Efficient Capture of Multispectral Reflectance of Complex Surfaces

Dissertation

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

vorgelegt von

Dipl.-Inform. Martin Rump

aus
Bergisch Gladbach

Bonn, August 2020

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

Dekan: Prof. Dr. Johannes Beck

1. Gutachter: Prof. Dr. Reinhard Klein
2. Gutachter: Prof. Dr. Matthias B. Hullin

Tag der Promotion: 27.11.2020
Erscheinungsjahr: 2020
Diese Dissertation ist auf dem Publikationsserver bonndoc der ULB Bonn elektronisch publiziert:
https://nbn-resolving.org/urn:nbn:de:hbz:5-60646

CONTENTS

ZUSAMMENFASSUNG

Die synthetische Erzeugung von photorealistischen Bildern ist eines der am längsten in der Computergrafik verfolgten Ziele. Die Anwendungen sind breitgefächert und reichen von Virtual Prototyping über Werbung bis hin zu visuellen Effekten in Filmen. Seit den Anfangstagen der Computergrafik wurde ein beträchtlicher Forschungsaufwand in die Entwicklung von Renderingalgorithmen und in die Repräsentation von Szenengeometrie, Lichtquellen und optischen Oberflächeneigenschaften investiert. Obwohl es in einigen dieser Gebiete schnelle Fortschritte gab, fanden die optischen Materialeigenschaften über lange Zeit wenig Beachtung.

Zunächst wurden nur einfache Texturen und phenomenologische Modelle mit wenigen Parametern, die auf einfachen Annahmen über die Oberfläche basieren, verwendet. Mit der Zeit wurden dann auch bessere, physikalisch basierte Reflexionsmodelle entwickelt, die die Reflexionseigenschaften von zumindest einigen Materialklassen schon gut beschreiben konnten. Als Durchbruch in der Beschreibung von optischen Materialeigenschaften kann dann die Entwicklung der sogenannten datengetriebenen Methoden bezeichnet werden. Dabei werden die Reflexionseigenschaften in einer großen Tabelle dicht abgetastet gespeichert. Diese Tabellen werden oft durch Messung von realen Oberflächen erzeugt. Datengetriebene Methoden haben den Realismus in der Wahrnehmung vor allem komplexer Oberflächen mit deutlicher Struktur oder örtlich variierenden Reflexionseigenschaften deutlich erhöht. Die Messung von optischen Materialeigenschaften hat aber auch den Reflexionsmodellen zu neuer Bedeutung verholfen, da sich die Parameter für diese Modelle automatisch so bestimmen lassen, dass das Modell die gemessenen Daten und damit die reale Oberfläche bestmöglich wiedergibt. Sind die Parameter bestimmt, so kommen die Vorteile der Reflexionsmodelle, nämlich Speichereffizienz und einfaches, semantisches Editieren, zum Tragen.

Heutzutage ist die Messung von Reflexionseigenschaften schon weiter verbreitet. Durch die Verwendung von digitalen Photokameras konnten effiziente Aufbauten zur Akquise von Reflexionsdaten mit hoher Orts- und Winkelauflösung konstruiert werden. In fast allen Fällen wurden jedoch RGB Kameras wegen ihres geringen Preises und der einfachen Handhabung verwendet. Leider führt die Diskretisierung des Lichtspektrum durch nur drei, sich eventuell sogar über-

lappende Filter zu einer falschen Farbdarstellung, wenn unterschiedliche Lichtquellen und Materialien in einer virtuellen Szene kombiniert werden. Der Grund dafür ist ein Effekt namens *Metamerismus*. Grundsätzlich lässt sich dieses Problem einfach durch eine bessere Abtastung der spektralen Dimension umgehen. Leider sind die dafür notwendigen Geräte teuer und benötigen einen vergleichsweise immensen Aufwand um hochqualitative Messdaten zu liefern. Aus diesem Grund ist die Messung von dicht abgetasteten Reflexionseigenschaften mit guter Spektralauflösung im Bereich der Computergrafik immer noch ein weitgehend ungelöstes Problem.

Die in dieser Dissertation vorgestellten Methoden stellen einen Schritt in die Richtung von einfach verwendbarer Technologie für spektrale Reflexionsmessung dar. Die Grundidee ist die Wiederverwertung von bestehender und bereits ausgereifter RGB-Technologie mit Zusatz von möglichst wenigen spektralen Messdaten. Die vollständigen, dicht abgetasteten spektralen Reflexionseigenschaften werden dann durch eine neue, in dieser Arbeit vorgestellte Methode rekonstruiert. Auf diese Weise wird die Benutzung von teuren Geräten zur Spektralaufnahme, wie Filter, Kameras und speziellen Lichtquellen, so gering wie möglich gehalten. Dies reduziert Preis und Komplexität der Messtechnik sowie die Messzeiten.

Dieses Ziel wird in drei Schritten angegangen: Im ersten Schritt wird ein Messaufbau konstruiert und kalibriert, der spektrale Reflexionseigenschaften von Materialmustern ohne weitere Optimierungen aufnehmen kann. Dieser Aufbau erlaubt uns Ground-Truth Daten zu akquirieren die wir dann in den folgenden Schritten zur Qualitätsbewertung von komplexeren Methoden heranziehen können.

Der zweite Schritt ist die Entwicklung eines neuen Rekonstruktionsverfahrens, das spektral aufgelöste Bilder aus dünn gemessenen Spektraldaten und dicht abgetasteten RGB Daten rekonstruiert. Unsere Methode basiert auf der Minimierung einer Energiefunktion, die einerseits Übereinstimmung mit den Messdaten bewertet und andererseits die Plausibilität mittels in einer neuer Funktion dargestellten Vorwissens sicherstellt.

Im letzten Schritt wird die reale Anwendung der Rekonstruktionsmethode in einem RGB-Spektral-Aufbau behandelt. Dafür stellen wir zunächst eine neue Methode vor, die RGB-Kameras spektral exakt zu charakterisieren. Danach evaluieren wir eine Methode, um schnell die zusätzlichen Spektraldaten zu akquirieren, indem eine zusätzliche Spektralkamera in einen bestehenden RGB Aufbau integriert wird. Wir demonstrieren, wie damit spektrale Reflexionseigenschaften von hoher Qualität mit geringem Aufwand gemessen werden können.

ABSTRACT

The synthetic generation of photo-realistic images is a long-standing goal in the area of Computer Graphics. The applications are numerous, including virtual prototyping, advertisement and visual effects in movies. Since the early days of Computer Graphics, a tremendous amount of work has been spent on rendering algorithms and on the representation of scene geometry, light, and surface reflectance. While some of these areas evolved rather quickly, the optical reflectance behavior of surfaces did not receive much attention over a long time.

In the first time, only simple textures and phenomenological models with few parameters based on rather simple assumptions about surface materials were used. Over the time, better physically-based reflectance models were developed which have the potential to describe at least some classes of materials faithfully. The introduction of so-called *data-driven* representations for reflectance can be considered a major breakthrough. Here, reflectance data is typically acquired from real-world samples and stored in large tables with dense sampling. Data-driven techniques led to a substantial increase in the quality of virtual material appearance especially for complex surfaces exhibiting a lot of structure and spatially varying reflectance behavior. Measured reflectance properties of real-world materials also helped to use reflectance models in a more sensible way, by finding parameters for which the model resembles the real-world sample as good as possible. Afterwards, the parametric model can unveil its strengths, namely memory-efficiency and simple, physically-based editing.

Nowadays, measuring reflectance properties of surfaces has become more commonplace. The use of digital cameras enabled the development of efficient setups to acquire reflectance data with a dense sampling of both angular and spatial domain. In nearly all cases, RGB or similar trichromatic cameras were utilized, because they are cheap and easy to use. Unfortunately, the discretization of the light spectrum using three filters, which might even overlap, leads to inaccurate colors by an effect called *metamerism* as soon as different light sources and materials are combined during rendering. In principle, this can be easily circumvented using a better sampling of the spectral domain. Unfortunately, the necessary spectral devices are expensive and require a lot of additional effort to provide high quality measurement data. For this reason, densely sampled reflectance mea-

surement of complex surfaces with good spectral resolution still remains an open problem in Computer Graphics.

The methods presented in this thesis are a step towards widely usable spectral reflectance capture. The basic idea is to re-utilize established and matured RGB technology and to add as few spectral measurement data as possible. The dense, spectral reflectance data is reconstructed afterwards using a novel method. This way, the amount of dedicated spectral hardware like filters, cameras, and special light sources and therefore cost and complexity of hardware are reduced, and the speed of spectral measurements is increased.

We divide the work towards that goal into three main steps: The first one is to build and calibrate a measurement setup that can acquire spectral reflectance data in a brute-force manner. This setup allows us to acquire ground-truth data which in turn enables to judge the quality of more sophisticated methods.
The second step is to develop and evaluate a novel reconstruction method for spectral images from sparse spectral and dense RGB data. This method is the foundation of our enhanced spectral reflectance capture. The spectral reconstruction is performed by minimizing an energy function that meters deviation from the measured data as well as compliance with a novel prior.
In the last step we aim at the real usage of our aforementioned method in a combined RGB-spectral measurement setup. We propose a novel solution to the practically relevant problem of obtaining a highly accurate spectral characterization of RGB cameras. We then evaluate a method to rapidly acquire the additionally required spectral data by integration of only one spectral camera into an existing RGB setup without any further modifications. We demonstrate first results for spectral reflectance capture using this setup.

ACKNOWLEDGEMENTS

Writing this thesis has been a long journey. It started in 2008, when professor Reinhard Klein asked me - having my diploma thesis half-finished - whether I would consider staying as a PhD student. I agreed quickly - one of the best decisions in my life. I want to thank him for his trust back then and also along the whole way, especially after I left the university to pursue other opportunities. He was and is an expiring good example for me, especially his sharp way of analyzing any kinds of mathematical and algorithmic questions.

During my work at the Computer Graphics group I learned how valuable a place with so many talented and open-minded persons actually is. In particular, I had the pleasure to publish and co-author papers with Gero Müller, Arno Zinke, Christopher Schwartz, Roland Ruiters, Dirk Koch, and Michael Weinmann. Special thanks go to Sebastian Merzbach for continuing with my ideas after I left the group and also supporting me towards finishing of this thesis. Also, I want to give Ralf Sarlette a special mention, as it wouldn't have been possible to build all the lab setups, that I used during my work, without his guidance and help. I also want to thank former office mates at the university, namely Marlon Mylo and Tobias Langenbucher. I have enjoyed working with all of them and have benefited from every discussion we had.

Moreover, it is necessary to mention the European Union, the Deutsche Forschungsgemeinschaft and the Rheinische Friedrich-Wilhelms-Universität Bonn for financing my work over the years.

And last but not least I want to thank my family for their love and support. It was not easy to continue this thesis while doing full-time work, but they stayed patient and helped me to finish.

CHAPTER 1

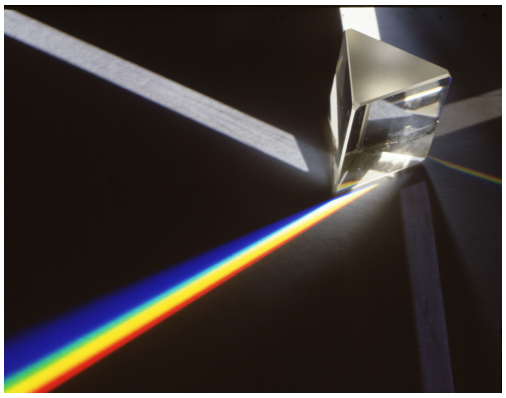
INTRODUCTION

One of the long-standing goals in Computer Graphics is the generation of images from virtual scenes that are indistinguishable from photographs of a corresponding, real scene. Photo-realistic image generation has a large impact in many areas. As an example, decisions in the design process of new products might be based solely on rendered images thus omitting the costly manufacturing of prototypes. Another example is advertisement, where computer generated images are cheaper to produce and offer greater artistic freedom compared to traditionally produced photographs or movies. Furthermore, teleshopping applications can offer a greater level of interaction between a virtual representation of the product and the customer than traditional photographs.

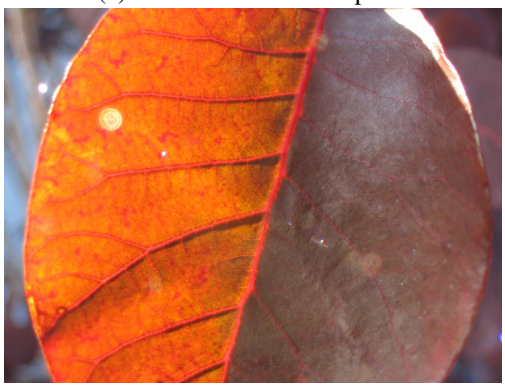
Generating virtual photographs from computer-stored virtual scenes requires the simulation of light generation and transport inside of the scene as well as the detection of light by a virtual camera. For this reason, photo-realistic rendering touches different areas of Computer Graphics, ranging from rendering algorithms to the representation and acquisition of geometry, light sources and materials.

This thesis focuses on the area of virtual materials. In the context of rendering, the important properties of a material are determined by its interaction with light. This includes in most cases reflection, internal scattering and transmission of light. Figure 1.1 shows some examples for light-material interaction. Over the time, a tremendous amount of work was spent on virtual materials and resulted in a rich variety of representations for surface reflectance or transmittance. Traditionally, analytical functions have been used to describe the reflectance or transmittance. Here, Computer Graphics has drawn a lot of knowledge from other areas like physics where light-matter interaction has been studied for a long time. The interaction is modeled at a microscopic level and the integral over all microscopic effects is described by an analytical function. Despite the fact that some of the analytical reflectance models describe certain material classes faithfully, their usage for photo-realistic rendering is limited. This is because it is often a non-trivial task to choose the right model and to find parameters for the chosen

(a) Metallic automotive paint



(b) A prism refracts light depending on the wavelength (photograph by Adam Hart-Davis)



(c) A leaf showing translucent behavior (photograph by Eric Skiff)

(d) Light scattering by fog (photograph by Sam Javanrouh)

Figure 1.1: Examples of light interacting with matter

model in a way that it resembles a given real material. When this process is done by hand it involves a lot of expert knowledge and trial iterations. This material design process gets easily infeasible with increasing surface complexity.

As a consequence, the measurement of surface reflectance or transmittance has received more and more attention in the area of photo-realistic rendering during the last two decades. Here, real samples are exposed to known lighting and the scattered light is captured by certain optical sensors. Afterwards, the acquired data can be used to automatically determine model parameters or, in case of so called data-driven techniques, be used for rendering directly. Due to the generality of data-driven techniques, they can represent the scattering behavior of a vast amount of real-world surfaces. Measured optical material properties combined with data-driven techniques offer a much higher visual fidelity achieved with less specialist and time effort. For these reasons, optical material measurements are one of the key techniques to make photo-realistic rendering a broadly available tool.

1.1 Reflectance Measurements

By now, a lot of measurement setups have been proposed that acquire reflectance or transmittance data of real surfaces. All rely on the principle described above: the material sample is lit by a known light source and scattered light is captured using a sensor. Differences between setups arise from the assumptions about the measured sample. While a diffusely reflecting, homogeneous surface like a piece of paper just needs one such measurement, most materials scatter light into preferred directions, also depending on the direction of incident light. This makes it necessary to place light source and light sensor in different positions with respect to the material sample. Here, one can distinguish between serially and parallel operating measurement setups. Serial setups just have one light source and sensor and need to relocate both with respect to the sample using a mechanical system. Parallel setups utilize multiple sensors and/or light sources to reduce capture time. Parallelizing the capture of scattered light is trivial, but parallel illumination requires to invert the used illumination patterns afterwards which might only be possible using additional assumptions about the material.

1.2 Spectral Reflectance Measurements

At the beginning of reflectance measurements, only homogeneous, structure-less materials have been considered and the scattered light was measured using classical spectrophotometers. However, with the advent of digital cameras based on CCD and CMOS imaging sensors, the efficiency of reflectance measurement setups has grown tremendously because cameras allow to take millions of reflectance samples at once. Unfortunately, nearly all such devices were built using cameras having only three color channels, most commonly red, green and blue (RGB) matching the established display technology. Such trichromatic cameras are cheap due to their large market and easy to use, leading to highly cost and time efficient reflectance measurement setups.

Unfortunately, the use of RGB or similar trichromatic reflectance data leads to wrong colors of the acquired samples when used in a rendering system under an illumination differing from the light source in the measurement setup. The reason for this is an effect called *metamerism*, which has been extensively studied before (e.g. [WS00], [HP11], [FANF06]). Metamerism is caused by the fact that an infinite number of spectral distributions of light map onto the same trichromatic values. So having the trichromatic values alone, the original spectral distribution cannot be easily inferred from it. In cases where trichromatic images are just used to be shown to an observer, this does not cause major problems. However, as soon as an interaction of light with matter has to be simulated, the correct spectral

distributions of incoming light and surface reflectance are required to calculate the spectral distribution of the scattered light and therefore the correct color.

While an enhancement from RGB to spectral is simple for rendering algorithms and many analytical scattering models, the fully spectral measurement of reflectance and also of complex lighting is a tough task. One reason for the difficulties is a pure physical one: incoming light needs to be filtered by narrow-band optical filters or split up by dispersing elements and thus less photons reach the sensor elements. Therefore, sensors with better quantum efficiency and less noise are required and the time to take one sample increases. Of course, there exist techniques using broadband filters combined with spectral reconstruction. These filter pass a larger amount of light through, but they require complex filter changing mechanisms that have to be integrated in front of light sources or sensors. These systems - typically huge filter wheels - are quite costly and bulky. Furthermore, the spectral reconstruction from broadband filters does not support high frequent features in the captured spectra, often.

These drawbacks make spectral reflectance capture an expensive and tedious task which has not reached the same practicability as trichromatic reflectance capture by now.

1.3 Main Contributions

The central contribution of this thesis is a method to allow for faster, simpler and cheaper measurement of spectral reflectance of real-world samples which allows for materials with anisotropic reflectance and nearly arbitrary complex surface structure. The main idea of the method is the re-utilization of matured RGB capture technology adding as few dedicated spectral measurement samples as possible. The full spectral reflectance should be reconstructed by a suitable algorithm from this data. The technique is also directly usable in other related areas, e.g. capturing of radiance maps for image-based illumination.

We achieve our goal in three steps and therefore structure this thesis accordingly:

1. **Ground-truth Data** - In a first step we have built a device to capture the reflectance of real material samples in a classical and non-optimized manner. The acquired data can be considered a ground-truth and it allows to simulate more sophisticated capture methods and to judge the reconstruction quality they achieve. Using this setup we have captured a small database of bi-angular, spatially varying, multi-spectral reflectance data-sets which also allows other researchers to evaluate novel methods for reflectance capture. Part I of this thesis contains detailed descriptions of the proposed setup and

its calibration and an evaluation concerning the accuracy of the acquired data.

2. **Spectral Reconstruction** - The central idea of our work is to acquire dense RGB reflectance data and as few spectral data as possible and to reconstruct the dense spectral information afterwards. Therefore, the second step of our work was to investigate a novel reconstruction method for spectral information using dense RGB and sparse spectral data. Our method is based on the minimization of an energy function that measures conformance of the unknown spectral images with the RGB and spectral measurements and with a novel prior on spectral images. In Part II this method and an evaluation of the reconstruction results on various datasets, especially in comparison with previous techniques, are presented.

3. **Practical Application** - To finally fulfill our main goal of efficient spectral reflectance measurement, an existing reflectance acquisition setup using RGB cameras has been extended with dedicated spectral hardware. In Part III we present methodology to calibrate such a setup and to acquire spectral data with minimal effort. For this, we will first introduce a novel method for the spectral calibration of RGB cameras and novel kind of radiometric calibration for cameras when operated in a range with low signal-to-noise ratio. Afterwards we describe the novel capture process and a first evaluation of reconstruction results based on the acquired data.

Moreover, the necessary background for this thesis is given directly afterwards this introduction in Chapter 2. This covers basics about light and colors as well as an introduction to photo-realistic rendering. Basic knowledge about measurement of light is presented as well as prior work on reflectance and transmittance capture. The thesis in concluded in Chapter 11 with a summary, conclusions and an outlook to possible future work.

As usual in the field of Computer Graphics, most of the presented, novel methodology has been previously published on several international conferences ([RSK10], [RK10], [RZK11], [MWRK17]).

CHAPTER 2

BACKGROUND AND PREVIOUS WORK

This thesis is mainly about light - measuring light and simulating the propagation of light. For this reason a background and previous work on all related topics is given. While in the first sections quite general background is given, the later sections continue to get more specifically related to the main topic of this thesis, namely reflectance measurement.

2.1 Light

In general, light is electromagnetic radiation (EMR) which is a type of energy exchanged by charged particles. Since an in-depth explanation on all details of the nature of EMR would clearly go beyond the scope of this thesis, a high-level introduction to the physical background is given. The further interested reader is referred to physics books like [NEBS04].

A lot of theories about the nature of EMR have been proposed throughout the history. While some physicists like Newton believed in a particle nature of light, others like Huygens and Descartes thought of it as a wave. After a lot of experiments on refraction, interference, diffraction and birefringence and the publication of the wave-based theory of Maxwell [Max63], which could explain all these effects, the wave theory was commonly accepted. In wave optics EMR is described by an electric and a magnetic field component. These two oscillate always in phase and perpendicular to each other and to the direction of energy propagation. Figure 2.1 shows a schematic view of electromagnetic radiation according to Maxwell.

However, in the early 20th century, experiments revealed the photoelectric effect or certain photo-chemical reactions which could not be explained by the intensity of light alone, but were dependent on the wavelength of light instead. Furthermore, other physicists, most notably Max Planck, worked on an explanation of the spectrum of thermal radiation from black bodies and found the EMR to be quantized [Pla01]. Albert Einstein then formalized the theory that light consists

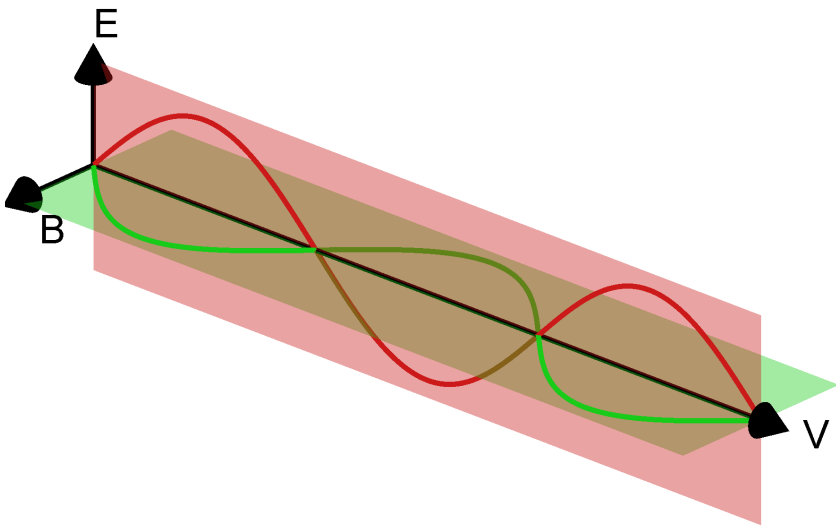


Figure 2.1: Schematic view of electromagnetic radiation according to wave-theory: It has both an oscillating electric field E and a magnetic field B being perpendicular to each other and to the direction of propagation V and being in phase.

of small particles, called *photons*. A photon has a specific energy Q_e determined by the wave frequency or the wavelength:

$$Q_e = h\nu = \frac{hc}{\lambda} \tag{2.1}$$

Here, ν is the frequency of light and λ the wavelength. $h = 6.62606957 \times 10^{-34} Js$ denotes the Planck constant and $c = 299,792,458 \frac{m}{s}$ the light speed in vacuum, which are both universal constants. Using the photon-theory, the effects in question could all be explained. Nowadays, it is commonly accepted that EMR exhibits the *wave-particle duality*, meaning that it has both properties of a wave and of a particle.

Our natural detector for EMR is the human eye (see also Section 2.3.1). It turns photons in approx. the range from 380 nm to 780 nm wavelength into signals interpreted by our brain where a color sensation is provoked. This wavelength range is called *visible spectrum* and EMR in this range is simply called *light*. However, there is a large range of other wavelengths / frequencies for which detectors and natural or man-made emitters exist and which also have various different applications in technology and science. Figure 2.2 gives an overview about the full range of EMR.

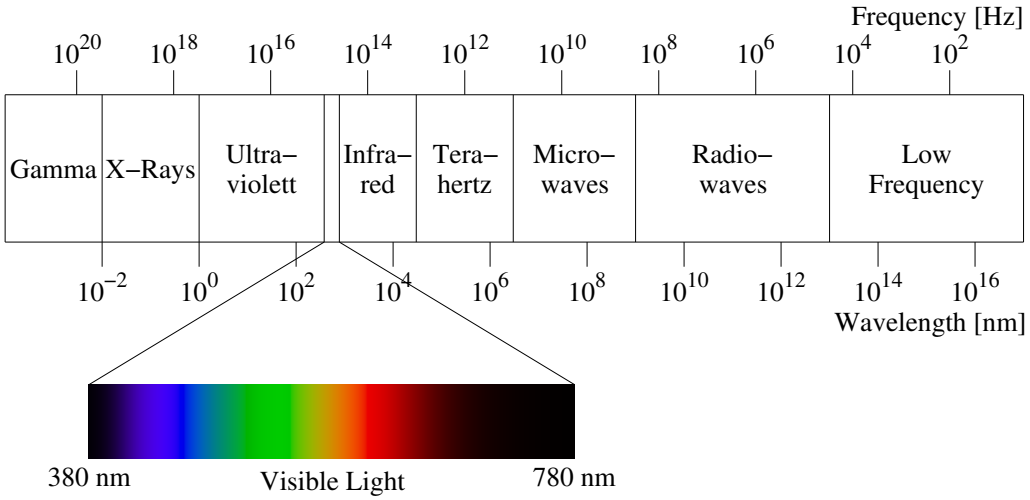


Figure 2.2: Overview of the whole spectrum of electromagnetic radiation. The (human) visible range covers a small portion only.

Equation 2.1 implies that a single beam of light or a single photon is *monochromatic* which was first noticed by Isaac Newton [New72] when experimenting with prisms. Newton concluded that all other colors including white have to be caused by a heterogeneous mixture of photons of different energies. One such mixture is called *spectral power distribution* or simplified also *spectrum*. We will use the word *spectrum* in this sense throughout this thesis.

2.1.1 Radiometry

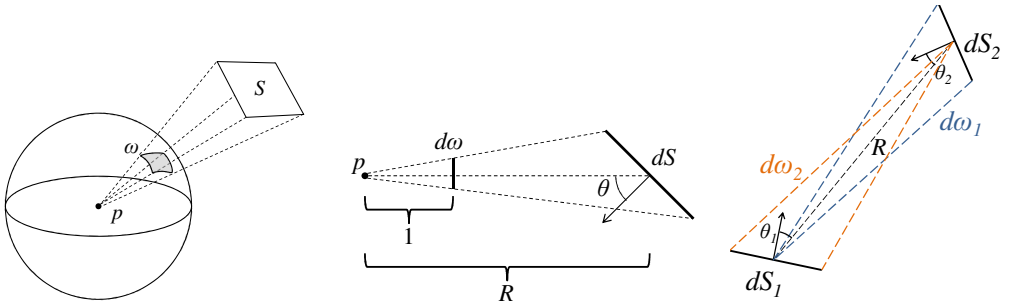
When propagation of light in space and interaction of light with matter is considered, physical quantities beyond the radiant energy Q_e have to be defined.

Radiant flux or *radiant power* Φ_e is radiant energy per time: $\Phi_e = \frac{dQ_e}{dt}$. The unit of radiant flux is therefore Watt (W).

Irradiance is the radiant power incident on a surface of unit area:

$$E_e = \frac{d\Phi_e}{dA} \tag{2.2}$$

Irradiance has a unit of $\frac{W}{m^2}$. When light is emitted from a surface one speaks of *radiant emittance*, *radiant exitance*, or *radiosity*.



(a) The solid angle is obtained by projecting the surface S onto the unit sphere around p and measuring the area of the projection.

(b) Computing $d\omega$ from dS and the distance R.

(c) Considering radiation transfer from dS_1 to dS_2.

Figure 2.3: Solid angle and radiation transfer

Solid angle When considering the transfer of radiant energy between surfaces, the *solid angle* ω has to be defined. It is a generalization of a two dimensional angle to three dimensions. The solid angle of a surface S with respect to a point p is the area of the projection of S onto the unit sphere around p. Despite the fact that the solid angle is physically and mathematically unit-less, it is often denoted to have the unit *steradian* (sr) for clarity. The solid angle of any surface completely enclosing p is $4\pi sr$ since the projection onto the unit sphere covers it completely. Figure 2.3(a) clarifies the concept of the solid angle. The differential solid angle $d\omega$ of a differential surface patch dS being in distance R from p is computed as:

$$d\omega = \frac{dScos(\theta)}{R^2} \tag{2.3}$$

Here, θ is the angle between the surface normal and the direction to p (see also Figure 2.3(b)).

Radiant intensity I_e or often simply called *intensity* is the radiant power per unit solid angle:

$$I_e = \frac{d\Phi_e}{d\omega} \tag{2.4}$$

The unit of intensity is $\frac{W}{sr}$. Care should be taken as the term *intensity* often refers to irradiance outside of the field of radiometry.

Radiometric Quantity	Unit	Photometric Quantity	Unit
Radiant energy	Joule [J]	Luminous energy	lumen second $lm \cdot s$
Radiant power	Watt [W]	Luminous flux	lumen [lm]
Irradiance	$\frac{W}{m^2}$	Illuminance	$\frac{lm}{m^2}$
Radiant intensity	$\frac{W}{sr}$	Luminous intensity	candela [cd]
Radiance	$\frac{W}{m^2 sr}$	Luminance	$\frac{cd}{m^2}$

Table 2.1: Overview of radiometric and photometric quantities and their units.

Radiance is power per unit solid angle per unit projected source area:

$$L_e = \frac{d^2\Phi_e}{d\omega dS\cos(\theta)} \tag{2.5}$$

with θ being the angle between the solid angle $d\omega$ and the surface normal of the surface patch dS emitting or receiving the radiance. $dS\cos(\theta)$ is therefore the projected differential surface. The unit of radiance is $\frac{W}{m^2 sr}$. When calculating irradiance from radiance one has to integrate over the whole hemisphere above dS:

$$E = \int_\Omega Lcos(\theta)d\omega \tag{2.6}$$

where Ω denotes the whole hemisphere.

2.1.2 Photometry

In photometry, light energy is measured with respect to the spectral sensitivity of the human eye, meaning that radiant power at a certain wavelength is weighted by the visual sensitivity function. There is a counterpart for all radiometric quantities in photometry. Table 2.1 contains an overview of the photometric quantities related to radiometric ones.

2.2 Light Transport

To generate images from virtual scenes, the propagation of light in the scene has to be simulated. Here, it is important to describe the transport of light between surfaces mathematically. This section explains the basics of light transport simulation and how they are implemented in nowadays commonly used rendering system.

Mostly, in computer graphics, only geometrical optics are used by now. That means that "wave effects" like interference or diffraction, polarization or effects

dependent on energy quantization are neglected in most rendering systems and light is represented in form of rays only. Though some research work like [OKG+10], [KFY+10], [TG17] deals with the rendering of single wave-based phenomena, there is currently no generic solution available. For this reason, we will restrict ourselves to ray optics in this section.

The first thing we have to consider is a radiation transfer from a differential surface element dS_1 to another element dS_2 with distance R (see Figure 2.3(c)). Seen from dS_1, dS_2 covers the solid angle $d\omega_2$ and vice versa.

When we assume that dS_1 and dS_2 are in vacuum, the radiant power $\Phi_{1\to 2}$ being emitted from element 1 towards element 2 has to be conserved during the transfer. Then it holds for the received radiant power $\Phi_{2\leftarrow 1}$ at dS_2:

$$\Phi_{1\to 2} = \Phi_{2\leftarrow 1} = \Phi \tag{2.7}$$

When inserting the definition of radiance from Equation (2.5) we yield for the radiance emitted from dS_1:

$$L_1 = \frac{d^2\Phi}{d\omega_2 dS_1 \cos(\theta_1)} \tag{2.8}$$

and for the radiance received at dS_2:

$$L_2 = \frac{d^2\Phi}{d\omega_1 dS_2 \cos(\theta_2)} \tag{2.9}$$

With definition of the solid angle ω from Equation (2.3) we get:

$$\begin{aligned} d\omega_1 &= \frac{dS_1 cos(\theta_1)}{R^2} \\ d\omega_2 &= \frac{dS_2 cos(\theta_2)}{R^2} \end{aligned} \tag{2.10}$$

Inserting the solid angle yields:

$$\begin{aligned} L_1 &= \frac{d^2\Phi R^2}{dS_1 \cos(\theta_1) dS_2 \cos(\theta_2)} \\ L_2 &= \frac{d^2\Phi R^2}{dS_2 \cos(\theta_2) dS_1 \cos(\theta_1)} \end{aligned} \tag{2.11}$$

We conclude therefore that:

$$L_1 = L_2 \tag{2.12}$$

and follow, that radiance is constant along rays. It is therefore an optimal choice to store this quantity for each ray of light traced through a virtual scene.

Now, we will consider the radiance that is reflected from an opaque surface into a specific direction. This is described by the so-called *reflection equation*:

$$L_r(x, \omega_o) = \int_\Omega \rho(x, \omega_i, \omega_o) L_i(x, \omega_i)(n \cdot \omega_i) d\omega_i \tag{2.13}$$

Here, L_r is the reflected radiance at point x into solid angle ω_o. Ω is the hemisphere at x in direction of the surface normal n. L_i is the incident radiance from solid angle ω_i and ρ is the *Bidirectional Reflectance-Distribution Function BRDF* as introduced by Nicodemus et al. [NRH$^+$77]. More about the BRDF can be found in section 2.4.

When additionally considering the light emitted by a surface, the *rendering equation* as proposed by Kajiya [Kaj86] can be formulated:

$$L_r(x, \omega_o) = L_e(x, \omega_o) + \int_\Omega \rho(x, \omega_i, \omega_o) L_i(x, \omega_i)(n \cdot \omega_i) d\omega_i \tag{2.14}$$

Here, L_e is added to the reflection equation and represents the radiance emitted at point x into direction ω_o.

It should be noted that the *rendering equation* is sufficient to describe light transport in scenes with opaque surfaces. Whenever transparent or translucent materials have to be considered, the more generic *radiative transfer equation* has to be considered. We refer the reader e.g. to the report by Novak et al. [NGHJ18].

2.2.1 Rendering

The calculation of photo-realistic images from virtual scenes requires to determine the incident irradiance at all sensor elements of a virtual sensor - mostly the pixels of a virtual camera. Considering Equation 2.2, this requires to determine the incident radiance L_i onto the sensor element from all possible solid angles. This, in turn, is equivalent to solving the rendering equation for all virtual scene points visible from each image sensor element. This way, all rendering techniques can be seen as solutions of Kajiya‘s rendering equation differing only in the additional assumptions made about the scene‘s geometry, camera, light sources or materials.

When considering the rendering equation and equation (2.12) we see that L_i is equal to a L_r at a different position x' visible from x in direction ω_i, revealing the recursive nature of the rendering equation. For this reason, it is seldom possible to solve this equation analytically. When substituting L'_r for L_i over and over again, we end up with a very high dimensional - in principle infinitely recursive - integral. When using simple numerical quadrature to solve this integral, a huge number of samples have to be evaluated to enable for a good sampling in all dimensions.

