



BACHELOR THESIS

Inverse Rendering of Wave Optical BRDFs

Bertan Karacora

Matriculation number: 3226384

Institute of Computer Science II - Computer Graphics
University of Bonn

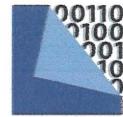
First reviewer: Prof. Dr. Matthias Hullin

Second reviewer: Prof. Dr. Reinhard Klein

May 13, 2022



Rheinische
Friedrich-Wilhelms-
Universität Bonn



Prüfungsausschuss
Informatik

universität bonn · Institut für Informatik · 53012 Bonn

Vorsitzender des Prüfungsaus-
schusses

Prof. Dr. Thomas Kesselheim

Prüfungsamt:
Judith König
Tel.: 0228/73-4418
Fax : 0228/73-4788
pa@informatik.uni-bonn.de
Postanschrift: 53115 Bonn
Endenicher Allee 19a

www.informatik.uni-bonn.de

**Erklärung über das selbständige Verfassen einer Abschlussar-
beit/ Declaration of Authorship**

Titel der Arbeit/Title:

Inverse Rendering of
Wave Optical BRDFs

Hiermit versichere ich, Karacora, Bertan,
Name / name, Vorname / first name

dass ich die Arbeit – bei einer Gruppenarbeit meinen entsprechend ge-
kennzeichneten Anteil der Arbeit – selbständig verfasst und keine anderen als
die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich ge-
macht habe.

R. Karacora Bonn, den 13.05.2022
Unterschrift/signature date

Acknowledgements

Authoring a bachelor thesis is almost always a memorable experience in a student's life. Especially at a time like this when online lectures and video calls dominate the daily routine at many universities. I have certainly gone through a formative stage of my studies paved with many different emotions and with both moments of euphoria and ones close to exasperation. I want to thank everybody who gave me their support throughout this period.

First and foremost, I wish to thank Prof. Dr. Matthias Hullin, head of the Digital Material Appearance Group in Bonn, for giving me the opportunity to carry out this project in an interesting research field while supporting me personally from the beginning to the end. I am thankful for his steady and resilient assistance in these important months of my curriculum as well as all the educational discussions which have become a rarity in these times and are as valuable as one. Not only did he supply me with useful material and exciting ideas, he also gave me confidence in times of self-doubt.

I would also like to thank Markus Plack who helped me to assess possibilities for a particularly difficult decision, even though we only talked once.

Definitely, I want to thank my family and friends who listened indulgently to my ramblings about this project and provided me with advice, whether or not my explanations were even understandable. Particularly, I want to mention my dear friend Arthi who is always there and looked out for me every time I forgot to do it myself.

Abstract

The accurate reconstruction of specular microgeometry is important for a broad range of applications that require modeling materials and surfaces with specific optical properties. Capturing the structure of object surfaces on a microscale requires the consideration of complex phase delay and diffraction effects that modern methods incorporate into wave optical BRDFs efficiently. While recently, there is an increasing interest in differentiable rendering, current advances might also be applied to analytic wave optical computations in order to handle the strongly intertwined influences of surface points on a discretized spatially-varying heightfield towards their reflectance. In this work, an inverse rendering system is implemented, tested, and evaluated, integrating wave optical BRDF generation and gradient-based optimization techniques. Although the original microgeometry usually cannot be retrieved because of local minima, given a starting estimation, the system is able to approximate a feasible candidate heightfield with improved similarity than the estimation if compared to specified reference BRDF slices.

Contents

1	Introduction	1
1.1	Application areas	2
1.2	Aims of this work	4
1.3	Contribution	5
2	Related work	8
3	Approach	10
3.1	Wave optical BRDFs	11
3.2	Differentiable rendering	14
3.2.1	Automatic differentiation	14
3.2.2	Radiative backpropagation	15
3.2.3	Finite differences	16
3.3	Gradient-based optimization	16
4	Implementation	18
4.1	Library interface	18
4.2	Enoki	20
4.3	Overall procedure	21
5	Evaluation	22
5.1	Experimental results	22
5.2	Ambiguity	24
5.3	Extrapolation of the concepts	27
5.4	Problem complexity	28
6	Conclusion	30
	Bibliography	32

1 Introduction

Accurately modeling the intrinsic characteristics of real-world objects and realistic lighting has been a fundamental motivation of computer graphics from the early days until now. Unlike traditional computer vision and image-processing tasks, *inverse rendering* does neither aim to obtain information at the highest level of human understanding nor to improve human perception. Instead, inverse rendering refers to retrieving the inherent physical properties of a scene, such as object geometry, reflectance, surface normals, material, and illumination, given only a set of photographs or photorealistic images.

As this is an inverse problem, naturally, all of these properties may be inserted as input parameters into a rendering algorithm. Apart from direct-solving methods, taking advantage of this prospect, the indirect analysis-by-synthesis approach has gained a lot of popularity by combining efficient forward rendering with modern optimization schemes. In their study, Iseringhausen and Hullin (2020) give an impressive example of the capabilities of inverse rendering, in particular, reconstructing shape that is out-of-sight, only through multi-bounce diffuse reflections. Frequently, at least one of geometry, lighting, or reflectance is known in advance, so the remaining scene attributes can be approximated with high accuracy, sometimes to the extent that the results are indistinguishable from the original. However, in many cases, inverse rendering still remains an ill-posed, i.e., ambiguous, and under-constrained problem that suffers from great complexity and technical demand (Loscos, Jacobs, et al., 2006).

A specific field of inverse rendering is *inverse surface design* (Patow and Pueyo, 2005). This exact problem deals with the estimation of surface geometry while lighting conditions and reflectance models are known. In the context of the fundamental rendering equation (Kajiya, 1986), such a reflectance model is given by the BRDF as in definition 1.1. Among further variants of the function, a *Spatially Varying BRDF (SVBRDF)* takes an additional two-dimensional parameter x specifying a point on the surface. Inverse surface design aims to utilize the BRDF and the lighting information for the reconstruction of surface structure.

Definition 1.1 (BRDF; Nicodemus et al., 1977). *Let ω_i and ω_o be the directions of incoming and outgoing light, respectively. All directions are parametrized by an azimuth angle ϕ and a zenith angle θ . Let L_o be the radiance reflected in direction ω_o and L_i the irradiance incident on the surface from direction ω_i . The Bidirectional Reflectance Distribution Function (BRDF) f_r quantifies the proportion of L_o and L_i as follows:*

$$f_r(\omega_i, \omega_o) = \frac{dL_o(\omega_o)}{L_i(\omega_i) \cos \theta_i d\omega_i}$$

1.1 Application areas

There is a broad range of applications for capturing the physical properties of scene geometry. On the one hand, it enables image-generating systems to reproduce photorealistic images and modify the appearance as desired. As a straightforward example, it is possible to recreate the same scene with different lighting conditions (relighting; K. Zhang et al., 2021; Loscos, Frasson, et al., 1999) or from another point of view (novel view synthesis; B. Liu et al., 2021). This also implicates the virtual insertion of new objects into a scene and the removal of existing ones (Karsch, 2020). Figure 1.1 outlines the concept of inverse rendering in the case of relighting. The ability to manipulate a scene while maintaining physical realism is of great value for both research and industrial purposes, especially with regard to interactive design tools like in architectural planning and luminaire design (Tourre et al., 2008). Generally, further applications of inverse rendering include image decomposition, shadow removal, post-production, augmented reality, and even photo forensics (Johnson and Farid, 2007).

On the other hand, especially for inverse surface design, there is a significant body of research with the purpose of improving construction and manufacturing techniques. Herein, fast image-based systems are created in order to fulfill the needs of 3D object assembly, material design, and the precise surface construction of optical components such as reflectors and lenses (Bielawny, 2019). For example, Kassubeck et al. (2021) insert an inverse renderer into a practical manufacturing procedure providing adaptive feedback during the production of 3D-printed glass with specified optical properties.

Regarding the versatility of application areas, a multitude of different methods has been established to tackle this important task. In spite of that, all of these methods are limited in their precision due to the theoretical basis of their computation. Bearing in mind the definition 1.1, most algorithms work on the model of ray-based light transportation and a geometric optics BRDF. Although there

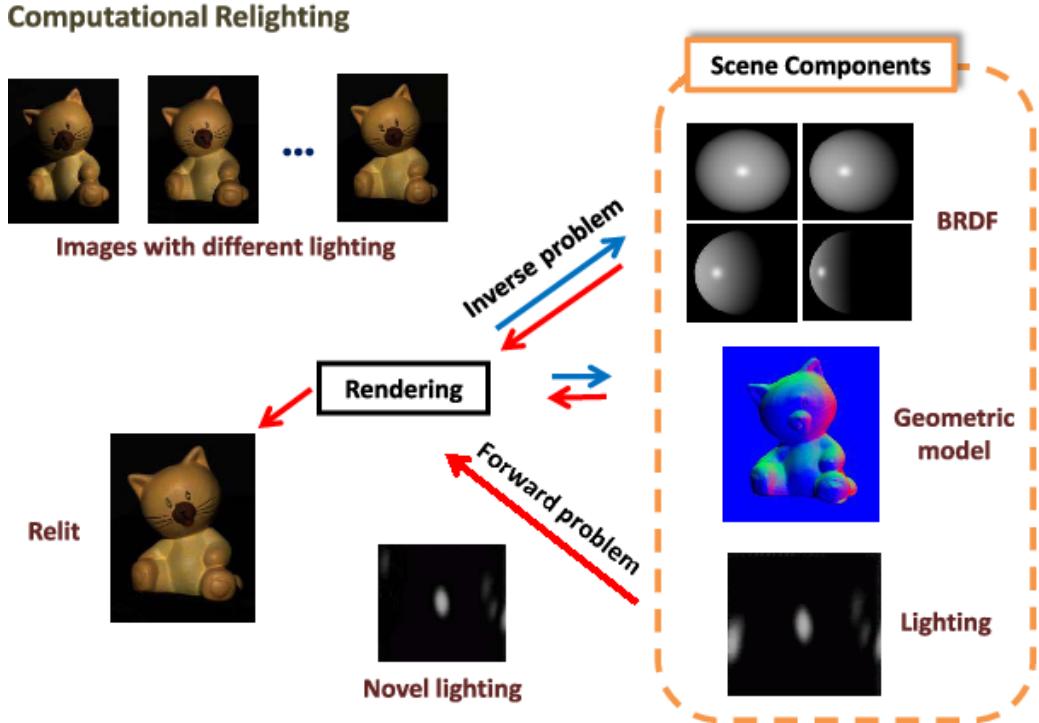


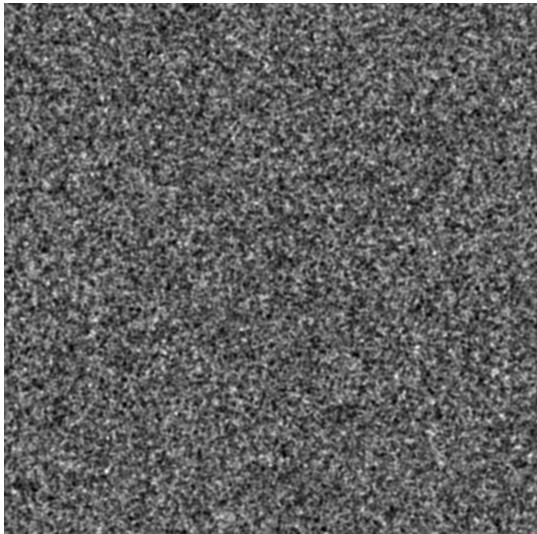
Figure 1.1: Application of inverse rendering for computational relighting. Any of the scene components BRDF, geometry, and lighting (right) can be inferred from the reference images (top left) using inverse rendering. For relighting, the lighting element is substituted in the forward rendering process. Figure by S. Liu and Do (2017).

has been a tremendous amount of progress on photorealistic rendering using geometric optics in the past, only little attention is paid to *wave optics*. Wave optical BRDFs induce calculations that are crucially more complex, but also incorporate more realistic surface interactions taking diffraction effects into account.

