol{r}'_0)}{\sigma} \\
c_0 &= \sqrt{\pi/2}\,\sigma\, e^{h+if} \\
h &= -\frac{|\boldsymbol{r}'_0|^2}{2\sigma^2} - 2\pi^2\sigma^2\xi'^2_1 + \frac{(\hat{\boldsymbol{t}}'\cdot\boldsymbol{r}'_0)^2}{2\sigma^2} \\
f &= 2\pi\xi'_1(\hat{\boldsymbol{t}}'\cdot\boldsymbol{r}'_0) - 2\pi(\boldsymbol{r}'_0\cdot\boldsymbol{\xi}')
\end{aligned} \tag{5.24}$$

5.B Scratch and mask profiles

We separate the spatial and the spectral component of the scratches which enables us to drive the reflection properties of a scratch by varying its 1d transfer function via the profile

$\mathcal{P}_{\text{scratch}}(b)$. As we need to compute the Fourier transform of $\mathcal{T}_{\text{scratch}}(b)$ it is convenient to choose profile functions that lead to a closed-form solution and are drivable by the geometric parameters width W and depth D. We note that it would in general be possible to replace this function by a lookup table. The transfer function for a specific profile in general reads:

$$\mathcal{T}(b) = A(b)\, e^{-2\pi i(\gamma_i+\gamma)\mathcal{P}(b)/\lambda} \tag{5.25}$$

where $A(b)$ contains the Fresnel term and b is the bitangential coordinate in the $(\hat{\boldsymbol{b}},\hat{\boldsymbol{n}})$-plane (c.f. fig 5.2).

The mask transfer function $\mathcal{T}_{\text{mask}}(b)$ and its Fourier transform, given in Eq. 5.14, for a single scratch reads

$$\begin{aligned} \mathcal{F}\left\{\mathcal{T}_{\text{mask}}(b)\right\}_{\xi_2'} &= \mathcal{F}\left\{A_{\text{base}}\,\text{rect}\left(\frac{b}{W}\right)\right\}_{\xi_2'} \\ &= A_{\text{base}}\, W \cdot \text{sinc}\left(\pi W \xi_2'\right) \end{aligned} \tag{5.26}$$

which is always the case for the mask profile. On the other hand we are able to choose an arbitrary scratch profile. For simplicity we concentrate on two different profile functions which are driven by the scratches' widths and depths. The simplest case is a rectangular profile with constant depth, i.e.,

$$\begin{aligned} \mathcal{F}\left\{\mathcal{T}^{\text{rect}}_{\text{scratch}}(b)\right\}_{\xi_2'} &= \mathcal{F}\left\{A_{\text{scratch}}\,\text{rect}\left(\frac{b}{W}\right)\right\}_{\xi_2'} \Phi_D \\ &= A_{\text{scratch}}\, W \cdot \text{sinc}\left(\pi W \xi_2'\right) \Phi_D \end{aligned} \tag{5.27}$$

and depth-phase

$$\Phi_D = e^{-2\pi i(\gamma_i+\gamma)D/\lambda} \tag{5.28}$$

For triangular profiles, we obtain

$$\begin{aligned} &\mathcal{F}\left\{\mathcal{T}^{\text{tri}}_{\text{scratch}}\right\}_{\xi_2'} \\ =&\mathcal{F}\left\{A_{\text{scratch}}\,\text{rect}\left(\frac{b}{W}\right)\cdot e^{2\pi i(\gamma_i+\gamma)D/\lambda\,(1-|\frac{b}{W/2}|)}\right\}_{\xi_2'} \\ =&B\cdot\left(1-e^{\pi i(\xi_2'-\frac{2D(\gamma_i+\gamma)}{\lambda})}\right) \\ &+C\cdot\left(e^{-\pi i(W\xi_2'+\frac{2D(\gamma_i+\gamma)}{\lambda})}-1\right) \end{aligned} \tag{5.29}$$

with

$$B = A_{\text{scratch}}\,\frac{i\,\Phi_D}{2\pi(\xi_2'-\frac{2D(\gamma_i+\gamma)}{W\lambda})} \tag{5.30}$$

$$C = A_{\text{scratch}}\,\frac{i\,\Phi_D}{2\pi(\xi_2'+\frac{2D(\gamma_i+\gamma)}{W\lambda})} \tag{5.31}$$

5.C Diffracted radiance

This supplemental material provides detailed derivations and parameters for recreation of results presented in the main body of our paper. In particular, we show a detailed derivation of the BRDF from Harvey's *diffracted radiance* on which our model is based. Additionally, we list all render settings and scratch parameters and the corresponding render times for all results obtained with our model.

Our model builds on tools from Fourier optics [Goo96], specifically the *angular spectrum* and the concept of *diffracted radiance* [HVK+00], which we review here for completeness. Being part of a *scalar* theory of light transport, these two tools assume that the electromagnetic field can be described by the (scalar) amplitude of the oscillations that make up the electromagnetic field, as opposed to the commonly used vectorial electric and magnetic fields. This approximation is accurate in the far-field and for diffracting apertures that are larger than the wavelength of the radiation. Without loss of generality, we restrict ourselves to monochromatic radiation at a wavelength of λ. The following discussion assumes that all spatial coordinates are expressed in units of λ, since this leads to simpler mathematical expressions. Let $U(x, y, z)$ denote the scalar amplitude at position $(x, y, z)^T$, and let $U_0(x, y) \coloneqq U(x, y, 0)$ denote a planar slice at position $z = 0$ (here called the aperture plane). A well-studied problem in this domain entails computing $U(x, y, z)$ for $z > 0$ given the amplitude distribution in the aperture plane $U_0(x, y)$. In the context of Fourier optics, solutions can be found by taking the Fourier transform of all quantities in the xy-plane, i.e.

$$V(\alpha, \beta, z) \coloneqq \mathcal{F}\big\{U(\cdot, \cdot, z)\big\}_{\alpha,\beta},\ V_0(\alpha, \beta) \coloneqq \mathcal{F}\big\{U_0(\cdot, \cdot)\big\}_{\alpha,\beta}, \tag{5.32}$$

and solving the Helmholtz equation $[\nabla^2 + 4\pi^2]U = 0$ analytically in terms of the frequency-space representation V. The latter has an intuitive physical interpretation: the amplitude $U(x, y, z)$ on any fixed z-slice can be described as a superposition of plane waves arriving from different directions. In this context, $V(\alpha, \beta, z) \in \mathbb{C}$ denotes both phase and amplitude of such a plane wave arriving from direction (α, β, γ) where $\gamma = \sqrt{1 - \alpha^2 - \beta^2}$ (Figure 3a, main paper). The variables of this parameterization are referred to as *direction cosines*. Evaluating the superposition of plane waves is equivalent to an inverse Fourier transform that recovers the original signal. Assuming that radiation travels undisturbed through the half-space $z > 0$, the Helmholtz equation has a particularly simple solution which states that the plane waves arriving at any z-slice correspond exactly to those at $z = 0$ except for a phase shift $V(\alpha, \beta, z) = V_0(\alpha, \beta)e^{i2\pi\gamma z}$. This solution is exact under the stated assumptions, but the resulting field $U(x, y, z)$ is prohibitively expensive to evaluate due to its definition in terms of a pair of Fourier transforms. We instead rely on a far-field approximation, which makes the reasonable assumption that the distance between the surface and the camera is much greater than the wavelength of light (Figure 3a andFigure 4(d), main paper). This far-field approximation, known as *diffracted radiance*, was introduced by Harvey et al. [HVK+00] and is defined as

$$L(\boldsymbol{\omega}) = \frac{\lambda^2}{A_s}\big|\mathcal{F}\left\{U_0(\cdot, \cdot)\right\}\big|^2_{\alpha,\beta} = \frac{\lambda^2}{A_s}\big|V_0(\alpha, \beta)\big|^2, \tag{5.33}$$

where $\boldsymbol{\omega} = (\alpha, \beta, \gamma)$. U_0 describes both the source for the field at $z > 0$ and the result of the radiation incident at $z = 0$, as the corresponding angular spectrum $V_0(\alpha, \beta)$ is given by

the superposition of plane waves from all directions. A change of the direction of incident radiance by direction cosine β_i results in a shift applied to all plane waves contributing to $V_0(\alpha, \beta)$, and the angle-shifted angular spectrum now reads $V_0(\alpha, \beta - \beta_i)$ (we show the one-dimensional case for simplicity, but the concept holds for the second dimension as well). As angular spectrum and complex amplitude are related by a Fourier transform, this can be interpreted in terms of the Fourier shift theorem as a linear phase shift applied to U_0. An additional attenuation factor, the third direction cosine γ_i [HVK+00], accounts for the decreased intensity at oblique incident angles and modulates Eq. 5.33 as

$$\begin{aligned} L(\boldsymbol{\omega}, \alpha_i, \beta_i) &= \gamma_i \frac{\lambda^2}{A_s} \left| V_0(\alpha - \alpha_i, \beta - \beta_i) \right|^2 \\ &= \gamma_i \frac{\lambda^2}{A_s} \left| \mathcal{F}\left\{ U_0(\cdot, \cdot)\, e^{2\pi i(\beta_i y + \alpha_i x)} \right\}_{\alpha,\beta} \right|^2. \end{aligned} \tag{5.34}$$

This influence of the angular distance in direction cosine space on the diffracted radiance is also known as *shift invariance*.