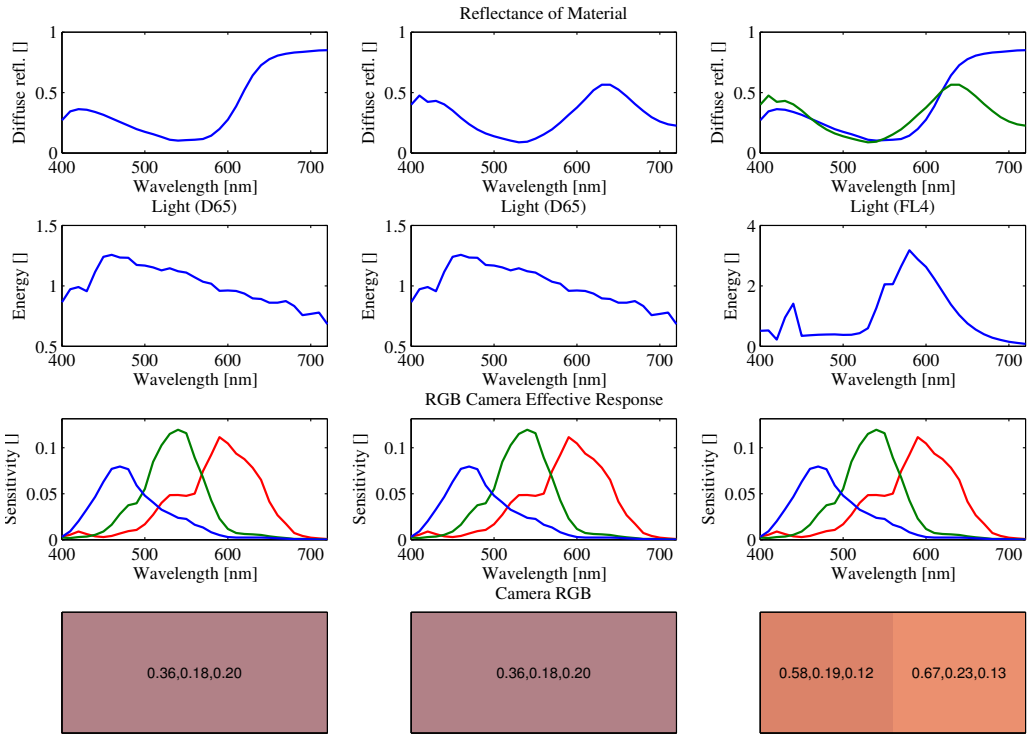


Figure 2.4: Metamerism and its effect on color accuracy in rendering: In the top row reflectance spectra of two different materials are shown which are lit by two different light sources having spectral power distributions plotted in the second row. The reflected light is then captured using a RGB sensor having the sensitivity plotted in the third row. The resulting RGB values are shown in the bottom row.
The two leftmost columns show that the same RGB values are observed for quite different material reflectances when lit by the same light source (D65). The third column shows that the same materials lit by a light source having a different spectral power distribution (CIE FL4) will then produce different RGB values when captured by the same sensor. This effect cannot be reproduced when relying on the RGB values captured under the D65 light source used in the two leftmost columns.

Instead of numerical quadrature, Kajiya proposed to use Monte Carlo integration (MCI) to solve the high-dimensional integral [Kaj86] and called this method *pathtracing*. MCI has the fundamental advantage, that the convergence speed with increasing number of samples does not depend on the dimension of the integral. Nowadays, all unbiased rendering systems, that are capable of providing photo realistic images from virtual scenes in a quite general way, are based on

MCI. However, a lot of other solutions to various simplifications of the rendering equation have been proposed over time. A thorough discussion would clearly go beyond the scope of this thesis. Therefore we refer the reader to text books like [PJH16] for a good overview on rendering techniques in general and on Monte-Carlo based methods in detail.

2.2.2 Color in Rendering

All equations above treat radiance as a scalar value neglecting the wavelength dependence. When color is introduced to the computation of light propagation, a spectral distribution of radiance has to be stored for each virtual light ray. In theory, one needs to store *spectral radiance*, which is radiance per frequency or wavelength and therefore has unit $\frac{W}{m^2\,Hz\,sr}$ or $\frac{W}{m^2\,nm\,sr}$. Of course, an infinitesimal sampling of the spectral domain is not possible and in practice a representation using a finite sampling has to be used.

First rendering systems stored and calculated with only three color bands corresponding to red, green and blue (RGB) light. This was inspired by display technology, which in turn was adapted to human vision (see also Section 2.3.1). While an RGB representation is sufficient for display purposes, it is not sufficient to calculate the interaction of light and matter in an exact way. As it can be seen in the reflection equation (equation (2.13)) the product of incoming light L_i and material reflectance ρ has to be calculated per wavelength band. When using three wavelength bands it is no longer possible to infer the real spectrum with sufficient accuracy. The reason for this is an effect called *metamerism*. Metamerism means that an infinite number of spectra is mapped to the same RGB triple. All spectra corresponding to the same triple are called *metamers*. In practice the product between L_i and ρ is therefore often simply calculated element-wise between the two RGB triples. Since in most RGB measurement devices, the three color filters highly overlap, additional bias is introduced, as this overlap is totally neglected by element-wise multiplication. Figure 2.4 shows examples of metamers and the errors introduced by calculating products using RGB values.

For applications where exact color reproduction is required, it is therefore inevitable to compute all interactions between light and materials with fine spectral resolution. This is implemented in modern commercial rendering systems like Maxwell[1] or FryRender[2] or freely available systems like LuxRender [3] or Mitsuba [4]. To fully utilize the accuracy of those spectral rendering engines, illumination L_e and material BRDF ρ have to be known with sufficient spectral resolution.

[1] http://www.maxwellrender.com/
[2] http://www.randomcontrol.com/fryrender
[3] http://www.luxrender.net
[4] http://www.mitsuba.org

Rendering which uses an unbiased algorithm and uses full spectral data is sometimes called *predictive rendering* since an exact prediction about object's appearance is possible. This kind of rendering systems can be used for all applications discussed in the introduction.

2.3 Measuring Light

In the context of this thesis, the measurement of amplitude and wavelength of light is an important issue.

Different kinds of interaction between light and matter can be utilized to transform incident light to other measurable signals. The energy discharged by a photon can cause the re-configuration of a molecule. Furthermore, electrons might be emitted from a piece of matter, especially metals, when they absorb enough energy from a photon. This is called the *photoelectric effect* and was first observed by Heinrich Hertz [Her87]. The so called *inner photoelectric effect* takes place in semiconductors, where an electron is raised from the valence band to the conduction band being free to move around in the semiconductor afterwards.

2.3.1 Devices

The effects described above are used in many different devices including human and animals eyes.

In the retina of human or animal eyes, so called *cone* and *rod cells* transform incoming light into electrical signals that are transmitted to the visual cortex. Inside of these cells the pigments *rhodopsin* and *photopsin* are the primary detectors of light. They consists of the protein *opsin* bound to a molecule called *retinal* which is closely related to vitamin A. In inactivated state the retinal is in the 11-cis form. When absorbing a photon, the retinal changes to all-trans state effectively re-configuring the whole large molecule. In a process called *visual phototransduction* further chemical reactions are triggered until the optical signal is finally transformed into an electrical signal. The all-trans-retinal is afterwards transformed back to 11-cis state. This is called the *visual cycle*.

In dependence of the kind of opsin the retinal is bound to, the absorbance characteristics of the molecule with respect to wavelengths is changed. Human rod cells contain rhodopsin for vision under low-light conditions having a peak absorbance at about 500nm. In the cone cells, which are responsible for color vision, different photopsins are bound to retinal leading (in case of a normal human) to three different absorbance characteristics with peak absorbances of about 420, 530 and 560nm corresponding to color stimuli of blue, green and red. This three dimensionality of human color vision was noticed early by Grassmann [Gra53].

The first man-made light detector was photographic film. It consists of silver halide crystals embedded in a gelatin film. When light strikes the surface of the crystals, it raises electrons from the valence to the conduction band. The electrons travel to so called *sensitivity specks*, points with crystal defects. Here, they recombine with silver ions to form deposits of metallic silver. This so called *latent image* is then made visible during development. The developer turns those entire crystals into metallic silver, which have enough silver already deposited on their sensitivity specks by incident light.

The photoelectric effect was utilized in video camera tubes. Here, the incoming image is focused on a metal plate causing an emission of electrons from the plate in relation to the intensity of incoming light. The outgoing electron image beam is focused and used to generate an electric signal in different ways.

Modern electronic imaging devices are all based on the inner photoelectric effect. Mostly the semiconductor silicon is used, which has a band gap of 1.12eV (at room temperature) corresponding to light of 1107 nm wavelength. This means, that incoming light of shorter wavelength (thus including the full visible range) will raise electrons from the valence to the conduction band making them freely movable inside of the silicon. By applying a voltage or by doping, the free electrons and holes can be moved, stored and read out. This process was used in photodiodes to convert light to voltages. Photodiodes and linear arrays of photodiodes have been used in *spectrophotometers* in conjunction with a diffractive element like a prism or grating to measure single spectral power distributions.

However, it remained unclear how an imaging device could be constructed based on photodiodes or other photoactive silicon elements. The solution was the *charge-coupled device* (CCD) invented by Boyle and Smith [BS70] while working on new types of computer memory. The CCD is an array of photoactive capacitors that accumulate charge (electrons) when exposed to light. The array allows to shift the charge from capacitor to capacitor in each row by applying alternating shift voltages. At the end of each row the charge of the last pixel can be converted to a voltage by floating diffusion and digitized by an analog-to-digital converter.

With the advent of CMOS technology more electrical circuits could be integrated into the pixels itself. This led to the development of CMOS *active pixel sensors* (APS) where floating diffusion and amplification is integrated into each pixel. Due to their much lower power consumption, CMOS sensors are nowadays often used in cell phones or similar mobile devices. More about CCD and CMOS imagers can be found in [HL07]. Since other measurement devices for light in the visible spectrum do no longer have large practical relevance, especially in the field of reflectance measurement, we will further on just consider CCD and CMOS imagers and respective spectrophotometers.

2.3.2 Color and Spectral Imaging

CCD and CMOS pixels are sensitive to a broad range of electromagnetic radiation. To acquire color images, optical filters or dispersive elements in front of the sensor have to be used to restrict the incoming light to certain wavelength bands. Here, a lot of different spectral resolutions and technical concepts exist.

Trichromatic cameras are the most commonly used ones. Since our human eye also has three different color sensors, trichromatic cameras like those that resemble the red, green and blue (RGB) retinal receptors are sufficient for color imaging with good color reproduction when displaying or printing the images. A perfect match can be delivered by those cameras satisfying the Luther condition [Lut27] meaning that the three effective spectral responses of the trichromatic camera are all different linear combinations of the human cone responses.

The desired filtering of incoming light by wavelength can be achieved in different ways. In trichromatic cameras so called *color filter arrays* (CFA) are most commonly used, meaning that the individual pixels of a CCD or CMOS imager chip are covered with small colored filters. Among these the Bayer filter mosaic is most frequently applied. It contains 50% green, 25% red and 25% blue pixels with the green pixels being diagonally distributed. The greater number of green pixels reflects the larger sensitivity of the human eye to green light. Figure 2.5 shows the layout of the Bayer pattern. To obtain a full RGB triple at each pixel location, interpolation techniques are employed. This process is called *demosaicing*. A thorough review on this topic was made by Li et al. [LGZ08].

Another option is to use a combination of beam splitters to project the same image onto three image sensors and to place an optical filter before each of the sensors. This has the advantage that the resolution is higher and no artifacts due to demosaicing can occur. The main disadvantage is the higher cost. Nowadays, this three-CCD layout is mainly chosen in better video cameras or in specialized industrial cameras.

When image display or printing is not the primary goal, or if more complex color operations shall be performed on images, three broad and overlapping color bands are often no longer sufficient. In the context of this thesis, especially photo-realistic image generation may suffer from bad color reproduction when RGB or similar trichromatic representations are used (see also Section 2.2.2). Here, acquisition systems with higher spectral resolution come into play. The acquisition process is then called *spectral imaging* or *imaging spectroscopy*. Here, a whole *spectral data cube* formed by the two spatial axes and the spectral axis has to be captured. One speaks of a *multi-spectral* system if the number of spectral bands is above 5 or 6 but below 100. All systems having a larger number of bands are called *hyper-spectral*. Since this nomenclature is a bit fuzzy, numbers being slightly different might be found in other literature. An older but thorough

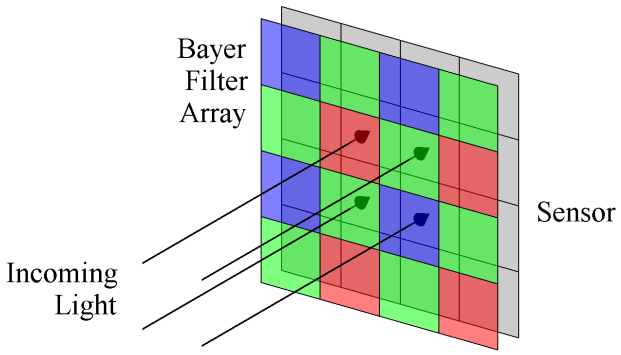
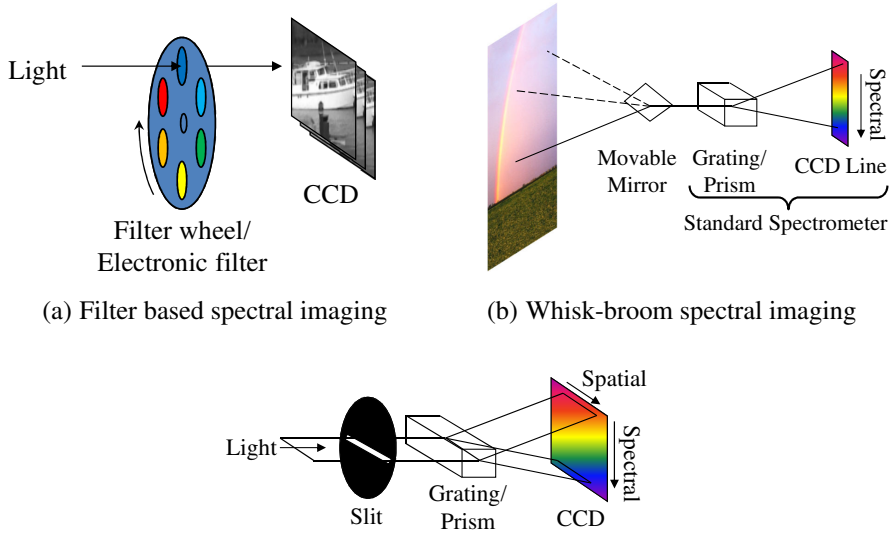


Figure 2.5: The bayer filter mosaic: the single pixels of the sensor (e.g. a CCD chip) are covered with red, blue and green filter elements in a fixed pattern. From the values of the single pixels a full RGB image can be recovered using a *demosaicing* technique.

overview about spectral imaging techniques can be found in [HBG+00].

Many techniques unroll the acquisition of the cube in time, assuming a static scene. Here, filter-based, whisk-broom and push-broom devices can be used (see Figure 2.6). In filter-based spectral imaging systems the spectral dimension of the data cube is serialized. A changeable optical filter is added in front of a monochrome image sensor to capture differently filtered light. Here, one has to distinguish between the type and the spectral transmittance of the filter. The filters can be interchanged using a mechanical system, often a wheel, or a so called *tunable filter* can be used. Tunable filters vary their spectral transmittance by a non-mechanical or micro-mechanical method and therefore allow for extremely fast change in comparison to mechanical filter wheels. Moreover, they are mostly lightweight. However, they only support transmittance with certain band-pass and also have often bad peak transmittance. A good overview about spectral imaging using tunable filters was made by Gat [Gat00]. Optical filters mountable in a mechanical filter wheel offer more freedom in the filter choice and therefore allow to use filter glasses with better optical properties. They also allow to use filters with broader, alternatively overlapping spectral transmittance. Then, the spectra have to be reconstructed from the imager‘s response after the acquisition. Examples for the use of mechanical filter wheels with overlapping filters are numerous in the color imaging literature. We would like to point to work by Hardeberg et al. [HSB+99b] in which a good overview on spectral reconstruction operators is given.

In whisk-broom devices (e.g. [MSI94]), both spatial dimensions are unrolled in time by scanning the image pixel-wise using a spectrophotometer. The incoming light from different viewing directions is transmitted to the spectrophotometer



(a) Filter based spectral imaging

(b) Whisk-broom spectral imaging

(c) Push-broom spectral imaging

Figure 2.6: Sketches of different approaches to spectral imaging

by using e.g. a freely rotatable mirror. This scheme is used in systems where acquisition time plays a minor role compared to spectral accuracy. These can be found e.g. in special satellites scanning the earth's surface.

In push-broom devices, one of the spatial dimensions of the spectral data cube is acquired serially. The respective capture devices are capable of imaging one line of a scene with high spectral resolution. This is achieved by using a slit aperture followed by a diffractive element like a prism or grating. The incoming light is masked by the slit, dispersed by the prism and then measured using a CCD or CMOS chip, where one of the image dimensions corresponds to the position on the line measured and one to the wavelength. This way a rather high spectral resolution is possible, which cannot be achieved using filter based approaches. Unfortunately, the whole device has to be moved with respect to the scene in order to capture a 2D image. Such devices are therefore often used in scanner setups like [AFOR04] or at industrial assembly lines where the target is moved anyway.

Another category of spectral imaging devices captures in a *snapshot* i.e. non-serial manner. Here, different techniques exist as well. The most simple solution is to extend the Bayer pattern with more filters. As an example, Kidono et al. [KN07] proposed to replace one of the green filters by an infrared filter for night-vision purposes. Another option is to used a prism or grating as in push-broom devices and then a hole-mask to prevent the dispersed spectra to overlap on the

image sensor. This approach has been implemented by Du et al. [DTCL09], [CDT+11]. A mathematically more involved technique is *Computated Tomography Imaging Spectroscopy* (CTIS) introduced by Descour et al. [DD95] where the incoming light is split by a diffraction grating and produces multiple diffraction patterns on the image sensor. In recent work Habel et al. [HKW12] showed how to construct a low-cost CTIS camera. The diffraction patterns can be seen as multiple projections of the spectral data cube onto a plane. From these projections the data cube can be reconstructed by the inversion of the projection matrix. Unfortunately, the reconstruction has no unique solution and is also prone to noise artifacts.

Another snapshot technique is *coded aperture snapshot spectral imaging* (CASSI) (e.g. [GJB+07], [WPSB08], [WPSB09]). Here, the image is dispersed onto a coded aperture, typically a LCD element, where spatial and spectral information mixes. A random binary pattern is used to select parts from the mixed information. Afterwards the dispersion process is reversed and the image is captured using a CCD or CMOS imager. The spectral data cube is reconstructed from the mixed information using algorithms like GPSR [FNW07]. Like in the case of CTIS, the solution is not unique and the reconstruction requires strong regularization.

2.4 Materials

Whenever light travels through matter or reaches matter boundaries, interaction takes place. To describe the large variety of effects in detail is clearly beyond the scope of this thesis. The reader is referred to [Nas01] for a more or less complete overview.

The interaction takes place at an atomic or molecular level. Incoming photons of appropriate energy can be reflected or refracted at medium boundaries or they can be absorbed by an atom or molecule and their energy is transformed. It may either be transformed to atomic vibrations, heating the matter, or an electron might be raised to a higher energy level or may even ejected from the atoms or molecules electron shells (ionization). By the reversion of these processes photons are emitted or re-emitted from matter. Colliding atoms or molecules in heated matter emit part of the energy in form of photons. Electrons falling back to lower energy levels emit a photon of frequency corresponding to the energy difference of the levels.

In the context of photo-realistic image generation modeling the light-material interaction at atomic level is too detailed, especially when considering the fact that mostly geometrical optics are used. For this reason, different levels of abstraction have been defined. The next section will give an introduction to the various

functions that have been defined to describe light-matter interaction. The section afterwards will give a detailed overview on the measurement of those functions from real material samples.

2.4.1 Formalization of Light-Matter Interaction

When considering a solid, opaque material, it is convenient to simplify the interaction of light with an object made of that material by computing interactions at the surface of the object only. The scattering behavior of light at the surface is then called *surface reflectance*. Informally, one wants to know how much light entering the object at a specific position x_i from a direction ω_i, having a wavelength λ_i at time t_i exits the object at a different position x_o into a direction ω_o with a different wavelength λ_o at time t_o. Figure 2.7(a) shows a sketch of the geometry and quantities. This would lead to a function on a 12 dimensional parameter space:

$$\rho(x_i, \omega_i, \lambda_i, t_i, x_o, \omega_o, \lambda_o, t_o) \tag{2.15}$$

Effects often neglected in computer graphics are *phosphorescence* - the ability of a material to re-emit incident light at a later point in time - and *fluorescence* - the transport of radiant energy from a wavelength λ_i to a (longer) wavelength λ_o. When we eliminate t_i, t_o and set $\lambda = \lambda_i = \lambda_o$ we end up with a 9D function. When assuming a finite discretization of the wavelength domain, we can also omit λ from the function and end up with the so called *Bidirectional Scattering-Surface Reflectance-Distribution Function* (BSSRDF) which has been formalized by Nicodemus et al. [NRH+77]:

$$\rho(x_i, \omega_i, x_o, \omega_o) = \frac{dL_o(x_o, \omega_o)}{d\Phi_i(x_i, \omega_i)} = \frac{dL_o(x_o, \omega_o)}{dE_i(x_i, \omega_i)dA_i} = \frac{dL_o}{L_i(x_i, \omega_i)cos(\theta_i)d\omega_i dA_i} \tag{2.16}$$

Here, L_o is the exitant radiance, $d\Phi_i$ the incident radiant flux through surface patch dA_i around x_i, dE_i the corresponding irradiance, L_i the incident radiance and θ_i the angle between ω_i and the surface normal at x_i. The BSSRDF has unit $m^{-2}sr^{-1}$.

When integrating the BSSRDF into the rendering equation (2.14), one has to add an additional integral over the surface of the respective object:

$$L_o(x, \omega_o) = L_e(x, \omega_o) + \int_\Omega \int_A \rho(x_i, \omega_i, x, \omega_o) L_i(x_i, \omega_i) dA_i d\omega_i \tag{2.17}$$

When considering light propagation in non-solid matter like fluids, fog or gases, classical rendering based on the rendering equation and the BSSRDF or BRDF is no longer applicable. Here, a scattering function has to be defined for

every point in the volume. This is formalized as a so called *phase function* which is a probability density function:

$$p(\omega_i, \omega_o) = prob(\text{photon scattered into } \omega_o | \text{photon incident from } \omega_i) \quad (2.18)$$

Using phase functions in photo-realistic rendering requires to compute scattering events not only at material surfaces but at regular or randomly chosen intervals when tracing light rays through a volume. Since this thesis focuses on surface reflectance, the whole area of *volume rendering* is out of scope and we refer the interested reader to [PJH16].

Except the restriction to geometric optics and the omission of phosphorescence and fluorescence, the BSSRDF contains no assumptions about the material or the kind of interaction between light and matter and is thus able to represent an extremely huge class of materials. Unfortunately, it is still a very high-dimensional function (8D) and therefore unhandy in terms of measurement, storage and rendering. For certain special cases, further simplifications have been defined. First, if one considers only nearly opaque surfaces, subsurface scattering can be neglected. Therefore, $x_i = x_o = x$ and the reflectance function simplifies to:

$$\rho(x, \omega_i, \omega_o) = \frac{dL_o(x, \omega_o)}{dE_i(x, \omega_i)} = \frac{dL_o(x, \omega_o)}{L_i(x, \omega_i) cos(\theta_i) d\omega_i} \quad (2.19)$$

This function now has unit sr^{-1} and is often called *Spatially-Varying Bidirectional Reflectance-Distribution Function* (SV-BRDF). It can be directly used in the rendering equation without any modifications.

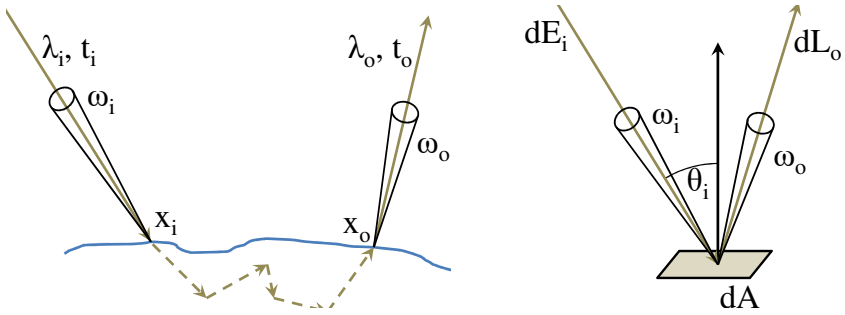
If homogeneous materials are considered, even the dependence on surface position x vanishes, ending up with the *Bidirectional Reflectance-Distribution Function* (BRDF) [NRH+77]:

$$\rho(\omega_i, \omega_o) = \frac{dL_o(\omega_o)}{dE_i(\omega_i)} = \frac{dL_o(\omega_o)}{L_i(\omega_i) cos(\theta_i) d\omega_i} \quad (2.20)$$

The BRDF is a four-dimensional function with unit sr^{-1} and a value range of $[0, \infty)$. Figure 2.7(b) clarifies the BRDF geometry.

The BRDF has some properties which are extremely useful in photo-realistic image generation:

- Superposition: Different light rays at a surface point do not affect each other. This makes is possible to integrate all contributions in the reflection equation (2.13).
- Reciprocity: $\rho(\omega_i, \omega_o) = \rho(\omega_o, \omega_i)$. The BRDF does not change its value when interchanging incoming and outgoing light direction. This is especially useful in rendering algorithms when tracing light rays in an "inverse" manner from the camera to the light sources.



(a) BSSRDF geometry (b) BRDF geometry

Figure 2.7: Geometry for the BSSRDF and BRDF

- Energy conservation: $\int_\Omega \rho(\omega_i, \omega_o) cos(\theta_i) d\omega_i \leq 1$ for all ω_o. No energy can be "generated" at a surface. For this reason the contribution of light paths scattering through a virtual scene decreases with the number of scattering events making rendering techniques like path-tracing possible, because the recursion of the rendering equation can be cut off at a certain maximum level or using a certain probability.

A further simplification is the assumption of so called *isotropy*. A material is called to be isotropic, when the value of its BRDF does not change when rotating the material around the surface normal. Otherwise the material is called *anisotropic*. Isotropic BRDFs are only three-dimensional since they can be parameterized over the two elevation angles θ_i, θ_o of ω_i and ω_o and the difference of the azimuth angles $\phi_i - \phi_o$.

Another way to represent surface reflectance is offered by the *Bidirectional Texture Function* (BTF) introduced by Dana et al. [DvGNK97],[DvGNK99]. A BTF is the extension of the *texture* concept as proposed by Blinn et al. [BN76] by bi-directionality. A BTF therefore consists of one texture map per pair (ω_i, ω_o) of incoming and outgoing light directions: $B(x, \omega_i, \omega_o)$. x is a point on a reference plane to which the real surface point is mapped in direction ω_o. Furthermore, it is assumed that the incident illumination comes from infinity and is therefore equal on all surface points on the material. This principle is illustrated in Figure 2.8. The BTF includes geometric information about the material's surface in its single textures. Furthermore, there is no additional assumption about the scattering process itself. Subsurface scattering, interreflections and shadowing are directly represented in the textures. For this reason, the BTF should not be confused with the SV-BRDF, which has the same parameter space but a rather different meaning. If we consider the properties of the single texels of a BTF, we see that, due to the inclusion of all these non-local effects, properties like energy conservation

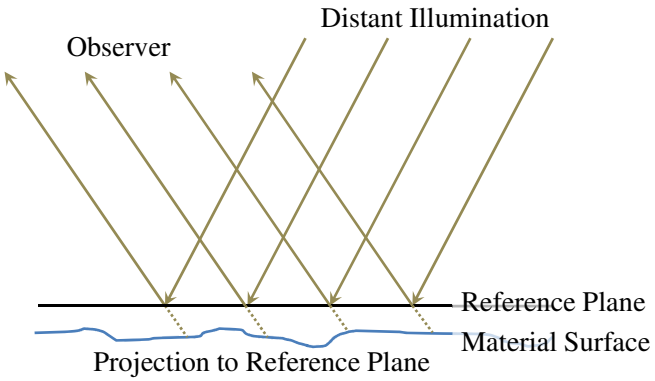


Figure 2.8: The BTF principle: The reflectance of an arbitrarily complex surface is parameterized on a reference plane. It is assumed that the light source is at infinite distance.

and reciprocity can be violated. Wong et al. therefore coined the term *Apparent BRDF* (aBRDF) [WHON97]. Due to its properties the BTF is thus well-suited for surfaces of high geometrical complexity which are difficult to model using traditional geometry representations and SV-BRDFs. However, it is mostly limited to materials being not too specular, as these would require a very fine resolution of the space of directions ω_i, ω_o and therefore a tremendous amount of texture images.

2.4.2 Representation of Reflectance Functions

Having introduced the conceptual functions, by which surface reflectance can be described, we will now consider the concrete representation of those functions. Here, two concepts exist:

1. representation by *analytical functions* and

2. *data-driven approaches*.

For some special material classes, the BRDF or SV-BRDF can be represented by simple analytical functions. One such example are perfect mirror surfaces, having the following BRDF:

$$\rho_{mirror}(\omega_i, \omega_o) = \frac{\delta([\phi_i - \phi_o]_{0..\pi} - \pi)\delta(\theta_i - \theta_o)}{cos(\theta_i)sin(\theta_i)} \tag{2.21}$$

where $\delta(x)$ is the Dirac delta function.

Another example are perfectly diffuse surfaces, equally scattering light into all directions:

$$\rho_{lambert}(\omega_i, \omega_o) = \frac{a}{\pi} \tag{2.22}$$

with a being the *albedo* of the surface in the range $0..1$. Diffuse surfaces are also called *Lambert-reflectors* after Johann H. Lambert.

Since most real world materials are neither ideal mirrors nor ideal diffusers, a lot of analytical BRDF models have been proposed to fill the gap in between. Some of them are purely phenomenological ones, meaning that they are not derived from an explicit model of the microscopic surface structure and reflectance. Among them the models from Phong [Pho75] and Blinn [Bli77] are most famous and still widely used. Another class of analytic BRDF models is derived by assuming a certain microscopic surface structure and optical parameters of the single surface elements. These so called *physically-based* BRDF models allow for higher accuracy for many material classes than the phenomenological ones. Most of these models are based on the assumption of *microfacets*, small, planar surface parts at microscopic level where the orientation of the facets with respect to the global surface normal is often described by a statistical function. The microfacets are assumed to have some rather simple reflectance behavior and the BRDF model function then approximates the reflectance of the overall surface consisting of the microfacets in a closed-form solution. Well known representatives of this kind of BRDF models are the Cook-Torrance model [CT82], where microfacets are assumed to be perfect smooth reflectors, and the Oren-Nayar model [ON94], where microfacets are assumed to be perfect diffusers. There is a large number of additional models proposed over the years that depend heavily on the Cook-Torrance model, replacing single parts of it with different functions or tables. A complete review of these models is out-of-scope of this thesis.

Data driven approaches follow a different paradigm. Here, the values of the reflectance function are stored in a high-dimensional table. Values in between the table‘s sampling points are interpolated. Especially the Bidirectional Texture Function is mostly stored in a data-driven manner since it is very cumbersome to represent the irregular structures inside of the aBRDFs using analytical reflectance models. Due to the fact that reflectance functions are high-dimensional and often a dense sampling is required for good approximation of real-world reflectance, the storage cost of the table easily exceeds practical limits. For this reason, compression techniques, e.g. based on statistical analysis, are used to reduce the storage size. For an overview of compression techniques for BTF the reader is referred to [FH09] or [Mül09].

In summary, the main advantage of data-driven approaches is their generality, as they make it possible to represent a huge range of materials with great accuracy. The main disadvantages are the large storage costs and the worse editability,

since there are no parameters having a intuitive physical meaning as with the analytical BRDF models. The advantages of analytical reflectance models are their compact storage cost and fast evaluation speed. The main drawback is their lack of accuracy for representing many real-world materials.

2.4.3 Application of Virtual Materials

As described above there is a large set of possible representations to chose from when assigning materials in virtual scenes. The fundamental problems are how to choose the correct function class, the correct representation and how to come up with the parameters for the chosen representation.

While the function class is solely determined by the kind of material, the other two choices still remain. Whether to choose an analytical reflectance model or a data-driven representation and how to obtain the necessary parameters or data heavily depends on the application. Of course, there are applications, where there is either no real material that has to be matched or there is no demand to match a given real material with high precision. Moreover, limited computational and memory resources may prevent the usage of data-driven approaches.

However, there is a lot of more demanding applications, where it is necessary to have a virtual representation of a material that enables for true photo-realistic image generation. Examples are virtual prototyping, advertisement, design or other applications requiring virtual product display like teleshopping.

Classically, virtual materials for demanding applications have been generated by manually choosing parameters for analytical reflectance models by hand. For this, a CG artist tries to match the look of the material from some photographs or a real sample on his table by creating a geometry with standard modeling tools and by trying around with the reflectance model parameters. Even with a lot of experience this is a tedious task. Moreover, the effort increases rapidly with surface complexity - imagine for example an artist who has to model every single thread of a fabric.

A possible solution to increase the productivity and quality in this setting is to capture the reflectance of real material samples and to represent the captured data in the desired representation. In case of analytic reflectance models this requires a model fitting process in which the model parameters are found by e.g. an energy minimization algorithm. In case of data-driven representation the process is much simpler since the acquired reflectance data can be used with an appropriate interpolation method directly. However, the huge amount of data might render a data reduction technique necessary.

In the following section we will deal with the measurement of surface reflectance, the central topic of this thesis.

2.5 Reflectance Measurement

Acquisition of surface reflectance requires the sampling of one of the functions defined in section 2.4.1. Let us consider the BRDF as an illustrative example. From equation 2.20 we see that the BRDF is $\frac{dL_o}{L_i cos(\theta_i) d\omega_i}$. To acquire one sample to approximate the BRDF for a certain pair of directions we have to measure the exitant radiance L_o in a certain solid angle ω_o while exposing the material to a defined radiance L_i from another solid angle ω_i. In practice this requires to place a light source of known power, (small) size and emission characteristics in direction ω_i and a light detector as introduced in section 2.3 in direction ω_o. It should be noted that we are not able to really acquire an ideal BRDF sample since real detectors or emitters having differential solid angle with respect to the sample do not exist, of course.

We will now first have a generic look at possible concepts to achieve this and will afterwards give a detailed overview about previous work in this area.

The number of measurement points, that have to be sampled, depends on the desired representation. If, for example, a homogeneous, diffusely reflecting material has to be measured, one sample to determine the albedo is sufficient (see equation 2.22). However, for more complex models a larger number of samples with different angles is required and for data-driven representations even a dense sampling of all parameter dimensions of the reflectance function is necessary.

When mounting light source, detector and material sample to a mechanical system that allows to place light source and detector in different directions with respect to the material sample, one speaks about a *gonioreflectometer*. Since gonioreflectometers need only one light source and one detector element they can be equipped with high quality components and are therefore often used to acquire data of high accuracy. However, since all direction pairs have to be sampled serially involving mechanical interactions, the measurement times are typically rather long. When also spatially varying reflectance should be captured, such a serial measurement process is getting highly impractical.

To overcome this problem, multiple detector elements can be utilized to capture the reflected radiance at several angles in parallel. Especially imaging sensors offer the opportunity to capture millions of reflectance samples in one picture. In case of spatially varying reflectance this could be samples of the surface of the material with similar light and view angles or in the case of homogeneous materials a larger range of incident and exitant light directions as long as the real material sample exhibits enough surface curvature.

Above methods still sample light directions serially and exploit only parallelism in view directions and/or spatial dimensions. Parallelizing capture of different light directions is not trivially possible. Only when targeting certain reflectance models or assuming that the reflectance function can be represented us-

ing some basis functions, parallelized light capture by utilizing extended or multiple light sources is possible. After the capture process the acquired data has to be processed to recover the reflected radiance per incoming light direction.