The main difference when using inverse rendering with wave optics is the capability to reconstruct microgeometry with close-to-wavelength accuracy manifesting in a more detailed surface information. This level of detail is becoming more and more relevant with the steady advances in fabrication techniques (Levin et al., 2013). Currently practiced methods of measuring microgeometry involve interferometry (Z. Dong et al., 2016), microscale photometry (Z. Li and Y. Li, 2011; Johnson, Cole, et al., 2011), and micro CT imaging (Zhao, Jakob, Marschner, et al., 2014). As one of several more capturing methods, Nam et al. (2016) propose the setup of a microscopic light dome for simultaneous acquisition of reflectance and structure. All of these require the availability of a physical object and are bound to preparatory effort and costly instruments. In consideration of this, the advantages of an image-based measuring system using inverse rendering seem apparent.

1.2 Aims of this work

In this work, I address the challenging problem of capturing specular geometry on a microscopic scale given the information of wave optical BRDFs. Therefore, an inverse renderer is implemented that approximates the microgeometry of a surface using analysis-by-synthesis. The microstructure is represented explicitly by a discretized heightfield encoded as a 2D texture with a resolution of one square micrometer per texel. This is necessary to model a general surface without prohibiting assumptions of its structure, while also granting sufficient detail for the forward renderer to generate the wave optical BRDF paying respect to diffraction effects. More precisely, the problem concerns a spatially-varying heightfield and a SVBRDF, respectively. The SVBRDF is sampled for an arbitrary amount of fixed surface points into a set of BRDFs under the assumption of known lighting conditions like incoming direction, etc., for every sample. Each BRDF itself is sampled for discretized outgoing directions covering the reflection hemisphere. Arguably, there might be no suitable solution at all, or multiple possible heightfields that induce the exact reference SVBRDF depending on the nature and the number of given samples.



(a) Isotropic heightfield.



(b) Wave optical BRDF.

Figure 1.2: An isotropic heightfield (a) and its corresponding BRDF (b) sampled at the center of the field with incoming light perpendicular to the surface. The height values are mapped to a normalized intensity distribution. Whereas the heightfield is given in spatial coordinates, the BRDF lobe for a fixed point x is displayed as a 2D slice of the projected hemisphere outgoing from point x on the surface. The heightfield data is provided by Yan, Hašan, Walter, et al. (2018).

To my knowledge, there has been no previous attempt on marrying inverse surface design with explicit fine-scale wave optics in related work. The merits of employing wave optics instead

of conventional geometric optics appear to be limited at a glance but benefit many application areas, as mentioned before. Above all, this work is meant to explore the essential ideas and the difficulties of combining these fields. In addition to the obstacles that inverse rendering presents by itself, the complexity is significantly increased in combination with wave optics.

This work is based on the studies of Yan, Hašan, Walter, et al. (2018) in their publication *Rendering Specular Microgeometry with Wave Optics*. Therein, the authors introduce a wave optical rendering algorithm that predicts the reflectance of a spatially-varying specular surface. In brief, they approximate the phase delay effect of the microstructure using Gabor kernels forming a grid. Then, the superposition of these influences is integrated over the coherence area, a limited area where the reflected light's phase is determined. Finally, the computed BRDF is rendered into an image. Figure 1.2 portrays a heightfield texture and a sample of the associated wave optical SVBRDF for a single point x and direction ω_i . Although the reflectance estimates are not compared against real-world measurements, e.g., obtained with a spherical gantry, it is shown that the method is able to produce realistic wave optical BRDFs on an arbitrary high-resolution heightfield. The forward rendering process uses the heightfield 1.2a as input and generates the BRDF lobe 1.2b as output. As one would expect, a full inverse rendering system reverts this relationship.

In order to achieve these aims, the chosen approach of this work is founded on analysis-by-synthesis. For that reason, the implementation consists of a generating and an optimizing component. The generation element is responsible for producing the BRDF samples from a given heightfield while collecting information and analyzing the calculations performed. Meanwhile, the optimization component modulates the parameters of the former with regard to an objective function after it gets forwarded the data extracted by analysis. Specifically, in this work, *differentiable rendering* is used to compute the gradient of the rendering function with regard to the input parameters inserted. Subsequently, it is employed for regression by gradient-descent. Since it is assumed that the lighting information is known, the only parameter determining the synthesized results is the heightfield itself. By iterating and alternating the generation and optimization passes, it is intended that any suitable starting heightfield converges into an approximation of a feasible solution to the inverse problem.

1.3 Contribution

Supplementary to their results, Yan, Hašan, Walter, et al. (2018) provide the source code necessary to sample a wave optical SVBRDF and render the BRDF slices into images. From

this initial situation, I rebuilt the program into a C++ library for differentiable rendering. The final implementation serves as an interface accessible in Python for gradient-based optimization, and, is tested with gradient-descent using Adaptive Moment Estimation (Adam; Kingma and Ba, 2017). In several experimental setups, the mean squared error (MSE) of the BRDF estimation can be reduced up to 99% if compared to the initial guess, but the original heightfield can not be retrieved. These observations will be elaborated on in the evaluation chapter of this thesis. Equally important, this work demonstrates the hurdles and the possibilities when reconstructing microgeometry with regard to complex wave optical computations and realistic handling of diffraction effects.

Creating a full and efficient inverse renderer on top of the source code of Yan, Hašan, Walter, et al. (2018) is a great challenge, hence, this work comes with several limitations. Because of restrictions in time and resources, in this thesis, I only investigate computer-generated heightfields as ground truth instead of taking real-world measurements. As a result of the generality when representing microgeometry as an explicit heightfield, this does not obstruct the methodology nor the effectiveness to any degree. On the contrary, it guarantees the existence of at least one exact solution to the simulated inverse problem and opens up further means of evaluation by comparison.

Moreover, due to heavy RAM usage, the final implementation accepts the heightfield input only on a minimal scale and samples each BRDF for a reasonably limited number of outgoing directions. In addition, any design specifications or structural constraints have not been taken into account. Despite all restrictions, this work presents a proof-of-concept for BRDF fabrication, as the reflectance of an arbitrary microstructure in a broken-down scenario can be imitated successfully.

Concretely, my contributions are:

- Building a functional C++ library from the provided source code by means of modularization and encapsulation.
- Extension of the library's interface for a Gabor kernel representation of the heightfield in addition to explicit texture. Complementing this with the inverse transformation (from Gabor representation back to height values).
- Extraction of gradients using automatic differentiation during the rendering process.
- Exposing the library to the optimization component written in Python. Implementation of gradient-descent algorithm based on Adam.
- Experimental analysis of the inverse rendering system.

CHAPTER 1. INTRODUCTION

The remainder of this thesis is organized as follows: First, there is a brief overview of related work. A selection of methods and the eventually chosen approach of this work are explored in chapter 3. Then, chapter 4 describes the implementation of this approach in more detail. Lastly, the results are presented and evaluated with the assistance of visualization means in chapter 5, followed by the conclusion in the end.

2 Related work

Additionally to the directly involved methods, this thesis is related to the following research fields.

Rendering specular microgeometry. The ever-rising standards for photorealism have led to a growing interest in the efficient prediction of microscale reflectance. On a larger scale or up to a certain degree of spatial variation, the preferred approach relies on geometric optics. Commonly for rendering specular microstructure, a microfacet BRDF is employed using normal distribution functions. Recent noteworthy contributions have been made by Jakob, Hašan, et al. (2014) using a statistical distribution of specular structural elements or conversely, Yan, Hašan, Jakob, et al. (2014) and thereupon Yan, Hašan, Marschner, et al. (2016) as well as Zhu et al. (2019) based on explicit normal maps or heightfields. Alternatively, a filtering approach as introduced by Gamboa et al. (2018) could dissolve any issues implied by the spatial variance at the expense of excluding the phase delay influence onto specular highlights. Although the normal map representation already suggests the investigation of the inverse problem, to achieve details and high-frequency variation, in this work, the accurate reflectance estimation of general heightfields needs to be handled by wave optics.

Diffraction effects. As early as the study of Stam (1999) microgeometry was modeled as heightfields whose discretized values affect the phase of light. This is how diffraction effects can be incorporated into a BRDF under costly computations. Many novel approaches since then focus on a particular type of microstructure, such as scratches and other constrained surface models. Z. Dong et al. (2016) were among the first to acquire a real heightfield using a special device based on interferometry called a profilometer for the exploration of its reflectance. However, they continue to evaluate its data for large-scale averaged BRDFs. In general, Pharr et al. (2016) can be consulted for state-of-the-art physically-based rendering techniques and Lange et al. (2021) show an implementation of these for wave optics. Besides Yan, Hašan, Walter, et al. (2018), the only method for generating a wave optical spatially-varying BRDF without discarding high-frequency information was proposed by Werner et al. (2017) and extended for real-time

CHAPTER 2. RELATED WORK

rendering (Velinov et al., 2018). In doing so, they start from a completely smooth surface and focus on one-dimensional scratches. The contribution of these scratches towards every pixel is calculated analytically and finally composed into the typical iridescence effects visible as glints. This model mainly differs from Yan, Hašan, Walter, et al. (2018) in its completeness as the irregularities of a plain rough surface structure with the desired spatial resolution are rarely limited to perfectly straight scratches.

Inverse surface design. A broad overview of micro-scale optics theory, their applications, common design methods, and practical fabrication techniques, such as plastic replication, all sorts of lithography, and even holographic exposure, is given in the book of Kress (2014). The developments of these techniques and the innovations of the manufacturing industry have aroused attention to inverse surface design. Patow and Pueyo (2005) survey the basic notions and summarize earlier research in this field. To my knowledge, there is no previous work on explicit inverse surface design based on the high-frequency directional patterns of wave optical BRDFs. However, similar aims can be found in the work of Levin et al. (2013). Both Papas et al. (2011) and Weyrich et al. (2009) are examples of surface fabrication methods from a desired reflectance model using geometric optics. In contrast, Levin et al. (2013) follow related concepts and examine the problem with regard to wave optics. Thus, their estimations reach a resolution of up to 100 micrometers squared, multiple orders of magnitude more than the results reached before. All these systems focus on fabrication rather than information extraction and can not be used to approximate a general heightfield with micrometer-level detail.

The latter approach is of particular interest since it involves the production of candidates for the solution to the inverse problem with a slightly inferior precision but without any iterative optimization. This is achieved by modeling the surface as piece-wise flat rectangles with zero-mean height and modulation of the piece's dimensions. The aim is to fabricate a wafer using photolithography, so only a small set of different height values for the whole surface is practical, as each requires a complete chemical etching pass. The desired reflectance properties are produced by sampling the step size to which the tiles are height-modulated randomly according to specific distributions. For illustration, a two-dimensional surface with anisotropic reflectance can be imitated by tiling wide on one axis but varying height in frequent steps on the other, resulting in a frequency spectrum that is rather narrow and broad, respectively, in Fourier space. It is shown that the expected reflectance of a generated surface area is directly proportional to the expected spectrum of step sizes. The optimal distributions can be acquired by minimizing the least-squares residuum of the expected and the target reflectance. However, this method offers a very restricted structural model and limited resolution; each block in the fabrication process requires covering a sufficient number of coherence areas to allow a sensible estimation of the expected reflectance.