BRDF model

Following Sec.3.2 (main paper) we repeat the well-known definition of the bidirectional reflectance distribution function (BRDF) and the accompanying notations for completeness;

$$f_r = \frac{dL_\mathrm{s}(\boldsymbol{x}, \hat{\boldsymbol{\omega}}_o)}{dE_i(\hat{\boldsymbol{\omega}}_i)}, \tag{5.35}$$

which relates differential irradiance to scattered radiance. $\boldsymbol{x}$ represents a position on the surface, $\hat{\boldsymbol{\omega}}_i$ is the direction from which this surface is illuminated and $\hat{\boldsymbol{\omega}}_o$ the direction from which it is observed. The radiance scattered by a diffracting aperture is given by Eq. 5.34 as a function in direction-cosine space using a coordinate system where all spatial variables are normalized to the optical wavelength. A change of variables enables us to rewrite the representation of the Fourier transform in a non-scaled coordinate system as

$$L_s(\boldsymbol{\xi}) = \gamma_i \frac{1}{A_s} \frac{1}{\lambda^2} \left| \mathcal{F}\left\{ U_0(\boldsymbol{x}) \right\} \right|^2_{\xi_{1,2}} \tag{5.36}$$

We can describe the complex wavefront $U_0(\boldsymbol{x})$ in the surface plane by the modulation of the wavefront of the incident light $U_i(\boldsymbol{x})$ with the so-called *transfer function* $\mathcal{T}(\boldsymbol{x})$ [LLL10; Goo96] of the diffracting plane as

$$U_0(\boldsymbol{x}) = U_i(\boldsymbol{x}) \cdot \mathcal{T}(\boldsymbol{x}). \tag{5.37}$$

Since the diffracting aperture is uniformly illuminated ([HVK+00]), we can neglect the position dependence of the complex amplitude of the incident light in the aperture plane. Thus, $U_i(\boldsymbol{x}) = U_i$ is a constant modulation factor. Substitution into Eq. 5.36 then yields

$$L_s(\boldsymbol{\xi}) = \gamma_i \frac{1}{A_s} \frac{1}{\lambda^2} \left| U_i \right|^2 \left| \mathcal{F}\left\{ \mathcal{T}(\boldsymbol{x}) \right\}_{\xi_{1,2}} \right|^2. \tag{5.38}$$

In the context of a BRDF definition, U_i corresponds to the differential incident radiance, i.e. $E_i = |U_i|^2$, which enables us to substitute Eq. 5.38 into Eq. 5.35; we can therefore write

$$f_r(\boldsymbol{\xi}) = \gamma_i \frac{1}{A_s} \frac{1}{\lambda^2} \left| \mathcal{F}\left\{ \mathcal{T}(\boldsymbol{x}) \right\}_{\xi_{1,2}} \right|^2. \tag{5.39}$$

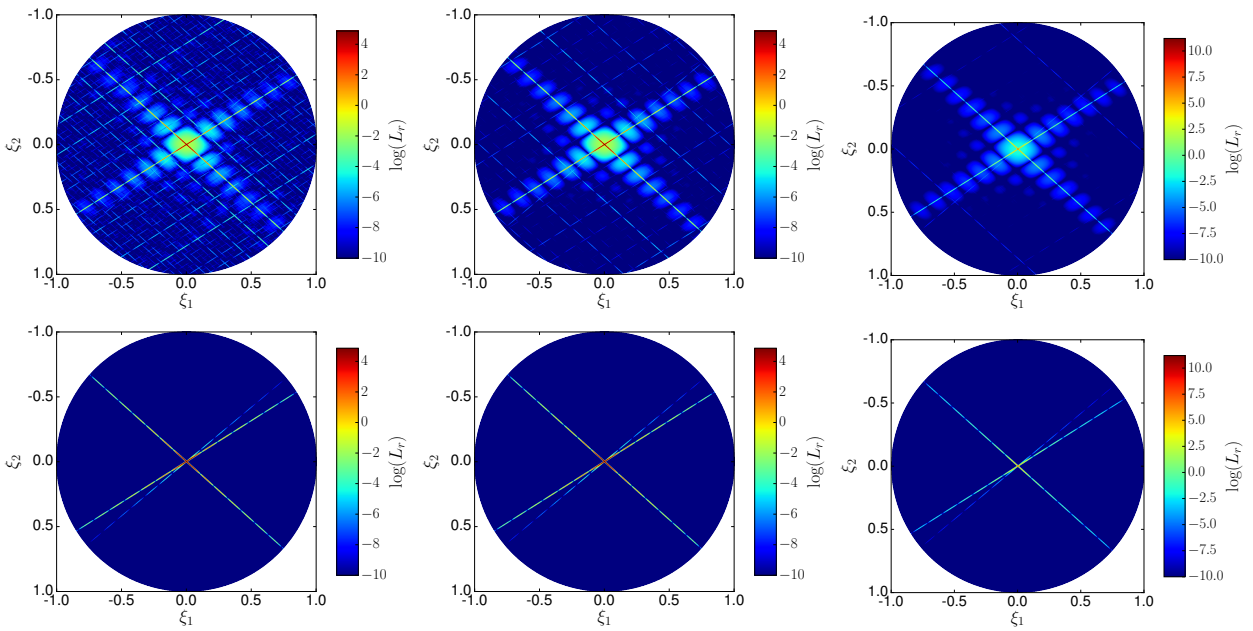

Figure 5.15: Comparison of radiance reflected of a scratched surface. The numerical solution (left) was computed using an explicit heightmap (c.f. Fig. 5.16), the analytical solution (right) was obtained using our model. The ghosting artefacts due to discretization are suppressed with increasing resolution of the heightmap whereas the parallelogram-shaped features originating from scratch-scratch intersections are not reproduced by our model. The (square) heightmaps are supplied with resolutions of 9 MP (upper row), 36 MP (middle row) and 144 MP (bottom row) with an area of 1 mm^2. This figure supplements Sec. 6:Approximation evaluation.

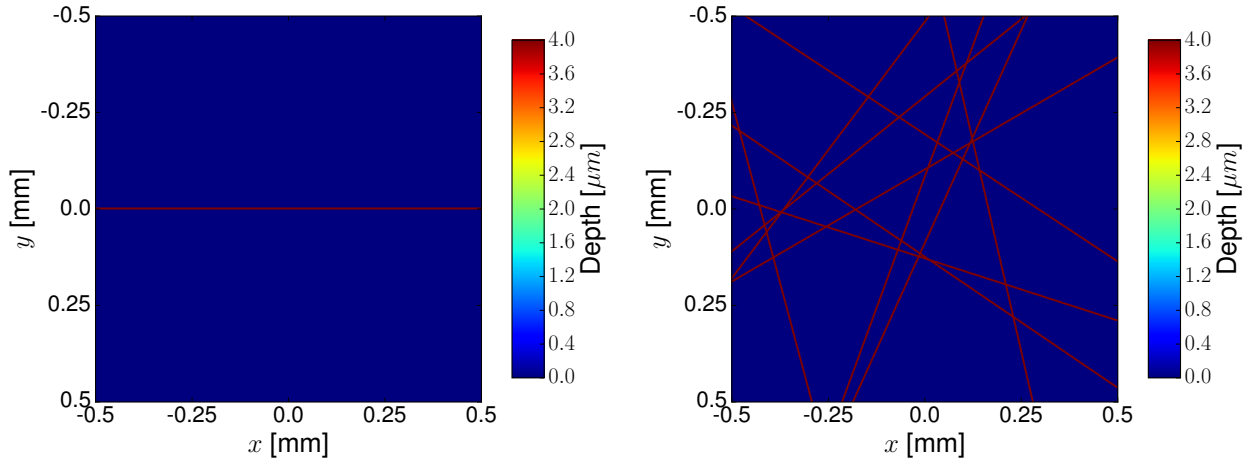

Figure 5.16: Surface heightmaps used as input for the numerical computation of reflected radiance for a single scratch (left) and ten randomly distributed scratches (right). This figure supplements Sec. 6:Approximation evaluation.