2.5.1 Practical considerations

Besides these fundamental principles a lot of practical issues have to be considered. The first problem arises from the high dynamic range of certain, especially specular materials. For them, many orders of magnitude lie between the reflected radiance for near specular compared to far off-specular angles. Since all light detectors have limited dynamic range (see Section 2.3.1), methods to extend dynamic range like exposure bracketing have to be applied.

Additional constraints are caused by the physical layout of the setup. For BTF capture or for gonioreflectometers one would like to have a perfectly directional light source as well as an orthographic sensor/camera projection. For this it would be advantageous to move light sources and sensors far away from the material sample and use a light source that is as small as possible. Of course, in nearly all practical scenarios the space for the setup is limited and therefore also the distance to the light sources and sensors. Moreover, larger distance and smaller light source size decrease the irradiance at the material sample and therefore cause longer exposure times at the sensor to record the reflected radiance.

There is also a trade-off in the light source itself: one would obviously want to have it as small as possible and at the same time as bright as possible to decrease exposure times and/or acquisition noise. Obviously both parameters are physically contradictory and a certain compromise has to be chosen.

When acquiring samples with a camera system, the aperture size of the optical system is another crucial parameter. On the one hand, when it is too large, the depth-of-focus will be too small to capture the sample sharply focused especially from shallow viewing angles. This could be circumvented by tilt-shift optics (Scheimpflug principle) which can be mechanically complicated. Additionally, with a large aperture the solid angle wrt the sample deviates more from the ideal case of a differential solid angle. This will lead to angular blurring of the captured reflectance which will be especially noticeable at highlights.

On the other hand, when the aperture is too small, the amount of light reaching the sensor is strongly reduced leading to longer exposure times or a higher noise level. Moreover one will run into the diffraction limit when choosing an aperture being to small.

Summarizing, all practical setups to acquire real-world reflectance have to make a lot of compromises and it is therefore often advantageous to optimize a setup for a special use-case.

2.5.2 Previous Work

After having introduced reflectance capture concepts we will now proceed to give an overview about real setups, that have been proposed by now.

Diffuse Reflectance: The most simple case of materials considering the reflectance measurement are those being homogeneous and (nearly) diffuse. As we can see in equation 2.22 only the spectral albedo has to be captured. The sample has to be illuminated in a defined manner and either a hemispherically integrated or a point-wise sampled spectrum of reflected light has to be taken. Different geometrical configurations have been defined and standardized. An overview on geometries and corresponding devices can be taken from [WH87].

Classical Gonioreflectometers: Classically, BRDFs are measured using gonioreflectometers using a spectrophotometer as light detector. Therefore, the reflectance is resolved with high spectral resolution. Examples for this approach are the setups by Proctor et al. [PB96] and Serrot et al. [SBBC98]. These setups just differ in the choice of mechanical system, light source and spectrophotometer they use.

Foo [Foo97], [LFTW06] present a setup that is restricted to isotropic reflectance. The light source can just be rotated around one axis going through the sample, thus lacking one degree of freedom. Otherwise, the setup is comparable to those above.

Multi-View BRDF Measurements: The serial nature of gonioreflectometers can be straightforwardly overcome by parallelizing the capture of reflected light.

An example is the LED setup of Ben-Ezra et al. [BEWW+08], which utilizes a hemisphere densely covered with LEDs both for illuminating the material sample as well as for measuring the reflected light in parallel.

Another, widely used technique is to employ CCD and CMOS imagers which can take millions of reflectance samples at once.

Ward [War92] use a CCD camera with a fish-eye lens to image the reflection of a sample in a hemispherical mirror. This way they capture the reflected light from the sample in parallel and with high resolution. By moving a light source inside of the hemispherical mirror, they provide different illumination directions.

Another possibility is to take images of a curved surface which is homogeneously covered by the material in question. Marschner et al. [MWL+99] take images of a sphere (or other curved surface) using a digital camera that can be rotated around the object with one degree of freedom while the light source is fixed with respect to the object. This way, isotropic BRDFs can be captured.

Matusik et al. [MPBM03a], [MPBM03b] also used a setup like this, but proposed to reduce the number of camera positions and to fill the holes using

the push-pull algorithm applied to the BRDF data in half-diff parameterization [Rus98].

Both Marschner and Matusik utilized standard RGB cameras in their setups therefore limiting the use of the acquired data to cases where no exact color reproduction is required (see also section 2.2.2). However, the extension to spectral measurement just requires the exchange of the camera system. Since only one camera is used, this can be done with reasonable effort. Such an extension was described by Kim et al. [KSKL10]. They utilize a HDR-CMOS camera (already introduced in [KKP+08]) and a tunable band-pass filter to capture images of painted spheres with 16 spectral bands.

When furthermore extending the approach of Kim to fluorescent materials, that re-radiate incoming light of short wavelength at a longer wavelength, band-pass filters in front of light source and camera have to be used. Such a setup was constructed by Hullin et al. [HHA+10]. To avoid sampling the whole bi-spectral domain for each camera position, they calculated a basis from a small set of this full bi-spectral slices and reconstructed the slices for the other camera positions by using a bi-spectral entry from the basis which matches sparsely sampled points from the respective camera position best.

Spatially-resolved Gonioreflectometers: When extending reflectance capture to spatially varying materials, digital imaging sensors are inevitable. Of course, whisk-broom approaches could be used, but with increasing spatial resolution such an approach renders infeasible.

First setups in this area also had a gonioreflectometer design, just exchanging the spectrophotometer by a color camera.

The capture setup constructed by Levoy and Hanrahan [LH96] is a first variant despite the fact that they wanted to capture whole light fields.

The first straightforward spatial gonioreflectometer was presented by Dana et al. [DvGNK97], [DvGNK99] to measure BTFs (see section 2.4.1). They utilized a robot arm, which brings the material sample in different orientations with respect to the light sources and a video camera. The video camera was moved manually to different positions. With this setup they captured 60 BTFs with 207 images per material sample and published the data in the CuReT database.

Later setups from McAllister et al. [McA02], Sattler et al. [SSK03] and Koudelka et al. [KMBK03] relied on the same principle but provided a better angular resolution, image quality and automated measurement procedures. Typical measurement times for one bi-directioanl sample are in the order of a few seconds. Due to the serial nature of the capture process, 10 or more hours are required to digitize one material sample even for a moderate angular sampling.

As for the image-based BRDF measurement devices, RGB cameras were uti-

lized limiting the use of the acquired data. Fortunately, only one camera is required, making an extension using a spectral imaging device possible with reasonable effort. Until now, there is only one such setup presented by Tsuchida et al. [TAN+05], [TSA+05]. They included a wheel of 16 spectral bandpass filters in front of a strong lightsource and used a sensitive, monochrome camera. With this setup they are able to measure multi-spectral, spatially-varying reflectance of samples up to 4x4 cm in size. For 6500 bi-directional samples they had to spend 13 hours of measurement time. However, only results for homogeneous materials were presented by now.

Multi-View Setups: For spatial gonioreflectometers, the long measurement times can only be circumvented when taking multiple samples in parallel.

One approach can be to use an optical system that projects reflected light from multiple directions onto a single image sensor. The setup from Han et al. [HP03] is an example. They placed a kaleidoscope on the material sample and took photographs from above the kaleidoscope. The resulting images contain multiple views of the material. Unfortunately, the resulting angular and spatial resolution is extremely low.

The more promising approach is to use multiple cameras mounted at different directions with respect to the sample.

Vertical arcs equipped with cameras and light sources were used by Matusik et al. [MPN+02]. The light arc and the sample could be turned around the vertical axis to allow for the scanning of the whole hemispherical reflectance.

The maximum level of view-parallelization is achieved by the device first described by Müller et al. [MBK05] and in extended form by Schwartz et al. [SWRK11], [SSW+14]. It consists of a hemispherical gantry densely covered with 151 consumer digital cameras facing the material sample. The flash light source of the cameras are used as light sources. Due to the maximal level of parallelization this device is able to capture a low-dynamic range BTF in about 30 minutes. The main drawback is the moderate quality of the cameras themselves, as they are only consumer grade.

Multi-Light Setups: Parallelizing the capture of reflected light is theoretically straightforward. Further optimization is only possible when the illumination is parallelized as well.

Schechner et al. proposed to use multiplexed illumination [SNB03]. Here, multiple light sources are switched on, increasing the brightness of the illumination and therefore reducing the required exposure time to capture the reflected light and reducing the dynamic range of the scene. To be able to recover the reflectance for a single light afterwards, images using multiple light patterns have

to be recorded. Devices based on this idea were presented by Wenger et al. [WGT+05] and Mukaigawa et al.[MSY07].

When assuming that the reflectance of the material can be represented using certain data-driven basis functions or analytical reflectance models, other light sources and acquisition processes are applicable. Linear light sources like tubes have been used by Gardner et al. [GTHD03] and Ren et al. [RWS+11]. The tube is moved across the material sample in one direction and a video is recorded by one camera. Gardner fitted a Ward BRDF model [War92] afterwards while Ren uses a basis of previously acquired BRDFs from which the BRDF of the single surface positions is reconstructed.

Nearly arbitrary, uncontrolled illuminations are allowed by the method from Dong et al. [DWT+10]. They also utilize a basis of BRDFs captured from the same material using a small handheld device in beforehand.

Ghosh et al. [GAHO07] assumed that the BRDF can be represented in a spherical basis. They utilized a projector to create illuminations corresponding to the positive and negative basis entries and measure the response over a large part of the hemisphere simultaneously using a digital camera. Furthermore, they can also fit analytical reflectance models to the acquired data.

It should be emphasized that all these multi-light methods except the light multiplexing ones are not capable of measuring surface reflectance in a reliable and generic manner as they all rely on certain assumptions on the material. Therefore, they have limited use in predictive rendering.

Subsurface Scattering: When considering materials like human skin, marble or plastics, scattering of light beneath the objects surface contributes significantly to the overall appearance of the object. Since it would be a very tedious task to densely sample a BSSRDF (equation (2.16)), one often assumes that the light being scattered subsurface leaves the material diffusely distributed. This is a quite reasonable assumption except for very thin surfaces like leafs.

When assuming diffuse re-emittance of light, only one sample per surface point has to be measured. For this purpose Nickell et al.[NHE+00] and Weyrich et al. [WMP+06] proposed to use a measurement head densely filled with optical fibers. One of these fibers is used to illuminate the respective surface point while the others transmit the emitted light from nearby surface points to an imaging sensor. By placing the measurement head onto a nearly planar surface point, samples of the BSSRDF can be acquired.

Spectral reconstruction: When summarizing previous work on reflectance acquisition of spatially varying surfaces, it is apparent that a tremendous amount of work has been spent on the optimization of capture processes to increase quality

and reduce capture time. Unfortunately, all optimized setups are only equipped with RGB or similar trichromatic cameras due to their low cost and easiness of use. When aiming at predictive rendering this data has limited use.

Other researchers have already worked on approaches to retrieve spectral reflectance data from RGB data (e.g. [HSB+99b], [IB99], [Ima98], [SH10], [DSD+16], [ABS16], [JZG+17], [NPB14], [AWT17]). Since this would allow to re-use existing, efficient capture setups, this is a good idea. We will discuss these approaches in greater detail in Chapter 5 and then present a thorough comparison of our novel method against theirs in Chapters 7 and 10.4.

2.6 Summary

At a first glance, it seems to be straightforward to extend some of the optimized setups to spectral domain by employing one of the spectral imaging techniques explained in section 2.3.2. Since materials exhibit in most cases only smooth spectral reflectance, even multi-spectral approaches should suffice.

Unfortunately, severe physical limitations hamper the practicality of such a straightforward approach. Even when using only a moderate spectral sampling of about 10 to 20nm, the corresponding optical bandpass filters have much less transmittance than e.g. traditional red, green or blue filters. This way, a lot of the photons is absorbed in the filter before reaching the image sensor. This has to be counteracted by more sensitive cameras, longer exposure times or stronger light sources. In cases where the setup just requires one light source and one camera as for the spatial gonioreflectometers, this might be done with acceptable additional costs. But as soon as a larger level of view parallelism is desired or as soon as e.g. multiple light sources are utilized to reduce the time for mechanical movements, simply using better hardware easily renders a spectral extension impractical.

Due to the presence of high quality and fast acquisition setups, RGB reflectance measurement has gained a rather high level of attention in both research and industry. In contrast to this, spectral reflectance of complex surfaces is seldom considered due to the absence of efficient capture setups, though it would be highly desirable to use spectral SVBRDFs or BTFs for predictive rendering.

We will now continue to describe our efforts to construct a setup to measure multi-spectral SVBRDFs or BTFs in a straightforward manner.

Part I

Ground-Truth Data

CHAPTER 3

SPECTRAL GONIOREFLECTOMETER DEVICE

3.1 Preliminaries

Our first step towards faster spectral reflectance capture of spatially-varying surfaces is to build a spatial gonioreflectometer with a multi-spectral imaging system. We restrict ourselves to materials exhibiting no fluorescence and phosphorescence because multi-spectral SVBRDFs or BTFs are already seven dimensional. When considering the measurement times reported in previous work with RGB gonioreflectometers like [DvGNK97], [McA02], [SSK03] and the multi-spectral gonioreflectometer of Tsuchida et al. [TAN+05] a lot of acquisition time is required even when neglecting fluorescence and phosphorescence.

Using the gonioreflectometer setup a database of spectral reflectance groundtruth data was acquired and published at the BTF database of the University of Bonn [BTF]. The acquired data was used to evaluate the quality of more sophisticated capture methods introduced later in this thesis and may hopefully also help other researchers to evaluate novel methods on spectral reflectance data.

The new device is an upgrade to the spatial gonioreflectometer of Sattler et al. [SSK03]. We integrated a monochrome CCD camera with high quantum efficiency and good signal-to-noise ratio. In front of the camera an electronically tuneable bandpass filter was used. When restricting to non-fluorescent materials, one could also use a bandpass filter in front of the light source. However, a camera-filter combination offers the opportunity to be used for other purposes as well. In our case, we used the same camera system for the capture of multi-spectral environment maps as well.

In the following chapters we first describe the hardware, the calibration and the measurement process. Afterwards we explain, how the acquired raw data is post-processed to be usable in a spectral rendering system and discuss results obtained

All text in Part I marked in gray is self-cited from M. RUMP, R. SARLETTE, AND R. KLEIN. Groundtruth Data for Multispectral Bidirectional Texture Functions. In proceedings of *CGIV 2010*, pages 326-330, June 2010.

with the proposed acquisition method.

3.2 Gonioreflectometer Device

Our multi-spectral gonioreflectometer is based on the hardware setup by Sattler et al. [SSK03]. We will therefore first describe their setup in detail and afterwards discuss the necessary changes.

3.2.1 RGB gonioreflectometer

The spatial gonioreflectometer from Sattler et al. is based on the basic idea of Dana et al. [DvGNK97] to measure reflectance of complex surfaces and represent them as a Bidirectional Texture Function (BTF). Unfortunately, the apparatus of Dana required a lot of manual work and also suffered from a bad image quality due to the usage of an analog video camera with a frame grabber and a low angular sampling. In contrast to this, the gonioreflectometer of Sattler should provide high quality with minimal manual effort.

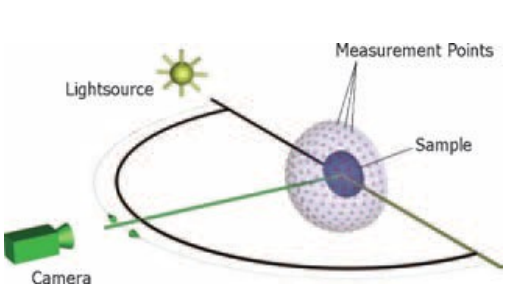
Common hardware used by both setups is a robot arm, which holds the material sample and can orient this sample in different angles with respect to camera and light source. The light source is at a fixed position with respect to the robot and the camera is moved to different positions on a half circle around the robot. Dana et al. mounted the camera on a standard tripod and had to place it at the different positions by hand.

Sattler used an Intellitek Scorbot ER-4u robot arm. This is an educational robot with five degrees of freedom able to carry about 2i̇ kg. The robot can be controlled via a USB connection using the accompanying software development kit on a PC.

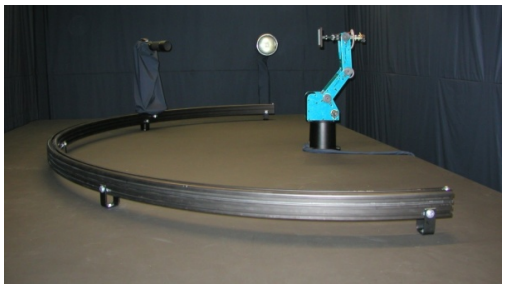
As a light source a 570W hydrargyrum medium-arc iodide (HMI) lamp is used. HMI lamps have a high luminous efficacy and a good *color rendering index* (CRI). The Broncolor F575 lamp used in the setup has a corresponding color temperature of about 6000K indicating that it emits nearly white light and has a CRI > 90 which means the spectral power distribution is suitable to acquire trichromatic reflectances with high accuracy.

The first conceptual difference to the setup of Dana was to use a digital still camera (DSC) instead of an analog video camera because professional DSCs offer a much higher resolution and image fidelity. In this case a Kodak DCS Pro 14n SLR camera was used. It allows full control of all functionality including download of raw images over a IEEE 1394 (”FireWire”) connection.

The second conceptual difference was to mount the camera onto a rail system which allows for fully automatic movement in a half circle around the robot. The



(a) Sketch of the setup and the angular sampling used by Sattler et al.



(b) Photograph of the setup: The sample is held by a robot arm, the light source is fixed and the camera is automatically movable on a rail around the robot.

Figure 3.1: The BTF measurement setup from Sattler et al. [SSK03].

rail and the dolly are custom designed. As a drive for the dolly another motor from the Intellitek Scorbot was used. Fortunately, the robot controller allows for the connection of additional motors that can then also be controlled via the PC.

Figure 3.1 shows a photograph of the RGB spatial gonioreflectometer in the lab at the University of Bonn.

3.2.2 Setup Modification

When extending the given RGB setup for spectral measurements, a lot of design decisions arise:

- Determine the kind of spectral imaging system to use
- Consider the mechanical limitations for the spectral imaging system and the light source
- Check, if power and spectral power distribution of the existing light source are appropriate for spectral measurements

Since we want to capture ground-truth data, it is necessary to chose a spectral imaging system that does not place any requirements on the scene like CTIS or CASSI systems (see section 2.3.2). It would be advantageous to use a snapshot technique like [DTCL09] to reduce capture time. Unfortunately, the spatial resolution is too low to capture complex surfaces with sufficient detail. Therefore only serial techniques remain. Among them the whisk-broom and push-broom techniques provide both excellent spatial and spectral resolution. However, they lead to impractically long measurement times since they need - in case of a push-broom technique - a few hundred scans per light/view direction pair and in case

of whisk-broom techniques even up to millions of data points per direction pair. Moreover, most reflectance spectra tend to be very smooth as they originate from superimposed, broad bands [Nas01]. For this reason, the high spectral resolution of whisk- and push-broom spectral imaging systems is not required. Therefore, the method of choice is filter-based spectral imaging providing high spatial and moderate spectral resolution.

For filter-based spectral imaging a monochrome camera is required which should have a good signal-to-noise and quantum efficiency to counteract the low incident photon count. Currently, the best performance is delivered by actively cooled CCD cameras which are for example used in astronomic applications (e.g. [RBdB$^+$99]). Unfortunately, since the camera has to be mounted on the rail dolly cooling designs requiring bulky equipment are not applicable. Currently, the most compact camera cooling system consists of a thermoelectric element directly mounted onto the CCD and a secondary cooling with airflow. These devices easily fit into a rather small camera housing and are therefore perfectly applicable in our case.

We utilize a Photometric CoolSnap K4. This is a CCD-based camera for scientific applications with 2048×2048 pixels resolution, a quite large pixel size of $7.4\mu m \times 7.4\mu m$. The thermoelectric cooling system holds the imager at a stable temperature of $-25°C$.

Additionally, a system of changeable spectral bandpass filters is required. Two different placements for the filter system are possible: either in front of the light source to filter incident light on the sample or in front of the camera to filter reflected light. Moreover, a filter system has to be chosen. Either mechanical filter wheels with glass-filters or electronic filters are applicable.

We decided to place the filter system in front of the camera because this enables the use of the resulting camera system beyond the reflectance capture scenario, e.g. for the acquisition of spectral environment maps. Moreover, a filter in front of the light source it technically quite complicated since it has to absorb high amounts of radiometric energy without getting destructed. Only few filters are available which are stable enough for this purpose. Since the camera has to be moved in the setup and is thus mounted onto the dolly, a compact and lightweight filter system is required. We therefore decided to use an electronically tuneable filter (ETF). Besides the mechanical advantages, these filters have the additional advantage of a fast switching time which cannot be achieved by mechanical systems.

Different ETF systems exist (see also [Gat00]). Among them *liquid crystal tunable filters* (LCTF) and *acousto-optical tunable filters* (AOTF) are most commonly used in practice. Unfortunately, AOTF offer only a very limited field of view [HSB99a] and are therefore not well suited for imaging purposes. In contrast, LCTFs have been used for spectral imaging before [Gat00], [HSB99a] and

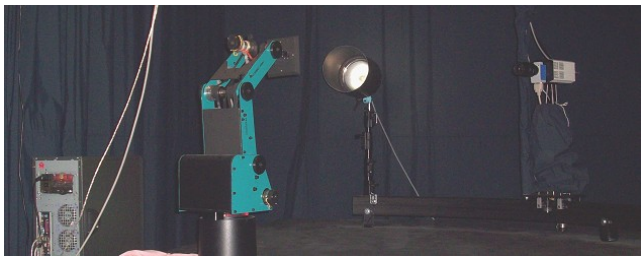
here especially those from Cambridge Research and Instrumentation[1]. Since they are compact and allow for direct integration into the camera‘s optical system, we can use them on the dolly of our gonioreflectometer as well. For the visible spectrum CRI offers three different models with 7, 10 and 20nm bandwidth. We decided to use the VariSpec-VS10 model with 10nm bandwidth to ensure sufficient spectral resolution for reflectance spectra with higher frequent flanks like those based on interference effects at the material‘s surface. The VariSpec is integrated into a special optical system from Schneider-Kreuznach with large focal length to capture an area of about 20x20 cm at 2m distance (our working distance due to the radius of the halfcircle rail). The first part of the optical system collimates the incoming light rays for the LCTF making the spectral transmittance not varying over the image. Behind the LCTF the second part of the optical system is projecting the image onto the sensor. As a result we have a very compact multispectral imaging system which we can also use for other tasks like environment map capture.

When considering the light source, we noticed that fortunately a quite strong light source (570W HMI) was already integrated into the RGB setup in order to achieve lower exposure times. The high color rendering index (CRI) indicates an overall good spectral power distribution. Unfortunately, the CRI is based on the tristimulus ΔE_{UVW} measure [Wys63] and can therefore hide certain spectral features like narrow peaks which are typical for gas discharge lamps. We therefore measured the emission spectrum of the HMI lamp using an Ocean Optics USB4000 spectrophotometer with high spectral resolution. The result is shown in Figure 3.4(a). The narrow peaks on the otherwise well distributed spectrum could make spectral calibration more difficult.

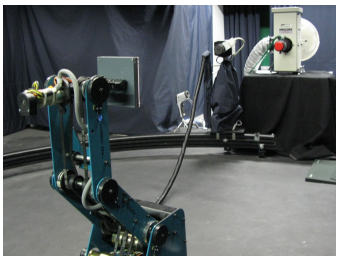
Since gas discharge lamps are often operated with pulses we also checked the spectrum of the HMI lamp with a high time resolution of 6ms. Unfortunately, higher time resolution was not possible with the Ocean Optics. However, the measurement series revealed that the three peaks in the lamp spectrum oscillate over time. A respective plot can be found in Figure 3.4(c). When performing a frequency analysis of this series using a Fourier transformation (see Figure 3.4(e)) we can clearly notice that the oscillation has a frequency of 15Hz. For comparison we also included the plots for a wavelength far off the peaks. One can notice that this part of the spectrum does not oscillate. Since the spread-out part of the HMI lamp spectrum is generated by thermal effects, this was the expected behavior. Due to the peak oscillations, all exposure times have to be multiples of 1/15s $\approx$ 66.6ms when using this lamp.

For this reason, we also considered a black-body like lamp. Here, we have

[1]Now a division of Caliper http://www.caliperls.com



(a) Modified setup with HMI lamp



(b) Modified setup with QTH lamp

Figure 3.2: Photographs of the modified gonioreflectometer setup.

chosen a quartz tungsten halogen (QTH) lamp from Newport[2]. These lamps have a black-body-like spectrum with a color temperature of about 3000K. Figure 3.4(d) shows the emission spectrum of the lamp. In contrast to arc lamps like HMI lamps they have a smooth spectrum with no peaks. Due to this fact and due to their stable and oscillation-free energy output they are well suited for measurement purposes. Unfortunately, due to their low color temperature, they emit little light in the "blue" part of the visible spectrum. We therefore decided to use the available model with maximum power, namely 1000W. The lamp is quite bulky and requires a lot of cooling, but this is acceptable since it is not moved during measurement.

Images of the modified gonioreflectometer are shown in Figure 3.2.

3.3 Gonioreflectometer Calibration

A careful calibration of all components of the setup was performed. For the geometric calibration we relied on the work of Sattler et al. since we did not change the mechanical system and its control. This means we had to perform a radiometric and spectral calibration of camera and light sources. Since we have two possible light sources to choose from, we have to determine which one provides best quality and measurement speed in combination with the camera system.

The radiometric calibration of the camera aims at recovering incident irradiance on the pixels from the pixel values. We assume an imaging model similar to those from Debevec et al. [DM97] and Robertson et al. [RBS03]:

$$p_x = f(tI_x) + d_x \tag{3.1}$$

where p_x is the final digital counts for pixel x, I_x the incident irradiance on pixel x, t the exposure time, f the so-called *response curve* and d_x the dark current counts

[2]http://www.newport.com/Quartz-Tungsten-Halogen-Lamps/378263/1033/info.aspx

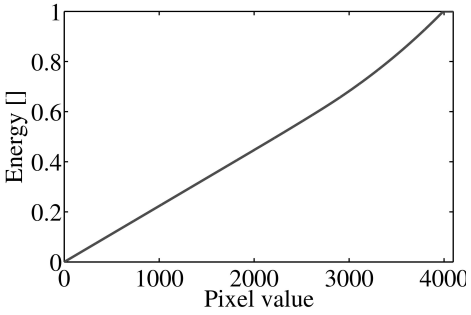
for pixel x. f is a - in general non-linear - function which summarizes all optical and electronic effects inside of the camera which determine how incident photon count on a pixel maps to the discrete pixel values. For the radiometric calibration of the CoolSnap camera, the per-pixel dark current and the *response curve* have to be determined. Then, incident irradiance can be recovered from pixel values by:

$$I_x = \frac{f^{-1}(p_x - d_x)}{t} \tag{3.2}$$

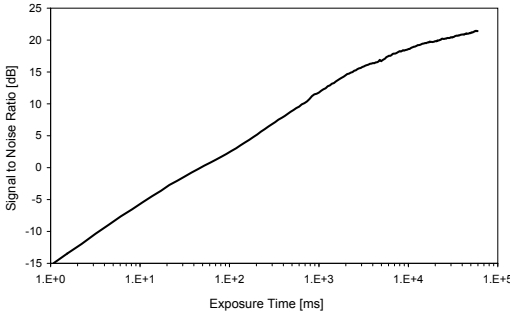
The dark current is mainly generated when electrons are raised to the conduction band by thermal effects or when there is a certain leakage in the electric circuits of a pixel [HL07]. Therefore, it depends on exposure time and on device temperature and is independent of the incident irradiance. There are two standard approaches to determine and subtract dark current. In the first variant, the pixel values of special, light-shielded pixels at the start of each image row are averaged by the CCD controller. This average is subtracted from the remaining "active" pixels of the row. This method however, does not consider the so called *fixed pattern noise* (FPN), the deviations in dark current leakage between the single pixels. Some pixels of a CCD are even so called warm or hot pixels that have such a large dark current leakage that they saturate only due to dark current. The second standard method is to capture a so called *dark frame* before or after the respective "real" exposure. Afterwards the dark frame is subtracted from the data frame. This way, fixed pattern noise and temperature dependence are correctly considered. Unfortunately, both frames are biased by dark current shot noise, reset noise, amplifier noise and potential other noise sources [HL07]. When subtracting, the variance increases as $\mu_{final} = \sqrt{2\mu^2_{invididual}}$. Moreover, capturing dark frames before or after each data frame roughly doubles the measurement time, which is crucial in our gonioreflectometer.

Fortunately, the CCD imager of the CoolSnap camera is actively cooled and has a constant temperature of about $-25°C$. Due to this, dark current is overall reduced and - for us important - depends on exposure time only. For this, we are fortunately able to capture the dark frames in beforehand which reduces both measurement time and noise. Noise reduction is achieved by capturing multiple dark frames for every exposure time and by averaging these dark frames. In our case, we captured 50 dark frames per exposure time. We preselected a set of exposure times and interpolate between the respective averaged dark frames linearly when recovering the incident irradiance using equation 3.2.

The calibration of the response curve was performed using the algorithm from Robertson et al. [RBS03] which recovers the inverse response curve f^{-1} up to a scale factor. We will call the resulting curve $\hat{f}^{-1}$ and refer to the return value of that function with the generic word "energy" in the following. To obtain a calibration with high quality, we took 200 images with steadily increasing exposure



(a) Inverse response curve of the CoolSnap camera with dark current subtracted. The response was calibrated using 200 images.



(b) Signal to noise ratio of the CoolSnap camera with increasing exposure time.

Figure 3.3: Radiometric calibration of the CoolSnap camera.

time, subtracted the dark current and run the Robertson algorithm. The resulting inverse response curve is shown in Figure 3.3(a).

To measure reflected radiance from the material sample, the missing scale factor has to be determined. This scale factor is moreover dependent on the wavelength since the VariSpec filter and the optical system have different transmission per wavelength band and the CCD has a wavelength-dependent quantum efficiency. To determine the scale factors it is common practice to place a sample with known BRDF in the setup, acquire the camera's response and determine the ratio between known reflectance and measured response. We use a Zenith UltraWhite Reflectance target from Sphere Optics, a nearly perfect white and diffuse surface made of barium sulphate. The factors are then calculated as:

$$F_\lambda = \frac{I_\lambda W_\lambda t}{\pi} \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \hat{f}^{-1}(p_{\lambda,i,white})} \tag{3.3}$$

where I is the spectrum of the light source used for calibration, W the manufacturer supplied reflectance spectrum of the Zenith UltraWhite (see Figure 3.5(d)), t the exposure time and $p_{\lambda,i,white}$ the n pixel values of the white standard. The first term is therefore the exitant radiance from the white standard (BRDF multiplied with light) and the second term is the reciprocal of the measured irradiance-proportional value using $\hat{f}^{-1}$.

The resulting factors are shown in Figure 3.4(b). The quantum efficiency of the whole camera system is very low in the short wavelengths corresponding to blue light due to the reduced quantum efficiency of the CCD and due to the worse transmission of the VariSpec filter for short wavelengths (see also Figure 3.4(f)).

Unfortunately, these differences in quantum efficiency increase the measured dynamic range of the scene. Two possible solutions exist to compensate for this.

The first option is to use constant exposure times for all wavelength bands and scale the measured energy values according to the inverse quantum efficiency. The second option would be to adjust the exposure time to the quantum efficiency. To check, which approach is superior for the CoolSnap camera, we measured the signal-to-noise ratio for a large range of exposure times up to 60s. The result is depicted in Figure 3.3(b), showing that the SNR is steadily increasing with exposure time, even for extremely long exposure times like 60s. This is due to the fact that the sensor is cooled producing few thermal electrons. We conclude, that longer exposure times deliver better results than scaling. Therefore, a different exposure time t_λ is calculated for every wavelength band with peak transmission at λ:

$$t_\lambda = \frac{t_{base} F_\lambda}{I_\lambda} \tag{3.4}$$

with I being the spectrum of the lamp used. This way, the camera system measures at similar signal levels in all wavelength bands when capturing a gray material.

When we calculate the resulting exposure times for both lamps (see Figures 3.5(b) and 3.5(a)) we realize that the QTH lamp is less suited for measurement with our camera system since it emits very low energy at short wavelengths. Please note that it was not even possible to determine a stable measurement of the necessary scaling factor for the 400nm band. We therefore decided to use the HMI lamp despite all the difficulties caused by the nature of that lamp. Due to the lamp osciallations we rounded all exposure times to the next multiple of 66.6ms (see discussion of HMI lamp in Section 3.2.2).

Having obtained the full calibration, the BRDF value for every pixel x on the material sample can be theoretically calculated as follows:

$$\rho(x, \lambda) = \frac{F_\lambda \hat{f}^{-}1(p_{x,\lambda} - d_{x,t_\lambda})}{I_\lambda t_\lambda} \tag{3.5}$$

where I is the spectrum of the illumination used during the measurement and $p_{x,\lambda}$ is the pixel value at x measured for wavelength λ. d_{x,t_λ} is the dark current value interpolated from the pre-recorded dark frames for exposure time t_λ.

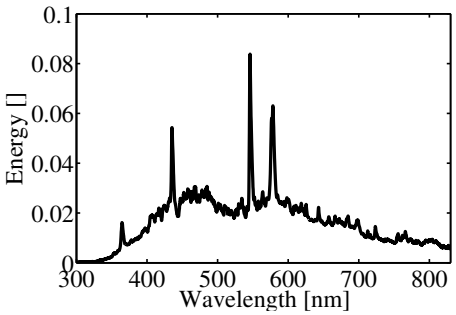
In practice, we determine the ratio $F'_\lambda = \frac{F_\lambda}{I_\lambda}$ prior to each measurement run because the light source has to moved to two different positions manually and the factors F' obviously also depend on the lamp distance which may vary by some centimeters. It is also important to note that both lamps also show certain aging effects and long-term instabilities over the whole capture time. To compensate for this we mount a small X-Rite I1 spectrophotometer opposite to the light source and capture the spectrum for each bi-directional measurement point. From the I1 measurements a single scaling factor $F_{i,long}$ per measurement point i is generated

in relation to the first point i_0. Figure 3.5(c) shows such a long-time variation of the HMI lamp.

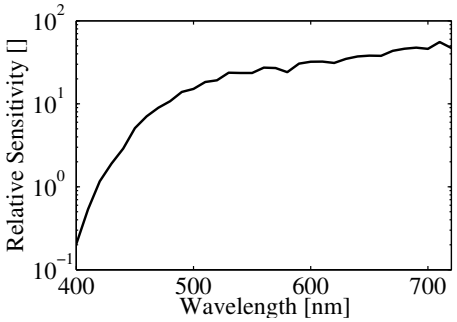
Summarizing, the reflectance for every point x is recovered by:

$$\rho(i, x, \lambda) = \frac{F'_\lambda F_{i,long} \hat{f}^{-}1(p_{i,x,\lambda} - d_{x,t_\lambda})}{t_\lambda} \tag{3.6}$$

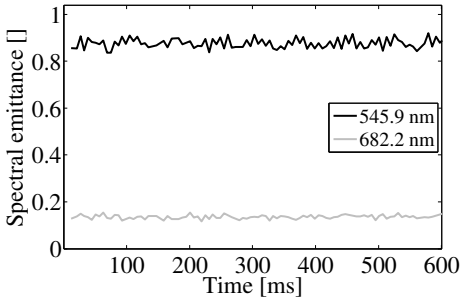
In the following chapter we will now describe the practical implementation of the measurement process and the post-processing of the acquired data. We will furthermore show results from the acquired ground-truth BTF data.