3 Approach

This chapter provides a detailed description of the used methodology. Initially, a reference SVBRDF is given as a set of BRDF slices that represent the desired reflectance of a specific heightfield. The reflectance can be realistically simulated by a physically-based renderer. An empiric measurement of microgeometry or real-world reflectance would be possible and provide more accurate ground truth than artificial images but is excluded from the scope of this thesis. Instead, a first forward rendering pass is performed on one of the heightfields provided by Yan, Hašan, Walter, et al. (2018). Now the inverse rendering system begins by initializing a starting hypothesis as an arbitrary heightfield. Exploiting the concept of analysis-by-synthesis iteratively, the final procedure approximates a microstructure with appropriate reflectance properties step-by-step.

Therefore, the BRDF of the hypothesis is synthesized and analyzed with regard to a loss function. The key idea of this approach is that the loss can be minimized if an optimization framework is instantiated and supplied with information about how the resulting error of the current hypothesis is correlated with each of the input parameters, i.e., the values of the heightfield texture. Many modern optimization techniques, as they can be found in any gradient-descent algorithm or neural network, resort to derivatives for this problem. Hence, all the rendering calculations need to be performed in a mathematically differentiable manner, so the presumed heightfield parameters can be modified according to the gradient to minimize the error specified by the loss function. Figure 3.1 depicts the complete sequence of the operational loop. As mentioned in section 1.3, the BRDF sampling density is set to a rather low number of outgoing directions leading to the moderate resolution of the projection.

Essentially, the approach can be structured into three stages: First, the prediction of wave optical BRDFs from microgeometry in the forward rendering process. Second, the technique of differentiable rendering and how it is utilized for inverse rendering. Finally, the gradient-based optimization of the input parameters.

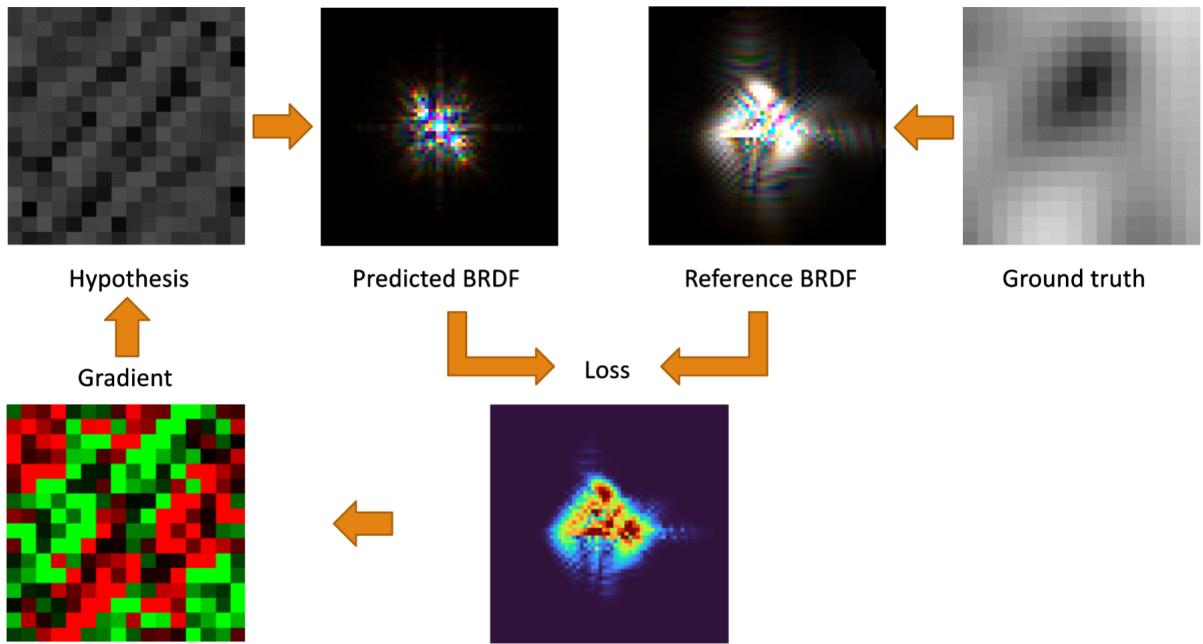


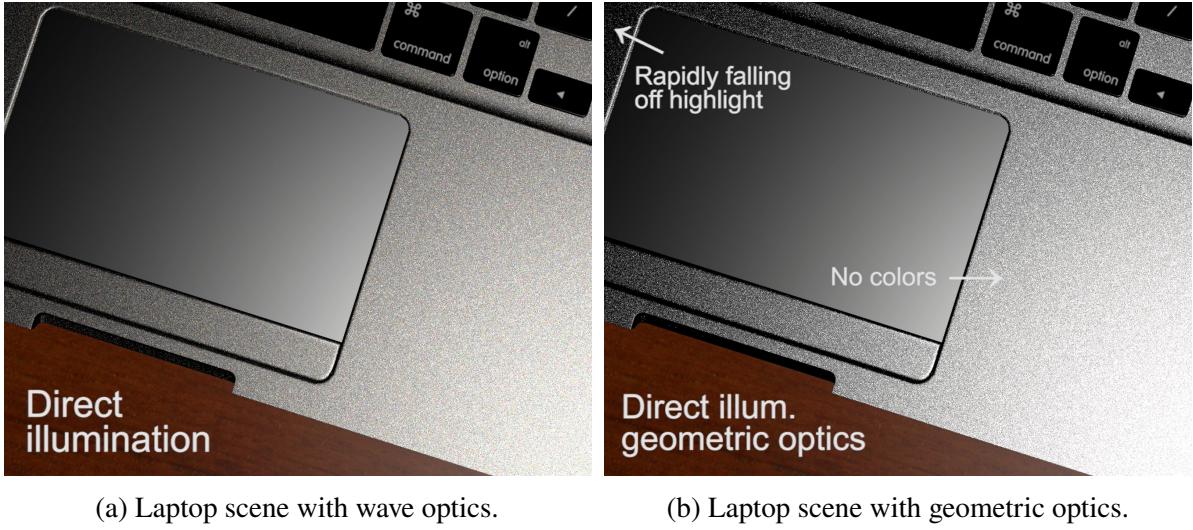
Figure 3.1: Visualization of the analysis-by-synthesis approach for approximating the heightfield. The inverse rendering loop consists of an iterative process alternating differentiable forward rendering and parameter optimization. From the hypothesis, i.e., the heightfield to be optimized, a BRDF slice is generated. This prediction is compared against the reference BRDF that results from the ground truth. With regard to the hypothesis, the gradient of a loss function is calculated and merged with it.

3.1 Wave optical BRDFs

In the intuitive understanding of everyday situations and in most applications of computer graphics, light is thought of as a geometric ray that interacts with objects through reflection and refraction. In general, this model is sufficient in almost all circumstances and even lets computers generate photorealistic images in real-time. However, at a closer inspection, one can observe a natural change throughout the whole spectrum of colors on specular surfaces that react dynamically to illumination changes. Especially the imperfections of specular surfaces such as scratched or brushed metal behave in a distinctive way (see figure 3.2). These kinds of optical effects cannot be physically explained with geometric optics but with wave optics, which results in more precise and realistic predictions in general.

The phenomenon of iridescence is caused by the complex interaction of diffraction effects that can only be simulated using wave optical reflectance models. When light waves meet an obstacle, e.g., tiny bumps of a rough surface that are not significantly larger than the light's wavelength, then the wavefield bends around its corners and propagates from there in a certain wave pattern.

CHAPTER 3. APPROACH



(a) Laptop scene with wave optics.

(b) Laptop scene with geometric optics.

Figure 3.2: Comparison of renderings of an exemplary scene with direct illumination using wave optics (a) and geometric optics (b). At a close look, the wave optical rendering features colors and longer tails of the highlights. Images from Yan, Hašan, Walter, et al. (2018).

These patterns interfere with each other and influence the light deflection in an unanticipated way. From a more global perspective, this means a complete area of the surface becomes the source of local reflectance. The area in which the light's phase is affected is called the *coherence area* and contains all surface points where the reflected light's phase might be coherent, i.e., possibly interfering. The local impact on the BRDF can not simply be added up linearly as optical intensity but only as the superposition of phase-interfering waves while global light transport in the rendering equation may remain ray-based. The further physical details of wave optical reflectance modeling are not relevant here, but some can be found, inter alia, in the work of Levin et al. (2013).

For computational purposes, the most crucial finding is that the reflectance can be obtained as the intensity of the Fourier transform of a signal whose phase is proportional to the height value. Usually, the modulation of a spatially resolved surface needs to be expressed analytically as locally-limited transforms, since numeric calculations integrating the whole coherence area are too expensive. This also is the foundation of the method of Yan, Hašan, Walter, et al. (2018), which is utilized in the forward rendering process of this work. Equivalently to this, for incoming light with wavelength λ , the authors define the signal R of point x on the surface plane with height $H(x)$ as follows:

$$R(x) = e^{-i\frac{2\pi}{\lambda}2H(x)} \quad (3.1)$$

The function describes the reflection of light as a complex value determining its spatially-varying phase. This version of the equation was chosen for reasons of convenience; it arises from the

CHAPTER 3. APPROACH

non-reciprocal Harvey-Shack diffraction model (Harvey, 1979), which is not physically accurate but may be created analogously for similar models (Levin et al., 2013).

Let R^* be the reflection function weighted by a coherence kernel centered at the coherence area and A_c the corresponding normalization factor. Moreover, let n be the surface normal, F an adequate term for reflectance, and $\bar{\Psi}$ the non-normalized half vector of ω_i and ω_o projected onto the surface plane. From there, the BRDF is derived from the squared magnitude of the Fourier-transformed reflection function:

$$f_r(\omega_i, \omega_o) = \frac{|\omega_o \cdot n|F}{\lambda^2 |\omega_i \cdot n|A_c} \left| \mathcal{F}[R^*] \left(\frac{\bar{\Psi}}{\lambda} \right) \right|^2 \quad (3.2)$$

To summarize the remaining of the methodology, Yan, Hašan, Walter, et al. (2018) approximate R^* using a weighted combination of Gabor kernels placed on every point in the coherence area.

Definition 3.1 (Gabor kernel (Yan, Hašan, Walter, et al., 2018)). *A Gabor kernel is a 2D wavelet consisting of a 2D Gaussian kernel $G_{2D}(x, \mu, \sigma) = \frac{1}{2\pi\sigma^2} \exp -\frac{\|x-\mu\|^2}{2\sigma^2}$ modulated by a complex exponential:*

$$g(x; \mu, \sigma, a) = G_{2D}(x, \mu, \sigma) e^{-i2\pi(a \cdot x)} \quad (3.3)$$

In the present context, it can be thought of as plane wave locally-limited by a Gaussian window function. One important property is that the Fourier transform of Gabor kernels can be acquired analytically and is a Gabor kernel itself. Since the approach of this thesis requires computing the derivatives of all performed calculations through the entire rendering pipeline, it is an useful prerequisite that the model is entirely analytical and continuous.