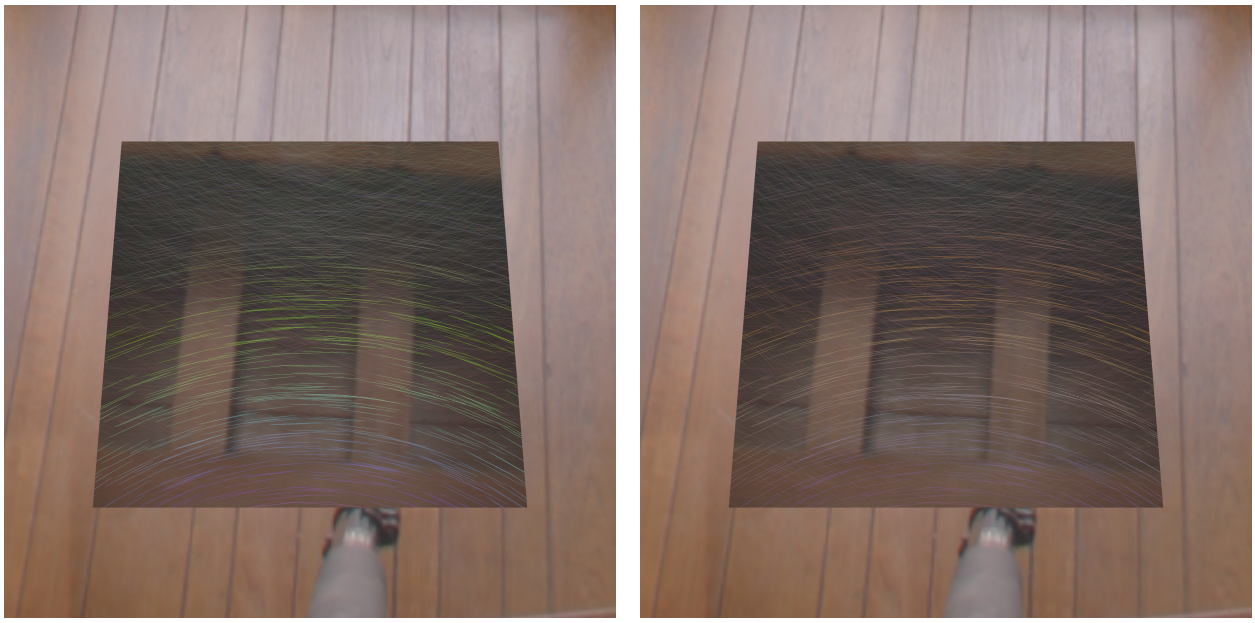

Figure 5.17: Rendering of a scratched plate with 3000 scratches uniformly distributed on the surface. Scratch parameters (depth, width) are kept constant, only the profile is changed from rectangular (left) to triangular (right). The change of the scratch profile results in a different scratch response function and thus reflectance distribution visible as a change in color of the scratches. This figure supplements Sec. 7:Other scratch profiles.

Figure	N_λ	Scene	Scratch count	Scratch parameters	Variation parameters	Render time [min]
Fig. 1 (middle) Fig. 1 (insert)	16	Dining table	300000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	–	93 197
Fig. 1 (right) Fig. 1 (insert)	16	Door	370000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	–	32 67
Fig. 6 (top, left)	16	Door	370000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	$D[1.1,0.04]$	33
Fig. 6 (top, middle) Fig. 6 (insert)	16	Dining table	300000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	$D[1.1,0.04]$	120 147
Fig. 6 (top, right)	16	CD	3128400	$D = 120nm$; $W = 1\mu m$		1029
Fig. 6 (bottom, left)	16	Spoon	100000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	$D[1.1,0.04]$	69
Fig. 6 (bottom, middle)	16					83
Fig. 6 (bottom, right)	16					368
Fig. 7 (b, @1MP)	16	Plate	3000	$G_D[250,80]nm; G_W[4,1.3]\mu m$	$D[1.1,0.04]$	15
Fig. 7 (d)	16	Disc	3700	$G_D[250,80]nm; G_W[5,1.6]\mu m$	$D[1.1,0.03]$	25
Fig. 8 (b)	16	CD	BRDF*	BRDF*	–	200
Fig. 8 (d)	16	CD	BRDF*	BRDF*	–	204
Fig. 9 (left, incoherent)	16	Grated plate	1600	$D = 1\mu$m; $W = 1\mu$m	–	14
Fig. 9 (right, coherent)						19
Fig. 10 (a)	16	Scratched plate	1500	$D = 250nm$; $W = 1\mu m$	$D[1.1,0.04]$	32
Fig. 10 (b, specular)						23
Fig. 10 (c, diffuse)						23
Fig. 11 (a)	16	Dining table	300000	$U_D[0,6]\mu m$; $U_W[0,6]\mu m$	–	266
Fig. 11 (b)					$D[1.1,0.04]$	284
Fig. 11 (c)					–	167
Fig. 11 (d)					$D[1.1,0.04]$	168
Fig. 12 (left, @1MP)	3	Ring	547381	$D = 500nm$; $W = 1\mu m$	–	23
Fig. 12 (right, @1MP)	3	Plate	481905	$G_{D,W}[4.0,0.01]\mu m$		31
Fig. 17:SuppMat (left)	16	Plate	3000	$D = 1\mu$m; $W = 2\mu$	$D[1.1,0.03]$	30
Fig. 17:SuppMat (right)	16	Plate	3000	$D = 1\mu$m; $W = 2\mu$	$D[1.1,0.03]$	31

Table 5.3: Performance values, scratch parameters and render settings for the renderings presented. Rendering was performed on 72 CPU cores (i7-5820K@3.30GHz) and, if not stated otherwise, corresponds to a resolution of 4MP@16384 samples per pixel(SPP). The high number of SPP results from the fact that we rely on Monte Carlo integration over the pixel footprint, which is inherently done by the ray-tracer, to achieve incoherent superposition of the coherent subsamples responsible for diffraction and interference. All scenes except Figures 7, 8, 9 and 10 were scratched using our editing tool.
N_λ gives the number of spectral samples used, the total number of scratches within the scene is given by scratch count.
Scratch parameters describes the parameters used to generate the scratch profiles on the surface, namely width and depth. $U[lower, upper]$ refers to a uniform, $G[mean, stddev]$ to a Gaussian distribution from which these parameters are drawn and the respective parameter is denoted via the index W for width and D for depth. For constant parameters the notation reads $index = value$.
Variation parameters are given as a tuple $[amplitude, frequency]$ describing the relative variation amplitude and frequency of the respective parameter. The frequency thereby is based on a simplex-noise function with 255 gradients, such that a frequency of $1/255$ corresponds to exactly one oscillation along the scratch. Notation for the respective parameter is given by the prefix $W[\cdots]$ for width and $D[\cdots]$ for depth.
Render time is given in minutes and is always set to the next larger integer value.
BRDF* relates to a CD-BRDF generated for periodic surfaces. We created an ensemble of parallel scratches within one circular coherent sample, their parameters depth and width as well as their separation are taken from Stam [Sta99]. Only view- and lighting conditions change within the rendering, the ensemble itself is constant. For the parameters taken from Stam, the number of scratches within the coherence area, and thus every coherent subsample, is $n = 2\frac{3\sigma}{\Delta_{\text{track}}+W_{\text{scratch}}} + 1 \approx 21$, where Δ_{track} is the track separation, W_{scratch} the scratch width and $\sigma = \delta_c/6$ the standard deviation of the Gaussian filter kernel.

CHAPTER 6

Discussion and insights

This final chapter discusses the projects presented in this thesis, provides possible directions for future research and concludes this cumulative thesis as a whole. We therefore split the discussion into summaries of the different projects in Section 6.1 and an overview-oriented conclusion in Section 6.2. For more details on the individual projects, we refer the reader to Sections 2.7, 3.4 and 5.7 respectively.

6.1 Discussion

Our work targets the building blocks of digital doubles directly: The shape and materials used for digital recreations. Per definition, our methods are restricted to the methodologies they are developed for: Our shape acquisition techniques rely on specific hardware that, although commonly available, is tailored to the task and offers characteristics we exploit for our approaches. The surface material we develop on the other hand is useful only in a physically-based rendering (PBR) framework [PJH16] and can only be employed to simulate materials that fulfill the model's assumptions.

Moment-based structured light In Chapter 2, we presented a frequency-domain approach to enhance the capabilities of available phase-shifting structured light shape acquisition setups. These approaches rely on the illumination of an unknown scene with specific spatially modulated patterns and use camera-projector-pixel correspondences to estimate the geometry of objects. Such methods suffer from complex light paths within the scene or even objects, induced by concave geometry or translucent materials for example. We utilize a closed-form spectral estimation to reconstruct per-camera-pixel reflectance profiles as if illuminated by a linesweep. This has the advantage of separating signals the camera would otherwise receive from different scene points along the sweep direction. A naive linesweep would require an amount of acquisitions equal to the number of projector pixels along the sweep direction. In contrast, we can reduce the number of acquisitions to only 20 for a reliable reconstruction in combination with a high quality confidence rating of individual pixel measurements and drastic noise reduction. In addition, we enable the user to semi-automatically edit the result in a data-driven fashion, for example to close gaps in the data. Our method performs well for scenes with and without global light transport contributions

by translucent objects or interreflections induced by complex scene geometry.