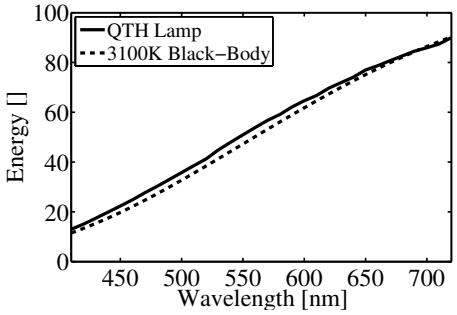
(a)HMI575 lamp: Mean emission spectrum



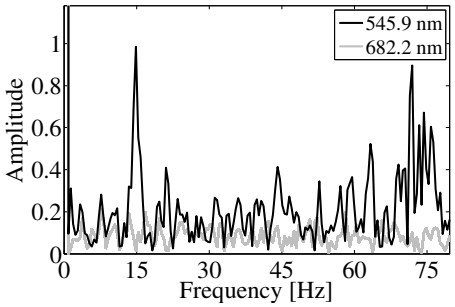
(b)Spectral sensitivity of the multi-spectral camera system.



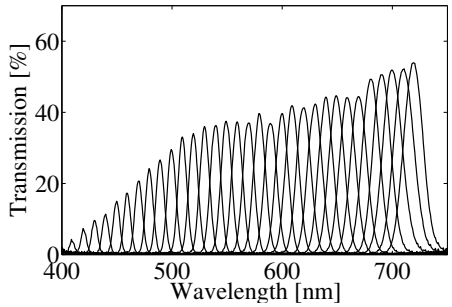
(c)HMI575 lamp: Spectral emittance of the 545.9nm peak over time with a time resolution of 6ms. The lamp controller generates regular electric pulses to the electrodes generating oscillating line emissions from the gas.



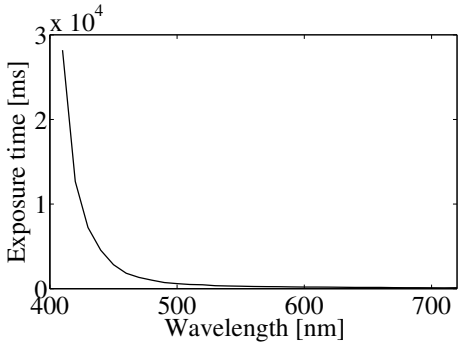
(d)Emission spectrum of the QTH lamp in comparison to a perfect black-body emitter. This lamp does not suffer from narrow spectral emission peaks or any oscillations over time as the HMI lamp does.



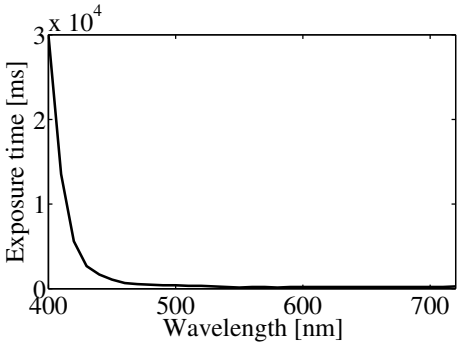
(e)HMI575 lamp: Frequency analysis of the oscillations. The 545.9 nm spectral peak shows a frequency peak at 15Hz (the peak at roughly 75Hz is due to the acquisition noise). The frequency analysis for the thermal part of the HMI spectrum shows a rather uniform distribution of amplitudes, meaning that this part does not oscillate.



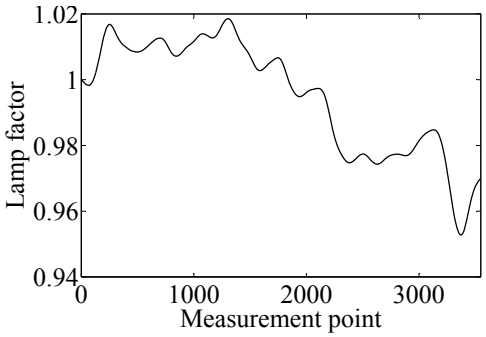
(f)Transmission spectra of the VariSpec filter tuned to peak transmission wavelengths from 400 to 720nm in 10nm steps according to manufacturer. The steadily decreasing transmittance with decreasing wavelength is notable.



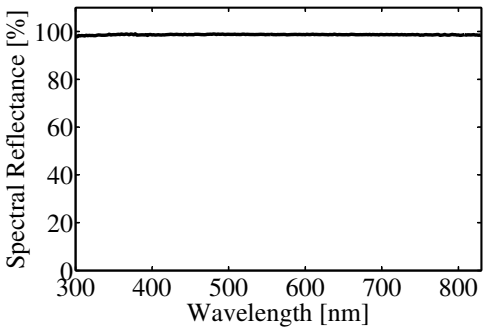
(a) Suitability of the QTH lamp in conjunction with the VariSpec bandpass filter: For an equal SNR in all wavelength bands very long exposure times would be necessary for the "blue" wavelengths to compensate for the faint lamp emission and the bad filter transmission. It was not even possible to determine a scaling factor for the 400nm band in a reliable way.



(b) Exposure times for all wavelength bands when using the HMI lamp. The exposure time for the "blue" wavelengths are still high to compensate for the bad filter transmission but the effect is much smaller compared to when using the QTH lamp (see nearby graph).



(c) Temporal change of HMI lamp emittance over long time.



(d) Spectral reflectance of the Zenith Ultra-White (provided by manufacturer).

Figure 3.4: Spectral calibration of the measurement setup: Spectral emittance of light sources, spectral sensitivity of the camera system and combination of camera system and light sources.

CHAPTER 4

GROUND-TRUTH MEASUREMENTS AND RESULTS

4.1 Ground-truth Measurements

4.1.1 Measurement Process

In order to capture the new kind of data the existing control program of the gonioreflectometer was altered to incorporate control for the new hardware, namely the VariSpec filter, the CoolSnap camera and the X-Rite I1 to track the light emitted by the HMI lamp. Since all these hardware parts have an accompanying SDK, extension of the program was simple and straightforward. Especially, the control code for the robot and rail was not touched.

The measurement process for each material is split into two parts since the lamp has to be moved to two different positions to allow for capture of directions all over the material's local hemisphere (otherwise a fully circular rail would be necessary). In front of each of these runs, the white reference target is held by the robot and calibration images are taken as explained in Chapter 3.3 to determine the factors F'_λ.

Afterwards the sample holder with the material sample itself is mounted to the robot arm. The sample holder is a metal frame with a handle for the robot and with fixing screws to mount plastic plates of about 6cm squared into the center. Onto the plastic plate materials can be glued or otherwise fixed. Due to this modular system it is possible to choose a plastic plate best suiting the material to measure without requiring to manufacture a completely new sample holder.

The two measurement runs itself are then fully automatized. For each directional sample, the mechanical system is moved to the respective position. Then the program waits for a short delay of one second for mechanical vibrations to die away. Afterwards the tunable filter is serially switched to peak transmission wavelengths from 400 to 720nm in 10nm steps and an image per wavelength band is taken with the CoolSnap camera using the exposure times t_λ shown in Figure 3.5(b). This image capture process takes about 68 seconds for all wavelength

bands, from which 63.6 seconds are exposure time, and 130 ms per image $\times$ 33 bands = 4.3 seconds are used for filter switching and image download to the control computer.

The angular sampling is unchangedly taken from Sattler's setup. He captured 81 view x 81 lights directions ending up with 6561 direction pairs. These angular sample points are ordered in a way that the average mechanical movement times are minimized. They take about 5 seconds in average and up to 20 seconds if the camera has to be moved on the rail. Therefore the total capture time for one material sums up to $6561 \times (68s + 5s + 1s) = 485514s = 134.9h = 5.6d$. Of course, such a long capture time is impractical for a practical use of the proposed setup in industrial processes but it suffices for the capture of a set of ground-truth datasets.

4.1.2 Post-Processing

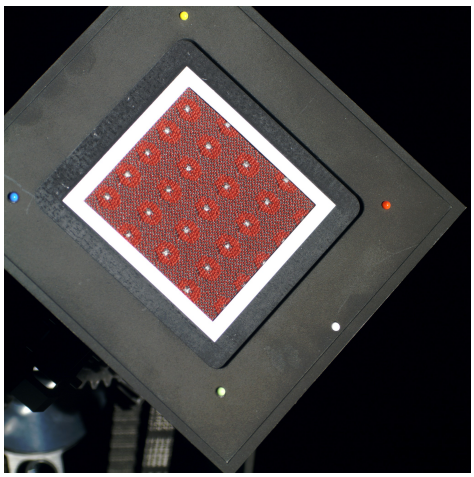
The acquired raw data has to be processed in order to recover final reflectance values for every point x and all direction pairs i on the material surface according to equation (3.5).

First, a mapping from raw image pixels x' to material surface positions x is determined. If surface geometry and camera parameters were known, this would be a trivial task. Since we do not have any knowledge about the material's surface geometry, we assume a planar surface in accordance with the BTF principle (see Section 2.4.1). The camera's intrinsic parameters could be calibrated by readily available methods like [Zha00]. Unfortunately, the extrinsic parameters of the camera, namely the position and rotation in relation to the materials surface, cannot be easily calibrated in our setting. The reason is the inaccuracy of the mechanical system. A mechanical system of the desired size which would be capable of positioning with sub-pixel accuracy would be too costly. Therefore, we employ computer vision methods to detect the material target in the raw images. For this, we enclose the material sample with a narrow border made of white paper. It is constructed by imprinting a larger black rectangle onto the paper using a inkjet printer leaving out the border area in the center. Inside of the border area the space for the material target is cut out. Figure 4.1 shows some raw images which clarify the construction of the white border. The corners of the white border are then detected in the images and a homography matrix can be computed, which maps the plane of the border in image space to the proxy/texture plane of the material. The main problem of detecting the border is the paper and the black print being not perfectly diffuse and therefore showing certain amounts of gloss as well as fresnel reflection artifacts. Therefore, a stable detection in the images is required. We achieve this with the following steps:

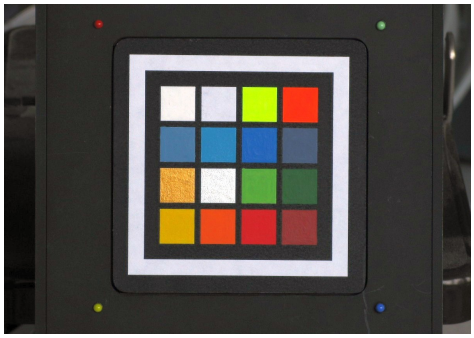
1. Generate a binary image by thresholding the raw image
2. Construct a list of all outer contours in the binary image
3. Remove all contours having a bounding box of inadequate size
4. For all contours:
   (a) Perform a local PCA [KL97] on the contour points using 6 clusters and one principal component. This corresponds to stably fitting 6 lines into the points.
   (b) Remove the two smallest clusters.
   (c) Compute intersections of the four remaining lines when angle between lines is larger than 10°.
5. Remove all rectangles having inadequate size
6. If no rectangles remain, reduce threshold and continue at step 1.

The two excess clusters in the local PCA help to better deal with small artifacts. After this coarse rectangle is found, a sub pixel accurate refinement of the corner positions has to take place to obtain a good registration. This is achieved by first computing the homography matrix H, mapping from the coarse image rectangle to the texture space $[0, 1]^2$, and then extracting small image pieces around the corners in texture space. This pieces have a width and height of 0.1 texture space units and are centered at the corners. The resolution can be freely chosen to adjust the desired accuracy. For each of the image pieces the 5 and 95 percentiles of the pixel values are calculated and the pixel values are linearly re-mapped (and clamped) so that the percentile values end up as black and white. The result are four small image pieces like those shown in Figure 4.1(d). The position of the corners in the small image pieces is determined by fitting a Chan-Vese model [CV01]. The ”inside” is assumed to have constant value of 1 and the ”outside” of 0 and the contour is assumed to be rectangular. Therefore, only the corner point in the image has to be optimized. The resulting target function is minimized by a brute-force scan using a GPU program written in CUDA and takes about one second for all four image pieces when using a resolution of 200x200 pixels. The resulting corner points are in texture space and are then mapped back to image space using the inverse homography matrix H^{-1}.

Since this process has to be performed for all view and light combinations, some of the images with shallow viewing angle and opposite light direction exhibit strong fresnel reflections on both the white paper and the black paint of the sample holder. Figure 4.1(c) shows an example of such a situation. In this cases, the algorithm based on thresholding does not provide any solution. For this, we

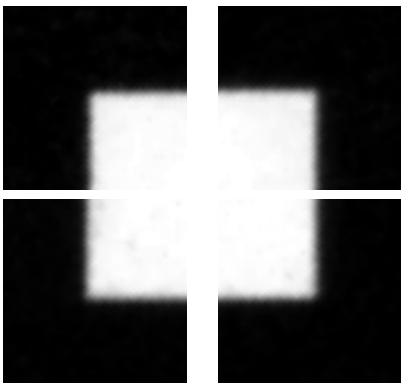


(a) Image from the red fabric dataset. One can clearly see the sample holder consisting of the metal frame and the white border made by printing on white paper.



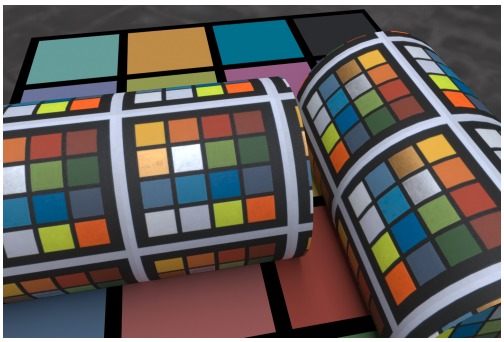
(b) Image from the "Colorchecker" dataset

(c) Image from the color checker dataset showing strong fresnel reflection: For these images the black ink and the white paper are barely distinguishable.

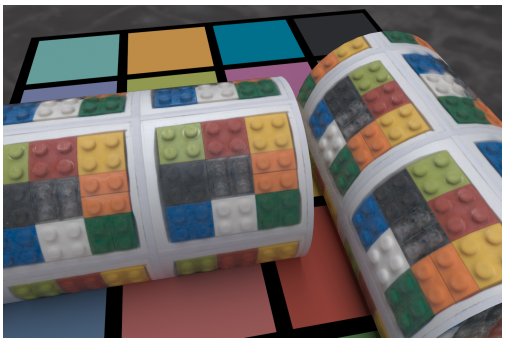


(d) Extracted and scaled corner image pieces to fit the Chan-Vese model.

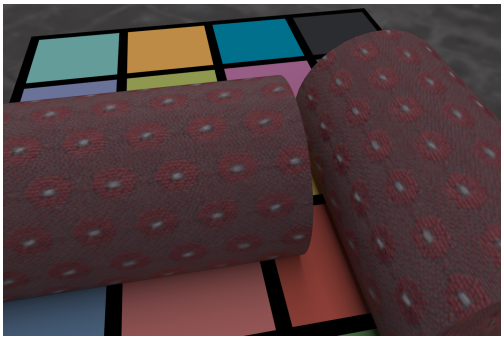
Figure 4.1: Example raw images from the gonioreflectometer setup. The images have been generated by applying the appropriate scale factors to the images corresponding wavelengths, then mapping the images from spectral space to sRGB and finally applying a gamma mapping.



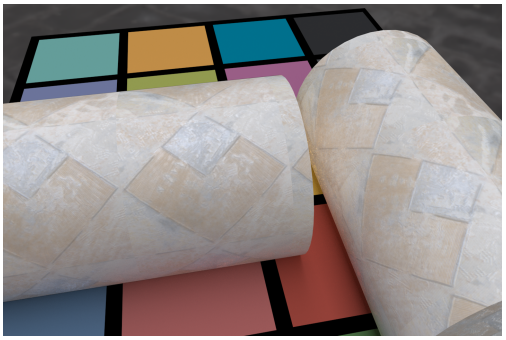
(a) The color chart under neutral, environmental illumination.



(b) The Lego bricks under neutral, environmental illumination.



(c) The red fabric under neutral, environmental illumination.



(d) The wallpaper under neutral, environmental illumination.

Figure 4.2: Some rendered images using the acquired multi-spectral groundtruth BTF data. The resulting spectral images have been converted to sRGB color space for display purposes.

implemented a small tool that enables for easy selection of the corners in a graphical user interface for all images where the automatic extraction did not work.

Having determined the image position of all corners of the white border in all raw images, textures for all direction pairs and wavelengths can be extracted. This is done by mapping all positions $x \in [0,1]^2$ in texture space to the respective raw image position x' and using linear interpolation to extract the pixel value. The resolution of the textures can be chosen freely, but more than 800×800 is not sensible due to the size of the material sample in the raw images corresponding to frontal views. To extract the BTF value for each of the pixels, equation (3.6) is applied.

This step is computationally relatively cheap and therefore highly limited by IO bandwidth. To maximize IO, the raw images are read by one CPU thread while multiple threads perform the processing and another thread is responsible for the

saving of the final images. On the processing computer (Intel Core i7 QuadCore and a recent HDD) this step takes about 6 hours.

Finally, one ends up with a BTF consisting of one registered, radiometrically and spectrally calibrated texture per measured direction pair. In principle, this data is directly suitable for rendering. However, in 800×800 pixel resolution and with 33 spectral bands, the 6561 textures require 516GB of memory. Therefore, the textures should be compressed using methods like those described in [MMS+05] or [FH09]. Following the findings of Müller [Mül09] about the compression of RGB BTFs we used a PCA based compression method that is a modification of the method first described in [KMBK03]. We first split the spectra of all texels into a sum (brightness) and a normalized spectrum and factorize the two parts individually. The brightness channel contains most of the complexity that arises from the mesostructure of the BTF. For this reason, we used 250 PCA components to represent it, whereas 60 components were sufficient to reproduce each single band of the normalized spectra. We end up with approx˙ 2.5 GB of compressed data without a large loss in accuracy.

If a decoding plugin can be supplied, this data can be directly fed into any spectral renderer. As a demonstration we rendered several images using the acquired data. These can be found in Figure 4.2.

4.2 Results

To build a first usable database of spectral BTFs, we selected four different materials:

- A hand-painted color checker with 16 fields, including fluorescent and specular ones
- A piece of complex fabric having four different threads in it (dark and light red, grey, and silver)
- A piece of wallpaper with complex structure and gloss
- A quadratic piece of Lego bricks of different colors including a transparent one

Figure 4.1 shows photographs of these four material samples. The post-processed images of these samples have been published in the BTF Database Bonn[1] to be used by other researchers to develop methods for efficient spectral BTF capture, compression, and rendering.

To check the accuracy of our datasets we made several comparisons. First, we compared the spectra of different parts of the material samples measured with

[1] http://cg.cs.uni-bonn.de/de/projekte/btfdbb/download/spectral/

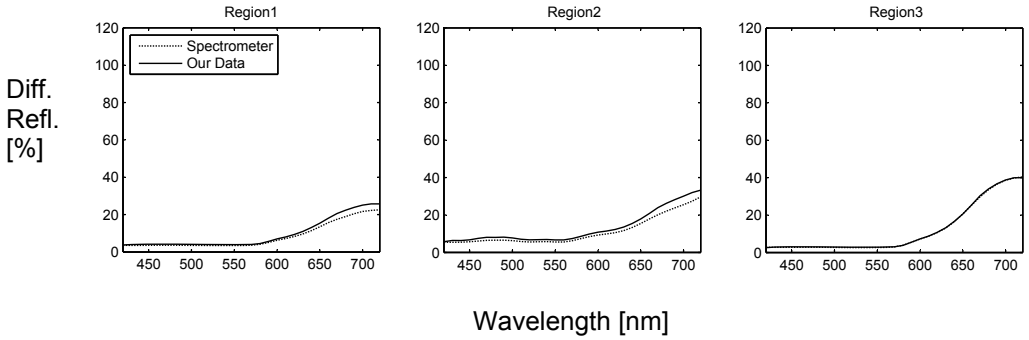


Figure 4.3: Comparison of our "red fabric" data (see also Figure 4.1(a)) integrated to mimic the spectrometer behavior with measurements acquired with the spectrometer: We tried to match the spots where we measured with the I1 as good as possible.

our setup to corresponding measurements using a X-Rite I1 handheld spectrophotometer. The I1 illuminates a circular spot of the sample using a ring-shaped light with elevation angle of $\theta = 45°$ and detects the reflected spectrum from the top (45/0 annular geometry). To simulate this, we integrated our data over circular spatial regions as well as all the images of the BTF that correspond to illumination directions of $\theta = 45°$ and viewing directions from the top. Since we do not known the exact geometry of the spectrometer we cannot hope to reproduce all of the data exactly. This is especially true for the specular fields since here a small angular deviation results in strong differences in brightness. The comparison result can be seen in Figures 4.4 and 4.3. Our setup is able to reproduce the spectra with high precision.The fluorescent colors on the colorchecker (please note that the paper contains some optical brightener as well) cannot be reproduced which results from the different amount of energy emitted by the light source of the I1 and our HMI lamp in the UV range. Since we did not sample the dimension of wavelength of incoming light, this effect cannot be reproduced. Most likely the differences for the specular materials are caused by the very simple I1 simulation, because it is done averaging all images with illumination from $45°$ elevation and top view. This matches only roughly the internal conditions of the I1 leading to larger errors for specular materials.

For the second comparison we took a multispectral photograph of the color checker in a room with complex lighting consisting of small thermal light sources (halogen lamps) as well as large fluorescent light sources (neon lamps). The resulting illumination spectrum has sharp peaks and is quite different compared to typically used daylight illuminants.

Then we captured the illumination conditions extending the method from De-

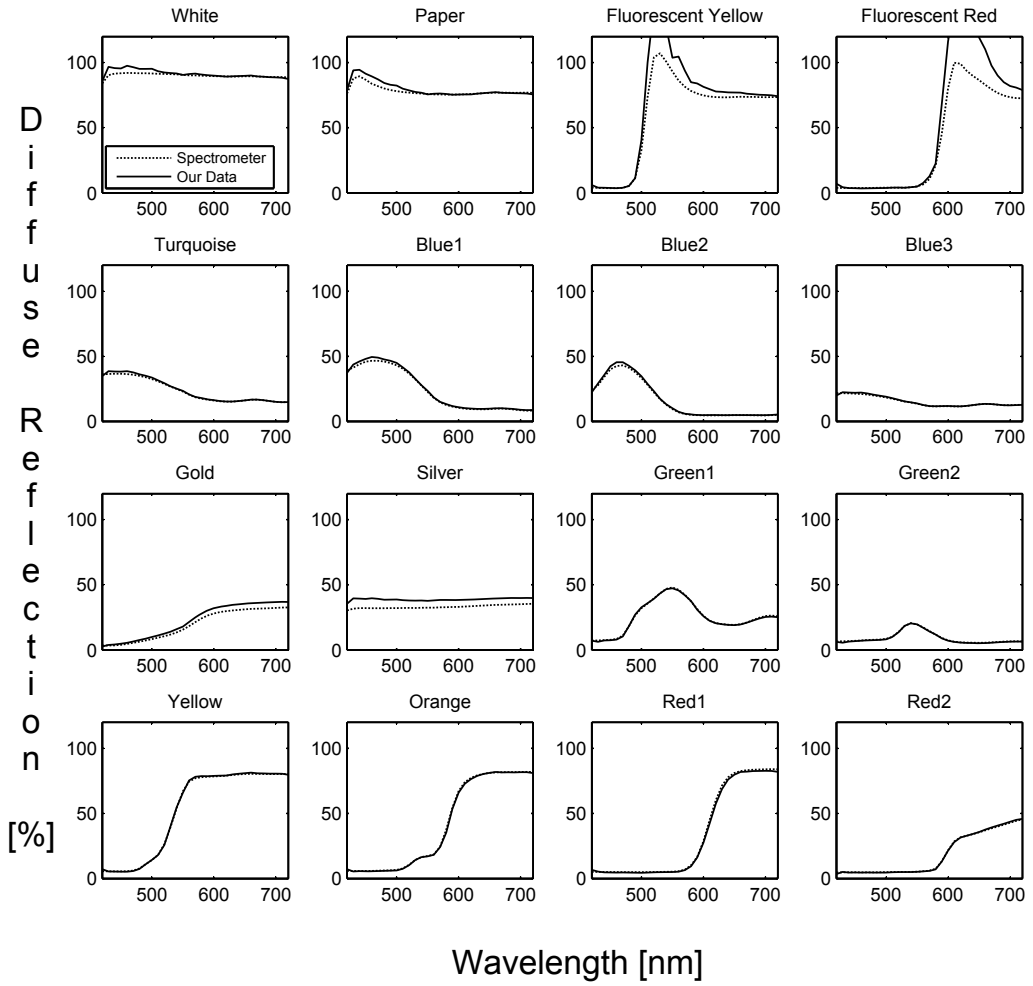


Figure 4.4: Comparison of our "Colorchecker" data (see also Figure 4.1(b)) integrated to mimic the spectrometer behavior with measurements taken with the X-Rite I1 spectrometer: All fields but the specular and the fluorescent ones are reproduced very good. Most likely the differences for the specular materials are caused by the very simple I1 simulation, because it is done averaging all images with illumination from $45°$ elevation and top view.

bevec et al. [DM97] by taking multispectral photographs of a chrome sphere mounted at the position of the material samples. The environment map is shown in Figure 4.5.

Using this image-based lighting data and the captured BTF data we re-rendered a photograph of the color checker using a spectral rendering system. A comparison of the rendering and the photograph can be seen in Figure 4.6. The multispectral images have been converted to RGB by convolving the spectra with the

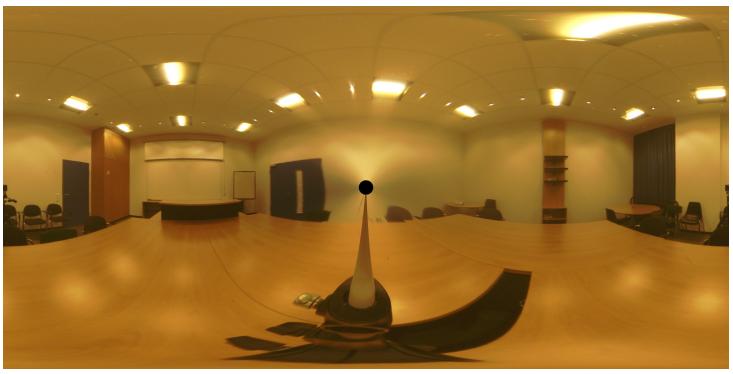


Figure 4.5: Environment map of a room with complex lighting consisting of large neon lamps and small halogen spotlights. The average spectrum has sharp peaks and the illumination in this room is therefore quite different from a typical daylight situation. The spectral image has been mapped to sRGB for display purposes.

CIE 1931 RGB color matching functions whereas the environment map and the BTFs for the RGB comparison were converted to RGB before the rendering. All three images have been white balanced using the same RGB factor derived from the environment map and tonemapped using a measured response curve of a RGB camera afterwards. Calibrated tonemapping and display of the multispectral images is out of scope of this paper and thus not further discussed. The multispectral rendering reproduces the photograph faithfully and shows the accuracy of our reflectance data. The third image of Figure 4.6 shows that RGB BTFs cannot be used for predictive rendering due to metamerism effects.

It should be noted that a BTF capture system based on trichromatic cameras would produce an even worse result. The reason for that is, that one would need to transform the system-dependent RGB output to some standard color space (here: CIE 1931 RGB). This would require an ICC profile or similar for the capture system. However, it is not possible to derive exact values in the standard color space from the system-dependent RGB values especially in such a complex case with many different spectra contained in one image (this issue is discussed later on in this thesis in greater detail). In contrast to this, we were able to calculate exact CIE 1931 RGB triples from the spectral data. The same issue is of course also true for the environment map. Thus, our comparison contains the bias introduced by the RGB rendering only.

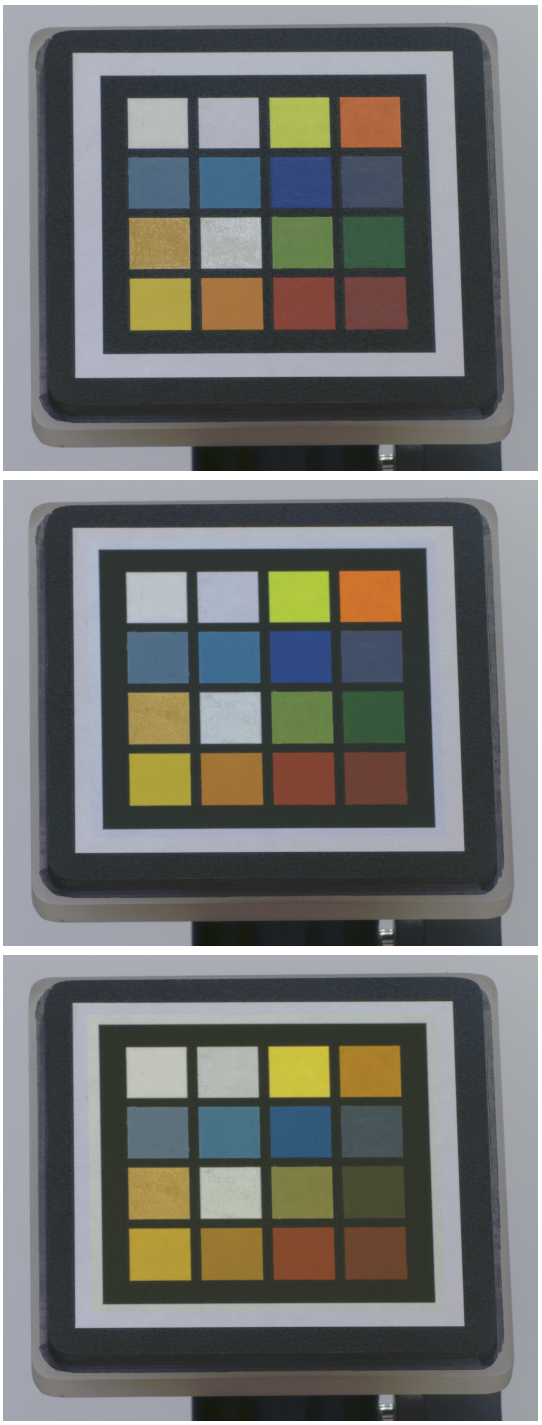


Figure 4.6: Comparison of a photograph (top) with a spectral rendering (center) and a RGB rendering (bottom) showing the accuracy of our data as well as the impacts of metamerism.

4.3 Summary

Summarizing, in Part I we have shown how to modify and calibrate an existing gonioreflectometer to capture multi-spectral, ground-truth BTFs and how post-processing of the raw images takes place. In detail:

- how to select an appropriate camera system, i.e. camera and filters.
- how to find a lighting system matching the camera system.
- how important an exact knowledge of the light source spectral power distribution is, especially also over time.
- how to do a spectral calibration of the whole system.
- how to apply calibration data at post-processing time.
- how to register all BTF images in the texture coordinate space.
- how existing BTF compression methods can be easily applied to spectral BTFs
- how to perform an inter-device matching of acquired spectral reflectance

Overall, this should give the reader the knowledge do build up his own multi-spectral BTF capture device and moreover it gives the opportunity to judge the properties and quality of the acquired and published data. The resulting data has been made publicly available for other researchers to e.g. test their algorithms.

Part II

Spectral Reconstruction

CHAPTER 5

INTRODUCTION TO SPECTRAL RECONSTRUCTION

Spectral reconstruction is the process of reconstructing a light spectrum or - in case of spectral imaging - an image of spectra from measurement device responses. Therefore, it is one of the central tasks in the area of spectroscopy and spectral imaging.

Spectral reconstruction is a comparably simple task when the device provides dense measurement data by using narrow-band optical filters or diffractive elements to split the incoming light into its spectral components (see also Part I). In this case, the device responses have to be transformed into an energy-linearized value and eventually a spectral bandpass correction has to take place. This is described in great detail in many standard text-books i.e. [Sch07]. Unfortunately, devices of this type require long acquisition times due to several reasons: First, they require long exposure times since low irradiance is incident on each sensor element due the filtering or splitting. Second, it requires multiple acquisition cycles to capture the whole spectral data cube (see also Section 2.3.2). It is therefore faster to acquire data with broadband filters like typical RGB filters or a small set of glass filters mounted in a filter wheel.

However, as soon as the original spectra have to reconstructed from this kind of data, the problem is getting more complicated and the reconstruction often has no unique solution. Nevertheless, broadband filters were already utilized for multi-spectral imaging [HSB+99b], [IB99], [Ima98], [SH10]. These approaches try to invert the mapping from spectral space to the low dimensional space of device responses by some numerical method like pseudo-inverse matrices, effectively assuming that all imaged spectra are located on a hyper-plane in spectral space having dimension equal to or less than the number of utilized broadband filters. It is easily imaginable that this is not true as soon as more different materials and/or light sources are present in the imaged scene.

All text in Part II marked in gray is self-cited from M. RUMP AND R. KLEIN. Spectralization: Reconstructing spectra from sparse data. In proceedings of *SR '10 Rendering Techniques*, pages 1347-1354, 2010.

The basic idea of our approach for multi-spectral capture - not limited to surface reflectance - is to utilize hybrid data. When reviewing the light measurement devices described in section 2.3.1 it becomes apparent that there exist a lot of devices - most of them being based on CCD or CMOS chips - which can either capture single 2D slices of or a certail integral over the spectral data cube with high resolution and in a single shot. It is therefore reasonable to try to combine different of these devices and fuse the acquired data to achieve a better constraint spectral reconstruction compared to single-mode data.

Of course, the choice of different devices is use-case dependent. Nevertheless, it can be noted that in many areas - and in computer graphics as well - a lot of effort has been spent on utilizing RGB cameras for various purposes like light and reflectance measurements. The reason is that RGB cameras are cheap, easy-to-use, provide a high spatial resolution and - due to their similarity to the human eye - a color accuracy that is acceptable in certain applications. Due to these large amounts of existing RGB hardware and knowledge on the usage and calibration of RGB devices, it is desirable to re-utilize and only extend existing devices to increase the spectral resolution from three (potentially overlapping) spectral bands to a much higher spectral resolution.

In this part of the thesis we will introduce a framework that allows to combine spatially dense data gathered by RGB devices with spatially sparse data having higher spectral resolution captured by devices like spectrophotometers, line-scan cameras, narrow-band filter based cameras (see section 2.3.1) or snapshot-spectral-imagers to reconstruct spectral datasets like images or reflectances. Our algorithm is based on energy minimization and allows for easy integration of many different types of devices. For easier understanding of the text, we will restrict ourselves to spectral images afterwards. Nevertheless, the algorithm is applicable to BRDFs or other types of reflectance data.

Hybrid data Hybrid data has been used by other researchers for spectral imaging in beforehand. Imai et al. [IB98] combine a high resolution grayscale scanning with a low-resolution multi-spectral image. They assume that inside of the footprint of a low-resolution pixel all spectra lie on a low-dimensional hyperplane in spectral space. Murakami et al. [MIYO07] combine a high-resolution RGB image and sparse, dot-wise measured spectra - possibly provided by a whisk-broom spectral imager - to a high-resolution spectral image. While they approach the reconstruction from a different view (Bayesian statistics) their method is to a certain degree similar to our one. They aim at reconstructing that spectral image which has the highest posterior probability with respect to the measured data. Their approach is theoretically sound but it lacks a good prior on spectral images. They simply assume that the entries of the spectral data cube have a gaussian distribu-

tion. Unfortunately, many real world images violate this assumption and we will give examples for this later on.

In contrast to this, we will introduce a novel prior in section 6.2 that is better suited for a broad range of real world images, can be easily adapted to specific use cases and can be easily integrated into the spectral reconstruction process.

Dong et al. [DSD+16] also use hybrid data to capture full spectra BTFs. They combine a full RGB BTF with a single spectral image of the material lit with spherical illumination. From the spectral image they compute a low dimensional linear basis (three-dimensional in case of the RGB system) using non-negative matrix factorization ([KP08]). In a second stage, this basis is used to reconstruct the spectral images for all different view and light combinations. While this method is very fast and a practical implementation is rather simple, the restriction to a low dimensional basis limits the use of this method to materials with few linearly independent spectra. We will make a comparison to their method in Chapter 10.4.3 where we deal with measurements of BTFs.