In conclusion, the algorithm integrates the phase delay of reflected light waves over a large pixel footprint for every directional sample of the BRDF. Since the BRDF is directly rendered into an image, in this case, the coherence area equals the pixel footprint. As a result, this method is associated with unavoidable costs depending on the spatial correlations of the reflections that are constricted by the size of the coherence area. Complementary, a single texel of the input heightfield contributes to the BRDFs of all other points of the area. In the practical system, the first step is preprocessing the heightfield into a representation as Gabor-related data structures that can be evaluated more immediately. This is why the library interface is extended to support the interim representation in a suitable data structure as it might be more compact and advantageous to save resources. This concept requires the inverse transform of the Gabor representation back into the height values in the end and is described in more detail in a later part of this thesis.

3.2 Differentiable rendering

Whereas the generation of wave optical BRDFs is the basis of my methodology, making the rendering process differentiable is its main effort. The naive attempt would be re-arranging equation 3.2 and the Gabor approximation into a long function term that takes all the height values as input. Clearly, the symbolic derivation of all the calculations by hand is not practical. For a heightfield texture with the dimensions 1000×1000 and a BRDF slice including the same amount of samples, even the memory footprint required for the gradients at x would induce 32 TB of floating-point data in a single iteration. For this reason, an objective function is established, compressing the overall loss of the current hypothesis towards the reference BRDF into a scalar. Thus, the gradient of each iteration is in the same dimensions as the input. Differentiable rendering aims to efficiently evaluate this gradient enabling analysis-by-synthesis.

Recently, physics-based differentiable rendering has gained significant interest as an effective instrument for inverse graphics tasks. Several different methods have been proposed, most of which are specialized for unique applications (Kato et al., 2020). Similar to the present aims, SVBRDF estimation encounters comparable requirements of fine-scale resolution and spatial variance for which differentiable rendering is widely practiced. Gao et al. (2019) reconstruct SVBRDF data from photographs using a convolutional neural network. To give one more example, Y. Dong et al. (2014) retrieve the normal distribution and SVBRDF from videos of rotating objects by tracing surface points and solving in the temporal gradient domain. The most promising technique considered for the present approach is automatic differentiation, given its versatility and the great interdependence of the BRDF samples with respect to the input parameters.

3.2.1 Automatic differentiation

The method of automatic differentiation (autodiff) addresses the problem in a practical manner (Margossian, 2018). Regardless of how much complexity wave optical predictions hold over ray-tracing and other common rendering approaches, the executed computations consist of a long sequence of elementary mathematical instructions such as addition, multiplication, trigonometric and exponential functions, etc., all of whose analytical derivatives are known. Unraveling and factorizing this sequence recursively by virtue of the chain rule, even the presented construct of indistinctively arranged reflection interactions can be differentiated automatically. Therefore, most autodiff systems keep track of the inputs to be derived in a structure similar to a computational

graph where variables are vertices and edges represent the operators (Zhao, Jakob, and T.-M. Li, 2020).

There are two opposing forms of autodiff, forward and reverse mode (Zhao, Jakob, and T.-M. Li, 2020). The former is more efficient for the differentiation of functions with a lower number of inputs and a larger amount of outputs, the latter vice-versa. During the forward mode computations, the variables in the intermediate node layers are derived from traversing the graph once per derivation of an input. Alternatively, parallel to the inputs the differential of every operation can be executed to its derivatives on-the-fly without the need of a transcript. Consequently, differentiating the rendering of a coherence area of m^2 points into a BRDF slice with resolution n^2 induces a time complexity of $O(m^2 \cdot m^2 n^2)$. On the contrary, the reverse mode first computes the output with insignificantly more time investment and then backpropagates the gradient to the inputs through the documented graph. Its space complexity is linear to the number of arithmetic operations. As we measure the error of the current estimation simply as a scalar, this alternative matches the present requirements most.

3.2.2 Radiative backpropagation

Motivated by the drawbacks of autodiff for rendering purposes, Nimier-David, Speierer, et al. (2020) introduce *radiative backpropagation*, a novel method that achieves the backpropagation of gradients efficiently without recording a graph of computations. They instigate that the idea of ray-tracing works on gradients backward as well as originally on radiance. The outputs, i.e., the discretized projection of the scene as pixels, are casted into the source of the derivative radiance energy with respect to the objective function. This energy is traced back through the scene, reflected, and refracted on surfaces like rays until received by and inflected into the scene parameters of relevance. As a result, radiative backpropagation separates the retrieval of gradients from the forward rendering completely. Vicini et al. (2021) proceed to reduce the computation time from quadratic to linear in the number of scattering occurrences by recording additional information to every evaluated path, named *path replay backpropagation*.

Both these methods could potentially yield two-fold advantages for this work; they offer improved efficiency and the original BRDF generation program can stay untouched. In fact, if one chooses a reciprocal diffraction BRDF model for directly rendering a BRDF lobe into an image, propagating backward and forward is equivalent. However, the parameters to be optimized, the receivers of the backpropagated loss, are not directly considered in the light transportation problem as interpreted by *radiative backpropagation*. It is unclear, how the height of each surface point manipulates

the emitted radiance since their contributions interfere inseparably as plane waves with different phase delays and not linearly as optical intensities. For this reason, radiative backpropagation or adapted ideas of separating the gradient retrieval from the rendering computations could not be included to overcome the issues of autodiff in the approach of this work.

3.2.3 Finite differences

As a replacement technique to the aforementioned, finite differences also achieve separability such that the implementation of Yan, Hašan, Walter, et al. (2018) would not have to be modified intensively. Here, the entire rendering system is handled as a black-box function. Complementing the naive symbolic, the automatic analytical, and the radiative method, finite differences can be used to approximate gradients numerically. Therefore, the function to be differentiated that is the rendering algorithm, is applied to slightly perturbed inputs trying to exploit the property $f'(x) \approx \frac{f(x+h)-f(x)}{h}$ for small h . There are many problems with this in the present scenario; the main restriction is that the rendering would need to be performed once for every modulation of each input parameter, even if one could determine a good value for h . This is also why numerical differentiation methods are traditionally not used for machine learning tasks.

3.3 Gradient-based optimization

Given the gradient of the wave optical BRDF generation function at the current position in parameter space, the hypothesis is optimized by descending along the opposite direction of the gradient. As a consequence, the parameters converge to one of the closest local minima of the loss function. One standard optimizer is Adam (Kingma and Ba, 2017), a parameter updating algorithm for stochastic gradient descent that is utilized as an all-around reasonable backbone in many different contexts. Adam includes the concepts of adaptive gradients and momentum acceleration. In this work, Adam is combined with a non-stochastic objective function, and its configuration is adjusted from experimental observations: $\alpha = 0.01$ was chosen as the learning rate, $\beta_1 = 0.9$ and $\beta_2 = 0.99$ for the exponential decay rates of the two momentum estimates, respectively. A slight exponential learning rate decay was added.

While treating a BRDF slice as a mini-batch, stochastic gradient descent for random surface points or randomized incoming light directions was not successfully put to work, as further discussed in chapter 5. In a simple fashion, both averaging the gradients of multiple random

CHAPTER 3. APPROACH

SVBRDF samples before enforcing them onto the inputs and a sequential setup similar to feeding on infinitely fabricated unseen samples result in noisy gradients and, ultimately, the divergence of the estimation. Keeping in mind the aims and restrictions of this work, the optimization was left rudimentary, offering many opportunities for improvements which are not depleted in this work.

The objective function to be minimized can be any differentiable function that measures the discrepancy between the generated outputs and the reference. To this end, this work makes use of a customized error measure between HDR images:

Definition 3.2 (Customized mean squared logarithmic error). *Let I and R be two HDR images with channels C and pixels \mathcal{P} . The customized mean squared logarithmic error (CMSLE) in this work is defined as follows:*

$$CMSLE(I, R) = \frac{1}{|\mathcal{P}|} \sum_{p \in \mathcal{P}} \sum_{c \in C} (\log(I_c(p) + \epsilon) - \log(R_c(p) + \epsilon))^2$$

with a constant ϵ symbolizing a weight shifting factor.

The definition of ordinary MSE between two multichannel images without logarithmic weighting can be deferred from this. Notably, the difference is evaluated in logarithmic space to counteract the HDR nature of the images. High intensities outweigh the subtle diffraction patterns at the external areas of the highlights when using ordinary mean squared error (MSE). The loss is estimated pixel-wise, with the difference equal to the Euclidean distance of channel-dimensioned pixel vectors. Considering the BRDF samples as variable units, instead of the channel values, the error is not normalized against the number of channels $|C|$. In practice, this has no significant effect.

One example of a more sophisticated loss function for HDR images is given in the work of Moriwaki et al. (2018), which also includes logarithmic error measures. From a visual point of view, an alternative of a pixel-wise cost function that describes the difference between BRDFs considering human perception can be found in the work of Bieron and Peers (2020).

4 Implementation

Following the chosen approach, I developed an effective inverse rendering system to demonstrate the capabilities of the methodology. Working on the basis of the source code provided by Yan, Hašan, Walter, et al. (2018) the BRDF generation program was transformed into a differentiable rendering library offering the exact interface required for the optimization algorithm to interact with. Notably, wave optical BRDFs can be inserted as usual into the rendering equation. In doing so, the program can be combined with rendering techniques such as path-tracing to form a full scene renderer. There is a selection of general-purpose differentiable rendering systems available, namely *Mitsuba 2* (Nimier-David, Vicini, et al., 2019) and *PSDR-CUDA* (C. Zhang et al., 2020) amongst others. Both these are based on *Enoki* (Jakob, 2019), the autodiff library that differentiates the forward operations in this implementation as well. The integration of the wave optical BRDF sampling into *Mitsuba 2* is an option, but requires a capable GPU and complicates the task inordinately, so this thesis settles for the inversion of the BRDF generation itself.

4.1 Library interface

In order to comply with the requirements of the optimization component, functionalities of the original source code need to be organized, modularized, extended and linked. On this account, the implementation makes use of a few helpful libraries for managing the computation-heavy renderer, written in C++, from Python, which is predestined for optimization tasks. One of which is *pybind11* (Jakob, Rhinelander, et al., 2017), granting this interoperability in a lightweight fashion. This is achieved by the encapsulation and the definition of each object's properties and functions to be Python-bound. Another constituent part is *Eigen* (Guennebaud, Jacob, et al., 2010), a template library for linear algebra with a native support of *pybind11*, which simplifies the transfer of *NumPy* arrays (Harris et al., 2020) between the two languages. For the optimizer, which is the main control flow unit at the same time, to access data and methods of the renderer,

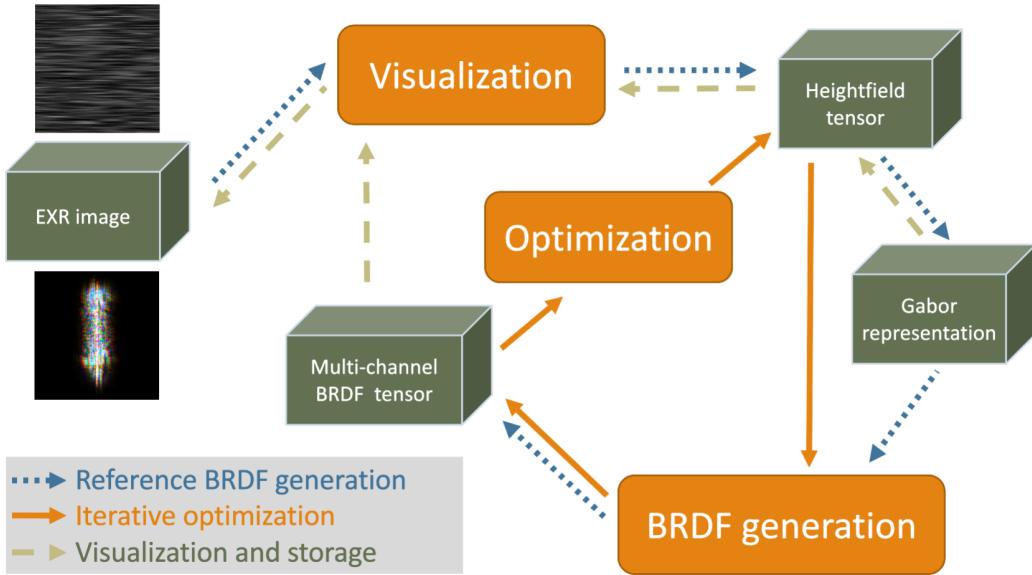


Figure 4.1: Diagram of modules and data structures. The inverse rendering loop consists of the BRDF generation and the optimization of input parameters. It is supplemented by a visualization module for storage and the generation of reference samples. The informal diagram type is not UML-conform, but was chosen for a more abstract view onto the mechanics.

the program is reshaped into a collection of classes with directly readable and writable attributes besides constructors and transforms untangled from the initial long sequence of instructions.