Currently, our method is restricted to acquisitions with integer multiples of a base frequency and an equal number of phase shifts per frequency, which introduces ringing artifacts due to undersampling. Measurements with higher frequencies mitigate this effect but drastically increase the amount of acquisitions for our method, posing a severe limitation to the approach. A major step with respect to future work would be to allow for arbitrary combinations of frequencies employed for the measurements as well as non-equal numbers of phase shifts per frequency. This is already available for standard phase-shifting techniques such as presented by Chen et al. [CLF+07] but does not allow for the maximum entropy spectral estimate which is our key component. Here, one way to achieve this goal would be to employ a weighting scheme that accounts for the different counts of phase measurements, to not overestimate the contribution of measurements with more phases. Introducing arbitrary frequencies on the other hand is a difficult task as the underlying mathematical theory of our approach does not allow for such alteration. However, the measurements are performed with integer multiples of a certain base frequency: Instead of acquiring data with a single base frequency, we could use a portfolio of such frequencies and their multiples to sample high-frequency content of the scene, leaving "only" a phase-unwrapping step to be solved.
Additionally, we rely on empirical values for the confidence map (cf. Sec. 2.5) which works well for our scenes. . Our approach utilizes relative measures to estimate the confidence per camera pixel, rendering it robust to scene-based amplitude alterations but might require more careful choice of thresholds for different, not yet considered materials. This in the end has to be done manually or by providing a catalog of possible values.
Our method only reconstructs a linesweep-like illumination with a specific sweep direction. Along this, signals that might otherwise mix in a single camera pixel are separated. For pixels contained in a single line illumination however, this is not the case. A major improvement with this regard would be the extension to two-dimensional measurements that resolve such superpositions, which in theory can be achieved by one of two ways: First, performing more measurements (at least two) with sinusoidal illumination patterns oriented at an angle (perpendicular for two directions) with respect to each other. This would further reduce the possible superpositions to a limited subset, but not fully remove them. Second, the usage of another illumination pattern that is spatially modulated in both dimensions. Although such patterns are readily employed, none allows for a closed-form reconstruction of per-pixel reflectance profiles yet.
Finally, a plausible extension of the approach would be a machine-learning based material characterization that utilizes the measured reflectance profiles. Such a method would also benefit from the aforementioned extension to 2D modulation.

Pulsed Correlation Time-of-Flight Chapter 3 presents our improvements for correlation-based Time-of-Flight range finding sensors. These consist of a laser illumination source (typically infrared) and a specialized sensor. Both are driven by a signal generator, inducing a time-dependent periodic amplitude modulation on the illumination signal and a gain modulation of each sensor pixel. As light traverses the unknown scene, the illumination, usually with sinusoidal or rectangular amplitude modulation, undergoes a phase shift that encodes the travel time and hence the distance. Per sensor pixel, this signal is then "compared" to

the original modulation which allows to reconstruct the depth of an object.
We improve available methods by introducing an additional mode of operation that seamlessly integrates into the standard acquisition scheme: Using the available acquisition mode we obtain a rough range estimate that indicates a certain depth of interest an object is positioned at. With a subsequent measurement combining pulsed illumination and rectangular sensor modulation, we are able to significantly improve the accuracy for this particular depth, introducing a focusing step much like for conventional cameras. The benefit of our approach is strongest for low frequency setups and bad (low signal-to-noise ratio) illumination conditions. These two characteristics combined render our method especially useful for mobile applications such as handheld devices.

A major limiting factor of our approach is the inherent dependence on an sufficiently accurate shape of illumination and sensor modulation signals: We require a delta pulse and a rectangular signal at best. This on the other hand proves difficult to realize in hardware, which limits the achievable accuracy gain.

Linked to this is the requirement of a lookup table that is the result of a rigorous calibration: As pulse shape and modulation signal do not necessarily match the theory, we cannot derive a closed-form solution for the depth estimation but rely on aforementioned lookup table. The key problem here is the inversion of an integral expression. A possible solution to this could be to not directly measure the lookup table but instead measure signal characteristics such as the pulse width and the rise time of the modulation signal. In combination with a more expressive theoretical model, this could allow for a numerical generation of the lookup table, replacing its detailed measurement.
Finally, our approach trades the ability to measure depth variations over a wide range of up to several meters, depending on the modulation frequency, for the highly increased sensitivity around a given depth of interest. This limits our approach to either operate in a two-step fashion as described before, or to task-specific setups where this depth is a priori known: We cannot perform measurement that cover several meters in depth with a single acquisition.

Scratch Iridescence Finally, in Chapter 5 we present an illumination model that simulates opaque materials with microscopic scratches. Due to diffraction, these give rise to iridescent colors that can be found on many everyday items such as kitchen tools for example. We develop this model to close the gap between distinct surface detail commonly rendered using geometrical optics and diffracting surfaces that are available only in models that neglect macroscopic details and focus on random surface defects. To this end, we develop a wave-optical spatially-varying bidirectional reflectance distribution function (BRDF), which is a representation commonly used in physically-based rendering frameworks. Our model thereby accounts for both, diffraction due to single scratches as well as mutual interference from scratch ensembles. To describe these effects, we base our BRDF on physically accurate models and methodologies.
The solution of the involved mathematical descriptions however is only possible due to assumptions that restrict the validity of our model: First, we assume our microscopic scratches to consist of linear segments with either triangular or rectangular profile. Second, our scratch representation does not include intersections of scratches nor self-shadowing effects within the profiles. Third and last, the solution of the diffraction integral makes use of the far-field

assumption, which requires the observer of a scratched surface to be at least centimeters away.
All these assumptions are necessary design choices that ultimately allow to solve the problem in the first place but pose restrictions to the applicability. While the far-field assumption is reasonable for our purpose (we do not consider microscopic observations), neglecting scratch intersections as well as only considering symmetric profiles is not: Such profiles are rather uncommon in real-world examples. However, our approach allows to replace the profiles by in principle arbitrary functions. These however might impose additional computational cost, increasing the already costly computation even more. Still, for specific cases we can exploit certain properties of the mathematical expression and achieve real-time performance, which has been published in Velinov et al. [VWH18] as a result of additional research from the authors of the paper contained in this thesis.
A further limitation is the restriction to only opaque materials. We showed that our model is able to represent the reflection properties of scratched surfaces. Our model however in principle generalizes to diffraction effects observed on transmissive, thin materials without internal scattering, like foils or very light curtains. This would be an interesting avenue to explore, as many real world diffraction effects are observed in transmittance rather than reflectance.
Finally, it would be great to utilize our model to infer parameters of scratched surfaces from real-world samples. Given real-world data, both, experimental validation of our model and fitting of model parameters remains a major challenge due to the large number of unknowns involved. As of now, it is not yet clear what kind of input data would be needed to provide sufficient constraints, how to acquire such data, how to represent the surface (individual scratches or distribution parameters) and how to determine the model parameters in finite time. Answering these questions provides an interesting topic for future research.

6.2 Conclusion and outlook

In this thesis we presented three approaches to widen the gamut and enhance the quality of available digital doubles. Digital doubles are omnipresent with respect to computer-generated imagery, as they encode the geometrical and optical properties required to generate a faithful digital recreation of real-world- and virtual objects alike. Our approaches build on available off-the-shelf hardware and target commonly used measurement setups and software frameworks to allow for reliable and accessible implementation. In particular, we presented two range finding methods that significantly enhance the quality of shape acquisition techniques, but do not require new hardware. Instead, we deliberately developed our approaches to make them available to a wide community. In Chapter 2 we present a modification for the acquisition and processing modalities of structured light methods that utilize a conventional camera and projector. In Chapter 3, we develop an additional mode of operation that relies on standard Correlation Time-of-Flight sensors that can be acquired at reasonable cost. Thereby, our approach is purely implemented in software, effectively reusing and enhancing the capabilities of available hardware solutions. On the other hand, our illumination model to simulate worn and scratched surfaces with microscopic surface defects as presented in Chapter 5 is built on a rigorous theoretical description that allows