Training-based approaches Moreover, training approaches have been used to reconstruct spectral signals from RGB or similar images. Here, the prior knowledge is not specified in terms of a (closed) mathematical expression, but it is learned based on similar example data that has to be present in a first phase - the training phase. In the second phase, the unknown spectral data can be guessed from RGB data based on the trained settings. The reconstruction quality is then dependent on the similarity between training data and the unknown data to be reconstructed. Multiple authors have focused on reconstruction of single images. Yet, their approaches could also be used in reflectance capture, of course.

In principle the method from Dong et al. [DSD+16] is also training-based, since there the reconstruction is learned based on a single spectral image. However, they benefit from the fact that the spectral image shows the *same* material. In general, other training-based approaches do not require spectral information from the same material but rely on a larger set of training data which needs to be sufficiently similar to the unknown data to reconstruct.

Ngyuen et al. [NPB14] use a radial basis function network to reconstruct spectral reflectance from RGB reflectance. In their method, the basis functions are recovered from the training data using a least squares optimization. For the reconstruction, the weights for the interpolation are then simply obtained by using the RBF kernel on the distance in RGB space.

Arad et al. [ABS16] and Aeschbacher et al. [AWT17] utilize a sparse, overcomplete dictionary of spectra obtained by running the K-SVD algorithm [AEB+06] on the training data. The dictionary entries are then projected to RGB space using the spectral sensor sensitivity. During the reconstruction Arad et al. find the

sparse interpolation weights for each entry of the dictionary using the orthogonal matching pursuit (OMP) algorithm [PRK93] in RGB space. The found coefficients are then used to reconstruct the spectra using the spectral dictionary. In contrast to this, Aeschbacher et al. use the regression method from Timofte et al. [TDSVG14]. By using a sparse, overcomplete basis they circumvent the problems of other approaches that fail with too many linearly independent spectra in a single image/dataset. We will perform a detailed comparison to this method in this chapter and also in Chapter 10.4.4.

Jia et al. [JZG+17] propose to utilize the dimensionality reduction of Zhao and Feng [ZF14] that also allows for reconstruction from low-dimensional data points. They reduce the training spectra to a three dimensional space and then train a compact neural network to map from RGB values to the 3D embedding. A given RGB image can then be mapped using the trained network and afterwards reconstructed using the method from Zhao and Feng.

In 2018, a whole NTIRE challenge [ABST18] focussed on neural-network-based approaches to reconstruct spectral images from RGB images. All methods are based on training a convolutional neural network to the training data. They only differ in the architecture of their networks. Here, the network of the HSCNN+ method from Shi et al. [SCX+18] showed the best results. However, those methods assume that they can distinguish between the different spectra in their training data using the RGB projection only. This causes problems with metameric - or near metameric - colors. Also, if not trained to spectral data, that is very similar to the actual data to reconstruct, they will fail. It is a very cumbersome process to train a CNN to example data (Shi et al. report 38-60 hours on a cluster of 8 GPUs) and also a lot of example data is required for a stable result. For that reason we argue that the deep learning solutions are not well suitable for the case of hybrid data where typically only sparse spectral data (but from the same sample) is available.

Therefore, we concentrate our evaluation on the methods that enable for sensible integration of the spectral data from a hybrid dataset:

- Hardeberg et al. [HSB+99b]
- Murakami et al. [MIYO07]
- Arad et al. [ABS16] and Aeschbacher et al. [AWT17]
- Dong et al. [DSD+16]

Colorization It should be noted, that the process of enhancing spectral resolution from three bands to e.g. 30 bands is highly related with the process of *colorization*. In colorization, all pixels of a grayscale image are given a color e.g. based on user defined example colors for some image regions [LLW04] or based on other, structurally similar color images [ICOL05]. Following this nomenclature, we call our method *spectralization*.

CHAPTER 6

SPECTRALIZATION

To reconstruct a spectral image S we will minimize an energy function $E(S)$. Like in many other approaches, our energy function will consist of two parts: a data term $E_{data}(S)$ and a prior term $E_{prior}(S)$. The data term is comparably simple, just metering compliance of S with the measured data. The prior, however, requires a lot more attention. Moreover, care has to be taken that the terms of E are just linear in the unknowns and only euclidean distance is used. Since S is a complete multi-spectral image, it consists of a tremendous amount of unknown variables rendering other minimization approaches than linear least squares methods infeasible on current computers.

We will hereafter first introduce a common notation for spectralization, then introduce and discuss our prior and afterwards state the exact energy function and explain, how to compute the minimum.

6.1 Preliminaries and Notation

We will now first shortly introduce the notation for this chapter. The unknown spectral dataset (image, BRDF, etc) is denoted by S. S is a $n \times k$ matrix, where n is the number of pixels/directions and k is the number of spectral bands. S_i denotes the i-th row of S and is therefore one spectrum.

The spectral response of the RGB (or similar) camera system describes the relation between RGB and spectral measurements. It is composed of multiple parts like transmission of the optical system of the camera, transmission of the filters and quantum efficiency of the image sensor. It can be summarized by one function $R^i_{cam}(\lambda)$ per color channel i, where λ denotes the wavelength. An incoming spectrum $x(\lambda)$ is mapped to device values by:

$$v_i = \int_\lambda x(\lambda) R^i_{cam}(\lambda) d\lambda \tag{6.1}$$

When discretizing to k wavelength bands, R^i_{cam} can be written as a vector and -

when combining these vectors for all b filters - R_{cam} as a $k \times b$ matrix. Then, the integral can be written as a dot product and therefore the device responses v corresponding to the discretized spectrum s are computed as a matrix product:

$$v = sR_{cam} \tag{6.2}$$

For simplicity of notation we will now drop the index cam from R. The complete spectral dataset S is therefore mapped by the matrix of the broadband camera system to a matrix $V = SR$.

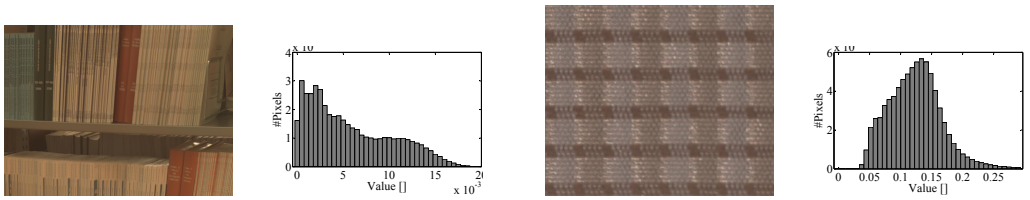
From this mathematical view it is again demonstrated, why there are infinitely many solutions for S when just R and V are known. Mathematically, an inverse matrix for R has to be calculated and an approximation S' to S could be reconstructed as $S' = VR^{-1}$. Due to the fact that $b < k$, the null-space of R is not empty and thus R^{-1} is not unique. One might choose e.g. the pseudo inverse R^{+}, which is the inverse matrix with minimal norm. This was already proposed by Imai and Berns [IB99]. A better solution was introduced by Hardeberg [HSB+99b], who computes an inverse matrix that minimizes the reconstruction error for a previously known set of example spectra. If the example spectra are well related to those being captured, the reconstruction accuracy can be increased. Unfortunately, regardless which inverse matrix R^{-1} is used for this simple reconstruction, the spectra in S' will all be located on a b-dimensional hyperplane in spectral space. If a scene is captured which contains more than b linearly independent spectra, the reconstruction using an inverse matrix introduces a systematic bias.

6.2 A prior on spectral images

As explained before, we will reconstruct spectral images from incomplete information. Without imposing any further restrictions on the image, the reconstruction process has infinitely many solutions. Therefore, we have to introduce additional knowledge and define a prior on spectral images. Some desirable properties are:

1. It should fit a huge class of images.
2. It should be of linear-least-squares nature to allow for efficient minimization.
3. It should be adaptable to special cases.

There have been significant amounts of research spent on (color) statistics of images, e.g. [HBS92], [Hei06], [Koe10], [CZ11]. These statistics can be used in the area of image denoising, reconstruction or restauration. However, it is not easy to integrate such statistical knowledge into a linear-least-squares error function.



(a) Example image 1 (b) Histogram for image 1 (c) Example image 2 (d) Histogram for image 2

Figure 6.1: Checking histograms of natural images: The values are typically not normal distributed. Therefore, the prior used by Murakami et al. [MIYO07] will not work well.

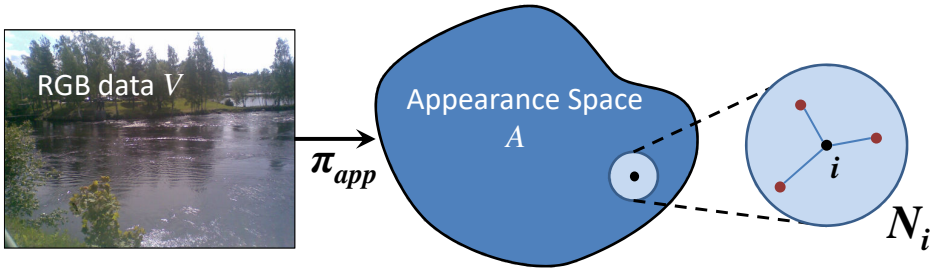


Figure 6.2: Principle of our appearance space: A projection of all pixels i of the RGB image V to the appearance space A is calculated. Afterwards, local neighborhoods in the appearance space are determined for all points, each corresponding to a pixel i.

Murakami et al. [MIYO07] assume that the entries of the spectral image are normal-distributed and simplify the resulting optimization problem in a way that a minimum can be computed easily. Unfortunately, images are in most cases not normally distributed, as it is shown on two example images in Figure 6.1, and therefore this prior introduces stronger bias.

We propose a quite different prior that is more related to typical colorization [LLW04] or editing approaches like AppWand [PL07] or AppProp [AP08]. We assume that the spectral image is - after being projected to an appropriate *appearance space* A - *locally smooth*. We furthermore assume that A only depends on V and not on S.

To be more precise, the pixels from V are mapped to A by means of a function $\pi_{app} : V \rightarrow A$ and then a nearest neighbor search is performed in A recovering a neighborhood set N_i for all pixels i of predefined size. This way, a neighborhood graph is built on all pixels. It is important to note that the implementation of the neighborhood search needs to ensure that $\forall j > i : j \in N_i \rightarrow i \notin N_j$. Otherwise,

the neighborhood graph would often have doubled connections and would also too easily break up in too many, locally connected clusters.

Having defined the neighborhoods, it is assumed that the spectra in each of these neighborhoods are similar to each other. This principle is visualized in Figure 6.2. We figured out, that 3 neighbors per pixels are already sufficient. The prior term does then consist of the distance from pixel i to all pixels in the neighborhood set:

$$E_{prior}(S) = \sum_{i=1}^{n} \sum_{j \in N_i} \gamma_{i,n} \left\| S_i - S_j \right\|^2 \tag{6.3}$$

where $\gamma_{i,n}$ is an additional weight, which might be derived from the distance between $\pi_{app}(V_i)$ and $\pi_{app}(V_n)$. However, it turned out that setting $\gamma_{i,n} = 1$ is a good choice for a small size of N_i. This prior term fulfills aim (2) because it is linear-least-squares and moreover easy to evaluate since only a small number of neighbors per pixel have to be taken into account.

The crucial part is the choice of the appearance projection function π_{app}. Fortunately, the choice does not affect the performance during minimization of $E(S)$ because the projection to A and the construction of the neighborhoods N_i are performed in a pre-processing step only once. It is therefore possible to choose quite complex functions π_{app} to obtain better results. Furthermore, a suitable π_{app} can be selected for each spectral image class, whenever it is known in beforehand. This makes our choice fulfill aim (3).

We will now first discuss some possible choices for the appearance projection function and will afterwards evaluate their quality on a database of spectral images to check whether it is suitable for a broad range of typical spectral images - or in other words - check fulfillment of aim (1).

The simplest choice for the appearance space is the RGB color space. The corresponding appearance mapping function is therefore the identity function:

$$\pi_{app}^{RGB} := id \tag{6.4}$$

Since the three RGB values can be seen as a rough sampling of the spectrum (despite the fact, that the three RGB filters might heavily overlap), it is natural, that in many cases similar RGB values also correspond to similar spectra. It should be noted that the RBF interpolation in [NPB14] has a quite similar assumption since it uses a distance in the RGB space between the measured RGB value and the cluster centers to derive the interpolation weights.

Of course, problems can occur with two or more metameric surfaces in the scene. Many appearance neighborhoods would likely cover pixels from more than one region leading to the reconstruction of an average spectrum there. Fortunately, the RGB image contains more valuable information than the individual

RGB triples. Simply using the RGB space as an appearance space discards all spatial and therefore textural information. One simple extension could be to include the spatial coordinates of the RGB pixels into the appearance space by defining:

$$\pi_{app}^{RGBXY}(V_i) := [V_i, x_i, y_i] \tag{6.5}$$

where x_i and y_i are the (appropriately scaled) x and y coordinates of pixel i. Then, two metameric regions, which are not spatially adjacent, would get separated in the corresponding 5D appearance space.

Another option is to include the RGB triplets of neighboring pixels into the appearance space e.g. by extracting small patches surrounding the pixel i. In the area of texture synthesis this is a well known approach to identify similar image regions. The corresponding appearance projection can be defined like this:

$$\pi_{app}^{patch,m}(V_i) := \begin{pmatrix} V_{idx(x_i-m,y_i-m)} & \cdots & V_{idx(x_i+m,y_i-m)} \\ \vdots & \ddots \quad \iddots & \vdots \\ \vdots & \cdots \quad \alpha_{center}V_i \quad \cdots & \vdots \\ \vdots & \iddots \quad \ddots & \vdots \\ V_{idx(x_i-m,y_i+m)} & \cdots & V_{idx(x_i+m,y_i+m)} \end{pmatrix} \tag{6.6}$$

where idx maps two pixel position x, y back to the pixel index. We weight the RGB value of the center pixel itself equally to the sum of all surrounding pixels by setting $\alpha_{center} = ((2m+1)^2 - 1)$.

It should be emphasized again, that other choices for the appearance space are also possible which might better fit to more specific image classes.

To evaluate, how much error is introduced by E_{prior} depending on π_{app} we evaluated a perceptual ΔE_{94}^* and a relative error in all neighborhoods N_i on real spectral images. For this, we used several multi-spectral image databases [NFF02], [FANA04], [CZ11] as well as a set of self-captured textures[1]. These image databases represent a wide class of objects and materials because they include natural, urban and indoor scenes. To enhance our multi-spectral texture database, we also extracted some stone and wood textures from the Harvard database [CZ11]. This wide variety can be seen in the selection of example images depicted in Figure 6.3.

For each image in the database we calculated $V = SR$. To determine the impact of the spectral response R of the RGB camera, we performed all tests using two different cameras. One was a industrial computer vision camera (SVS Vistek 4022) and one a professional SLR camera (Kodak DCS 760). The spectral responses of these cameras are shown in Figure 6.4. The DCS 760 shows more

[1] http://cg.cs.uni-bonn.de/en/projects/spectral-btf-measurement/

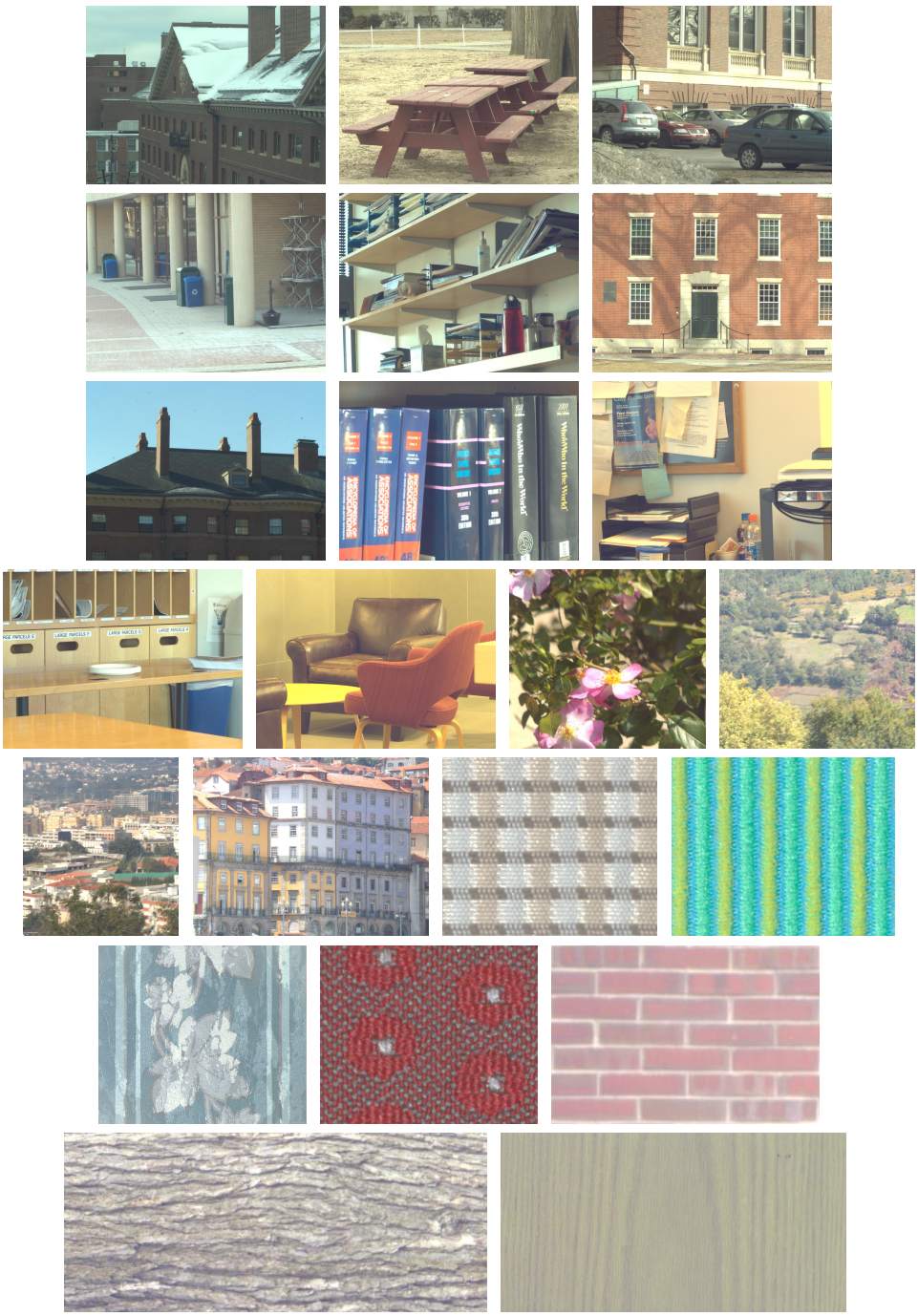
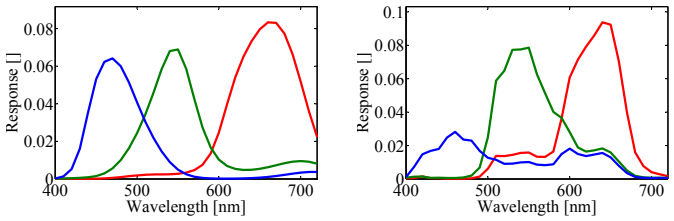


Figure 6.3: Example images from the image databases use throughout our evaluation. The databases cover many classes of images of outdoor and indoor man-made environments or vegetation as well as textures of different materials.



(a) SVS Vistek 4022 (b) Kodak DCS 760

Figure 6.4: Spectral response of two different RGB cameras. The DCS 760 shows much stronger overlap of the filters, especially for the blue one.

overlap between the filters - especially the ”blue” filter is very broad. After calculating $\pi_{app}(V_i)$ for all i we determined the three nearest neighbors N_i for all pixels i. This search can be speed up using a space partitioning technique like a kd-tree. Then we calculated the average error for all pixels i:

$$\Delta_i = \frac{1}{|N_i|} \sum_{j \in N_i} \Upsilon(S_i, S_j) \tag{6.7}$$

We used two different error functions Υ to measure the distance of two spectra in S. The first one is the well known $\Delta E^*_{94} =: \Upsilon_{\Delta E}$, which is perceptually linear but reduces to trichromatic space for color comparison. Since the illumination has to be known for this formula, we were not able to compute it on the Harvard database [CZ11]. For the textures, we evaluate the ΔE^*_{94} for a set of selected CIE lamps covering a broad range of everyday illumination conditions (A, D65, D50, D75, FL1, FL4, FL8, FL12, HP1, HP3) and compute an average from the resulting, single ΔE values.

The second error function is the mean relative error in percent:

$$\Upsilon_{rel}(x, y) = \frac{100}{k} \sum_{i=1}^{k} \frac{|x_i - y_i|}{|x_i| + \epsilon} \tag{6.8}$$

where ϵ is a small constant (we have chosen 10^{-3}) to avoid divisions by zero and to down-weight dark and therefore noisy pixels. We will refer to this error measure as *relative error* in the discussion of the results.

Figure 6.6 and Figure 6.7 show error histograms for the ΔE^*_{94} distance function on non-texture and texture images respectively. Figure 6.8 and Figure 6.9 contain the same kind of results for the relative error. The respective upper-left plot of each group shows the error when assuming full knowledge about the spectral image S and setting $\pi_{app}(V_i) = S_i$. There are three effects that lead to errors on the ground-truth data:

- noise on the acquired data
- pixels who do not have similar pixels in the image at all (e.g. small highlights)
- the nearest neighbor search in the space of image spectra to determine the N_i is run using euclidean distance by using a approximate kd-tree search (using FLANN [ML09]). This is of course worse compared to searching for neighbors having lowest ΔE^*_{94} or lowest relative error. That would require a brute-force search being prohibitively slow. Especially, the whole effect can be noticed in Figure 6.6 where the RGB appearance space with the Kodak DCS 760 used for plot (b) is even producing a better result than the search on the spectral space.

Small highlight pixels are extremely seldom and most of the error is introduced in dark regions where acquisition noise has larger impact in relation to the signal. This can be observed in the example images in Figure 6.5. The errors are generally lower for our textures since they were recorded using multiple exposures (HDR imaging) leading to less acquisition noise.

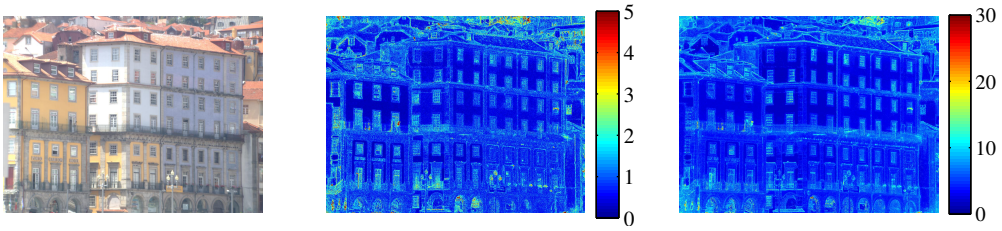
The mean errors for both kinds of images and for both error measures are summarized in Figure 6.10. It is remarkable, that the simple RGB space of the Kodak DCS 760 is performing extremely well in the ΔE^*_{94} error, despite of the fact that the spectral sensitivity (see Figure 6.4(b)) has strong overlap. When we compare the results to those evaluated with the relative error, where the DCS 760 performs worse, we see that this effect is caused by a better match between the XYZ primaries and the cameras response. All results for the ΔE^*_{94} are only provided to gain a vivid feeling on the overall level of errors. It cannot be used to really judge the errors on the whole spectrum since the spectrum is reduced to a point in a tristimulus space ($L^*a^*b^*$) before evaluating the difference. We will therefore continue to just evaluate errors using the relative error in the following parts of this thesis.

Looking at the relative errors (second row of Figure 6.10) one can observe that the restriction to RGB introduces additional error compared to the ground-truth case. Nevertheless, that additional error is not very significant compared to the inherent errors. Furthermore, we deduce that the spatial structure of the RGB image helps to improve the prior quality, because the errors get reduced when either the XY coordinates of the pixel i are included in the appearance space or where a small patch around pixel i is considered. However, since this gain is very small and since taking into account patches is computationally expensive, it is mandatory to use the π_{app}^{RGBXY} appearance space projection.

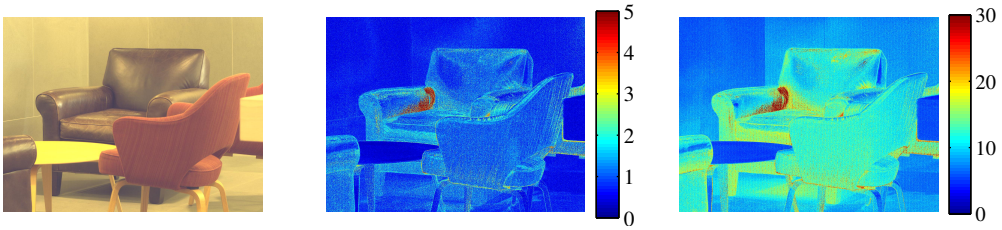
Furthermore, it is remarkable that the error added by the prior in comparison to the ground-truth case is similar for both textures and non-textures. This is

astonishing at a first glance, because we do not expect the same level of self-similarity in non-textures. However, we do only assume *local* smoothness and make no assumption about the global image structure (like Murakami et al.). Since the non-texture images consist of multiple surfaces each in turn exhibiting self-similar content, our prior is applicable in both cases.

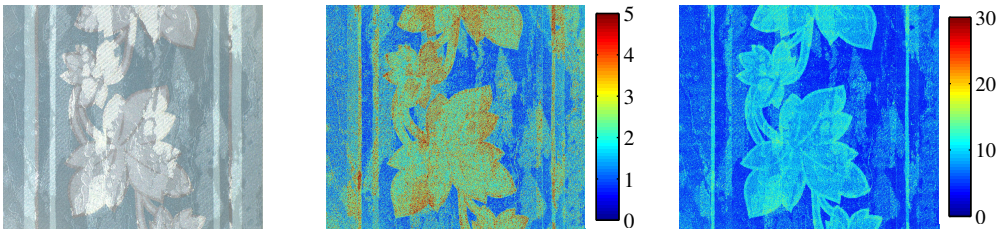
Summarizing, the prior based on the local smoothness assumption introduces some additional error compared to the ground-truth case. Nevertheless, the measurement noise stays the primary source of errors and we therefore conclude that our choice for the prior is viable. Moreover, it works on texture and non-texture images, allows for adaption to special cases and furthermore sticks to the linear-least-squares restriction allowing for efficient minimization.



(a) Outdoor image from [FANA04]

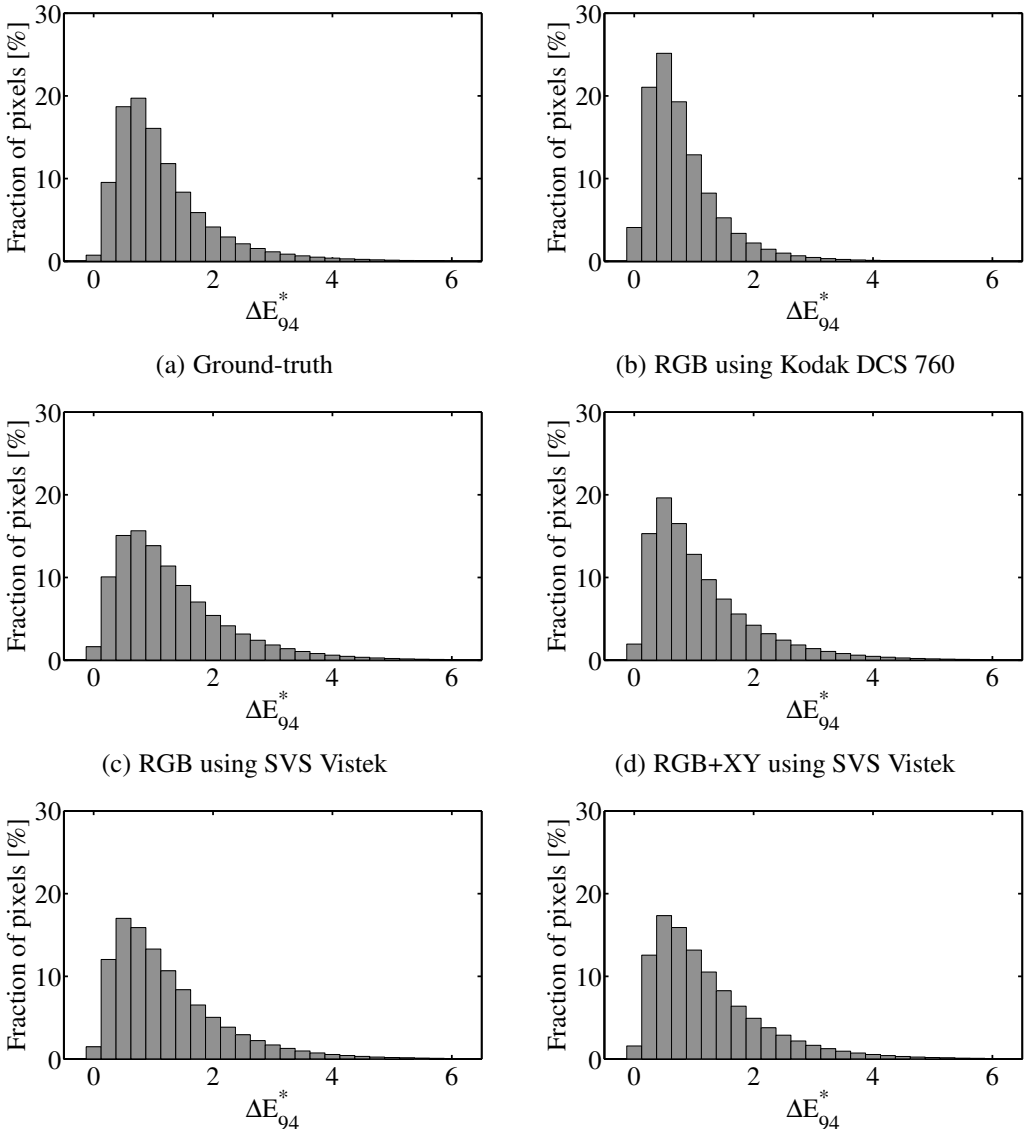


(b) Indoor scene from [CZ11]



(c) Texture from our own database

Figure 6.5: Three example images from the evaluation of the prior quality from three different categories (outdoor, indoor and textures) for the Vistek camera on simple RGB appearance space. Left column: original image mapped to sRGB, middle column: ΔE^*_{94}, right column: relative error.
The error introduced by the prior is small on most surfaces. In darker regions the error is dominated by the acquisition noise in the data. When comparing to the light surfaces it suggests that the primary source of inherent local non-similarity in the appearance space results from the acquisition noise.



(a) Ground-truth

(b) RGB using Kodak DCS 760

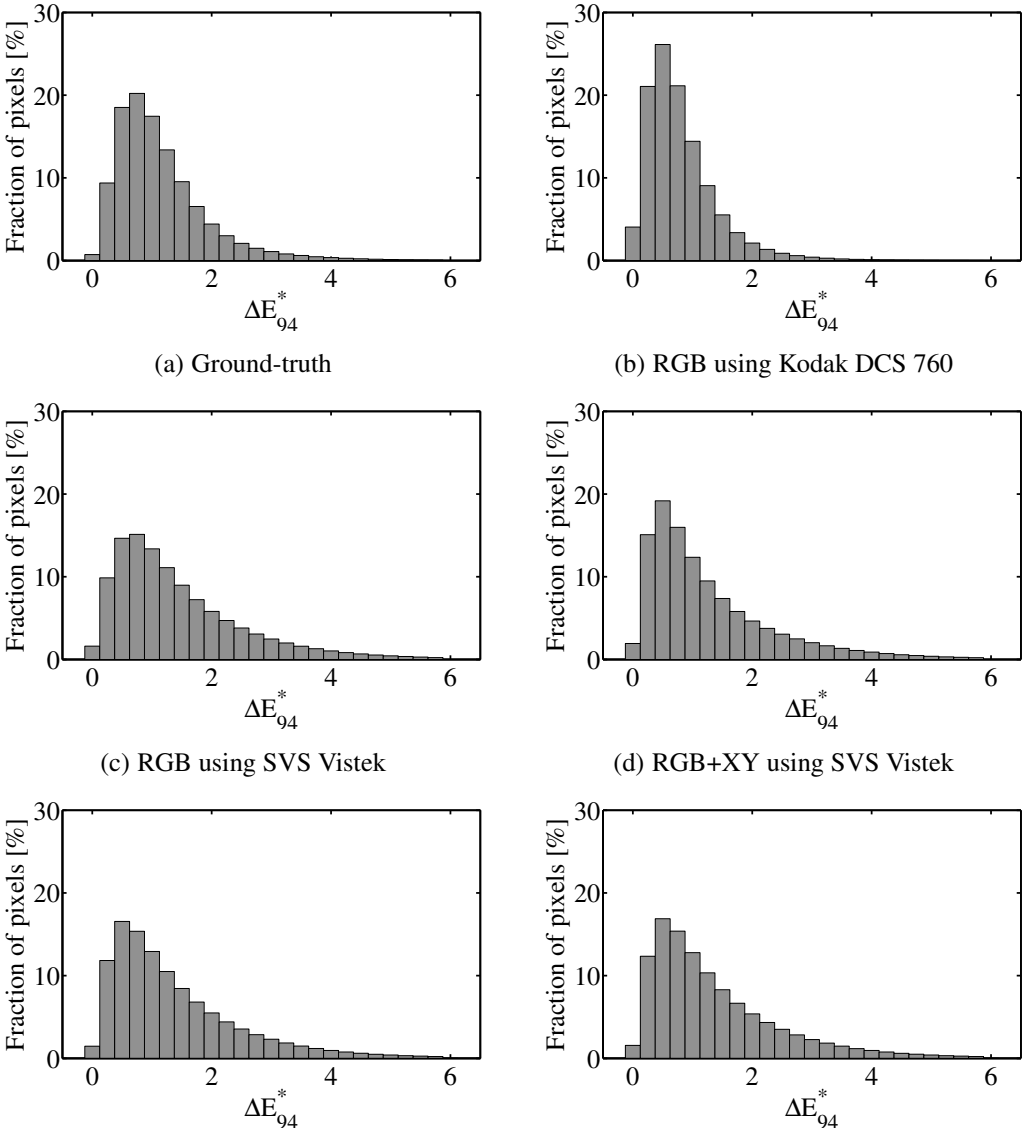
(c) RGB using SVS Vistek

(d) RGB+XY using SVS Vistek

(e) 3x3 patch using SVS Vistek

(f) 5x5 patch using SVS Vistek

Figure 6.6: Checking the smoothness prior on non-texture images: Accumulated ΔE^*_{94} error histograms. The upper left histogram shows that the ground-truth data does already contain some noise and local non-similarities. The reduction to trichromatic space is introducing nearly no additional error when using the ΔE^*_{94}. The RGB appearance space with the Kodak DCS 760 used for plot (b) is even producing a better result. See the respective text for a discussion on this effect.