Previously, the program accepts a heightfield texture along with some rendering specifications as arguments. After the restructuring step, it is compiled into a dynamic library that offers a more versatile interface for BRDF generation as well as some necessary utilities for reading and writing data to images. In figure 4.1, it is shown how the parts of the interface and the most crucial data structures interact during different stages of the execution. For data storage, support of the *OpenEXR* standard (Kainz et al., 2004) for HDR images is integrated.

More importantly, many computations that can be performed in advance are decoupled from the main generation process. A specially parametrized representation is introduced containing the preprocessed information of the Gabor kernel transform in prior to the determination of wavelength and lighting directions. However, this representation is incompatible with the optimization. The rendering function is not plausibly differentiable with respect to this parametrization since several values follow from the same surface point's height and autodiff cannot consider their correlation as dependent variables. In other words, manipulating a value induces the adaptation of a few others, so the gradients are inapplicable without indefinite oscillatory interactions trying to retain a valid state. Otherwise, the inverse transform of this representation back into heightfield values would be required in the end, which is why this transform was added to the functionalities of the

library if a proper solution for handling the interdependence is found.

4.2 Enoki

Much the same as Mitsuba 2, Enoki (Jakob, 2019) is the autodiff library of choice in this implementation. Enoki features practical array types as well as basic vectorized operations and translates code into SIMD (single instruction, multiple data) instructions. In the context of differentiable rendering, parameters, that should be factored into the computational graph as nodes, need to be available in certain data types provided by Enoki. Similarly, arithmetic expressions need to be replaced by designated autodiff instructions. Whereas Enoki supports the dynamic vectorization of these explicitly marked parameters on modern GPUs, for example, based on the *CUDA* (NVIDIA et al., 2020) framework, in this implementation every distinct heightfield value is marked separately, devoid of any GPU requirements. Scrutinizing the wave optical BRDF model, major parts of the code basis are rewritten to make autodiff possible.

Moreover, Enoki expects the outputs of the differentiable function to be linked in the graph to every input marked via data type. If there is no trail in the graph from each input to every output, the backpropagation fails. This has two important consequences: The pruning mechanisms accelerating the complex reflection approximation using Gabor kernels need to be disabled if autodiff is requested. Secondly, whether a surface point contributes to the outputs at all, need to explicitly specified before the computations to be differentiated begin. Practically, this is determined in a straightforward fashion by predicting the coherence area's coverage. Usually, the idea of gradient backpropagation in graph-style constructs can be found in deep learning frameworks. The modalities of neural networks like convolutions include a manageable amount of operations. Rendering algorithms differ from this, as they are inherently unstructured and face the additional challenge of discontinuities (Zhao, Jakob, and T.-M. Li, 2020). Without any further interventions, approximating the Fourier integral of equation 3.2 implicates a heavily cluttered computational graph and massive memory resources to keep track of every operation.

As the BRDF is directly sampled into an image, the pixel footprint equals the coherence area, and the size of this area is the most decisive factor dictating the computational effort of the inverse rendering process. Adjustments of the coherence kernel's weights to limit the contributions of more distant Gabor kernels and reduce the interdependence are waived since these effects are crucial to wave optics and, subsequently, their inversion to the aims of this thesis. Instead, in this implementation the heightfield is cropped into 16×16 texels, a fractional part of the original. The corresponding coherence area is $16 \mu\text{m}^2$ small covering all inputs. This size is roughly equivalent

to a light source with a subtended angle of 1.8° or the coherence area of a common light bulb a few meters away from the surface (Levin et al., 2013). In contrast, the sun subtends an angle of approx. 0.52° which would result in a proportionally larger area (Mandel and Wolf, 1995).

4.3 Overall procedure

Finally, a full inverse rendering system was composed of the library and the main unit controlling the overall procedure and optimizing the hypothesis in order to approximate the reference BRDF. Putting everything together, the main Python component takes all necessary arguments as command line parameters and delegates the sequential steps to the interface provided. Hence, the procedure can be easily adapted for different optimizers or small adjustments in the approach. A greatly simplified version of the overall procedure is shown in listing 4.1. The query object embodies lighting conditions and the surface point where the BRDF is requested. The hypothesis is initialized as a plane surface with height zero and modulated iteratively according to the retrieved gradients. The BRDF generation consists of the aforementioned preprocessing step prior to the actual wave optical computations, which can be performed either with autodiff enabled or like in the original source code. Not shown in the listing is the visualization of the reference, the hypothesis, its corresponding BRDF estimate, and the gradient per epoch.

Listing 4.1: Simplified Python code for the overall procedure.

```
from wavebrdf import Heightfield , GaborRep , Query , brdf_func , mse

# Generate reference
query = Query(args)
gr_truth_hf = Heightfield(gr_truth_exr)
gr_truth_gabor = GaborRep(gr_truth_hf)
ref_brdf = brdf_func(gr_truth_gabor , query , autodiff=False)

# Initialize hypothesis
hypo_hf = Heightfield(zeros(gr_truth_hf.size()))

# Optimize
for i in range(epochs + 1):
    if stochastic :
        ref_brdf = brdf_func(gr_truth_gabor , rand_query() , autodiff=False)
        hypo_gabor = GaborRep(hypo_hf , autodiff=True)
        brdf = brdf_func(hypo_gabor , query , autodiff=True , ref=ref_brdf)
        loss = mse(brdf , ref_brdf , autodiff=True)
        grad = backward(loss)
        hf.values = optimizer.step(grad)
```

5 Evaluation

In the following chapter, the implemented inverse rendering system is tested and evaluated in a series of experiments. Therefore, the computed gradients are inspected, the progression of the current estimation is measured, and the overall performance is presented for the restricted scenario of a single positional sample x of the SVBRDF and a single incoming light direction ω_i . Thereafter, the extrapolation of the concepts in this work is discussed, from this restricted scenario to more general settings.

5.1 Experimental results

In all experiments, the input heightfield texture consists of 16×16 texels, and the light hits perpendicular to the surface. The optimization is performed for at least 300 epochs, i.e., iterations in the deterministic case, with an initial learning rate of $\alpha = 0.01$ and an exponential decay rate of 0.005 per epoch. This learning rate was found effective, although it is significantly higher than the original suggestions of Kingma and Ba (2017). To some extent, this is due to the bias implicated by the initialization as a plain surface. The actual height values range from -1.0 up to 1.0 height units. The presumed heightfield hypothesis needs a considerable number of modulation steps to exploit even a fraction of this range. In the first presented experiment, the output resolution is set to 32×32 pixels in three color channels. The reference data is considered in its entirety in each epoch; one BRDF slice or, equivalently, a batch of distinct outgoing directions $\omega_o \in \Omega_o$ where $|\Omega_o| \leq 32^2 = 1024$, each of which entailing three scalars to be approximated.

Figure 5.1 presents the state of the inverse rendering process throughout its course in this first experiment. The reference BRDF and the heightfield it originated from are displayed at the top, the BRDF estimates below, thereafter the squared error between the above as a more visually apparent version of the objective function, and finally the hypotheses and the gradients.

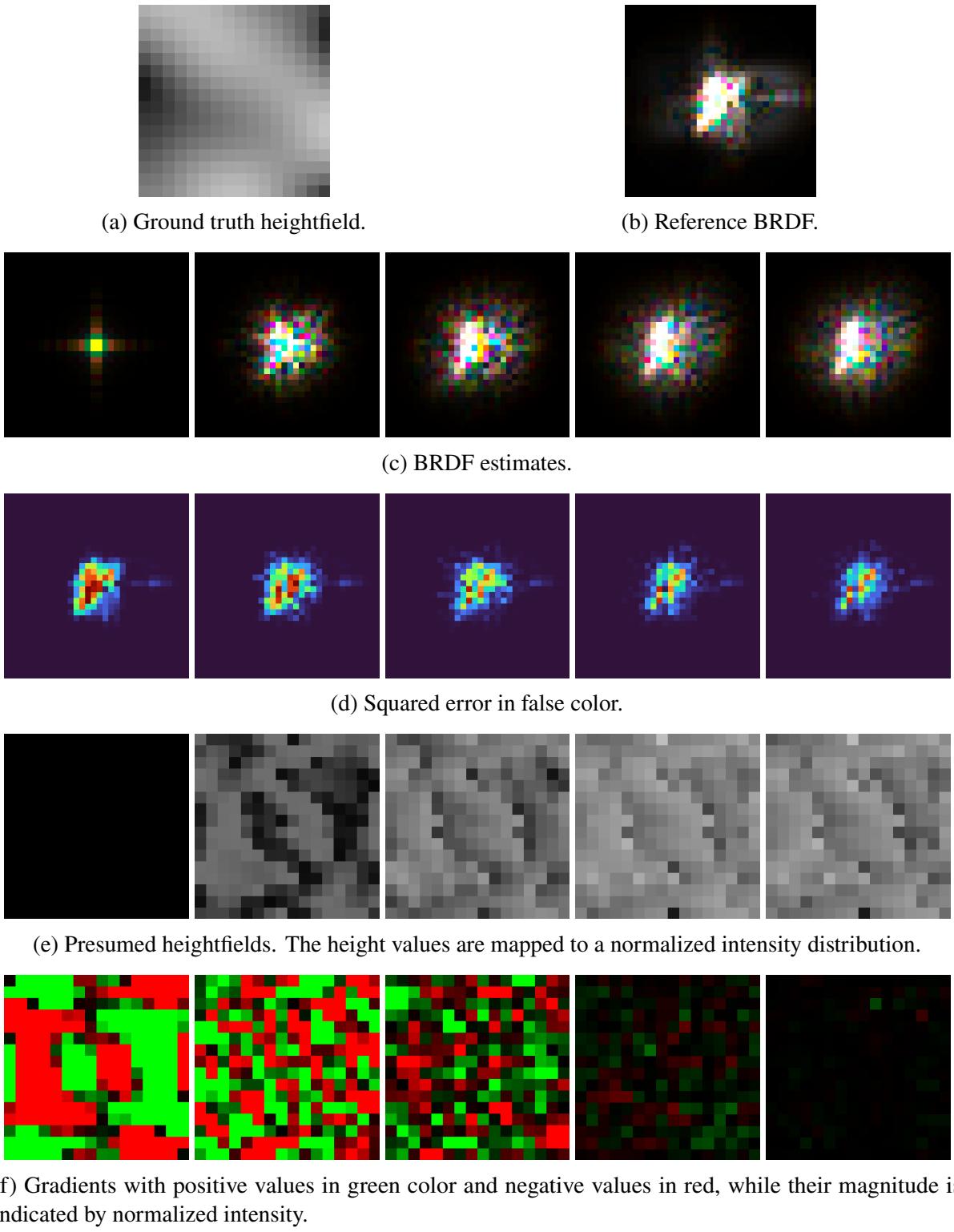


Figure 5.1: Visualization of crucial data during the optimization process. Whereas (a) and (b) picture the invariant ground truth and reference BRDF, the image series (c), (d), (e), and (f) depict the current state in succession. From left to right, the columns represent the epochs: 0 (initial guess), 10, 30, 100, and 300. Created by means of *The EXR Viewer* (Müller, 2017).