for its use in available physically-based rendering frameworks.
Our methods are based on results from a variety of different disciplines, including computer graphics, physics (especially wave-optics) and the mathematics of signal processing. Thereby, each of our publications targets a specific problem and either solves previously open challenges, or improves the existing state-of-the art. Apart from these improvements and additions to the spectrum of digital doubles we see future challenges that can make use of our approaches.
Shape acquisition for instance has an ever-growing need for high-quality range-finding techniques that are fast and resilient with respect to complex light transport. We have shown that with available, low-cost hardware, structured light methods can be realized that exhibit a significantly improved measure of confidence with respect to measurements in combination with data-driven editing for error correction. Our approach does not require any additional laboratory equipment at all and can be employed for industrial use-cases as well as hobbyist projects alike. However, the closed-form reconstruction of per-pixel profiles at this point cannot separate contributions from a virtual line of illumination pixels. We here see the opportunity to improve the spectral estimation technique to 2D to shift the focus of the approach from shape acquisition towards the acquisition of reflectance fields. This would open opportunities especially in the fields of material characterization, virtual relighting and dual photography [SCG+05; SD09].
Our pulsed Correlation Time-of-Flight range finding technique on the other hand requires specific hardware tailored to the task. Such modules however are commonly available at low cost, often in the form of developer boards with an accompanying interface for programmers, which we rely on. Here again, we offer a software solution for improvements. Our method scales with new generations of ToF sensors as long as the working principle remains unchanged. As a two-step procedure however, it relies on a first measurement to obtain a rough notion of depths present in a scene. Here, we see the possibility to draw on ideas from other *sensor fusion* approaches [CEJ15; ECJ17], where different kinds of sensors, such as ToF and conventional cameras are combined to provide high-resolution depth maps for example [FRR+13; FAT11]. Here we see the opportunity to combine ToF approaches such as ours with spatial modulation, combining the advantages of ToF and structured light range finding.
Our model to simulate iridescent scratches on worn surfaces is a purely parameter-driven BRDF model that utilizes an abstract description of scratches by their profile, width and depth. This way, the process remains data-independent and no high-cost measurement device is required. As recent approaches enable real-time performance of a simplified version of our method, the high computation cost of the original is no longer a drawback. Although being targeted at a niche effect, we feel that our publication shares renewed interest in the computer graphics community with concurrent research that aims at wave-optical effects [YHW+18]. We expect future work in this area to move from surface-based considerations to the simulation of wave-optical effects such as diffraction from three-dimensional objects.

Digital doubles, as a key component of CGI, will become more and more important as the amount of applications grows, ranging from entertainment to education and training of specialized personnel with virtual- and augmented reality setups. Likewise, the demand for

ever more realistic digital materials and more accurate geometry acquisition increases, leading to the emergence of new technologies. As with our approaches today, we look forward to see these new technologies be improved, to craft more realistic digital doubles.

List of Figures

Bibliography

[Bat84] DR Bates. Rayleigh scattering by air. *Planetary and Space Science*, 32(6):785–790, 1984.

[BB17] Laurent Belcour and Pascal Barla. A Practical Extension to Microfacet Theory for the Modeling of Varying Iridescence. *ACM Transactions on Graphics*, 36(4):65, July 2017. DOI: 10.1145/3072959.3073620. URL: https://hal.archives-ouvertes.fr/hal-01518344.

[BBN+14] Pascal Bérard, Derek Bradley, Maurizio Nitti, Thabo Beeler, and Markus Gross. High-quality capture of eyes. *ACM Transactions on Graphics (TOG)*, 33(6):223, 2014.

[BBP15] P. Barla, L. Belcour, and R. Pacanowski. In praise of an alternative brdf parametrization. In *Proceedings of the Third Workshop on Material Appearance Modeling: Issues and Acquisition*, MAM '15, pages 9–13, Darmstadt, Germany. Eurographics Association, 2015. ISBN: 978-3-905674-83-5. DOI: 10.2312/mam.20151197. URL: http://dx.doi.org/10.2312/mam.20151197.

[BDW81] FO Bartell, EL Dereniak, and WL Wolfe. The theory and measurement of bidirectional reflectance distribution function (brdf) and bidirectional transmittance distribution function (btdf). In *Radiation scattering in optical systems*, volume 257, pages 154–160. International Society for Optics and Photonics, 1981.

[BL99] John W. Buchanan and Paul Lalonde. An observational model for illuminating isolated scratches. *Proc. Western Computer Graphics Symposium 1999 (SKIGRAPH'99)*, 1999.

[Ble18] Blender Online Community. *Blender - a 3D modelling and rendering package*. Blender Foundation. Blender Institute, Amsterdam, 2018. URL: http://www.blender.org.

[Bli77] James F Blinn. Models of light reflection for computer synthesized pictures. In *ACM SIGGRAPH computer graphics*, volume 11 of number 2, pages 192–198. ACM, 1977.

[Bli78] Jim Blinn. Simulation of wrinkled surfaces. In *Siggraph 1978*, siggraph 1978 edition. Association for Computing Machinery, Inc., January 1978. URL: https://www.microsoft.com/en-us/research/publication/simulation-of-wrinkled-surfaces/.

[BPK+07] Mario Botsch, Mark Pauly, Leif Kobbelt, Pierre Alliez, Bruno Lévy, Stephan Bischoff, and Christian Röossl. Geometric modeling based on polygonal meshes, 2007.

[BPM+04] Carles Bosch, Xavier Pueyo, Stéphane Mérillou, and Djamchid Ghazanfarpour. A physically-based model for rendering realistic scratches. *Computer Graphics Forum*, 23(3):361–370, 2004. ISSN: 1467-8659. DOI: 10.1111/j.1467-8659.2004.00767.x. URL: http://dx.doi.org/10.1111/j.1467-8659.2004.00767.x.

[BPM+08] C. Bosch, X. Pueyo, S. Merillou, and D. Ghazanfarpour. A Resolution Independent Approach for the Accurate Rendering of Grooved Surfaces. *Computer Graphics Forum*, 2008. ISSN: 1467-8659. DOI: 10.1111/j.1467-8659.2008.01342.x.

[BRB+19a] Thomas Buffet, Damien Rohmer, Loïc Barthe, Laurence Boissieux, and Marie-Paule Cani. Implicit Untangling: A Robust Solution for Modeling Layered Clothing. In volume 38 of number 4 in *Proc. ACM SIGGRAPH*, Article No. 120:1–12. Association for Computing Machinery, July 2019. DOI: 10.1145/3306346.3323010. URL: https://hal.archives-ouvertes.fr/hal-02129156.

[BRB+19b] Thomas Buffet, Damien Rohmer, Loïc Barthe, Laurence Boissieux, and Marie-Paule Cani. Implicit Untangling: A Robust Solution for Modeling Layered Clothing. *ACM Transactions on Graphics*. Proc. ACM SIGGRAPH, 38(4):Article No. 120:1–12, July 2019. DOI: 10.1145/3306346.3323010. URL: https://hal.archives-ouvertes.fr/hal-02129156.

[BS08] B. Buttgen and P. Seitz. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 55(6):1512–1525, July 2008. ISSN: 1549-8328. DOI: 10.1109/TCSI.2008.916679.

[Bur79] J.P. Burg. *Maximum Entropy Spectral Analysis: A Dissertation*. University Microfilms, 1979. URL: https://books.google.de/books?id=LWecPwAACAAJ.

[BW13] Max Born and Emil Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[CEJ15] Dmitrij Csetverikov, Iván Eichhardt, and Zsolt Jankó. A brief survey of image-based depth upsampling. In *KÉPAF 2015*, pages 279–294. [s. n.], 2015. URL: http://eprints.sztaki.hu/8543/.

[CGS06] Tongbo Chen, Michael Goesele, and Hans-Peter Seidel. Mesostructure from specularity. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 1825–1832. IEEE, 2006.

[CHB+12] Tom Cuypers, Tom Haber, Philippe Bekaert, Se Baek Oh, and Ramesh Raskar. Reflectance model for diffraction. *ACM Trans. Graph.*, 31(5):122:1–122:11, September 2012. ISSN: 0730-0301. DOI: 10.1145/2231816.2231820. URL: http://doi.acm.org/10.1145/2231816.2231820.

[CKK05] Manmohan Krishna Chandraker, Fredrik Kahl, and David J Kriegman. Reflections on the generalized bas-relief ambiguity. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 788–795. IEEE, 2005.

[CLF+07] Tongbo Chen, Hendrik PA Lensch, Christian Fuchs, and Hans-Peter Seidel. Polarization and phase-shifting for 3D scanning of translucent objects. In *2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2007.

[CSL08] Tongbo Chen, Hans-Peter Seidel, and Hendrik Lensch. Modulated phase-shifting for 3D scanning. In *2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8. IEEE, 2008.

[CT09] Eugene L Church and Peter Z Takacs. Surface scattering. In Bass, M. et al., editor, *Handbook of Optics*, chapter 7, pages 7.1–7.14. McGraw-Hill, 2009.

[CT82] Robert L Cook and Kenneth E Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1(1):7–24, 1982.

[dFH+11] Eugene d'Eon, Guillaume Francois, Martin Hill, Joe Letteri, and Jean-Marie Aubry. An energy-conserving hair reflectance model. *Computer Graphics Forum*, 30(4), 2011. ISSN: 1467-8659.

[DFP+10] Francesca De Crescenzio, Massimiliano Fantini, Franco Persiani, Luigi Di Stefano, Pietro Azzari, and Samuele Salti. Augmented reality for aircraft maintenance training and operations support. *IEEE Computer Graphics and Applications*, 31(1):96–101, 2010.

[DJ18] Jonathan Dupuy and Wenzel Jakob. An adaptive parameterization for efficient material acquisition and rendering. In *SIGGRAPH Asia 2018 Technical Papers*, page 274. ACM, 2018.