(a) Ground-truth

(b) RGB using Kodak DCS 760

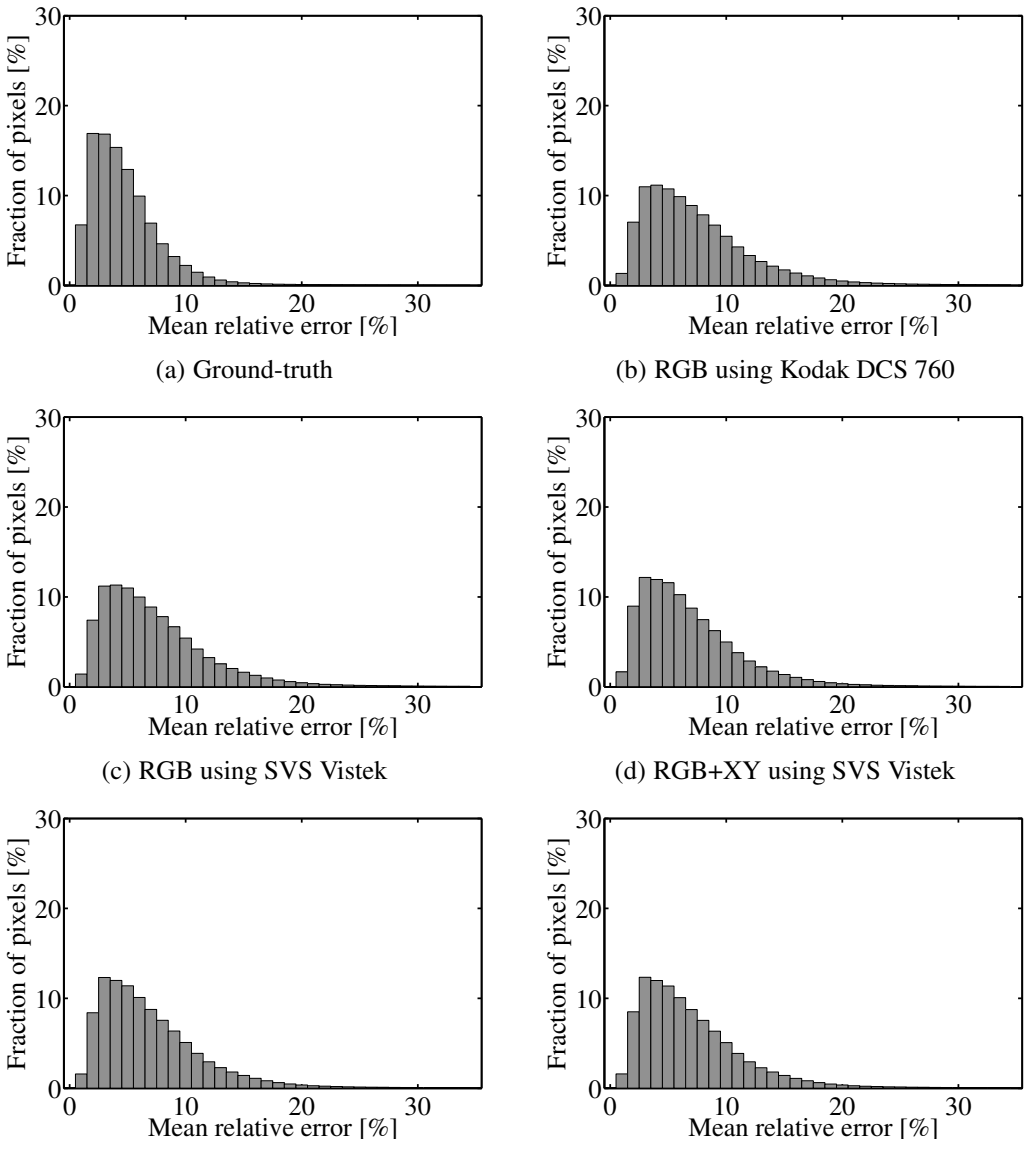
(c) RGB using SVS Vistek

(d) RGB+XY using SVS Vistek

(e) 3x3 patch using SVS Vistek

(f) 5x5 patch using SVS Vistek

Figure 6.7: Checking the smoothness prior on texture images: Accumulated ΔE^*_{94} error histograms. The upper left histogram shows that the ground-truth data does already contain some noise and local non-similarities. The reduction to trichomatic space is introducing some additional error, which may be reduced using the spatial information inside of the data.



(a) Ground-truth

(b) RGB using Kodak DCS 760

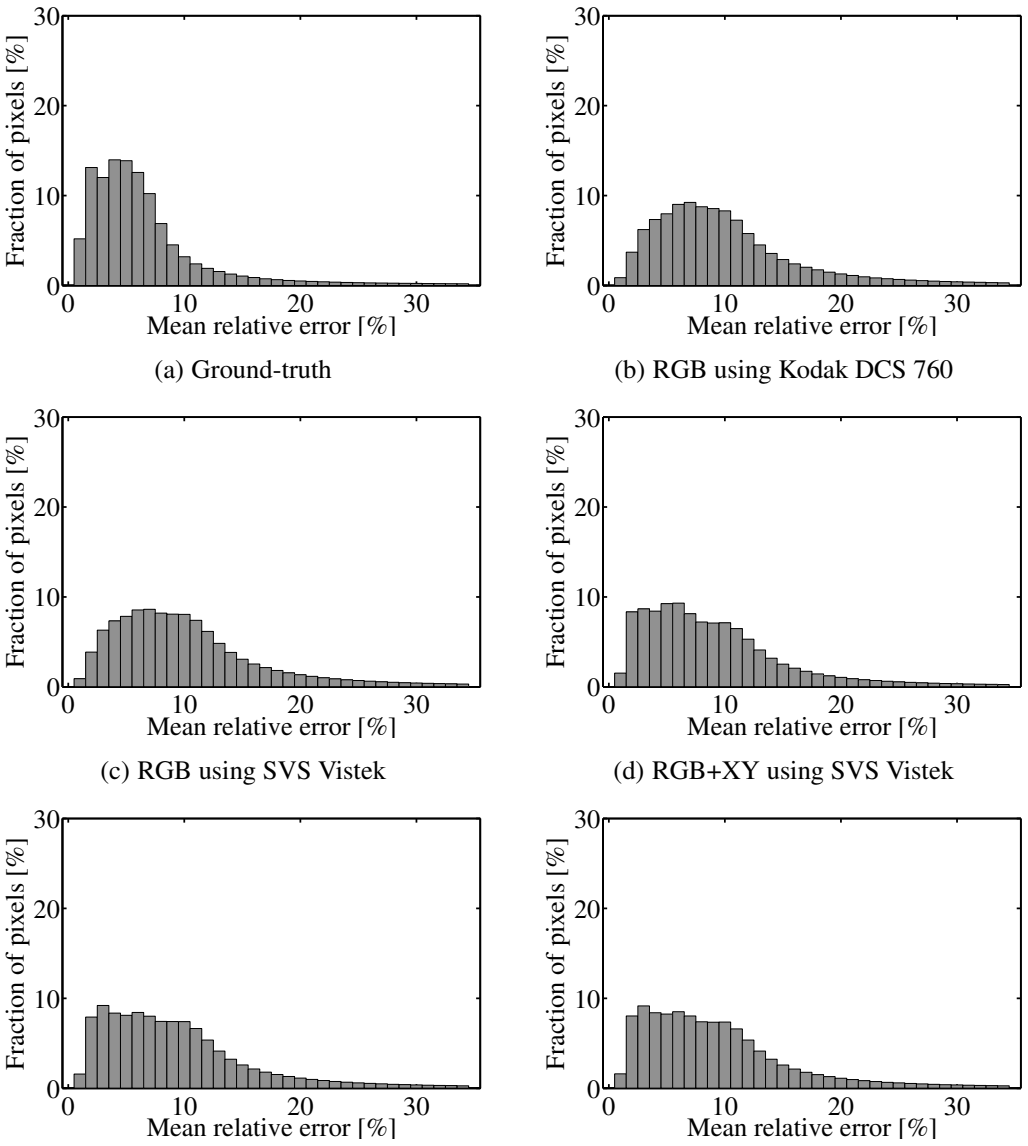
(c) RGB using SVS Vistek

(d) RGB+XY using SVS Vistek

(e) 3x3 patch using SVS Vistek

(f) 5x5 patch using SVS Vistek

Figure 6.8: Checking the smoothness prior on non-texture images: Accumulated relative error histograms. The upper left histogram shows that the ground-truth data does already contain some noise and local non-similarities. The reduction to trichomatic space is introducing some additional error, which may be reduced using the spatial information inside of the data.



(a) Ground-truth

(b) RGB using Kodak DCS 760

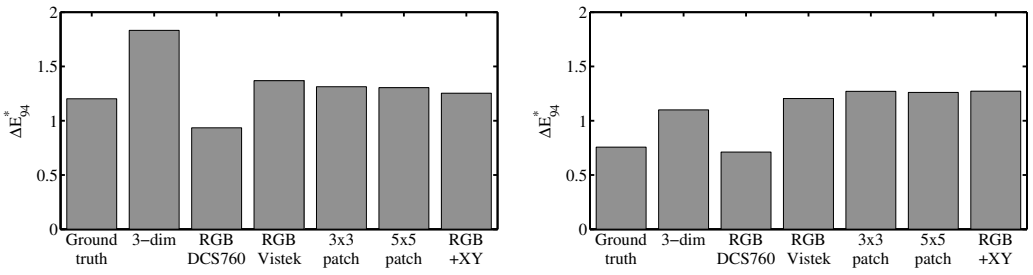
(c) RGB using SVS Vistek

(d) RGB+XY using SVS Vistek

(e) 3x3 patch using SVS Vistek

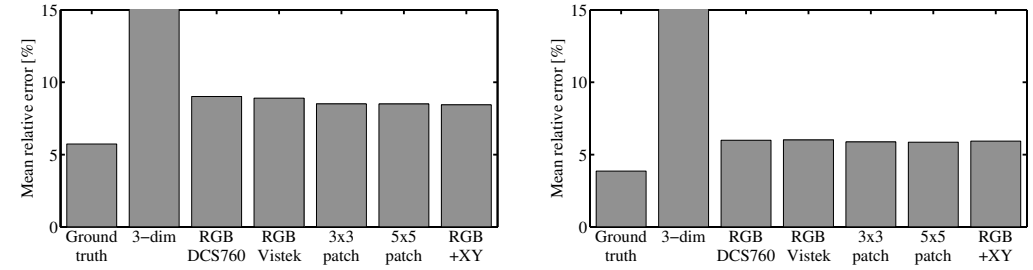
(f) 5x5 patch using SVS Vistek

Figure 6.9: Checking the smoothness prior on texture images: Accumulated relative error histograms. The upper left histogram shows that the ground-truth data does already contain some noise and local non-similarities. The reduction to trichomatic space is introducing some additional error, which may be reduced using the spatial information inside of the data.



(a) ΔE^*_{94} on non-textures

(b) ΔE^*_{94} on textures



(c) Relative error on non-textures

(d) Relative error on textures

Figure 6.10: Summary smoothness prior: The reduction to RGB introduces only small additional errors when our prior is used. For comparison purposes, we included the error on a reconstruction assuming that the spectra inside of the images are on a 3-dimensional hyperplane in the spectral space (second bar). When measuring in tristimulus space using the ΔE^*_{94}, the errors seem to be acceptable. However, when metering the relative error on the spectral data, the problems with this assumption become obvious (the errors are out of the display range here, being 150% on textures and 250% on non-textures). For our prior, both ΔE^*_{94} and relative errors are only moderately higher than the noise-introduced errors on the ground-truth data. It can be seen that using the spatial information inside of the RGB data generally helps to reduce the errors a bit. The errors on the textures are generally lower because this data contains less noise. The ΔE^*_{94} errors for the DCS 760 are generally lower because of the good match between the CIE XYZ primaries and the camera‘s sensitivity.

6.3 Energy function

After having introduced and discussed our prior term, we are now able to formulate the whole energy function for our reconstruction technique. For this, the data term has still to be defined. One part of the data term has to include the available RGB (or similar) data and has to measure conformance of the unknown spectral image with the measured RGB values. The second part has to incorporate the additional spectral knowledge which we will label as D having m scalar-valued entries. Any measurement of a device can be understood as an integral over the spectral data cube (in case of an image) multiplied with a sensitivity function for the respective device or device pixel. For example, the pixel values of a monochrome camera would correspond to integrals over the spectral dimension of the cube multiplied with the overall spectral sensitivity of the camera system. In this sense, all kinds of (spectral) measurement devices provide integrals and can therefore incorporated into the energy function in the same manner.

The most general form of our energy function can be stated as follows:

$$\begin{aligned} E(S) &= \sum_{j=1}^{m} \alpha_j \left\| \left(\sum_{i=1}^{n} \sum_{l=1}^{k} w_{i,l,j} S_i \right) - D_j \right\|^2 \\ &+ \beta \sum_{i=1}^{n} \|S_i R_{cam} - V_i\|^2 \\ &+ \gamma \sum_{i=1}^{n} \sum_{j \in N_i} \|S_i - S_j\|^2 \end{aligned} \tag{6.9}$$

Here, the first term accounts for the spectral data, the second for the RGB data and the third one is the previously introduced prior term. α, β and γ are factors to weight the terms. The $w_{i,l,j}$ are the sensitivity weights for the (spectral) measurement point D_j, pixel i and wavelength band l. Obviously, the RGB-term could also be expressed in the form of the first term, but we will keep it separated to emphasize the denseness of V.

We will now first consider a special case for spectralization by assuming, that the spectrum of certain pixels is known with the desired target resolution - e.g. by using a line-scan or a whisk-broom spectral imaging system as in [MIYO07]. In

this case the energy function simplifies to:

$$\begin{aligned} E(S) &= \alpha \sum_{j=1}^{m} \left\| S_{P(j)} - D_j \right\|^2 \\ &+ \beta \sum_{i=1}^{n} \left\| S_i R_{cam} - V_i \right\|^2 \\ &+ \gamma \sum_{i=1}^{n} \sum_{j \in N_i} \left\| S_i - S_j \right\|^2 \end{aligned} \tag{6.10}$$

Here, the P function retrieves the pixel index for the j-th spectral measurement point. In the rest of this part of the thesis we will analyze the performance of our spectral reconstruction for the cases of a whisk-broom device like in [MIYO07] as well as a push-broom device scanning one or more lines of each image.

6.3.1 Choice of Weights

For a good reconstruction quality, a viable choice of the three weights α, β and γ is crucial. Since the three terms of the energy function are of very different scale, we propose the following scheme to weight all three terms equally in a first step:

$$\begin{aligned} \alpha &= \alpha' \frac{mk + 3n + k \sum_{i=1}^{n} |N_i|}{mk} \\ \beta &= \beta' \frac{\bar{D}}{\bar{V}} \frac{mk + 3n + k \sum_{i=1}^{n} |N_i|}{3n} \\ \gamma &= \gamma' \frac{mk + 3n + k \sum_{i=1}^{n} |N_i|}{k \sum_{i=1}^{n} |N_i|} \end{aligned} \tag{6.11}$$

The large fractions compensate for the differences in the number of terms. Since the first and the third term measure distance in spectral space directly, we have to add further adjustment only to the second one. To bring the second term to the same scale, we divide by the mean of the RGB data $\bar{V}$ and multiply with the mean of the spectral data $\bar{D}$.

Then, the new weights α', β' and γ' determine the total weight of each term. A good choice is to weight all three equally by setting $\alpha' = \beta' = \gamma' = 1$. All our results have been obtained by this weighting.

6.4 Minimization

Our energy function (see Equation 6.9) is designed to be linear in the entries of the spectral image S and uses only squared error terms. The minimization $S_{recon} =$

$argminE(S)$ can therefore be expressed using matrix notation:

$$S_{recon} = argmin \left\| M_E S - b_E \right\|^2 \tag{6.12}$$

where M_E and b_E are a matrix and a vector respectively constructed from the energy function and $\|.\|$ denotes the standard euclidean distance. The absolute minimum can be found using closed solutions e.g. based on QR factorization. Unfortunately, the matrix M_E is of size $(m+bn+k\sum_i |N_i|)\times(kn)$ when considering the general form of the energy function. For an image having $n = 1$ Megapixel resolution, using $k = 30$ spectral bands, having $m = \frac{1}{100}n = 10000$ and using 3 appearance neighbors per pixel this results in a matrix of size 93,010,000 $\times$ 30,000,000. This matrix size currently prohibits the use of explicit solvers.

Therefore, an iterative solver has to be used. We implemented our method both using a conjugate gradient (CG) [HS52] method and the LSQR algorithm [PS82]. In both cases the matrix M_E does not need to be stored explicitly, it is sufficient to supply a function that computes the multiplication of M_E with a vector. Fortunately, the matrix is extremely sparse and therefore a multiplication can be computed efficiently. In case of the CG algorithm the matrix has to be quadratic and positive definite. Therefore we have to solve the *normal equation* instead: ${M_E}^T M_E S = {M_E}^T b$. ${M_E}^T M_E$ fulfills the desired properties but the normal equation has a worse condition when evaluated on actual computers having limited numerical accuracy. For the implementation an additional function has to be supplied that performs a multiplication of ${M_E}^T$ with a vector.

We used an own implementation of both algorithms in a MATLAB[2] environment. Both algorithms would have been readily available in MATLAB, but we tailored an implementation for our specific needs. First of all we avoided storing multiple vectors of result size since this is in our case a whole spectral image. Furthermore, both algorithms have to compute dot products of vectors of large length, more specific $k \times n$. To ensure numerical stability we implemented those using a tree-reduction scheme.

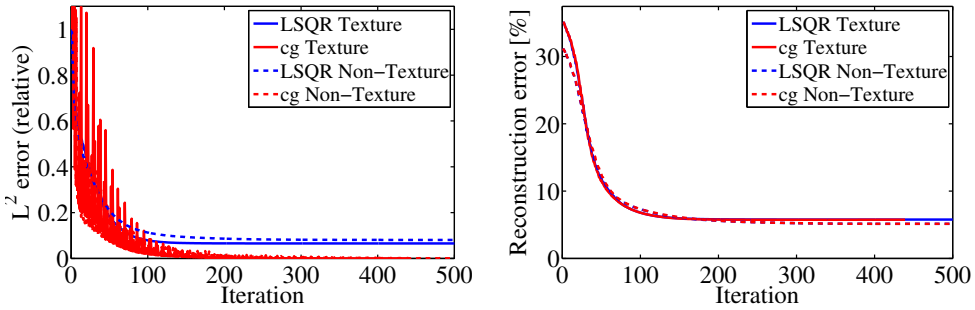
6.4.1 Convergence

To decide which minimization algorithm provides better results in terms of accuracy and computing speed we evaluated our implementation using one texture and one non-texture image. Figure 6.11 shows the convergence behavior using the LSQR and the cg solver for both cases. It is notable, that the cg algorithm does not converge strictly whereas the LSQR algorithm reduces L^2 error continuously. This is mainly due to the fact that the numerical stability when solving the normal equation with cg is worse than solving the optimization problem directly in LSQR.

[2] http://www.mathworks.com

However, it can be seen in (b) that both algorithm converge to approximately the same final solution. Moreover, it is notable that the convergence behavior is nearly identical for both images, even though being of very different content. Finally, the results indicate that convergence in the L^2 error is highly correlated with convergence in the relative reconstruction error. We can therefore rely on standard detection for convergence in the algorithms.

Since both algorithm require roughly the same amount of computation time per iteration in our MATLAB implementation, no recommendation can be made. We used the LSQR algorithm in all our further results.



(a) L^2 (b) Relative error

Figure 6.11: Convergence of spectralization using two different solvers: (a) shows the L^2 error with solver iterations. For comparison purposes, the curves for cg and LSQR have been normalized to start at one. In (b) the resulting reconstruction error for each solver iteration is depicted. There is a good correspondence between convergence in the optimizer's L^2 error and the reconstruction error.

6.4.2 Computation Speed

The speed per iteration is highly determined by the matrix-vector multiplications $M_E S$ (or ${M_E}^T M_E S$ for the cg algorithm). Here, the part ensuring similar spectral neighborhood in appearance space requires most time. The setting for the size of the N_i in Equation 6.9 is therefore crucial and should be chosen wisely. Through experiments we determined that a neighborhood size of 3 is already sufficient for good reconstruction results and does also not waste too much computing time.

In our implementation, the multiplications are implemented in C++ and included in MATLAB as a module. We used a Intel Core i7-3770 CPU with 3.4 GHz clock. With this setup an image having one megapixel resolution and 33 spectral bands can be reconstructed in about 840 seconds. Of this, 44 seconds are used to construct the appearance space neighborhoods and the remaining approx. 800 seconds are spent on 500 LSQR iterations.

CHAPTER 7

EVALUATION OF SPECTRAL RECONSTRUCTION QUALITY

We analyzed the reconstruction quality of our method using the same image databases as in 6.2. Furthermore we added 2 BRDFs where we assume spectral knowledge for just a very small subset of directions. The BRDFs are extracted from the Colorchecker dataset measured during our ground-truth experiments (see Figure 4.1(b)) and are equally sampled over the hemisphere at 81x81 directions.
For all datasets the complete multi-spectral ground truth was available so that the reconstruction quality can be reliably rated.The RGB device values are generated artificially using the camera response R_{cam} by:

$$V = SR_{cam} \tag{7.1}$$

The spectral data values D are generated by averaging over the respective image spots. Afterwards gaussian noise with $\sigma = 1\%$ of mean signal is added to the synthesized RGB data V. Since the spectral data for all our test datasets is captured using real cameras, it is inherently biased with noise. For this analysis we will assume perfect registration of RGB and spectral data. In Chapter 10 we will discuss the issue of possible misalignment.

We evaluated our method as well as the methods from Imai [IB99], Smits [Smi99], Hardeberg [HSB$^+$99b], Murakami [MIYO07], Dong [DSD$^+$16], and Aeschbacher [AWT17] on all test datasets.

The methods from Imai and Hardeberg are inverse matrix methods whereas the method from Smits is heuristic and tries to combine precomputed smooth spectra for red, green, blue, cyan, magenta, yellow and white to get the desired match with the measured RGB triple as well as a smooth spectrum. It is therefore only applicable in cases where material reflectance is measured. The method from Hardeberg uses prior knowledge of imaged spectra to find a best inverse matrix. Dong et al. create a similar three dimensional basis using a non-negative matrix factorization. This is in contrast to Hardeberg, who uses a generic matrix factorization. Spectral A+ from Aeschbacher et al. works rather similar to the methods of Hardeberg and Dong in the sense that they also create a linear basis

of the imaged spectra. However, Spectral A+ uses a K-SVD to create an overcomplete basis and is therefore not limited to three dimensional subspaces of the spectral space as the Hardeberg, Imai, and Dong methods. In our evaluation we pass the known spectral curves into all of the methods except for the Imai and Smits method, who cannot make use of them at all.

None of the aforementioned methods makes use of the spatial location of the known spectral data. This is incorporated into the Murakami method. It is - though being derived in a different way - an extension of the Hardeberg method and adds an additional correction term to improve results exploiting that knowledge about spatial positions.

As explained in section 6.2 we will use the relative errors to judge and compare the reconstruction quality and we will use the same spectral image databases, here. To simplify comparison, we evaluated our method for the spectral response of the Vistek camera only. Results for other cameras are very similar.

7.1 Whisk-Broom Data

One possibility to supply sparse spectral data is to use a whisk-broom imager that scans just certain spots distributed over the image. This was also proposed by Murakami et al. [MIYO07]. For our analysis we distributed 250 spectral spots in a regular grid over the image. The spots are square having a size of 0.5% of the image width.

Figure 7.1 shows the resulting average reconstruction errors of all methods mentioned above as well as our spectralization method. For our method we used the simple RGB appearance space. The errors were evaluated on the non-texture images of the database - though we will see that the performance on texture images is very similar.

We first notice that the two very simple methods from Imai and Smits deliver a very poor reconstruction quality. Furthermore, the Spectral A+ method did also not perform very well. This is rather astonishing since it is technically very sound and should also not suffer from problems of high dimensionality due to the overcomplete basis. We have tried various parameter sets for this method, but the result got worse if not using the parameters proposed by Aeschbacher et al. [AWT17] (sparsity $k = 8$, regularization $\lambda = 0.1$). We assume that the 250 spectral points are not enough to create a good basis. We also refer the reader to Section 10.4.4, where we compare reconstructed BTF data results.

Figures 7.2 and 7.3 go into more detail on the error evaluation. The respective upper left figures show the comparison of mean reconstruction errors whereas the others contain error histograms for the best performing methods. We notice that the Hardeberg method, though not using the spatial position of sparse spectral

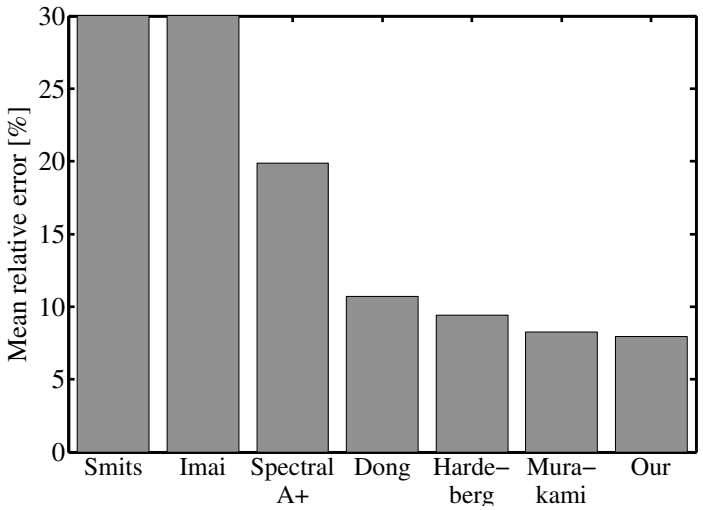


Figure 7.1: Reconstructing non-texture images using whisk-broom data: Mean relative error. 250 spectral data points on a regular grid have been used as sparse spectral input.

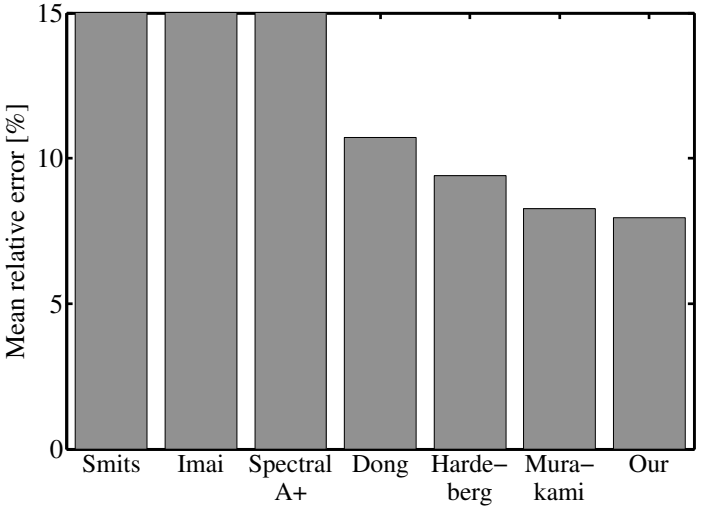
information, performs already quite well on average. The method from Dong et al. is rather similar to it, but it uses non-negative matrix factorization to build the spectral basis functions. It seems that this is - though being a sensible idea from a physical standpoint - a slightly worse choice. The Murakami method improves - as is expected due to the additional correction term that utilized the position of the spectral measurement spots - a bit on the Hardeberg and Dong methods. Finally, our method even outperforms the Murakami method in average and also in error distribution.

When comparing texture and non-texture results we notice, that the error on textures is generally lower than on non-textures for all methods, but that the shape of the histogram looks very similar. The reason for the lower mean error is again (see also Section 6.2) the lower noise level in the spectral reference data of the textures. All methods tend to produce spectra that are smoothed to some degree and therefore more noise-free. In real-world applications, this is a valuable property. The reason for the similar histogram shape is - as explained in Section 6.2 - the fact that non-texture images consist of locally similar image content, as well.

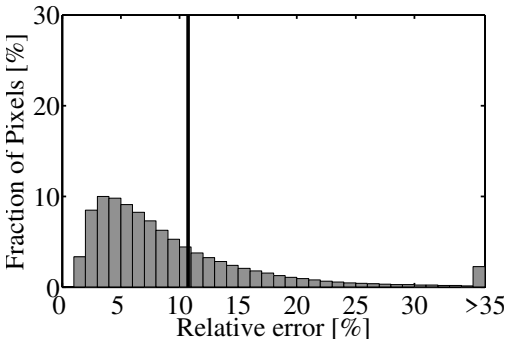
To deepen our insight into the reconstruction achievable by the different methods we will analyze some of the images in greater detail here. We will first concentrate on the texture of a complex fabric from our spectral textures database. The resulting reconstruction errors of the different reconstruction methods can be seen in Figure 7.4. The methods from Imai and Smits produce very high errors as they do not incorporate additional spectral knowledge. Spectral A+ does also not

perform well - to our very surprise as mentioned earlier. The methods from Hardeberg and Dong perform already quite well, but they can only handle three of the four kinds of threads with high accuracy. That is due to the limitation to a three dimensional subspace of the spectral space. The Murakami method can improve on this result due to an additional correction term incorporating the position of the spectral data points. However, our method performs even better on this dataset which is most likely due to our better image prior.

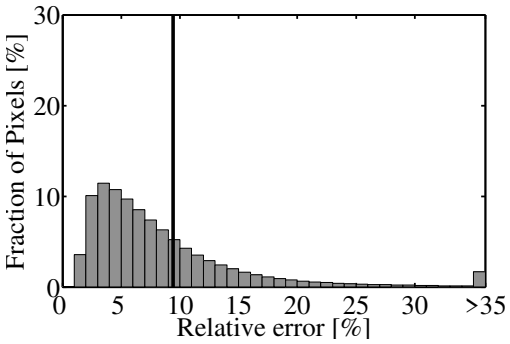
Another example is depicted in Figure 7.5. Again, a texture with more than three different basis materials is used and the Hardeberg and Dong methods fail to reconstruct one of the materials whereas our method performs well on the whole texture. The only difference to the previous example is that the Murakami method fails to improve on the result of the Hardeberg method and produces a worse result on average. The reason for this is most likely the prior used in the Murakami method. It is assumed, that the pixel values have a normal distribution, globally. As the histogram in 7.5(g) shows, the histogram of this texture is highly non-normal.



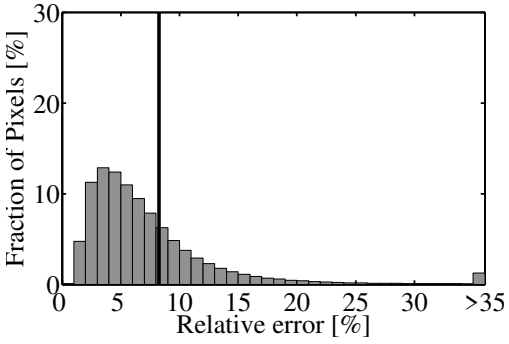
(a) Summary



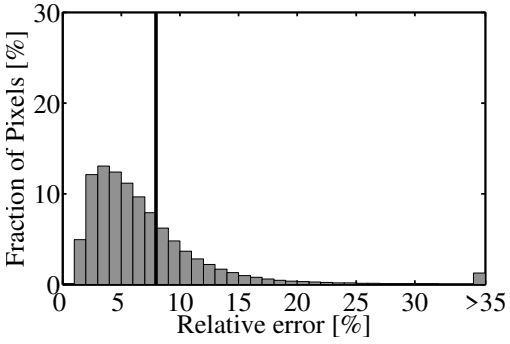
(b) Dong



(c) Hardeberg

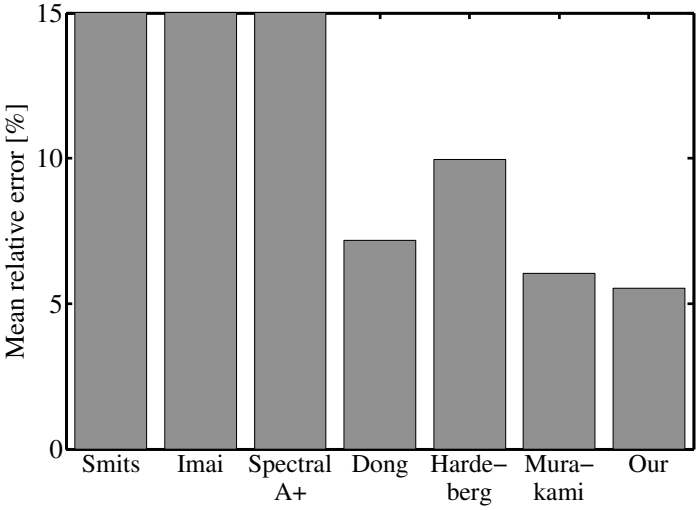


(d) Murakami

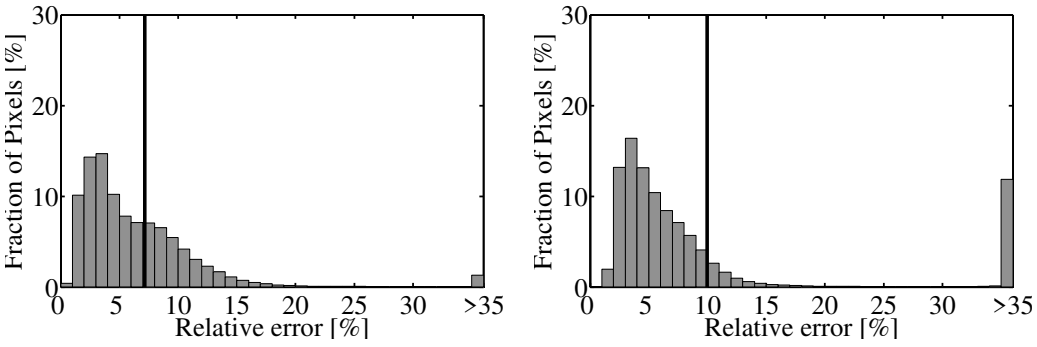


(e) Our method (RGB appearance space, SVS Vistek)

Figure 7.2: Reconstructing non-texture images using whisk-broom data: Accumulated relative error histograms. 250 spectral data points on a regular grid have been used as sparse spectral input.

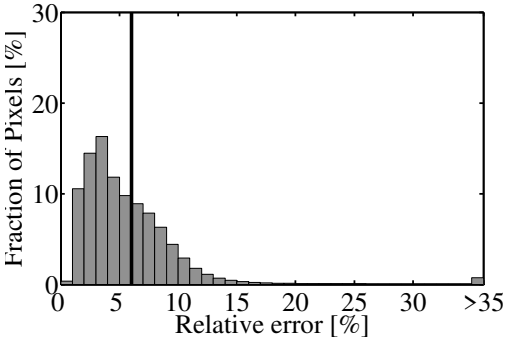


(a) Summary

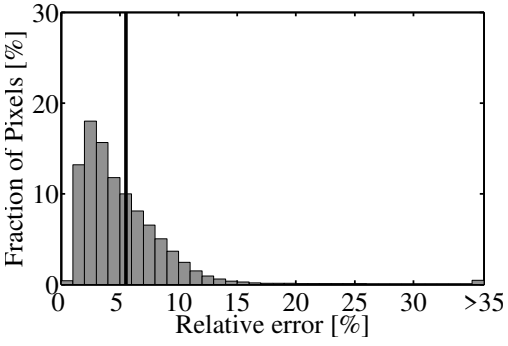


(b) Dong

(c) Hardeberg

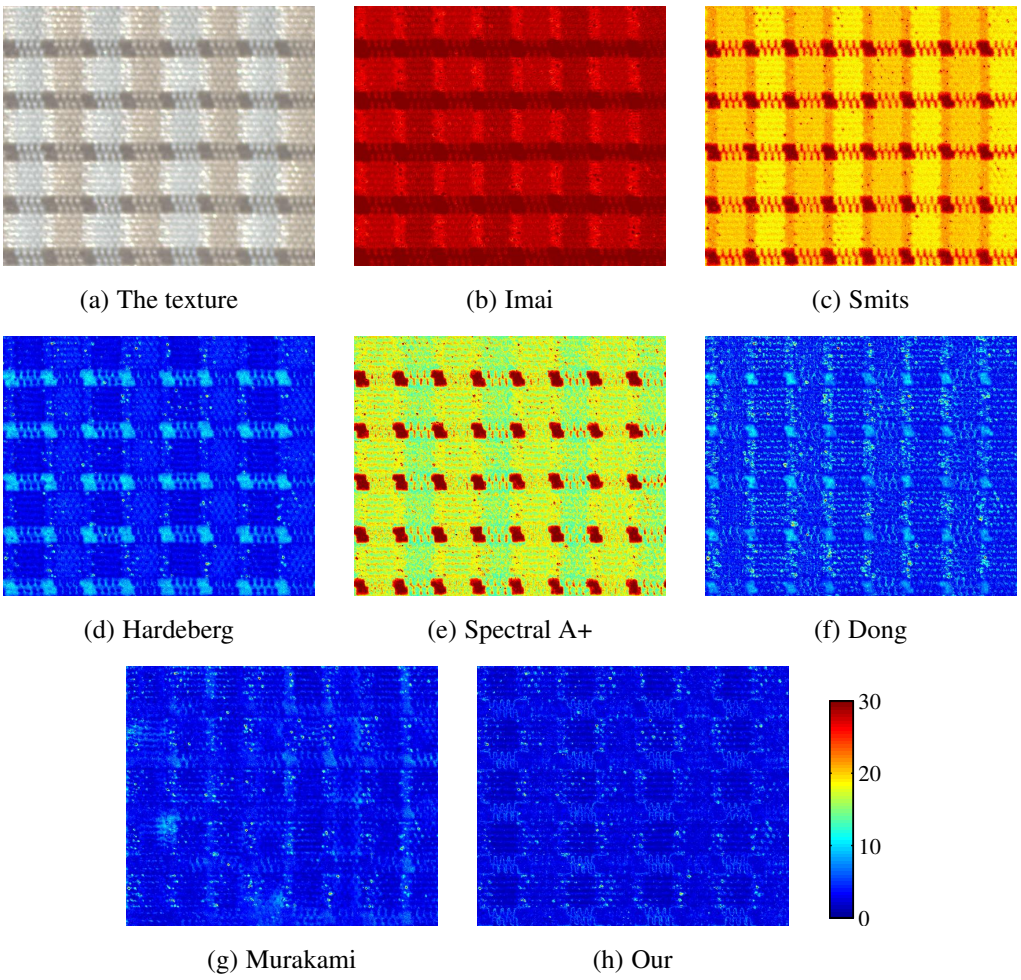


(d) Murakami



(e) Our method (RGB appearance space)

Figure 7.3: Reconstructing texture images using whisk-broom data: Accumulated relative error histograms. 250 spectral data points on a regular grid have been used as sparse spectral input.