The inverse renderer is effective in approximating microgeometry whose BRDF resembles the reference, mainly in the center of the highlight. Regularly, the gradient reveals an antisymmetric pattern similar to a plane wave in the first few epochs. A likely explanation is that predominant differences in the early stages of gradient descent are resolved by stretching the reflectance lobe and adding low-frequency features first. It might be that the variance of the gradient correlates with increasing frequencies of the Fourier transform or the level of detail in the estimation. In general, consecutive gradients behave seemingly unpredictably and dissolve in increasingly volatile and irregular structures, free of any paradigm. As it can be seen, the magnitude of the gradient diminishes steadily to a greater degree, and the presumed microgeometry changes less. Remarkably, as it is constructed from the superposition of antisymmetric gradients, the heightfield is antisymmetric itself at all times, which is consistent with the prevailing lighting conditions.

In this exemplary experiment, the optimization algorithm does not consider multiple surface points or incoming light directions for sampling the SVBRDF. This implicates that the task is simplified into approximating a feasible heightfield that induces a single reference BRDF slice, neglecting the correlation of multiple samples on the same surface. In fact, the presumed heightfield approximates a local minimum of the objective function, a heightfield much different from the original ground truth. This effectively means that completely unrelated microstructures can be the source of eminently similar BRDFs, at least if the desired reflectance defined by the reference is sufficiently restricted in its degree of detail. Ultimately, the implemented inverse renderer only aims to reconstruct a heightfield with reflectance properties that are as similar as possible to the reference SVBRDF from an unknown heightfield. It is not able to retrieve the ground truth if there are alternatives that share the same set of properties requested. To address this ambiguity, the reference can be extended to include other points on the surface, i.e., other BRDFs sampled from the surface's SVBRDF, or additional incoming light directions Ω_i or an increased number of outgoing directions Ω_o .

5.2 Ambiguity

First, it is examined how the quality of the local minima obtained by the inverse renderer changes upon varying sizes of $|\Omega_o|$, the set of distinct outgoing directions. Therefore, two other experiments are conducted; other than the resolution of the BRDF slices, none of the settings are changed. To be pointed out, these results are not affected by any decisions of the implementation but only demonstrate the suitability of how the approach of this work was designed. In figure 5.2,

the BRDF estimate after 300 epochs of optimization and the reference BRDF slice are depicted with up to 64^2 samples in $|\Omega_0|$ and up to 16^2 samples, respectively. With the increased precision, no appropriate candidate heightfield is found. On the contrary, if the BRDF slice is smaller, the estimated BRDF resembles the reference almost perfectly. It is striking that the quality strongly correlates with the number of dependent variables to be approximated, even though the amount of input parameters is unchanged. In both these examples, the presumed heightfield and the ground truth have no similarities apparent from visual inspection.



Figure 5.2: Comparison of reference (left) and final estimate (right) for alternative resolutions with 64^2 (a) and with 16^2 (b) samples, respectively.

Furthermore, figure 5.3 illustrates how the optimization progresses according to different loss functions. This example refers to the experiment where the output resolution is 16×16 pixels. The inverse rendering system always aims to minimize the CMSLE loss from definition 3.2, so the other functions do not affect the optimization. In the experiments so far, the CMSLE measure with $\epsilon = 1.0$ is applied. Therefore, the inverse renderer still focuses on high intensities but prevents outliers in the HDR data to dominate the error value. For comparison, more error functions are included giving additional insights into the optimization process:

- the MSE between the presumed heightfield and the ground truth,
- the CMSLE with $\epsilon = 0.001$,
- the non-logarithmic MSE between the estimated BRDF and the reference,
- the MSE in RGB space,
- the MSE in RGB space with gamma correction.

Every measure converges smoothly to independent limits. For every loss function to occupy a comparable interval, the values are normalized against their starting entries. These measures reveal the following observations: The hypothesis does not converge to the global minimum that is the ground truth but, according to the MSE, finds a local minimum of the objective function where the BRDF estimate is 99.73% more similar to the reference than a flat plane. Notably, a perfectly reflecting surface is a special case where the BRDF lobe is extremely narrow and

focused. Whereas the applied CMSLE loss with $\epsilon = 1.0$ is reduced close to the MSE value, CMSLE with $\epsilon = 0.001$ converges to a higher limit, which means that the lower weighted intensities at the tails of the highlight are less precisely approximated. This is also indicated by figure 5.2b. The MSE in RGB space exposes the same result since a gamma correction operation with $\gamma = 2.2$ increases the loss significantly. Thereby, lower intensities are mapped to a larger interval of the spectrum and thus, weighted more than before.

The transformation into RGB images makes use of *Reinhard tonemapping* (Reinhard et al., 2002) for mapping the HDR values that can become indefinitely large into a finite interval. Since tonemapping is not part of the original wave optical BRDF generation, it is rejected for use in the inversion system, so none of these additional measures are incorporated into the objective function.

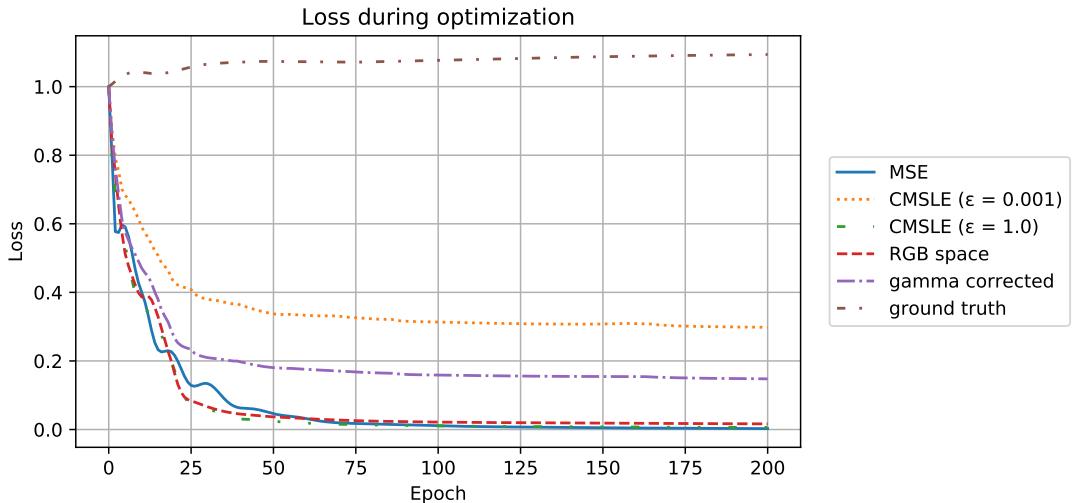


Figure 5.3: The error progression per epoch evaluated using different loss functions. The loss value is normalized against the initial value (epoch 0).

Based on these results, the choice of the objective function is examined in further tests. It could be argued that the value of ϵ might be decreased in the 16×16 case to consider less vigorous reflections more thoroughly, at the expense of an insignificant loss of accuracy in the center of the highlight. Analogously, if the output resolution is lifted, the CMSLE could be replaced by the MSE to concentrate the intended reflectance properties of the reference and obtain a slightly better approximation of only these. To summarize all these tests conducted: If the objective function is adapted, the focus of the approximated features shifts accordingly, but the overall quality is the same. Every improvement brings inaccuracy elsewhere with it. For instance, if MSE is used as the objective instead, its precision is improved, but the CMSLE loss shows the results are not more accurate overall. For this reason, it can be stated that the choice of the

objective function can draw attention to different aspects of the reflectance properties as desired, but the task to approximate a small number of outgoing light samples accurately has about the same difficulty as for the full high-resolution slice with only moderate accuracy.

Definitely, the capabilities of analysis-by-synthesis with straightforward gradient-descent are limited because of the ambiguity of the inverse problem in the search space. The inverse renderer attempts to retrieve the original heightfield from guessing a candidate, i.e., a local minimum of the objective function. The idea of extending the reference data to make it more restrictive in such a way that optimally, the only feasible heightfield is the ground truth, inevitably increases the complexity of the candidate-finding process prohibitively. In the following, the addition of multiple directional and positional samples of the SVBRDF is discussed.

5.3 Extrapolation of the concepts

The effective approximation of a single BRDF despite its limitation in the step size of its lobe discretization suggests the extrapolation of the present approach onto more universal problem statements. Instead of expanding the reference data once more by adding arbitrarily predetermined data to the reference, and deteriorating the reflectance loss of any local minimum, this time the dependent variables of the output are selected in a stochastic manner. Therefore, the same concepts as before are applied to stochastic gradient-descent with varying positional and directional samples of the original SVBRDF.

In further experiments, either the points on the spatially-varying surface microstructure, the incoming directions, or both are sampled randomly from the entire domain of the function. In contrast, one BRDF slice of any resolution is considered a mini-batch consisting of a fixed number of directional samples. Instead of sampling pixels using Monte Carlo importance sampling strategies, here, each direction is sampled exactly once. The idea is that every Gabor kernel in the coherence area contributes to the whole output, and all kernels are considered simultaneously anyway. Again, if the BRDF model is reciprocal (as in all preceding experiments), there is no difference in bundling a set of incoming or outgoing directions into a mini-batch. Stochastic sampling inside a mini-batch could reduce the technical requirements of the implementation, but the effect would be similar to the reduction of the output resolution as mentioned before. On the contrary, considering multiple surface points requires more reflectance measurements, if the reference is based on real-world data.

There are two ways how stochastic gradient descent was tested: randomly sequencing mini-

batches in an online-learning-style setup and applying the gradients continuously, or wrapping several reference BRDFs into a batch of a larger dataset while averaging the gradients from the BRDF generation passes before enforcing them all-together onto the presumed heightfield each epoch. Unfortunately, in all variations, the estimation does not converge in a foreseeable time. Due to frequent and unrelated updates, the reference has unlimited variance and the gradient steps in the search space are noisy. In all experiments, the stochastic gradient descent algorithm diverges for a long time or becomes anchored in a local minimum because of the non-convex, high-dimensional, and highly irregular nature of the problem.