[DN15] Shawn Divitt and Lukas Novotny. Spatial coherence of sunlight and its implications for light management in photovoltaics. *Optica*, 2(2):95–103, February 2015. DOI: 10.1364/OPTICA.2.000095. URL: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-2-2-95.

[DRS10] Julie Dorsey, Holly Rushmeier, and François Sillion. *Digital modeling of material appearance*. Morgan Kaufmann, 2010.

[DTS+14] D.S. Dhillon, J. Teyssier, M. Single, I. Gaponenko, M.C. Milinkovitch, and M. Zwicker. Interactive diffraction from biological nanostructures. *Comput. Graph. Forum*, 33(8):177–188, December 2014. ISSN: 0167-7055. DOI: 10.1111/cgf.12425. URL: http://dx.doi.org/10.1111/cgf.12425.

[DWM+15] Zhao Dong, Bruce Walter, Steve Marschner, and Donald P. Greenberg. Predicting appearance from measured microgeometry of metal surfaces. *ACM Trans. Graph.*, 35(1):9:1–9:13, December 2015. ISSN: 0730-0301. DOI: 10.1145/2815618. URL: http://doi.acm.org/10.1145/2815618.

[DWT+10] Yue Dong, Jiaping Wang, Xin Tong, John Snyder, Yanxiang Lan, Moshe Ben-Ezra, and Baining Guo. Manifold bootstrapping for SVBRDF capture. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 29(4):98:1–98:10, 2010.

[ECJ17] Ivan Eichhardt, Dmitry Chetverikov, and Zsolt Janko. Image-guided tof depth upsampling: a survey. *Machine Vision and Applications*, 28(3-4):267–282, 2017.

[FAT11] Sergi Foix, Guillem Alenya, and Carme Torras. Lock-in time-of-flight (tof) cameras: a survey. *IEEE Sensors Journal*, 11(9):1917–1926, 2011.

[FLR+98] Henry Fuchs, Mark A Livingston, Ramesh Raskar, Kurtis Keller, Jessica R Crawford, Paul Rademacher, Samuel H Drake, Anthony A Meyer, et al. Augmented reality visualization for laparoscopic surgery. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 934–943. Springer, 1998.

[FRR+13] David Ferstl, Christian Reinbacher, Rene Ranftl, Matthias Rüther, and Horst Bischof. Image guided depth upsampling using anisotropic total generalized variation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 993–1000, 2013.

[FSK+14] Daniel Freedman, Yoni Smolin, Eyal Krupka, Ido Leichter, and Mirko Schmidt. SRA: fast removal of general multipath for ToF sensors. In David Fleet, Tomas Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision – ECCV 2014*, pages 234–249, Cham. Springer International Publishing, 2014. ISBN: 978-3-319-10590-1.

[GAV+11] Mohit Gupta, Amit Agrawal, Ashok Veeraraghavan, and Srinivasa G Narasimhan. Structured light 3D scanning in the presence of global illumination. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 713–720. IEEE, 2011.

[GBR+01] Guy Godin, Jean-Angelo Beraldin, Marc Rioux, Marc Levoy, Luc Cournoyer, and Francois Blais. An assessment of laser range measurement of marble surfaces. In *Proc. Fifth Conference on optical 3-D measurement techniques*, 2001.

[GD67] B. S. Goldstein and G. F. Dalrymple. Gallium arsenide injection laser radar. *Proceedings of the IEEE*, 55(2):181–188, February 1967. ISSN: 0018-9219. DOI: 10.1109/PROC.1967.5437.

[GGG+16] Darya Guarnera, Giuseppe Claudio Guarnera, Abhijeet Ghosh, Cornelia Denk, and Mashhuda Glencross. Brdf representation and acquisition. In *Computer Graphics Forum*, volume 35 of number 2, pages 625–650. Wiley Online Library, 2016.

[GNH+15] Mohit Gupta, Shree K. Nayar, Matthias B. Hullin, and Jaime Martin. Phasor imaging: a generalization of correlation-based time-of-flight imaging. *ACM Trans. Graph.*, 34(5):156:1–156:18, November 2015. ISSN: 0730-0301. DOI: 10.1145/2735702. URL: http://doi.acm.org/10.1145/2735702.

[Goo96] J.W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill Series in Electrical and Computer Engineering: Communications and Signal Processing. McGraw-Hill, 1996. ISBN: 9780070242548.

[GVN+18] Mohit Gupta, Andreas Velten, Shree K. Nayar, and Eric Breitbach. What are optimal coding functions for time-of-flight imaging? *ACM Trans. Graph.*, 37(2):13:1–13:18, February 2018. ISSN: 0730-0301. DOI: 10.1145/3152155. URL: http://doi.acm.org/10.1145/3152155.

[HHG+13] Felix Heide, Matthias B. Hullin, James Gregson, and Wolfgang Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Trans. Graph.*, 32(4):45:1–45:10, July 2013. ISSN: 0730-0301. DOI: 10.1145/2461912.2461945. URL: http://doi.acm.org/10.1145/2461912.2461945.

[HL11] Michael Holroyd and Jason Lawrence. An analysis of using high-frequency sinusoidal illumination to measure the 3D shape of translucent objects. In *CVPR 2011*, pages 2985–2991. IEEE, 2011.

[HP16] Nicolas Holzschuch and Romain Pacanowski. A Physically-Based Reflectance Model Combining Reflection and Diffraction. Research Report RR-8964, INRIA, October 2016. URL: https://hal.inria.fr/hal-01386157.

[HP17] Nicolas Holzschuch and Romain Pacanowski. A two-scale microfacet reflectance model combining reflection and diffraction. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 36(4):66:1–66:12, 2017.

[HTS+91] Xiao D. He, Kenneth E. Torrance, François X. Sillion, and Donald P. Greenberg. A comprehensive physical model for light reflection. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '91, pages 175–186, New York, NY, USA. ACM, 1991. ISBN: 0-89791-436-8. DOI: 10.1145/122718.122738. URL: http://doi.acm.org/10.1145/122718.122738.

[HvdH81] Hendrik Christoffel Hulst and Hendrik C van de Hulst. *Light scattering by small particles*. Courier Corporation, 1981.

[HVK+00] James E. Harvey, Cynthia L. Vernold, Andrey Krywonos, and Patrick L. Thompson. Diffracted radiance: a fundamental quantity in nonparaxial scalar diffraction theory: errata. *Appl. Opt.*, 39(34):6374–6375, December 2000. DOI: 10.1364/AO.39.006374. URL: http://ao.osa.org/abstract.cfm?URI=ao-39-34-6374.

[HWR+10] Christian Hansen, Jan Wieferich, Felix Ritter, Christian Rieder, and Heinz-Otto Peitgen. Illustrative visualization of 3d planning models for augmented reality in liver surgery. *International journal of computer assisted radiology and surgery*, 5(2):133–141, 2010.

[HZ03] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge university press, 2003.

[IY01] G. J. Iddan and Giora Yahav. 3D imaging in the studio (and elsewhere...) In *Three-Dimensional Image Capture and Applications IV*, volume 4298, 2001.

[JBP+11] A. P. P. Jongenelen, D. G. Bailey, A. D. Payne, A. A. Dorrington, and D. A. Carnegie. Analysis of errors in ToF range imaging with dual-frequency modulation. *IEEE Transactions on Instrumentation and Measurement*, 60(5):1861–1868, May 2011. ISSN: 0018-9456. DOI: 10.1109/TIM.2010.2089190.

[JCP+10] A. P. P. Jongenelen, D. A. Carnegie, A. D. Payne, and A. A. Dorrington. Maximizing precision over extended unambiguous range for ToF range imaging systems. In *2010 IEEE Instrumentation Measurement Technology Conference Proceedings*, pages 1575–1580, May 2010. DOI: 10.1109/IMTC.2010.5488178.

[JHY+14] Wenzel Jakob, Miloš Hašan, Ling-Qi Yan, Jason Lawrence, Ravi Ramamoorthi, and Steve Marschner. Discrete stochastic microfacet models. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 33(4):115:1–115:10, July 2014. ISSN: 0730-0301. DOI: 10.1145/2601097.2601186. URL: http://doi.acm.org/10.1145/2601097.2601186.

[KBK+09] Andreas Kolb, Erhardt Barth, Reinhard Koch, and Rasmus Larsen. Time-of-flight sensors in computer graphics. In M. Pauly and G. Greiner, editors, *Eurographics 2009 - State of the Art Reports*. The Eurographics Association, 2009. DOI: 10.2312/egst.20091064.

[KGS00] Terry A King, Francis Graham-Smith, and Sir Frances Graham Smith. *Optics and photonics: an introduction*. Wiley, 2000.

[KHD+09] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar. Looking around the corner using transient imaging. In *2009 IEEE 12th International Conference on Computer Vision*, pages 159–166, September 2009. DOI: 10.1109/ICCV.2009.5459160.