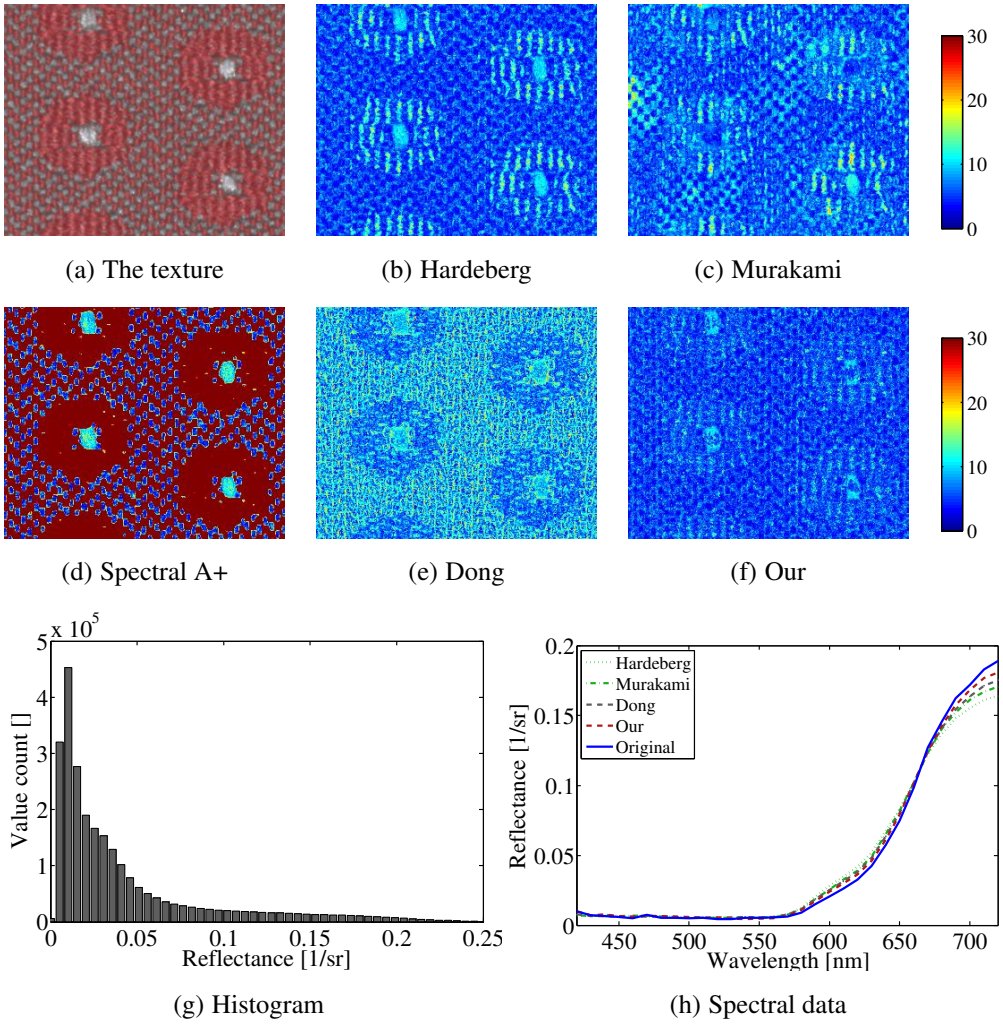


(a) The texture (b) Imai (c) Smits

(d) Hardeberg (e) Spectral A+ (f) Dong

(g) Murakami (h) Our

Figure 7.4: Comparison of pixel-wise relative reconstruction error (in %) for different methods on a fabric texture image: The methods from Imai and Smits produce very high errors as they do not incorporate additional spectral knowledge. The methods from Hardeberg and Dong performs already quite well, but it can only handle three of the four kinds of threads with high accuracy. That is due to the limitation to a three dimensional subspace of the spectral space. The Murakami method can improve on this result due to an additional correction term incorporating the position of the spectral data points. However, our method performs even better on this dataset.



(a) The texture (b) Hardeberg (c) Murakami

(d) Spectral A+ (e) Dong (f) Our

(g) Histogram (h) Spectral data

Figure 7.5: Comparison of pixel-wise relative reconstruction error (in %) for different methods on another fabric texture image: As in Figure 7.4 the Hardeberg and the Dong methods fail to reconstruct one of the four different kinds of threads with good accuracy whereas our method is not limited to three materials. Due to the non-gaussian distribution of pixel values (see (g)) the correction term of the Murakami method leads to a worse result compared to the Hardeberg and Dong methods. In contrast, our method is limited neither by the dimensionality of the image's spectra nor by the distribution of values and it therefore reconstructs all pixels with nearly the same accuracy. The pixel spectra from one exemplary pixel of the red base thread is plotted in (h), demonstrating the quality improvement.

7.2 Push-Broom Data

Using push-broom spectral imaging (see Chapter 2.3.2) is another way to supply sparse spectral data. It is easily possible to just scans a single or very few lines of the image. In our analysis we assumed two horizontal lines at 33% and 66% vertical positions.

Mean reconstruction errors as well as error histograms on non-texture and texture images are shown in Figures 7.6 and 7.7. The results are very similar to those obtained with whisk-broom input: The Hardeberg and Dong methods perform worse, the Murakami method improves a bit and our method is best on average. However, for this kind of input, the errors are generally higher and the different methods do not provide significantly different reconstruction quality.

This is most likely due to the fact that the two horizontal lines miss important spectral features in many images. The more evenly distributed points from the whisk-broom simulation provide a better sampling of all different spectral shapes contained in the images.

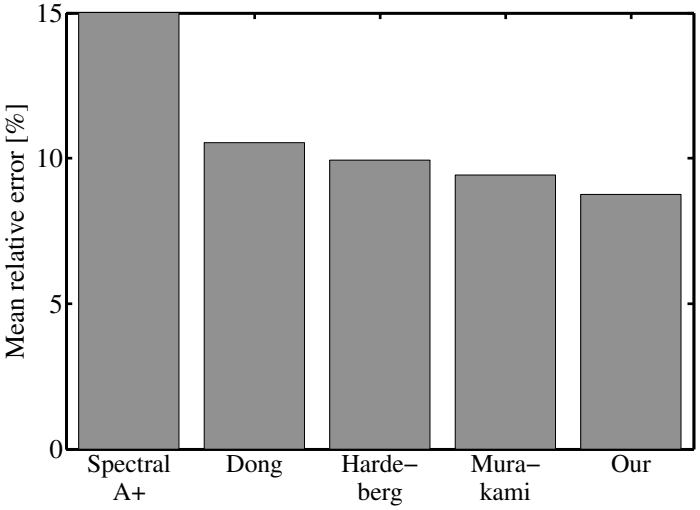
7.3 BRDF Data

To evaluate reconstruction quality on rather simple BRDFs we randomly sampled an increasing number of direction pairs for which spectral data is assumed to be provided. The top light and top view is always contained.

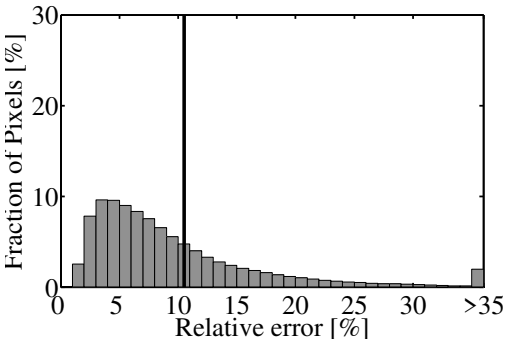
The resulting errors can be taken from Figure 7.8. We excluded the Murakami method as it is not straightforwardly applicable in this case. The results indicate that for BRDFs of rather simple materials like the solid paints our method does not improve upon the results of the trained inverse from Hardeberg. This might be different in case of more complex materials like pearlescent paints that show angular dependent color flops, but no such experiments have been performed for this thesis.

We follow, that for simple BRDFs a good reconstruction is already possible using a few direction pairs with spectral information, only. Therefore, our method may be used to capture high-resolution spectral BRDFs using existing RGB measurement techniques and dedicated multi-angle spectrophotometers[1] providing a few view/light combinations only. This helps to increase measurement speed and simplicity drastically compared to traditional gonioreflectometer setups.

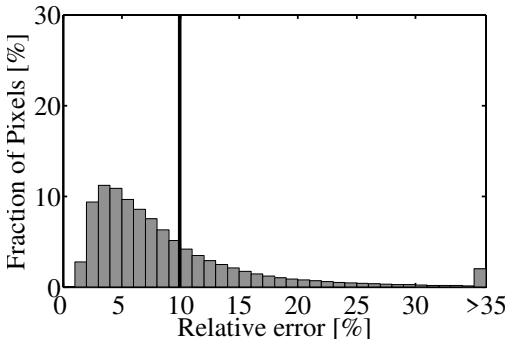
[1] like the X-Rite MA series or the BYK-mac



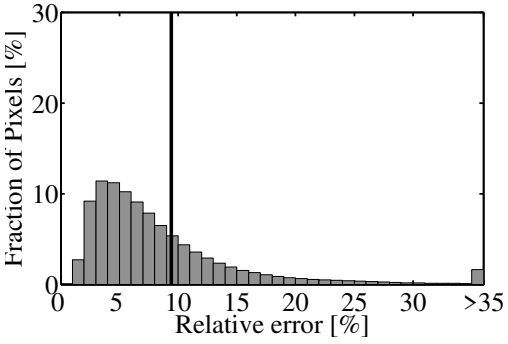
(a) Summary



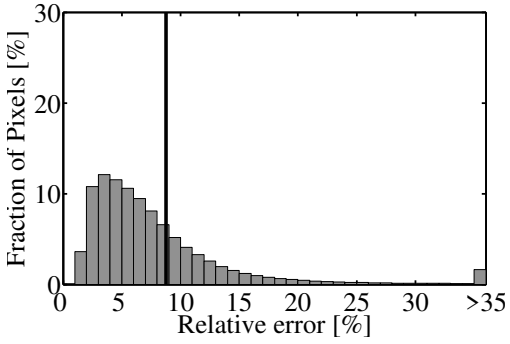
(b) Dong



(c) Hardeberg

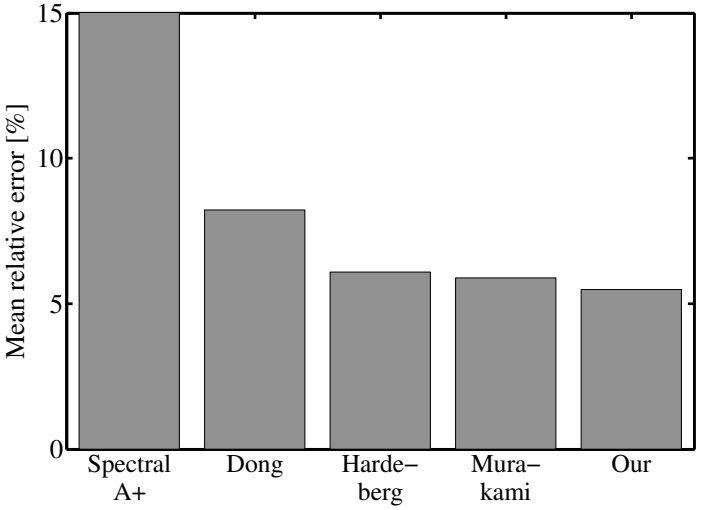


(d) Murakami

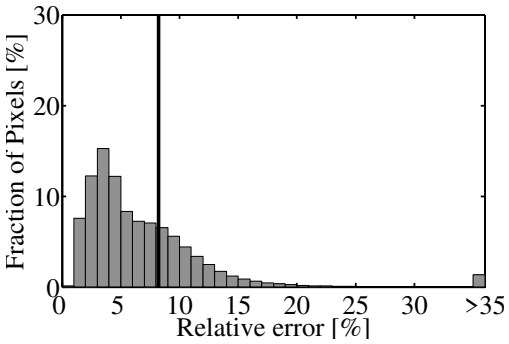


(e) Our method

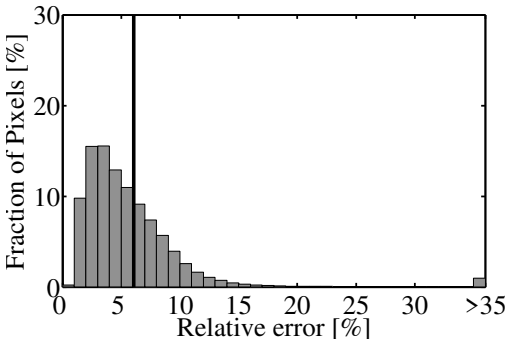
Figure 7.6: Reconstruction error on non-texture images using simulated spectral input from a push-broom device: The results confirm those of the whisk-broom input (see Figure 7.2).



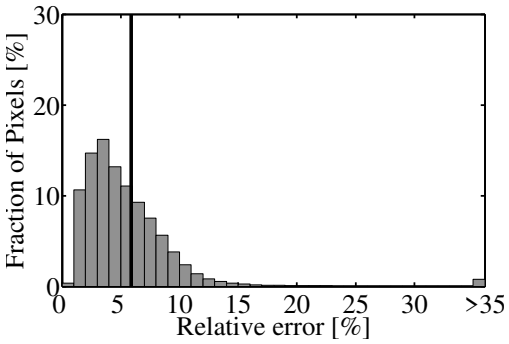
(a) Summary



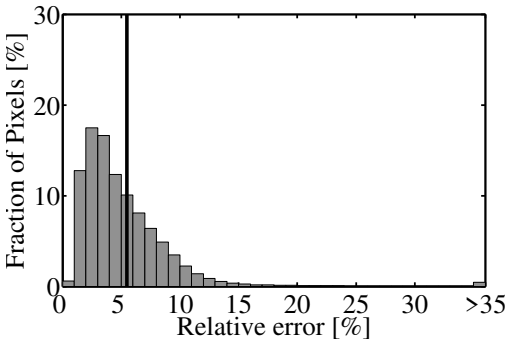
(b) Dong



(c) Hardeberg

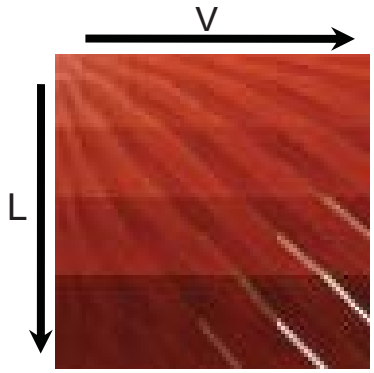


(d) Murakami

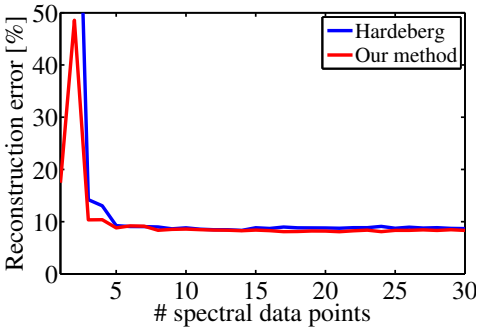


(e) Our method

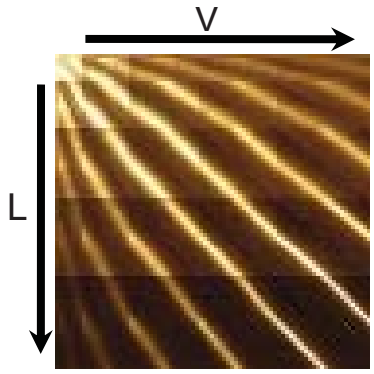
Figure 7.7: Reconstruction error on texture images using simulated spectral input from a push-broom device: The results confirm those of the whisk-broom input (see Figure 7.3).



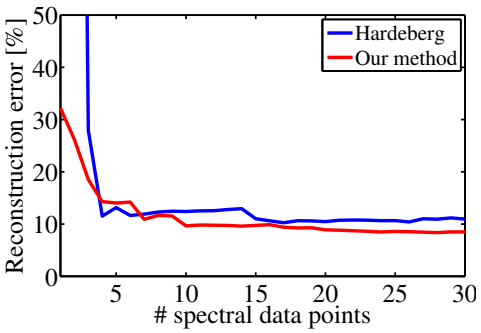
(a) Red BRDF



(b) Reconstruction for red BRDF



(c) Gold BRDF



(d) Reconstruction for gold BRDF

Figure 7.8: Comparison of reconstruction error for simple BRDFs: Our method compared to the Hardeberg method with increasing number of known direction pairs. For visualization purposes the BRDFs are converted to images (a) and (c) by unrolling the view and light directions into x and y directions, respectively. In (b) and (d) the reconstruction error with increasing number of spectral data points - i.e. number of (V,L) direction pairs sampled with full spectrum - is shown. For such simple examples our method does not help a lot to improve the result.

7.4 Reconstruction Error with Varying Amount of Spectral Data

An important property of a spectral reconstruction method is its stability with respect to the amount of spectral information supplied. We therefore performed the whisk-broom simulation from Section 7.1 with a varying number of spectral data points. The resulting reconstruction errors for our method and the methods from Hardeberg, Dong, and Murakami are plotted in Figure 7.9. We denote a general drop of errors with increasing number of spectral data points, as it is expected. However, the methods from Hardeberg, Dong and Murakami do not benefit any more, when the number of data points surpasses a certain limit, whereas our method provides steadily better results with more spectral information added. In the case of the Hardeberg and Dong methods, the reason is that it is restricted to a three dimensional spectral subspace. As soon as enough information is present to choose the best possible space for the respective image, the reconstruction quality cannot be increased by adding more data. Interestingly, we see an error increase for the Dong method for 250 data points and we have not found an explanation for this behavior. The Murakami method benefits a bit from more spectral information as it has an additional correction term added to the result of the Hardeberg method. This correction term takes into account the spatial position of the spectral data. We can see that this helps a lot for non-texture images but not so much for texture images. The reason is, that the dimensionality of the spectra for non-texture images is higher, in general, and the correction term helps to locally correct spectra. We also see that it helps to add more data to have more pixels corrected. However, the method is still outperformed in all cases by our method.

7.5 Reconstruction Error with Varying Amount of Acquisition Noise

In all our analysis above we always assumed that the synthesized RGB data (see Equation 7.1) is biased by gaussian noise having σ at 1% of the mean signal. As the spectral data itself originates from camera measurements in all our test datasets, it is noise biased by itself. To evaluate, how our reconstruction method performs with different levels of acquisition noise, we performed a whisk-broom simulation with 250 spectral datapoints (see Section 7.1) with different noise levels having a σ ranging from 0.5% up to 5% of the mean signal. The resulting reconstruction errors using our method as well as the Hardeberg, Dong and Murakami methods are depicted in Figure 7.10. As expected, the reconstruction errors of all methods increase with more noise in the input data. For the Hardeberg

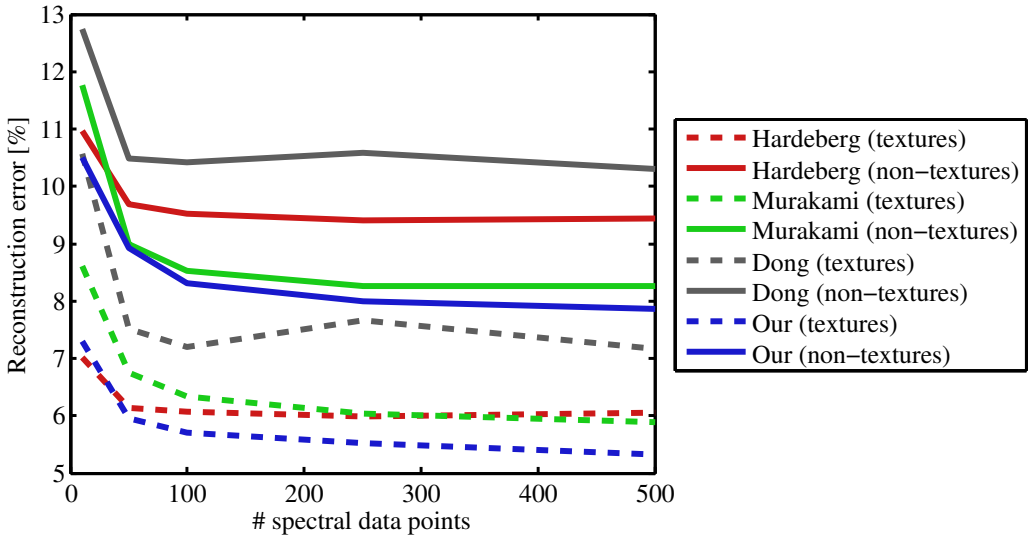


Figure 7.9: Checking reconstruction error with varying amount of spectral input data: Results for our method and the two best previous methods on the whole image database. Especially the Hardeberg and Dong methods do not benefit when more and more spectral data is added. The additional correction of the Murakami method added to the results of the Hardeberg method does help a bit, when enough spectral data is added. However, it can be seen that our method is capable of reducing the reconstruction error steadily with increasing amount of spectral information.

and Murakami methods, the error increase is very similar to that of our method. Only for textures, the Murakami method seems to behave more robust against noise. While having much higher reconstruction error for lower noise, it performs like our method for high noise. The reason for this is not directly clear and would require more analysis on the Murakami method itself, what is out of scope of this thesis. For the Dong method we see a clear advantage with respect to robustness to noise. Though it is outperformed by our method in the noise range that was analyzed here, we see a less steep increase of reconstruction error with RGB data noise. For applications with high noise in the RGB acquisition process, the Dong method might therefore be a preferable choice.

7.6 Incomplete Spectral Information

Another critical case is when incomplete spectral information available, i.e. for a whole part of an image or a material on a texture no spectral data is contained

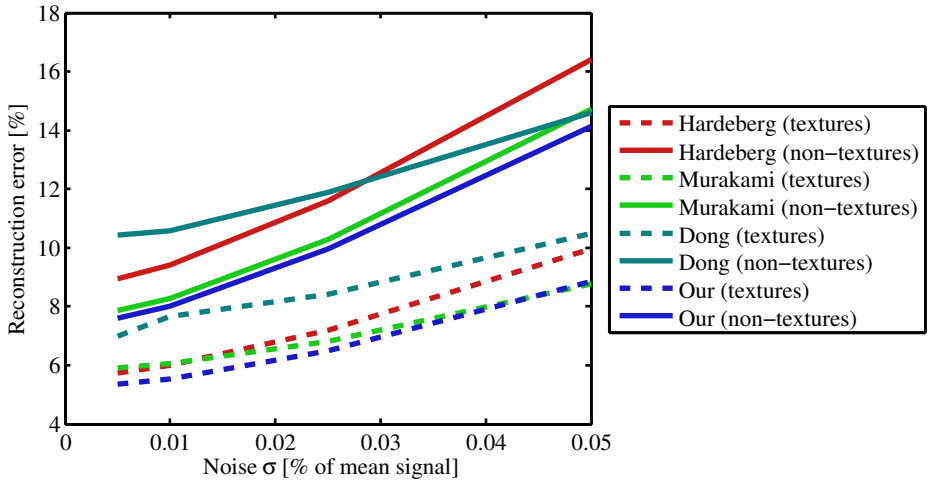
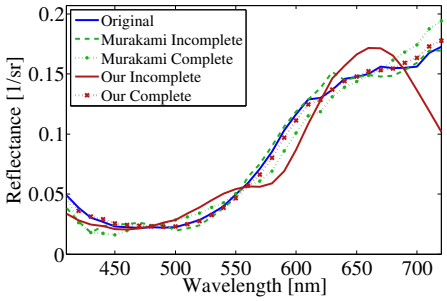


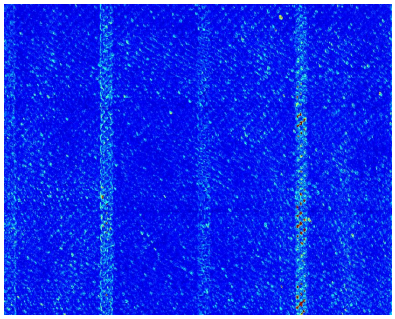
Figure 7.10: Checking reconstruction error with varying noise in the input data: Results for our method and the three best previous methods on the whole image database with gaussian noise of varying σ added to the synthesized RGB data. As expected, reconstruction errors increase with more and more noise. The method from Dong et al. behaves most robustly with increasing noise. However, our method does still perform best for all investigated noise levels. However, in the presence of even more noise, the Dong method might become superior.

in the sparse set. To analyze the behavior of spectralization in comparison to the other previous methods in this case, we have a closer look at one particularly suitable example from the texture database: a dark fabric having colorful stripes on it. Figure 7.11 shows relative reconstruction error maps for this texture where the colorful stripes have (c), (e) and have not been sampled (d), (f). When having data from the colorful stripes, they are reconstructed rather well using both methods. Our algorithm performs a bit better due to the aforementioned dimensionality problem of the Murakami method (see also Figure 7.5). As soon as this data is missing, quality on the stripes drops in both methods as expected. However, as our method has more freedom to shape the spectral curves, it performs worse in this case. Example spectra from the orange stripe are shown in (b), demonstrating this behavior. It should be pointed out, that a different choice of the appearance space and therefore a different construction of neighborhood could have prevented this behavior.

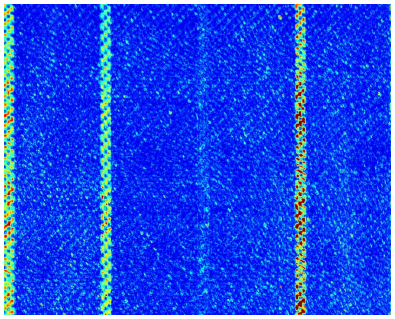
(a) The texture



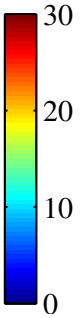
(b) Spectra from the orange stripe

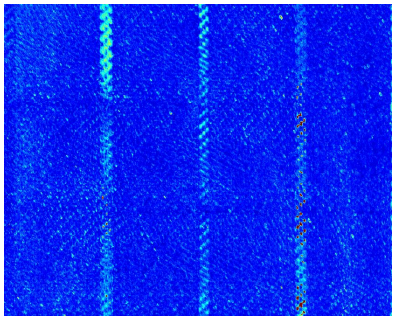


(c) Enough spectral information (our method)

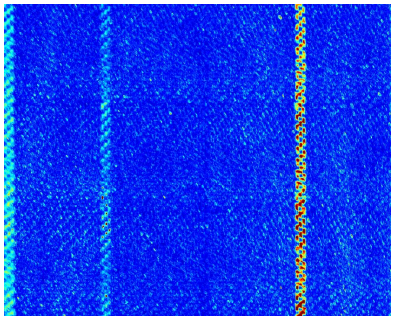


(d) Incomplete spectral information (our method)





(e) Enough spectral information (Murakami)



(f) Result with method from Murakami and incomplete spectral information

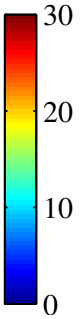


Figure 7.11: Incomplete data: Color coded relative error maps for the "dark fabric" dataset with different levels of spectral information. When having data from the colorful stripes, they are reconstructed well (c) and (e), our algorithm performs a bit better due to the aforementioned dimensionality problem of the Murakami method (see also Figure 7.5). As soon as this data is missing in (d) and (f) both methods drop in quality on the stripes as expected. However, as our method has more freedom to shape the spectral curves, it performs worse in this case. Example spectra from the orange stripe are shown in (b), demonstrating this behavior. A different choice of the appearance space could have prevented this behavior.

CHAPTER 8

SUMMARY

In Part II of this thesis we have introduced a novel method to reconstruct dense spectral data from dense RGB and sparse spectral data.
For this, we introduced a novel smoothness prior on spectral images, stating that one can deduce a similarity of the spectrum from a similarity in the RGB space. This might not work on a simple per-pixel basis, of course, but our method allows to define nearly arbitrarily complex metrics to compare locations in an RGB image for similarity, e.g. comparing a whole surrounding patch. Since the method only performs a neighborhood search based on the defined metric in a pre-processing step, the spectral reconstruction algorithm itself remains untouched by the choice of this metric. This way, our method allows for the adaption to special cases of images or datasets by the selection of an appropriate metric. We have shown the accuracy of the smoothness prior on a wide variety of datasets and also with a selection of RGB space metrics.

Based on the aforementioned prior, we have then introduced a novel algorithm to reconstruct dense spectral data from dense RGB and sparse spectral data. This algorithm is based on a linear-least-squares optimization and is quite generic with respect to the kind of data that needs to be reconstructed. Especially, it is not required that the data-set to reconstruct is an image - i.e. it might be a tabulated BRDF, a slice of a BTF, or something even less structured. The energy function for the optimization is first introduced in a quite generic manner (see Equation 6.9). We then presented a thorough evaluation of the algorithm based on a more specific formulation of the energy function (see Equation 6.10) with respect to convergence, speed, and reconstruction accuracy. The accuracy is compared to prior work in great detail in Chapter 7, showing that our method outperforms prior work in nearly all cases. Compared to the - on our test datasets - best performing prior-work method from Murakami et al. [MIYO07] our algorithm is much more generic though being more or equally accurate.

With these findings we will conclude this whole part of this thesis and will continue to discuss another major challenge: the spectral calibration of trichromatic cameras.

Part III

Implementation

CHAPTER 9

RGB DEVICE CALIBRATION AND MATCHING

To use the *Spectralization* technique in practical applications, two different matching problems have to be solved: the geometric matching, where the location of sparse spectral information with respect to the dense RGB data has to be determined, and spectral matching, where the relation between RGB pixel/reflectance values and spectral data is of interest. In this chapter we will deal with the latter problem and propose and evaluate a novel method for spectral characterization of RGB or similar trichromatic imaging devices. Geometric calibration is discussed in the next chapter.

Calibration of camera devices is of fundamental importance for all kinds of color accurate photography beyond image based reflectance measurement. Commonly used standard software (e.g. [Gil11]) for color calibration aims only at *colorimetric* calibration, that is for transformation from the device dependent trichromatic values to some device independent trichromatic color space like XYZ or L*a*b*. Recovering this transformation is often called *profiling* and the transformation *device profile*. The International Color Consortium[1] (ICC) has standardized such device profiles. An ICC profile is sufficient for color reproduction under the same or very similar lighting conditions compared to the profiling conditions and is therefore widely applied for simple photographic applications by pre-recording profiles for some standard illumination conditions.

However, knowledge about the spectral response of the imaging system offers greater flexibility and often greater accuracy compared to standard colorimetric profiles. It has been proven [CTNM06] that standard color correction and color processing purposes like white balancing can benefit from the effective spectral response. Moreover, standard ICC profiles can be generated for arbitrary lighting conditions when the effective spectral response is known.

All text in Chapter 9 marked in gray is self-cited from M. RUMP, A. ZINKE, AND R. KLEIN. Practical Spectral Characterization of Trichromatic Cameras. In: *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)*, 30:6, 2011

[1] http://www.color.org

Finally, there are a lot of application for which the effective spectral response of the imaging system is crucial. This includes designing image-based measurement or photographic setups [FOP08] and of course all spectral imaging techniques that utilize trichromatic cameras like our Spectralization technique or e.g. [IB99].

For digital cameras the imaging process is influenced by many different factors affecting the pixel value $P(x_i)$ of a pixel x_i. Assuming an idealized system the response may be written as:

$$P(x) \propto \int_{N_x} C(x, x') \int_{\Lambda} E(\lambda) T_c(x', \lambda) T_o(\lambda) L_{scene}(\lambda, x') d\lambda dx', \tag{9.1}$$

where λ denotes wavelength, Λ the respective range of integration, L_{scene} is the scene radiance imaged in pixel x', T_o the transmission of the optical system, T_c the transmission of the color filter array, E the quantum efficiency of the sensor and C the cross-talk between pixels x and x' in the neighborhood N_x of x. In equation (9.1) a potential non-linearity between incoming radiance and pixel value is not explicitly considered. If required, methods like [DM97] or [RBS03] can be applied to compensate for such effects by estimating the response curve.

Parameters for certain camera components, like spectral transmittance or quantum efficiency, are sometimes published by the manufacturer. However, for the complete system usually no such data is available. Unfortunately, to use a camera system for image based measurement or other applications requiring a deeper understanding about how pixel values are related to scene radiance, the total *effective spectral response* R^{eff} is required (called R_{cam} in Chapter 6). This effective spectral response comprises all effects influencing spectral imaging as summarized in Equation 9.1:

$$P(x) \propto \int_{\Lambda} R^{eff}(\lambda) L_{scene}(\lambda, x) d\lambda \tag{9.2}$$

Assuming R^{eff} to be independent of x, pixel cross-talk effects are neglected. This is obviously a simplifying assumption because R^{eff} might depend on brightness differences between nearby pixels.

To make a spectral calibration method usable in practice and to make it accessible to a wide range of users simplicity is a key factor. To date no such practical methods for estimating the spectral response of cameras are available - existing approaches for measuring it require costly hardware such as either tunable filters or light sources (e.g. monochromators) or they need strictly controlled acquisition conditions limiting their applicability and leading to a rare consideration of the spectral response.

9.1 Previous Work

The most common method to measure the effective spectral response of cameras is to capture the response of the camera under tunable monochromatic illumination ([VFTB97], [FOP08]). This approach is also suggested as EMVA Standard [EMV10]. Unfortunately, the technical requirements are high hampering broad application of this technique. To reduce the effort of monochromator-based approaches, several methods have been proposed, that recover the effective spectral response from color chart images. Hubel et al. [HSF94] take photographs of a Macbeth color chart using 16 narrowband and 8 broadband filters in front of a dedicated light source. While the coarse shape of the spectral response is determined by the narrowband illumination, the broadband illumination in conjunction with the different color fields is used to recover finer detail. Relying upon the same basic idea further methods have been proposed since then.
Cheung et al. [CLH+05], Solli et al. [SALK05] and Vrhel and Trussel [VT99] used image color charts or different color fields under broadband illumination to calibrate cameras. They exploit the fact that recovering the effective spectral response of a camera with known spectral reflectances is essentially the same as recovering spectral reflectance from camera responses with known camera characteristics. Therefore, they employ methods for spectral reflectance estimation to this problems, which mostly rely on linear methods like least-squares pseudo-inverse matrices.
Shen and Xin [SX04a], [SX04b], [SX06] performed spectral characterization of flatbed scanners by acquiring color charts. From the response of the scanner to the known color fields they optimized a transformation matrix from scanner response to spectral space. To this end, they either used Wiener estimation on training data in a local surrounding of the candidate response or they employed constrained least squares optimization. Ebner [Ebn07] followed a similar approach for camera calibration but used evolution strategies so solve for the energy minimization problem.