5.4 Problem complexity

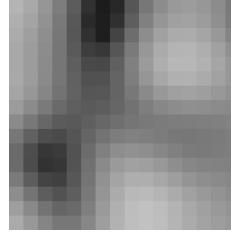
Exploring the complexity of the inverse problem is a key aspect of this work. There are $n = 16^2$ parameters to be approximated, with infinite possibilities of their contributions towards the reference. The aforementioned experiments suggest that there is a multitude of heightfields without any comparable pattern that induce equivalent reflectance properties depending on the specificity of those. The rendering function mapping these parameters to their according BRDF estimation is n-dimensional and non-linear, because of how the phase delay is superimposed from the heights. Alternatively, the computational complexity might be bypassed by handling the rendering function as a black box and use deep learning to train a neural network estimating heightfields from a set of BRDF samples.

From chapter 3, it is evident that the generation of wave optical BRDFs is continuous in all height values of the coherence area. Thus, the optimization via gradient descent is able to find local minima in the parameter space. However, the more attainable the reference is, the harder it is to identify global minima from the multitude of candidates. On the other hand, with overly restrictive demands, the quality of the candidates found is decreased. The height-modulation of one surface point affects all output BRDF samples. This is why it is unclear how the impact on the light wave's phase can be localized with regard to the degree of detail required during rendering. For the heightfield to be general, a representative model for arbitrary surface structure, there is no assumption of spatial correlation between the surface points made.

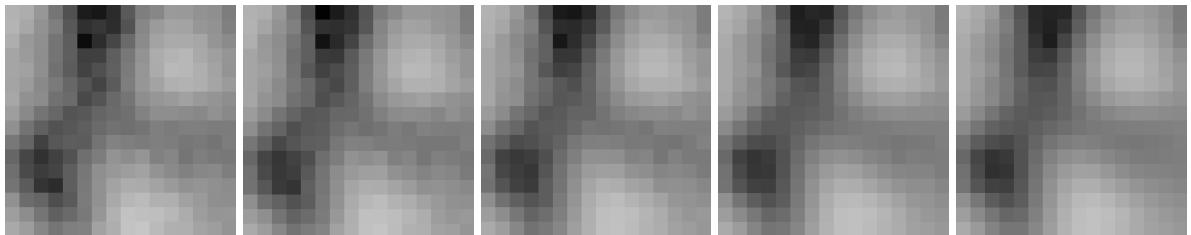
The problem can also be regarded as multivariate non-linear regression. The forward rendering function $f(X, \beta)$ takes the independent variables X including the surface position as well as directions ω_i and ω_o to predict the reflectance using the heightfield parameters β . Gradient descent is a usual approach for this kind of problem. To prevent the optimization to converge to

CHAPTER 5. EVALUATION

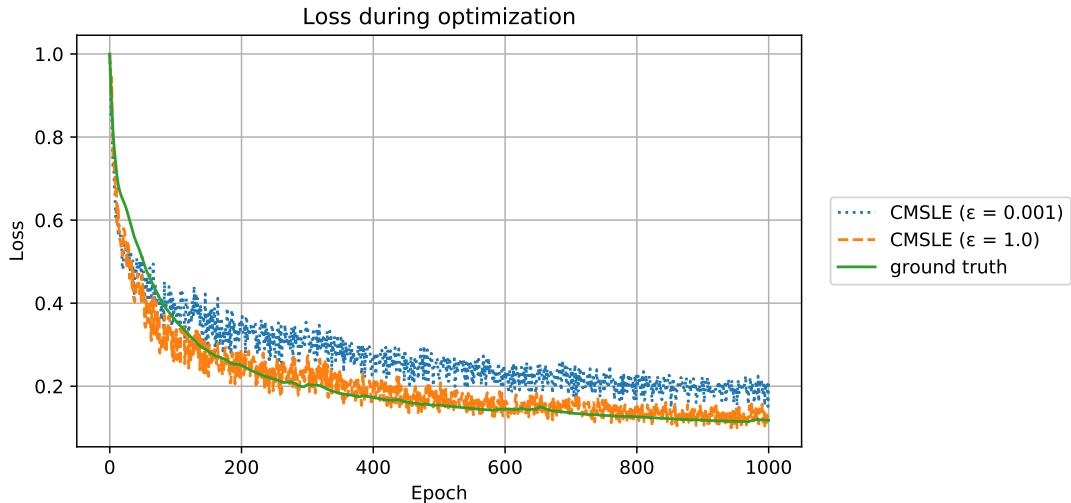
local minima instead of the global optimum, these might have to start from a rough estimate. Figure 5.4 shows the implemented inverse renderer starting from an estimated hypothesis. This estimation is created by adding a slight random noise onto the ground truth heightfield, sampled uniformly from an interval equivalent to approx. 10% of the range that the original height values take. The learning rate is reduced to 0.0015 and the CMSLE with $\epsilon = 0.01$ is utilized as the objective function. In 1000 epochs with random positional samples the noise is resolved by approx. 89%, which might be improved even further if the procedure would be continued.



(a) Ground truth heightfield.



(b) Presumed heightfields. From left to right, representing the epochs: 0 (estimation), 10, 50, 200, 1000.



(c) Progression of CSMLE loss as well as the MSE between presumption and ground truth. The loss value is normalized against the initial value (epoch 0).

Figure 5.4: Visualization of the results when starting from an estimation.

6 Conclusion

To conclude, it is shown that inverting the realistic rendering of wave optical BRDFs to retrieve the general microgeometry of specular surfaces from reflectance data is a challenging task that inherits many possibilities and difficulties to consider. Using the analysis-by-synthesis approach and gradient-based optimization techniques, the appearance of arbitrary microstructures can be imitated by iteratively approximating a feasible heightfield starting from an initial guess. This discretized heightfield represents a candidate for the surface geometry whose properties suffice a reference set of specified SVBRDF samples to a certain degree.

It is demonstrated that different variations of the gradient descent method can obtain an imitation only with limited accuracy because of the reciprocal effect between ambiguity and difficulty of the problem when defining the reference. Inside the coherence area, the modulation of a single surface point's height causes changes in every directional or positional SVBRDF sample. More importantly, it fails to reconstruct the original microgeometry because of the complex interdependence and the density of local minima in the high-dimensional parameter space. However, by simplifying the task and assuming an infinite number of random reference measurements, the inverse renderer successfully improves geometry significantly using stochastic gradient descent and starting from an estimation of the ground truth.

In order to accomplish this task, the efficient forward renderer provided by Yan, Hašan, Walter, et al. (2018) was rebuilt into an adapted library for differentiable rendering. Hence, the rendering function is combined with an objective function for the extraction of gradients regarding the presumed height values of the surface. To handle the number of computations for the interference effects and diffraction patterns, the implementation makes use of automatic differentiation, a method that creates a transcript of all the actual arithmetic operations of the rendering process and propagate the loss backward through this construct to the contributors responsible. As the objective function, a logarithmic version of the standard mean squared error loss was applied to measure the current error between the HDR images.

CHAPTER 6. CONCLUSION

In future work, the technical shortcomings might be overcome efficiently using capable hardware and a vectorized implementation. The vectorization of the forward rendering might not only accelerate the process but also resolve the redundancy of the autodiff transcript, which turned out to be the bottleneck of the implementation of this work. Since any pruning mechanisms need to be disabled or specifically handled otherwise for automatic differentiation, in any case, the estimation of the contributions from all surface points should be executable in a parallel manner. The space complexity of the computational graph would be reduced tremendously. As mentioned before, a seamless extension of the forward rendering into a full scene renderer with Mitsuba 2 is possible. As a concern of this work, it was ensured that this extension does not require difficult re-arrangements of the methodology since Mitsuba 2 also relies on pybind11 and Enoki. Arguably, in the full inverse rendering system, vectorization might only be practical to a limited extent: Let m be the size of the coherence area and n the number of BRDF samples to be determined simultaneously for rendering the scene. A parallel computation would require operating on a vector of n^2m^2 complex numbers, which might exceed the RAM of common GPUs, and there is no reusability of sub-vectors beyond the step of preprocessing. However, even a partial vectorization, e.g., of the m^2 Gabor kernels, could accelerate the propagation through the computational graph of the rendering greatly.

Besides any improvements of the implementation, the approach and its individual concepts should be reconsidered in detail. The stochastic gradient descent algorithm, its configuration, or the development of more sophisticated loss functions do not appear to change the inherent nature of the problem and the search space, avoiding suboptimal local minima. For this reason, there might be a need of additional regularisation techniques. Since there is a specific interest of finding the one true global optimum, simulated annealing could be a suitable addition to the optimization methodology. The experiments show that the rough localization of the global minimum heightfield is difficult, but, once a promising neighborhood of is found, the ground truth can be approximated smoothly from there. At the same time, simulated annealing is bound to a more efficient implementation to be practical, as the search space is extremely large and the corresponding BRDF predictions of neighboring heightfield candidates are noisy.

Alternatively, a different approach might be to ignore the complications of physically accurate forward rendering methods and make use of deep learning techniques. Kuznetsov et al. (2019) utilize conditional generative adversarial networks to produce BRDF samples indistinguishable from a dataset of wave optical BRDF slices. Their results implicate that, with some effort, a neural network is able to learn the underlying features of the reflectance of a spatially-varying surface. In future work, it might be a considerable approach to develop a suitable deep learning model learning the heightfield from which the BRDFs originate as well.

Bibliography

- Bielawny, A. (2019). “Reflectors in lighting design: Reflector-based non-imaging optics for lighting applications”. In: *Advanced Optical Technologies* 8.6, pp. 469–481. URL: <https://doi.org/10.1515/aot-2018-0052>.
- Bieron, J. C. and P. Peers (2020). “An Adaptive Metric for BRDF Appearance Matching”. In: *MAM@EGSR*.
- Dong, Y., G. Chen, P. Peers, J. Zhang, and X. Tong (Nov. 2014). “Appearance-from-Motion: Recovering Spatially Varying Surface Reflectance under Unknown Lighting”. In: 33.6. URL: <https://doi.org/10.1145/2661229.2661283>.
- Dong, Z., B. Walter, S. Marschner, and D. P. Greenberg (Dec. 2016). “Predicting Appearance from Measured Microgeometry of Metal Surfaces”. In: *ACM Trans. Graph.* 35.1. URL: <https://doi.org/10.1145/2815618>.
- Gamboa, L. E., J.-P. Guertin, and D. Nowrouzezahrai (Dec. 2018). “Scalable Appearance Filtering for Complex Lighting Effects”. In: *ACM Trans. Graph.* 37.6. URL: <https://doi.org/10.1145/3272127.3275058>.
- Gao, D., X. Li, Y. Dong, P. Peers, K. Xu, and X. Tong (July 2019). “Deep Inverse Rendering for High-Resolution SVBRDF Estimation from an Arbitrary Number of Images”. In: *ACM Trans. Graph.* 38.4. URL: <https://doi.org/10.1145/3306346.3323042>.
- Guennebaud, G., B. Jacob, et al. (2010). *Eigen v3*. <http://eigen.tuxfamily.org>.
- Harris, C. R., K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant (Sept. 2020). “Array programming with NumPy”. In: *Nature* 585.7825, pp. 357–362. URL: <https://doi.org/10.1038/s41586-020-2649-2>.