[Koe68] W. Koechner. Optical ranging system employing a high power injection laser diode. *IEEE Transactions on Aerospace Electronic Systems*, 4:81–91, January 1968. DOI: 10.1109/TAES.1968.5408936.

[KRP+15] Oliver Klehm, Fabrice Rousselle, Marios Papas, Derek Bradley, Christophe Hery, Bernd Bickel, Wojciech Jarosz, and Thabo Beeler. Recent advances in facial appearance capture. In *Computer Graphics Forum*, volume 34 of number 2, pages 709–733. Wiley Online Library, 2015.

[Kry06] Andrey Krywonos. *Predicting Surface Scatter Using a Linear Systems Formulation of Non-paraxial Scalar Diffraction*. University of Central Florida, 2006. ISBN: 9780542975868. URL: http://etd.fcla.edu/CF/CFE0001446/Krywonos_Andrey_200612_PhD.pdf.

[KWB+13] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (TOG)*, 32(6):167, 2013.

[Lee12] Kangdon Lee. Augmented reality in education and training. *TechTrends*, 56(2):13–21, 2012.

[LGS+09] C. Lauterbach, M. Garland, S. Sengupta, D. Luebke, and D. Manocha. Fast BVH Construction on GPUs. *Computer Graphics Forum*, 2009. ISSN: 1467-8659. DOI: 10.1111/j.1467-8659.2009.01377.x.

[LGX+13] Anat Levin, Daniel Glasner, Ying Xiong, Frédo Durand, William Freeman, Wojciech Matusik, and Todd Zickler. Fabricating BRDFs at high spatial resolution using wave optics. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 32(4):144:1–144:14, July 2013. ISSN: 0730-0301. DOI: 10.1145/2461912.2461981. URL: http://doi.acm.org/10.1145/2461912.2461981.

[LKK00] Rong Lu, Jan J. Koenderink, and Astrid M.L. Kappers. Specularities on surfaces with tangential hairs or grooves. *Comput. Vis. Image Underst.*, 78(3):320–335, June 2000. ISSN: 1077-3142. DOI: 10.1006/cviu.2000.0841. URL: http://dx.doi.org/10.1006/cviu.2000.0841.

[LKY+12] Joakim Löw, Joel Kronander, Anders Ynnerman, and Jonas Unger. BRDF models for accurate and efficient rendering of glossy surfaces. *ACM Trans. Graph.*, 31(1):9:1–9:14, February 2012. ISSN: 0730-0301. DOI: 10.1145/2077341.2077350. URL: http://doi.acm.org/10.1145/2077341.2077350.

[LLL10] Ariel Lipson, Stephen G Lipson, and Henry Lipson. *Optical physics*. Cambridge University Press, Leiden, 2010. URL: https://cds.cern.ch/record/1338386.

[LNM10] Siying Liu, Tian-Tsong Ng, and Yasuyuki Matsushita. Shape from second-bounce of light transport. In *European Conference on Computer Vision*, pages 280–293. Springer, 2010.

[LSB+00] Robert Lange, Peter Seitz, Alice Biber, and Stefan C Lauxtermann. Demodulation pixels in ccd and cmos technologies for time-of-flight ranging. In *Sensors and camera systems for scientific, industrial, and digital photography applications*. International Society for Optics and Photonics, 2000.

[Luc19] Lucid Vision Labs, Inc. Sony Depth Sense IMX556. https://thinklucid.com/tech-briefs/sony-depthsense-how-it-works/, 2019. Accessed: 2019-11-20.

[MDG01] S. Merillou, J.M. Dischler, and D. Ghazanfarpour. Surface scratches: measuring, modeling and rendering. *The Visual Computer*, 17(1):30–45, 2001. ISSN: 1432-2315. DOI: 10.1007/s003710000093. URL: http://dx.doi.org/10.1007/s003710000093.

[MHP+07] Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 183–194. Eurographics Association, 2007.

[Mik08] Morten Mikkelsen. Simulation of wrinkled surfaces revisited, 2008.

[MMR+13] A. Musbach, G. W. Meyer, F. Reitich, and S. H. Oh. Full wave modelling of light propagation and reflection. *Computer Graphics Forum*, 32(6):24–37, 2013. ISSN: 1467-8659. DOI: 10.1111/cgf.12012. URL: http://dx.doi.org/10.1111/cgf.12012.

[MW95] Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995. ISBN: 9780521417112.

[MWA+05] Stephen R. Marschner, Stephen H. Westin, Adam Arbree, and Jonathan T. Moon. Measuring and modeling the appearance of finished wood. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 24(3):727–734, July 2005.

[MWL+99] Stephen R Marschner, Stephen H Westin, Eric PF Lafortune, Kenneth E Torrance, and Donald P Greenberg. Image-based brdf measurement including human skin. In *Rendering Techniques' 99*, pages 131–144. Springer, 1999.

[NDM05] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of BRDF models. In *Proceedings of the Sixteenth Eurographics Conference on Rendering Techniques*, EGSR '05, pages 117–126, Konstanz, Germany. Eurographics Association, 2005. ISBN: 3-905673-23-1. DOI: 10.2312/EGWR/EGSR05/117-126. URL: http://dx.doi.org/10.2312/EGWR/EGSR05/117-126.

[NIK91] Shree K Nayar, Katsushi Ikeuchi, and Takeo Kanade. Shape from interreflections. *International Journal of Computer Vision*, 6(3):173–195, 1991.

[NKG+06] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics (TOG)*, 25(3):935–944, July 2006. ISSN: 0730-0301. DOI: 10.1145/1141911.1141977. URL: http://doi.acm.org/10.1145/1141911.1141977.

[NRH+92] Fred E Nicodemus, Joseph C Richmond, Jack J Hsia, IW Ginsberg, and T Limperis. Geometrical considerations and nomenclature for reflectance. *NBS monograph*, 160:4, 1992.

[OMK14] Matthew O'Toole, John Mather, and Kiriakos N Kutulakos. 3D shape and indirect appearance by structured light transport. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3246–3253, 2014.

[ON94] Michael Oren and Shree K Nayar. Generalization of lambert's reflectance model. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 239–246. ACM, 1994.

[PDC+08] Andrew D Payne, Adrian A Dorrington, Michael J Cree, and Dale A Carnegie. Improved linearity using harmonic error rejection in a full-field range imaging system. In *Three-Dimensional Image Capture and Applications 2008*. International Society for Optics and Photonics, 2008.

[PDC+10] Andrew D. Payne, Adrian A. Dorrington, Michael J. Cree, and Dale Anthony Carnegie. Improved measurement linearity and precision for amcw time-of-flight range imaging cameras. *Applied optics*, 49 23:4392–403, 2010.

[PDC11] Andrew D Payne, Adrian A Dorrington, and Michael J Cree. Illumination waveform optimization for time-of-flight range imaging cameras. In *Videometrics, Range Imaging, and Applications XI*. International Society for Optics and Photonics, 2011.

[Per02] Ken Perlin. Improving noise. In *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '02, pages 681–682, San Antonio, Texas. ACM, 2002. ISBN: 1-58113-521-1. DOI: 10.1145/566570.566636. URL: http://doi.acm.org/10.1145/566570.566636.

[Pho75] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 18(6):311–317, 1975.

[PJH16] Matt Pharr, Wenzel Jakob, and Greg Humphreys. *Physically based rendering: From theory to implementation*. Morgan Kaufmann, 2016.

[PK08] Johnny Park and Avinash Kak. 3D modeling of optically challenging objects. *IEEE Transactions on Visualization and Computer Graphics*, 14(2):246–262, 2008.

[PKH+15] Christoph Peters, Jonathan Klein, Matthias B. Hullin, and Reinhard Klein. Solving trigonometric moment problems for fast transient imaging. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)*, 34(6), November 2015. DOI: 10.1145/2816795.2818103.

[PVB+11] R. Pandharkar, A. Velten, A. Bardagjy, E. Lawson, M. Bawendi, and R. Raskar. Estimating motion and size of moving non-line-of-sight objects in cluttered environments. In *CVPR 2011*, pages 265–272, June 2011. DOI: 10.1109/CVPR.2011.5995465.

[RGB16] Boris Raymond, Gaël Guennebaud, and Pascal Barla. Multi-scale rendering of scratched materials using a structured SV-BRDF model. *ACM Trans. Graph.*, 35(4):57:1–57:11, July 2016. ISSN: 0730-0301. DOI: 10.1145/2897824.2925945. URL: http://doi.acm.org/10.1145/2897824.2925945.

[RRC12] Dikpal Reddy, Ravi Ramamoorthi, and Brian Curless. Frequency-space decomposition and acquisition of light transport under spatially varying illumination. In *European Conference on Computer Vision (ECCV)*, pages 596–610. Springer, 2012.

[RRT+16] Sean Ryan Fanello, Christoph Rhemann, Vladimir Tankovich, Adarsh Kowdle, Sergio Orts Escolano, David Kim, and Shahram Izadi. Hyperdepth: learning depth from structured light without matching. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5441–5450, 2016.