BIBLIOGRAPHY

- Harvey, J. E. (1979). “Fourier treatment of near-field scalar diffraction theory”. In: *American Journal of Physics* 47, pp. 974–980.
- Iseringhausen, J. and M. B. Hullin (Jan. 2020). “Non-Line-of-Sight Reconstruction Using Efficient Transient Rendering”. In: *ACM Trans. Graph.* 39.1. URL: <https://doi.org/10.1145/3368314>.
- Jakob, W. (2019). *Enoki: structured vectorization and differentiation on modern processor architectures*. <https://github.com/mitsuba-renderer/enoki>.
- Jakob, W., M. Hašan, L.-Q. Yan, J. Lawrence, R. Ramamoorthi, and S. Marschner (July 2014). “Discrete Stochastic Microfacet Models”. In: *ACM Trans. Graph.* 33.4. URL: <https://doi.org/10.1145/2601097.2601186>.
- Jakob, W., J. Rhinelander, and D. Moldovan (2017). *pybind11 – Seamless operability between C++11 and Python*. <https://github.com/pybind/pybind11>.
- Johnson, M. K., F. Cole, A. Raj, and E. H. Adelson (July 2011). “Microgeometry Capture Using an Elastomeric Sensor”. In: *ACM Trans. Graph.* 30.4. URL: <https://doi.org/10.1145/2010324.1964941>.
- Johnson, M. K. and H. Farid (2007). “Exposing Digital Forgeries in Complex Lighting Environments”. In: *IEEE Transactions on Information Forensics and Security* 2.3, pp. 450–461.
- Kainz, F., R. Bogart, and D. Hess (Jan. 2004). “OpenEXR image file format”. In: *GPU Gems: Programming Techniques, Tips, and Tricks for Real-Time Graphics*. URL: <https://developer.nvidia.com/gpugems/part-iv-image-processing/chapter-26-openexr-image-file-format>.
- Kajiya, J. T. (1986). “The rendering equation”. In: *Computer Graphics*, pp. 143–150.
- Karsch, K. (2020). “Inverse Rendering Techniques for Physically Grounded Image Editing”. In: *CoRR* abs/2001.00986. arXiv: 2001.00986. URL: <http://arxiv.org/abs/2001.00986>.
- Kassubeck, M., F. Bürgel, S. Castillo, S. Stiller, and M. Magnor (Jan. 2021). “Shape from Caustics: Reconstruction of 3D-Printed Glass from Simulated Caustic Images”. In: *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 2877–2886.
- Kato, H., D. Beker, M. Morariu, T. Ando, T. Matsuoka, W. Kehl, and A. Gaidon (2020). “Differentiable Rendering: A Survey”. In: *CoRR* abs/2006.12057. arXiv: 2006.12057. URL: <https://arxiv.org/abs/2006.12057>.

BIBLIOGRAPHY

- Kingma, D. P. and J. Ba (2017). *Adam: A Method for Stochastic Optimization*. arXiv: 1412.6980 [cs.LG].
- Kress, B. (2014). *Field Guide to Digital Micro-optics*. Field Guides. SPIE Press. URL: <https://books.google.de/books?id=V8WwoQEACAAJ>.
- Kuznetsov, A., M. Hašan, Z. Xu, L.-Q. Yan, B. Walter, N. K. Kalantari, S. Marschner, and R. Ramamoorthi (Nov. 2019). “Learning Generative Models for Rendering Specular Microgeometry”. In: *ACM Trans. Graph.* 38.6. URL: <https://doi.org/10.1145/3355089.3356525>.
- Lange, T., A. Völl, M. Berens, J. Stollenwerk, P. Loosen, and C. Holly (2021). *Rigorous approach for unifying wave optics and state of the art rendering techniques*. URL: <https://publica-stage.fraunhofer.de/handle/publica/269683>.
- Levin, A., D. Glasner, Y. Xiong, F. Durand, W. Freeman, W. Matusik, and T. Zickler (July 2013). “Fabricating BRDFs at High Spatial Resolution Using Wave Optics”. In: *ACM Trans. Graph.* 32.4. URL: <https://doi.org/10.1145/2461912.2461981>.
- Li, Z. and Y. Li (2011). “Microscopic photometric stereo: A dense microstructure 3D measurement method”. In: *2011 IEEE International Conference on Robotics and Automation*, pp. 6009–6014.
- Liu, B., J. Lei, B. Peng, C. Yu, W. Li, and N. Ling (2021). “Novel View Synthesis from a Single Image via Unsupervised learning”. In: *CoRR* abs/2110.15569. arXiv: 2110.15569. URL: <https://arxiv.org/abs/2110.15569>.
- Liu, S. and M. N. Do (2017). “Inverse Rendering and Relighting From Multiple Color Plus Depth Images”. In: *IEEE Transactions on Image Processing* 26.10, pp. 4951–4961.
- Loscos, C., K. Jacobs, G. Patow, and X. Pueyo (2006). “Inverse Rendering: From Concept to Applications”. In: *Eurographics 2006: Tutorials*. Ed. by N. Magnenat-Thalmann and K. Bühlert. The Eurographics Association.
- Loscos, C., M.-C. Frasson, G. Drettakis, B. Walter, X. Granier, and P. Poulin (1999). “Interactive Virtual Relighting and Remodeling of Real Scenes”. In: *Proceedings of the 10th Eurographics Conference on Rendering*. EGWR’99. Granada, Spain: Eurographics Association, pp. 329–340.
- Mandel, L. and E. Wolf (1995). *Optical Coherence and Quantum Optics*. Cambridge University Press.

BIBLIOGRAPHY

- Margossian, C. C. (2018). “A Review of automatic differentiation and its efficient implementation”. In: *CoRR* abs/1811.05031. arXiv: 1811.05031. URL: <http://arxiv.org/abs/1811.05031>.
- Moriwaki, K., R. Yoshihashi, R. Kawakami, S. You, and T. Naemura (2018). “Hybrid Loss for Learning Single-Image-based HDR Reconstruction”. In: *CoRR* abs/1812.07134. arXiv: 1812.07134. URL: <http://arxiv.org/abs/1812.07134>.
- Müller, T. (2017). *tev — The EXR Viewer*. <https://github.com/Tom94/tev>.
- Nam, G., J. H. Lee, H. Wu, D. Gutierrez, and M. H. Kim (Nov. 2016). “Simultaneous Acquisition of Microscale Reflectance and Normals”. In: *ACM Trans. Graph.* 35.6. URL: <https://doi.org/10.1145/2980179.2980220>.
- Nicodemus, F. E., J. C. Richmond, J. J. Hsia, I. W. Ginsberg, and T. Limperis (1977). “Geometrical considerations and nomenclature for reflectance”. In.
- Nimier-David, M., S. Speierer, B. Ruiz, and W. Jakob (July 2020). “Radiative Backpropagation: An Adjoint Method for Lightning-Fast Differentiable Rendering”. In: *Transactions on Graphics (Proceedings of SIGGRAPH)* 39.4.
- Nimier-David, M., D. Vicini, T. Zeltner, and W. Jakob (Dec. 2019). “Mitsuba 2: A Retargetable Forward and Inverse Renderer”. In: *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 38.6.
- NVIDIA, P. Vingelmann, and F. H. Fitzek (2020). *CUDA, release: 10.2.89*. URL: <https://developer.nvidia.com/cuda-toolkit>.
- Papas, M., W. Jarosz, W. Jakob, S. Rusinkiewicz, W. Matusik, and T. Weyrich (Apr. 2011). “Goal-based Caustics”. In: *Comput. Graph. Forum* 30, pp. 503–511.
- Patow, G. and X. Pueyo (2005). “A Survey of Inverse Surface Design From Light Transport Behavior Specification”. In: *Computer Graphics Forum*.
- Pharr, M., W. Jakob, and G. Humphreys (2016). *Physically Based Rendering: From Theory to Implementation*. 3rd. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Reinhard, E., M. Stark, P. Shirley, and J. Ferwerda (July 2002). “Photographic Tone Reproduction for Digital Images”. In: *ACM Trans. Graph.* 21.3, pp. 267–276. URL: <https://doi.org/10.1145/566654.566575>.

BIBLIOGRAPHY

- Stam, J. (1999). “Diffraction Shaders”. In: *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’99. USA: ACM Press/Addison-Wesley Publishing Co., pp. 101–110. URL: <https://doi.org/10.1145/311535.311546>.
- Tourre, V., J.-Y. Martin, and G. Hégron (2008). “An Inverse Daylighting Model for CAAD”. In: *Proceedings of the 24th Spring Conference on Computer Graphics*. SCCG ’08. Budmerice, Slovakia: Association for Computing Machinery, pp. 83–90. URL: <https://doi.org/10.1145/1921264.1921283>.
- Velinov, Z., S. Werner, and M. B. Hullin (2018). *Real-Time Rendering of Wave-Optical Effects on Scratched Surfaces*.
- Vicini, D., S. Speierer, and W. Jakob (July 2021). “Path Replay Backpropagation: Differentiating Light Paths Using Constant Memory and Linear Time”. In: *ACM Trans. Graph.* 40.4. URL: <https://doi.org/10.1145/3450626.3459804>.
- Werner, S., Z. Velinov, W. Jakob, and M. B. Hullin (2017). “Scratch iridescence: Wave-optical rendering of diffractive surface structure”. In: *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 36.6, 207:1–207:14.
- Weyrich, T., P. Peers, W. Matusik, and S. Rusinkiewicz (July 2009). “Fabricating Microgeometry for Custom Surface Reflectance”. In: *ACM Trans. Graph.* 28.3. URL: <https://doi.org/10.1145/1531326.1531338>.
- Yan, L.-Q., M. Hašan, W. Jakob, J. Lawrence, S. Marschner, and R. Ramamoorthi (July 2014). “Rendering Glints on High-Resolution Normal-Mapped Specular Surfaces”. In: *ACM Trans. Graph.* 33.4. URL: <https://doi.org/10.1145/2601097.2601155>.
- Yan, L.-Q., M. Hašan, S. Marschner, and R. Ramamoorthi (July 2016). “Position-Normal Distributions for Efficient Rendering of Specular Microstructure”. In: *ACM Trans. Graph.* 35.4. URL: <https://doi.org/10.1145/2897824.2925915>.
- Yan, L.-Q., M. Hašan, B. Walter, S. Marschner, and R. Ramamoorthi (July 2018). “Rendering Specular Microgeometry with Wave Optics”. In: *ACM Trans. Graph.* 37.4. URL: <https://doi.org/10.1145/3197517.3201351>.
- Zhang, C., B. Miller, K. Yan, I. Gkioulekas, and S. Zhao (2020). “Path-Space Differentiable Rendering”. In: *ACM Trans. Graph.* 39.4, 143:1–143:19.
- Zhang, K., F. Luan, Q. Wang, K. Bala, and N. Snavely (2021). “PhySG: Inverse Rendering with Spherical Gaussians for Physics-based Material Editing and Relighting”. In: *CoRR* abs/2104.00674. arXiv: 2104.00674. URL: <https://arxiv.org/abs/2104.00674>.

BIBLIOGRAPHY

- Zhao, S., W. Jakob, and T.-M. Li (2020). “Physics-Based Differentiable Rendering: From Theory to Implementation”. In: *ACM SIGGRAPH 2020 Courses*. SIGGRAPH ’20. Virtual Event, USA: Association for Computing Machinery. URL: <https://doi.org/10.1145/3388769.3407454>.
- Zhao, S., W. Jakob, S. Marschner, and K. Bala (Oct. 2014). “Building Volumetric Appearance Models of Fabric Using Micro CT Imaging”. In: *Commun. ACM* 57.11, pp. 98–105. URL: <https://doi.org/10.1145/2670517>.
- Zhu, J., Y. Xu, and L. Wang (2019). “A Stationary SVBRDF Material Modeling Method Based on Discrete Microsurface”. In: *Computer Graphics Forum* 38.7, pp. 745–754. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13876>.