[RVR+17] Sean Ryan Fanello, Julien Valentin, Christoph Rhemann, Adarsh Kowdle, Vladimir Tankovich, Philip Davidson, and Shahram Izadi. Ultrastereo: efficient learning-based matching for active stereo systems. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2691–2700, 2017.

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

[SD09] Pradeep Sen and Soheil Darabi. Compressive dual photography. *Computer Graphics Forum*, 28(2):609–618, 2009.

[SFD+00] Yinlong Sun, F. David Fracchia, Mark S. Drew, and Thomas W. Calvert. Rendering iridescent colors of optical disks. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, pages 341–352, London, UK, UK. Springer-Verlag, 2000. ISBN: 3-211-83535-0. URL: http://dl.acm.org/citation.cfm?id=647652.732138.

[SFP+10] Joaquim Salvi, Sergio Fernandez, Tomislav Pribanic, and Xavier Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.

[Shu04] Jeffrey H Shuhaiber. Augmented reality in surgery. *Archives of surgery*, 139(2):170–174, 2004.

[SK05] Yoav Y Schechner and Nir Karpel. Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of Oceanic Engineering*, 30(3):570–587, 2005.

[SK12] Christopher Schwartz and Reinhard Klein. Acquisition and presentation of virtual surrogates for cultural heritage artefacts. In *EVA 2012 Berlin*, pages 50–57, Volmerstraße 3, 12489 Berlin. Gesellschaft zur Förderung angewandter Informatik e.V., November 2012. ISBN: 978-3-942709-06-4.

[Smi98] Brian Smits. Efficiency issues for ray tracing. *J. Graph. Tools*, 3(2):1–14, February 1998. ISSN: 1086-7651. DOI: 10.1080/10867651.1998.10487488. URL: http://dx.doi.org/10.1080/10867651.1998.10487488.

[SNN03] Yoav Y Schechner, Srinivasa G Narasimhan, and Shree K Nayar. Polarization-based vision through haze. *Applied Optics*, 42(3):511–525, 2003.

[SNT+98] Yoshinobu Sato, Masahiko Nakamoto, Yasuhiro Tamaki, Toshihiko Sasama, Isao Sakita, Yoshikazu Nakajima, Morito Monden, and Shinichi Tamura. Image guidance of breast cancer surgery using 3-d ultrasound images and augmented reality visualization. *IEEE Transactions on Medical Imaging*, 17(5):681–693, 1998.

[SOB+04] Tobias Sielhorst, Tobias Obst, Rainer Burgkart, Robert Riener, and Nassir Navab. An augmented reality delivery simulator for medical training. In *International workshop on augmented environments for medical imaging-MICCAI Satellite Workshop*, volume 141, pages 11–20, 2004.

[Sta99] Jos Stam. Diffraction shaders. In *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '99, pages 101–110, New York, NY, USA. ACM Press/Addison-Wesley Publishing Co., 1999. ISBN: 0-201-48560-5. DOI: 10.1145/311535.311546. URL: http://dx.doi.org/10.1145/311535.311546.

[SXH+97] Rudolf Schwarte, Zhanping Xu, Horst-Guenther Heinol, Joachim Olk, Ruediger Klein, Bernd Buxbaum, Helmut Fischer, and Juergen Schulte. New electro-optical mixing and correlating sensor: facilities and applications of the photonic mixer device (pmd). In *Sensors, Sensor Systems, and Sensor Data Processing*. International Society for Optics and Photonics, 1997.

[TAH+07] Eino-Ville Talvala, Andrew Adams, Mark Horowitz, and Marc Levoy. Veiling glare in high dynamic range imaging. *ACM Transactions on Graphics (TOG)*, 26(3):37, 2007.

[TG17] A Toisoul and A Ghosh. Practical acquisition and rendering of diffraction effects in surface reflectance. *ACM Transactions on Graphics*, 2017. DOI: 10.1145/3012001. URL: http://dx.doi.org/10.1145/3012001.

[THP+19] Jiapeng Tang, Xiaoguang Han, Junyi Pan, Kui Jia, and Xin Tong. A skeleton-bridged deep learning approach for generating meshes of complex topologies from single rgb images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4541–4550, 2019.

[Tor16] Shane Torbert. *Applied Computer Science*. Springer, 2016.

[TR75] T. S. Trowbridge and K. P. Reitz. Average irregularity representation of a rough surface for ray reflection. *J. Opt. Soc. Am.*, 65(5):531–536, May 1975. DOI: 10.1364/JOSA.65.000531. URL: http://www.osapublishing.org/abstract.cfm?URI=josa-65-5-531.

[VG95] Eric Veach and Leonidas J Guibas. Optimally combining sampling techniques for monte carlo rendering. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 419–428. ACM, 1995.

[VWG+12] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature communications*, 3:745, 2012.

[VWH18] Zdravko Velinov, Sebastian Werner, and Matthias B Hullin. Real-time rendering of wave-optical effects on scratched surfaces. In *Computer Graphics Forum*, volume 37 of number 2, pages 123–134. Wiley Online Library, 2018.

[WIC+19] Sebastian Werner, Julian Iseringhausen, Clara Callenberg, and Matthias Hullin. Trigonometric Moments for Editable Structured Light Range Finding. In Hans-Jörg Schulz, Matthias Teschner, and Michael Wimmer, editors, *Vision, Modeling and Visualization*. The Eurographics Association, 2019. ISBN: 978-3-03868-098-7. DOI: 10.2312/vmv.20191315.

[WK15] Michael Weinmann and Reinhard Klein. Advances in geometry and reflectance acquisition. In *SIGGRAPH Asia 2015 Courses*, 1:1–1:71. ACM, 2015. Article No. 1.

[WML+07] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques*, pages 195–206. Eurographics Association, 2007.

[WOV+12] D. Wu, M. O'Toole, A. Velten, A. Agrawal, and R. Raskar. Decomposing global light transport using time of flight imaging. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 366–373, June 2012. DOI: 10.1109/CVPR.2012.6247697.

[WVJ+17] Sebastian Werner, Zdravko Velinov, Wenzel Jakob, and Matthias B Hullin. Scratch iridescence: wave-optical rendering of diffractive surface structure. *ACM Transactions on Graphics (TOG) (Proc. SIGGRAPH Asia)*, 36(6):207, 2017.

[WWL+19] Zhijie Wu, Xiang Wang, Di Lin, Dani Lischinski, Daniel Cohen-Or, and Hui Huang. Sagnet: structure-aware generative network for 3d-shape modeling. *ACM Transactions on Graphics (TOG)*, 38(4):91, 2019.

[WZT+08] Jiaping Wang, Shuang Zhao, Xin Tong, John Snyder, and Baining Guo. Modeling anisotropic surface reflectance with example-based microfacet synthesis. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 27(3):41:1–41:9, 2008.

[XRi19] XRiteVLB. XRite Virtual Light Booth. `https://www.xrite.com/categories/appearance/virtual-lightbooth`, 2019. Accessed: 2019-10-29.

[XZJ+19] Yang Xu, Huijie Zhao, Hongzhi Jiang, Yunfan Wang, and Xudong Li. 3d shape measurement in the presence of interreflections by light stripe triangulation with additional geometric constraints. In *Optical Measurement Systems for Industrial Inspection XI*, volume 11056, 110563N. International Society for Optics and Photonics, 2019.

[YHJ+14] Ling-Qi Yan, Miloš Hašan, Wenzel Jakob, Jason Lawrence, Steve Marschner, and Ravi Ramamoorthi. Rendering glints on high-resolution normal-mapped specular surfaces. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 33(4):116:1–116:9, July 2014. ISSN: 0730-0301. DOI: `10.1145/2601097.2601155`. URL: `http://doi.acm.org/10.1145/2601097.2601155`.

[YHM+16] Ling-Qi Yan, Miloš Hašan, Steve Marschner, and Ravi Ramamoorthi. Position-normal distributions for efficient rendering of specular microstructure. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 35(4):56:1–56:9, July 2016. ISSN: 0730-0301. DOI: `10.1145/2897824.2925915`. URL: `http://doi.acm.org/10.1145/2897824.2925915`.

[YHW+18] Ling-Qi Yan, Miloš Hašan, Bruce Walter, Steve Marschner, and Ravi Ramamoorthi. Rendering specular microgeometry with wave optics. *ACM Transactions on Graphics (TOG)*, 37(4):75, 2018.

[YIM07] G. Yahav, G. J. Iddan, and D. Mandelboum. 3D imaging camera for gaming application. In *2007 Digest of Technical Papers International Conference on Consumer Electronics*, pages 1–2, January 2007. DOI: `10.1109/ICCE.2007.341537`.

[You81] Andrew T Young. Rayleigh scattering. *Applied optics*, 20(4):533–535, 1981.

[Zha18] Song Zhang. High-speed 3D shape measurement with structured light methods: a review. *Optics and Lasers in Engineering*, 106:119–131, 2018.