Common to all the previous methods is the need for controlled lighting conditions and/or very specialized hardware. While this is acceptable for dedicated scientific measurements it is prohibitively costly to be established as standard for everyday applications. Building on prior work that attempt to estimate spectral response from color chart images, we present a novel fast and simple method that is robust enough to handle uncontrolled acquisition conditions. The only assumption about lighting we make is that a spectral characterization of the dominant illuminant is available, either by using spectral data (e.g. obtained with low-cost spectrophotometers) or measured color temperature in certain cases.

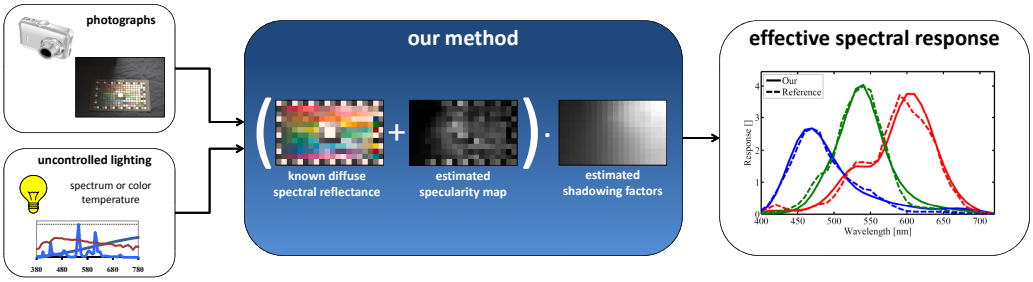


Figure 9.1: Illustrating the workflow of our spectral calibration approach. Our method is capable of reconstructing the effective spectral response of trichromatic cameras from photographs of a color chart captured under uncontrolled lighting conditions. The only assumption is that the spectral power distribution of the dominant illuminant is given. Alternatively, in many relevant cases even knowing a color temperature of the illumination is sufficient.

9.2 Calibration Overview

Our method is capable of reconstructing the effective spectral response R^{eff} of trichromatic cameras from photographs of a color chart[2] captured under largely uncontrolled lighting conditions. Furthermore the spectral power distribution of the dominant illuminant must be given. Alternatively, in many relevant cases even knowing a color temperature of the illumination is sufficient. By introducing a novel imaging model accounting for specularity and spatially varying illumination, our approach is robust with respect to uncontrolled, out-of-lab capturing conditions. It is important to note that even carefully designed lab-setups can suffer from these problems. As a consequence, our method can lead to better results, even when applied to data acquired under lab conditions.

A schematic overview of our approach is shown in Figure 9.1. In the following sections the different steps of our algorithm are explained in further detail.

9.3 Data Acquisition

To estimate the effective spectral response of a camera, radiometrically calibrated photographs of a color chart[3] as well as the relative spectral power distribution of the respective illumination are required.
In contrast to existing techniques our methods yields reliable results, even in conjunction with highly uncontrolled indoor and outdoor conditions. The only fun-

[2]We use the big Gretag Macbeth ColorChecker DC, the quasi standard for color correction in photography.

[3]The reflectance spectra of the color fields are assumed to be known.

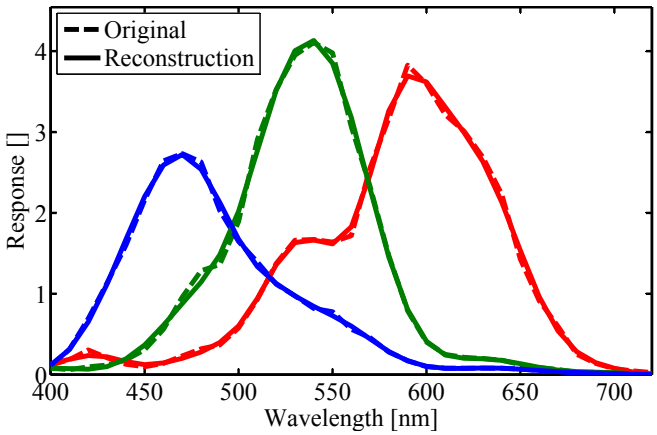


Figure 9.2: Synthetic example: Exemplary reconstruction of effective spectral response R^{eff} minimizing E^{basic} for a synthetic dataset. In this example, the camera response was generated assuming perfectly Lambertian color fields and isotropic daylight illumination.

damental assumption made is that the lit color fields contain enough information for reconstructing spectral camera response.
For measuring the illumination, standard spectrophotometer hardware like the EyeOne from X-Rite or similar cost effective devices suffice. Alternatively, for certain lighting conditions (e.g. halogen bulb lamps, filament lamps or cloudy sky), the illumination spectrum may be also inferred by measuring color temperature (e.g. by using chroma-meters).
Ideally, several photographs of a color chart are taken under different lighting conditions.

9.4 Estimating Effective Spectral Response

The basic idea of the estimation process is to optimize for effective spectral response curves R^{eff} that explain the measurements in a best possible manner. More precisely, the difference between the trichromatic colors of the color fields captured by the camera and the known reflected spectral radiance projected to trichromatic color space using R^{eff} is minimized.
After introducing the basic energy formulation in the next section we explain how it needs to be extended to be usable with data captured under uncontrolled conditions.

9.4.1 Basic Approach

All existing methods for estimating spectral response from color chart images assume that only light spectrum and measured spectral reflectance is required to model the color of a color field. These assumptions are commonly formalized along the lines of the following objective function to be minimized [SX04a, SX04b, SX06, Ebn07]:

$$E^{basic}(R^{eff}) = \sum_{c=1}^{3} \left(E_d^{basic}(R_c^{eff}) + \alpha E_{sm}^{basic}(R_c^{eff}) \right) \tag{9.3}$$

with

$$E_d^{basic}(R_c^{eff}) = \sum_{j}^{J} \left\| \sum_{i=1}^{k} \left[S_{i,j} L_i R_{i,c}^{eff} \right] - D_{j,c} \right\|^2, \tag{9.4}$$

$$E_{sm}^{basic}(R_c^{eff}) = \sum_{i=1}^{k-1} \left\| R_{i,c}^{eff} - R_{i+1,c}^{eff} \right\|^2. \tag{9.5}$$

Here, $R_{i,c}^{eff}$ are the unknown effective spectral responses for color channel c and a discrete wavelength λ_i. k is the number of spectral bands, j the field on the color checker, J the total number of color fields, L the spectral power distribution of the

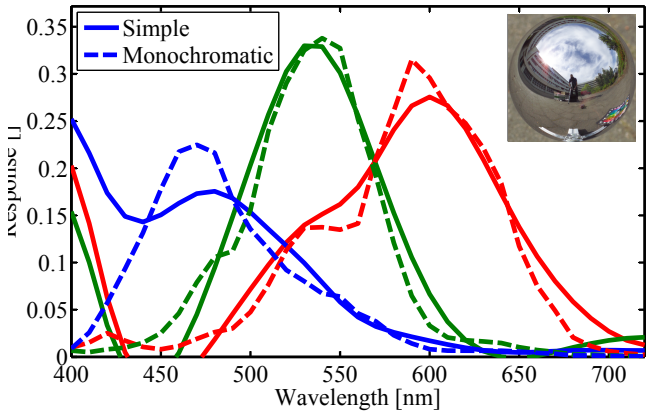


Figure 9.3: Using real data: the reconstruction suffers from serious artifacts if E^{basic} is used as objective function, neglecting spatial non-uniformity of illumination and specularity of color fields. The image in the upper right shows the illumination conditions under which the image was captured. For comparison, the effective spectral response captured using a monochromator is plotted as dashed curves.

illuminant, S the known reflectance spectra of the color fields and D the average camera response to the respective color field in the photograph.
The data term E_d^{basic} minimizes color differences between observed and reconstructed data while E_{sm} enforces smoothness on the spectral response curves which is controlled using the factor α. A global optimum of the above can be easily found by standard linear least squares optimization techniques like a QR-factorization.
Using E^{basic} with synthetic data works as expected. For an example we refer to Figure 9.2, where a synthetic example with 33 spectral bands and measured daylight illumination is shown. As can be seen, the reconstruction error is negligible. The results indicate, that the information provided by the different color fields of the color chart is in principle sufficient for accurate reconstruction which was also reported in previous work. Informally speaking, there are enough slopes in the spectral power distribution of the light reflected by the color fields to resolve the effects of R^{eff} for the given spectral resolution.

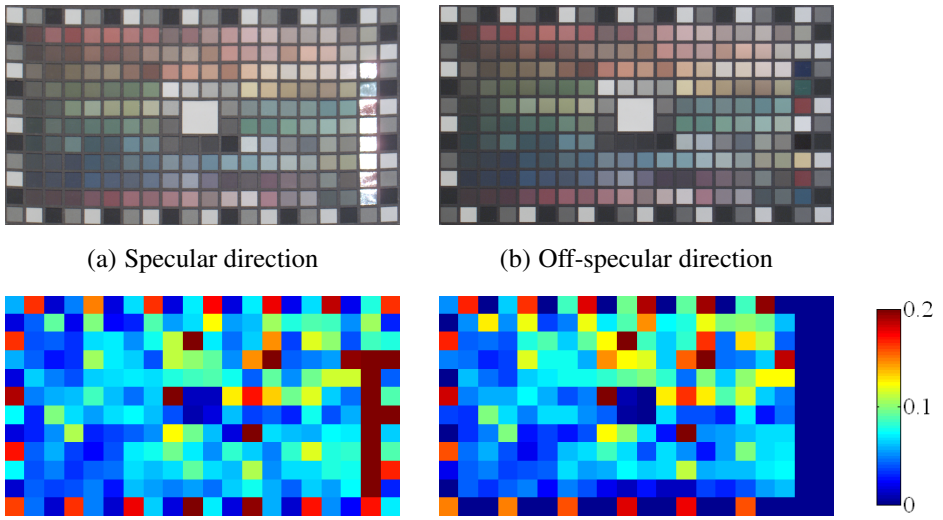
9.4.2 Using Real Data

When applying the basic algorithm introduced above to real data, the method fails as can be seen in Figure 9.3. Obviously the original idealized model assumptions — that only light spectrum and spectral reflectance of the color chart are required for reconstruction — do not hold in practice. There are mainly two sources of bias:

1. Contrary to common belief, the color fields of the Macbeth color chart include substantial specularity as illustrated in Figure 9.4. We found this to be true also for all other color charts we considered.

2. In out-of-lab conditions the illumination across the color chart is often non homogeneous, mainly due to occluding objects or due to incident angle when using a nearby light source. Additionally, some kinds of color charts do not have a solid housing and are therefore prone to slight bending. This is especially true for the larger color charts. Due to the bending, the incident angle is slightly different for each color field biasing the results.

As will be shown, accounting for the above issues is essential to obtain usable results. In particular, specularity can be a serious problem, even in the case of controlled conditions in a laboratory. Of course, specularity can be reduced in a laboratory using for example cross-polarization, but it continues to pose a serious problem in out-of-lab conditions. Spatial inhomogeneity has already been considered in previous work (e.g. in [CLH+05]) but was always corrected by hand



(a) Specular direction (b) Off-specular direction

(c) Specular data (d) Fit using our method

Figure 9.4: Color chart specularity: comparing two rectified color chart images taken under spot light illumination. **Top row:** The fields in (a) are considerably brighter and less saturated than the ones in (b), revealing the amount of specularity of the color fields. **Bottom row:** (c) shows the normalized difference between (a) and (b), (d) is the specularity estimated by our method as discussed in section 9.4.2. Please note the close match between estimated and measured specularity maps.

before estimating the spectral response. In contrast, we integrate the spatial homogeneity as well as the specularity correction directly into the estimation process and completely eliminate manual work.

To account for both specularity and shadowing effects the basic imaging model used in equation (9.4) needs to be extended. In particular we assume, that the BRDFs of the color fields include a diffuse and a white specular component (caused by Fresnel reflection on a dieletric surface). Furthermore, we make the idealized assumption that incident illumination has a dominant component whose relative spectral distribution is independent of the viewing angle. However, the illumination may be partially occluded generating smooth lighting variations across the color chart. The proposed imaging model is illustrated in Figure 9.5.
Under these assumptions, shadowing is approximated by a single scaling factor per color field and specularity by adding an individual wavelength independent constant to the diffuse reflectance spectrum of each color field. As the model assumptions given above might be violated in practice, we propose to capture

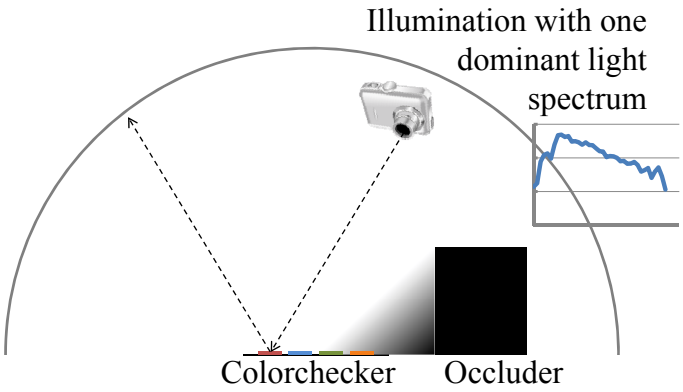


Figure 9.5: Illustrating the proposed imaging model considering specularity and spatially varying illumination: Assuming a dominant illuminant, specularity is approximated by adding an individual wavelength-independent constant to the reflectance spectrum of each color field. Spatial non-uniformity is modeled by a scaling factor per field.

multiple images under different illumination conditions or at least by rotating the camera around the color chart in place seeing different parts of the environment in the specular reflection.
However, according to our experience it is often sufficient to take as few as one photograph to obtain usable results.

The new imaging model including specularity and shadowing was implemented by the following objective function:

$$\begin{aligned} E(R^{\mathit{eff}}, F, \sigma) = & \sum_{c=1}^{3} \left(E_d(R_c^{\mathit{eff}}, F, \sigma) + \alpha E_{sm}(R_c^{\mathit{eff}}) \right) \\ & + \sum_{c=1}^{3} \left(\gamma E_{\mathit{damp}}(R_c^{\mathit{eff}}) + \delta E_{\mathit{border}}(R_c^{\mathit{eff}}) \right) \\ & + \beta E_{sm}^{\mathit{illu}}(F) \end{aligned} \tag{9.6}$$

Data term:

$$E_d(R_c^{\mathit{eff}}, F, \sigma) = \sum_{j}^{J} w_{j,c} \left\| \sum_{i=1}^{k} \left[(S_{i,j} + \sigma_j) L_i R_{i,c}^{\mathit{eff}} \right] - F_j D_{j,c} \right\|^2 \tag{9.7}$$

For the computation of the data term $E(R^{\mathit{eff}}, F, \sigma)$ additional scaling factors F_j determine shadowing of each of the color fields of the color chart. The parameter

σ_j is the unknown specularity per color field j and is controlling the amount of white added to the spectral reflectance S of a color field. Furthermore, we introduce weights $w_{j,c}$ reflecting the importance of different color fields. A detailed explanation how to compute these weights is given below.

Regularization:

$$E_{sm}(R_c^{\mathit{eff}}) = \sum_{i=1}^{k-1} \left(\frac{\left\| R_{i,c}^{\mathit{eff}} - R_{i+1,c}^{\mathit{eff}} \right\|}{R_{i,c}^{\mathit{preveff}} + R_{i+1,c}^{\mathit{preveff}}} \right)^2 \tag{9.8}$$

$$E_{sm}^{\mathit{illu}}(F) = \sum_{j}^{J} \sum_{n \in N_j} \|F_j - F_n\|^2 \tag{9.9}$$

$$E_{\mathit{damp}}(R_c^{\mathit{eff}}) = \sum_{i=1}^{k} \left\| R_{i,c}^{\mathit{eff}}(\lambda_i) - R_{i,c}^{\mathit{preveff}}(\lambda_i) \right\|^2 \tag{9.10}$$

$$E_{\mathit{border}}(R_c^{\mathit{eff}}) = \left(R_{1,c}^{\mathit{eff}} + R_{k,c}^{\mathit{eff}} \right) \tag{9.11}$$

We extended the basic regularization term E_{sm}^{basic} from equation (9.5) by iteratively dividing through the result from the last step. The term in the denominator leads to stronger regularization in the lower (nosier) parts of the spectral response curves and to weaker regularization in the "more interesting" peaks that exhibit quite a lot variation. An extra term E_{sm}^{illu} enforces the scaling factors F_j used in the data term to be comparable to those in its spatial neighborhood N_j. Essentially this means that smooth shadow gradients across the color chart are preferred. The E_{border} term fixes the start and end of the spectral curves to zero, since we assume no substantial contribution from outside the spectral range we use for optimization.

Unfortunately, the above energy functional is bilinear in R^{eff} and σ and is thus no longer solvable using linear least-squares optimization techniques.
For that reason an alternating least squares strategy was employed: The parameters R^{eff} and F are fit in one iteration of the optimization, whereas in a subsequent step R^{eff} and F are kept fixed and σ is estimated.
In the first step, the responses R_c^{eff} are constraint by an additional term E_{damp} to the previous result R^{preveff} to avoid oscillations in the alternating least squares algorithm.

Moreover, to increase stability and to accelerate the fitting process multigrid optimization was used: We start with a low spectral resolution with 8 spectral components, that is successively increased by 4 once the alternating least squares algorithm for this level has converged. At every resolution increase R^{preveff} is

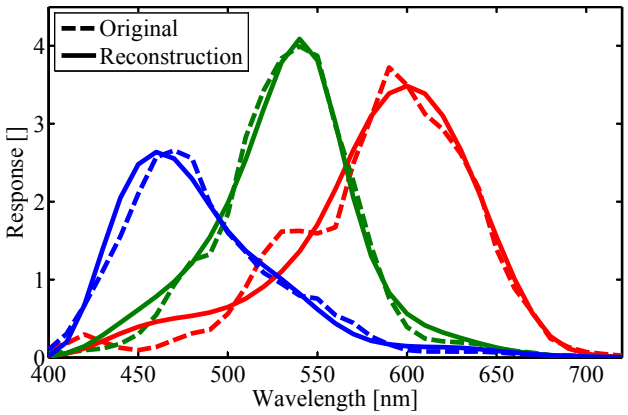


Figure 9.6: Synthetic example: Reconstruction results on disturbed synthetic data using Equation 9.6.

resampled accordingly and E_{damp} ensures fitting consistency with respect to the previous result.

The effectiveness of the modified functional was tested on synthetic data sets with scaling factors given by a smooth gradient field in the range of $[0.8, 1.2]$ and with a random specularity in the range of $[0, 0.1]$. The estimated spectral response for this synthetic example is shown in Figure 9.6. Opposed to the result obtained with the basic energy functional when used with real data (see Figure 9.3), the extended version succeeds in reconstructing the shape of original spectral response curves.

Generally, the method gave stable results for various lighting conditions such as overcast sky, blue sky and direct sunlight even in the presence of (partial) shadowing and specularity.

Choosing field weights The color fields on the color chart are optimized for photographic applications. However, as they are not evenly distributed in spectral domain (in particular when additionally considering the illumination spectrum) and since not all of the fields are equally relevant for guiding the reconstruction process field weights $w_{j,c}$ are used for compensation. Choosing viable weights is crucial for obtaining reliable results. We base the choice of the weights on the following three ideas:

1. The sum spectral power distribution of the product of illuminant and reflectance of the fields should be optimized with respect to equal spectral distribution across all spectral bands. This is particularly important for the shorter wavelengths that typically have less support.

2. For each of the effective response curves we have to select fields with a

spectral reflectance that has steep slopes and good support in those bands where the response itself has high support.

3. Extremely high weights of certain fields should be avoided to have a good noise to signal ratio and to minimize reconstruction bias due to the assumptions made.

The above objectives are modeled by a sum of two weighting terms:

$$w_{j,c} = w_j^{\textit{distrib}} + w_{j,c}^{\textit{slopes}} \tag{9.12}$$

Here, the term $w_j^{\textit{distrib}}$ is used to enforce spectral equalization, to have approximately the same support in all wavelength bands:

$$w_{1..J}^{\textit{distrib}} = \underset{\{w_1,..,w_J\}}{\operatorname{argmin}} \ \operatorname{var}_i \left(\sum_{j=1}^{J} w_j S_{i,j} L_i \right), \tag{9.13}$$

where var_i denotes the variance across spectral bands i. To this end we fit the set of $w_j^{\textit{distrib}}$ simultaneously. The other term $w_{j,c}^{\textit{slopes}}$ attempts to model the slope constraints. To this end, we calculate the derivatives of $\partial SL_{i,j} = \frac{\partial (S_{i,j} L_i)}{\partial i}$ and $\partial R_{i,c}^{\textit{eff}} = \frac{\partial R_{i,c}^{\textit{eff}}}{\partial i}$ with respect to wavelength by finite differencing characterizing the slopes:

$$w_{j,c}^{\textit{slopes}} = \sum_{i=1}^{k} \Big((\partial SL_{i,j})(R_{i,c}^{\textit{eff}}) + (\partial SL_{i,j})(\partial R_{i,c}^{\textit{eff}}) + (S_{j,i} L_i) R_{i,c}^{\textit{eff}} \Big) \tag{9.14}$$

Figure 9.7 shows a weight map for halogen lighting. In this example, the weights for the blue fields are higher to compensate for the small amounts of energy in the blue part of the spectrum.

Parameter settings. The proposed objective function $E(R^{\textit{eff}})$ is controlled by four different weights α, β, γ and δ. Consistently good reconstruction results were obtained with the following values that have been kept fixed for all examples presented in this work: $\alpha = 3$, $\beta = 500$, $\gamma = 1$, $\delta = 10$. We found that our method is not very sensitive to moderate variations of these parameters. Please note that both, photographic device RGB data and light spectra, are normalized to have a mean of one and that diffuse reflectance is assumed to be in $[0, 1]$.

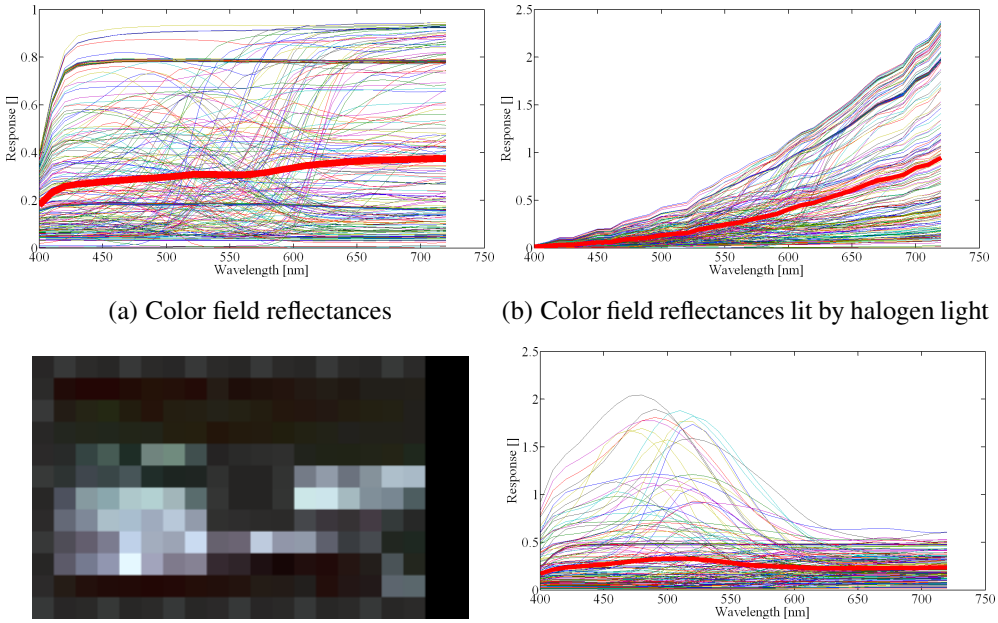
9.5 Results

9.5.1 Acquired Datasets

For testing the effectiveness of our methods, the effective spectral response of five different camera models was estimated:

1. Kodak DCS 760 SLR (Kodak760)
2. Kodak DCS Pro 14n SLR (Kodak14n)
3. SVS Vistek svs4022COGE industry video camera (Vistek4022)
4. Canon PowerShot G9 (consumer level DSC)
5. CCam BCi-6600-USB CMOS industry video camera

For all camera models raw image data without demosaicking was acquired. We took images of the Macbeth Color Checker DC as discussed in Section 9.3 for a



(a) Color field reflectances

(b) Color field reflectances lit by halogen light

(c) Weight Map

(d) Scaled color field reflectances

Figure 9.7: Spectral equalization using field weights: **(a)** Reflectance spectra of all 240 color fields of ColorChecker DC together with the mean spectrum (bold red). **(b)** The result when illuminated with halogen light. In this case short wavelengths are clearly underrepresented. **(c)** The corresponding weight map according to equation (9.12). **(d)** Weights applied to the spectra.

Figure 9.8: Example images for the reconstruction of effective spectral camera response: top row: Kodak DCS Pro 14n at daylight, bottom row: SVS Vistek svs4022COGE at halogen lighting. For display purposes, the images have been demosaicked and a gamma of 2.2 has been applied.

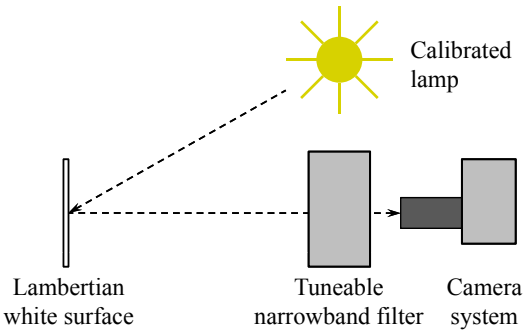


Figure 9.9: Sketch of our measurement setup for measuring the reference spectral responses.

wide range of different lighting conditions, including outdoor and indoor illumination. Some example images are shown in Figure 9.8. Based on the captured data we analyzed the behavior of the novel method regarding accuracy, stability and limitations.

Reference data. For comparison purposes the effective spectral response of the above camera models was also measured under laboratory conditions using near-monochromatic lighting as suggested in previous work ([VFTB97], [FOP08]).

We'd like to stress the fact that cross-talk effects affect the effective spectral response and that the relative magnitude of these effects increases with the strengths of local contrasts on the Bayer pattern. Since monochromatic light yields a very strong contrast between the different color channels, a spectral response curve measured using monochromatic light may differ from a curve determined under more natural broadband illumination used for our measurements. As such the acquired reference data should not be considered a ground truth.

A sketch of our measurement setup can be found in Figure 9.9. It consists of a calibrated lamp illuminating a white Lambertian surface and a liquid crystal tunable filter (the VariSpec-VS10 used during BTF measurements, see Chapter 3.2) in front of the measured camera device. The filter has a narrowband transmission of approx. 10 nm bandwidth and can be tuned to wavelengths between 400 and 720nm, thus produces nearly monochromatic light. Beforehand we carefully measured the transmittances of the tunable filter. We acquired images with varying exposure time tuning the filter throughout its operable range in 10nm steps and extracted and averaged pixel values $\Delta_{i,c}$ of the Lambertian surface as seen by the camera for each wavelength band i and color channel c.

Adapting equation (9.2) yields:

$$\Delta_{i,c} = \int_{\lambda} \tilde{R}_c^{\mathit{eff}}(\lambda)\rho(\lambda)T_{b,i}(\lambda)L(\lambda)d\lambda \tag{9.15}$$

$$\Delta_{i,c} = \sum_{j=1}^{k} \tilde{R}_c^{\mathit{eff}}(j)\rho(j)T_{b,i}(j)L(j) \tag{9.16}$$

Here, $T_{b,i}$ is the transmittance of the bandpass filter tuned to a peak wavelength λ_i, ρ is the reflectance of the white Lambertian reflector and $\tilde{R}_c^{\mathit{eff}}$ is the effective spectral response for color channel c. Using matrix notation gives $\Delta_c = T_b\left(\tilde{R}_c^{\mathit{eff}} \odot \rho \odot L\right)$ with transmission matrix T_b and $A \odot b$ being the element wise product of a vector b with all columns of a matrix A.

Like Hubel et al. [HSF94] we reconstruct $\tilde{R}_c^{\mathit{eff}}$ using a pseudo inverse of T_b, dividing by $\rho \odot L$ and applying a gaussian filter with $\sigma = 7nm$ to avoid ringing.

9.5.2 Comparison to Reference

In Figure 9.10 a comparison between our reconstruction results and the reference data obtained as described above is shown. As it can be seen, the overall shape of the response curves is consistent with the reference for the broadband illuminations in the "sunlight" dataset and the "blue sky" dataset indicating a good reconstruction of the effective spectral response here. Unfortunately, the reconstruction

results for the "halogen" datasets are less good in the blue part of the spectrum where especially the red and green curves show bumps Since halogen lamps emit very small amounts of energy in the blue part of the spectrum, the support for the data term in the energy functional is quite small here and not even the weights can compensate for this. The fourth column of Figure 9.10 shows reconstructions results if all firsts images of the datasets are used in our algorithm. Here, the reconstruction of the effective response is accurate again. In essence this results indicate that our algorithm is well suited to reconstruct the effective spectral response when illumination conditions are used with a dominant broadband light source present or when images from different environments are collaboratively used as an input.

9.5.3 Cross Validation

The comparison of our results to the reference data provides already good evidence that our method works as desired. However, since no ground truth data is available, we also employed cross validation to compare spectral response curves obtained with our method to reference measurements. Here, the main objective was to test consistency and accuracy of fitting results with respect to different lighting conditions as well as to identify potential overfitting issues. For this purpose camera responses under different illumination conditions were predicted using acquired data. We took photographs with all five cameras in a laboratory, where we explicitly tried to avoid any specularity and spatial non-homogenity. Unfortunately, this was not really possible and so we manually eliminated most of the non-homogenity by looking at the surrounding grey fields of the chart to enable for better comparison and error analysis. As illuminants we used neon lamps, a HMI lamp and a halogen lamp. Then, after estimating the effective response of all cameras using our technique (R^{eff}) under sunlight (see Figure 9.10) and using the reference method with controlled monochromatic lighting ($\tilde{R}^{\mathit{eff}}$), we predicted the camera response of the datasets using both R^{eff} and $\tilde{R}^{\mathit{eff}}$.

To this end, we calculated the difference d_j for a color field j between expected and measured device response:

$$d_j = \left| (S_j \odot L)\, R^{\mathit{eff}} - D_j \right| \tag{9.17}$$

with D_j being the measured pixel value. The same variable was calculated for $\tilde{R}^{eff}$.

The average errors using our technique and the much more complicated monochromator measurements can be taken from Figure 9.11. As an example, Figure 9.12 shows the detailed prediction result for the neon lighting and the Pro14n camera. In essence, our reconstruction performs even slightly better on most cameras

to predict the camera images for different illuminants without the need for any special and costly hardware.

In Figure 9.11, is apparent that the perceptual errors for the CCam camera are much higher than those for the other cameras. The higher perceptual errors result from the very broad spectral filters of the CCam camera (see Figure 9.10). Using these filters, a high-quality transformation of device RGB values to the XYZ color space is not possible and therefore the perceptual errors are very high. However, the errors for the monochromator and our approach are similar for this camera as well.

9.5.4 Validation of Imaging Model

Our assumptions regarding specularity and shadowing were validated by reconstructing the spectral response using challenging controlled test cases:

- Using very directional illumination created by a halogen spotlight causing strong specularity effects.
- Placing occluders near the color chart to provoke shadowing.

A reconstruction of the specular term σ for an image captured under halogen spotlight illumination is presented in Figure 9.4(d). Please note how closely it matches the reference specularity map (c) generated by comparing acquired in-specular and off-specular images of the color chart (a) and (b) shown in Figure 9.4. As for shadowing we refer to Figure 9.13 indicating that our method is able to reconstruct shadowed regions accurately. For the images in this figure we placed a black box as an occluder between the color chart and a huge window which was the main light source in this scene.

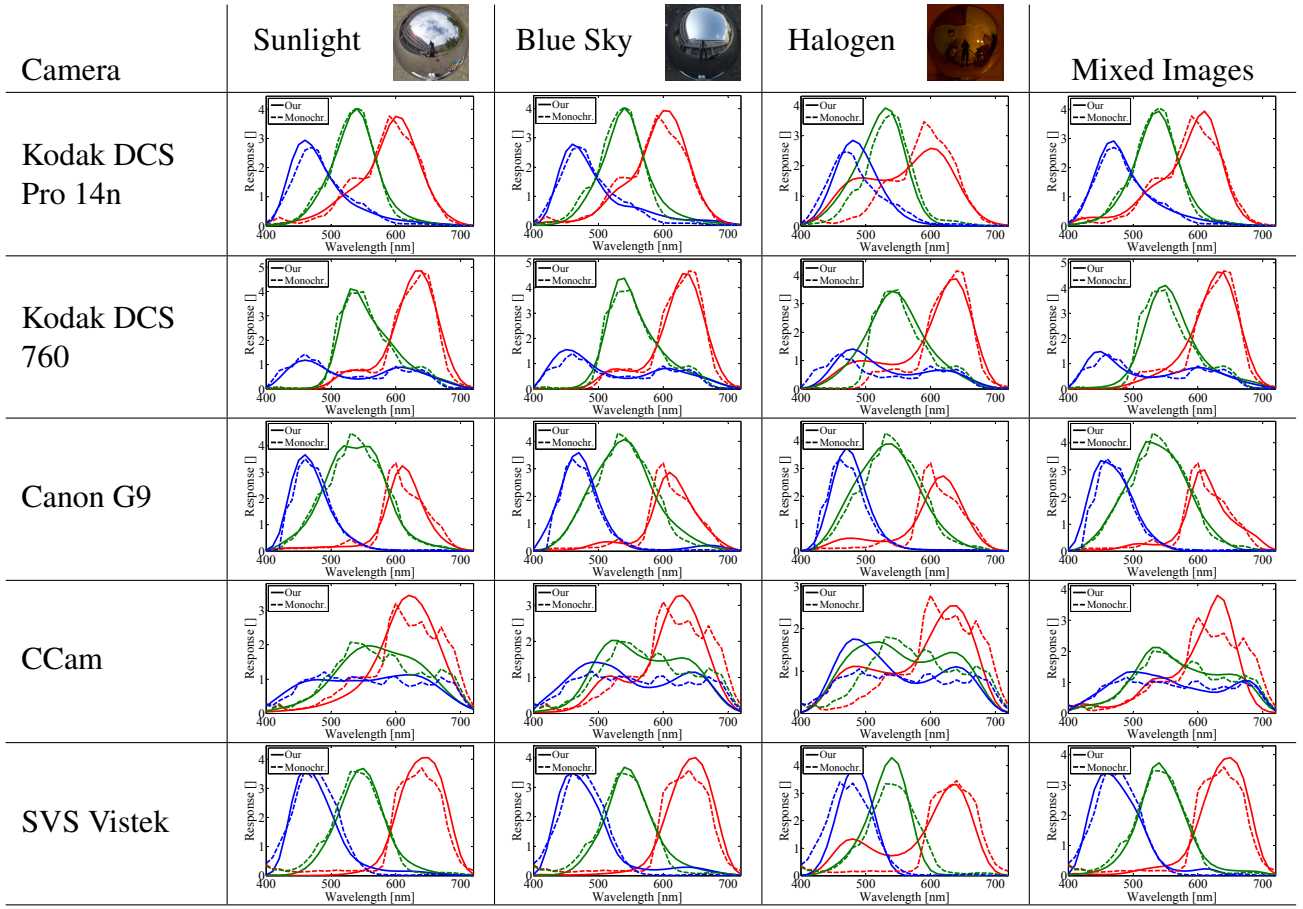
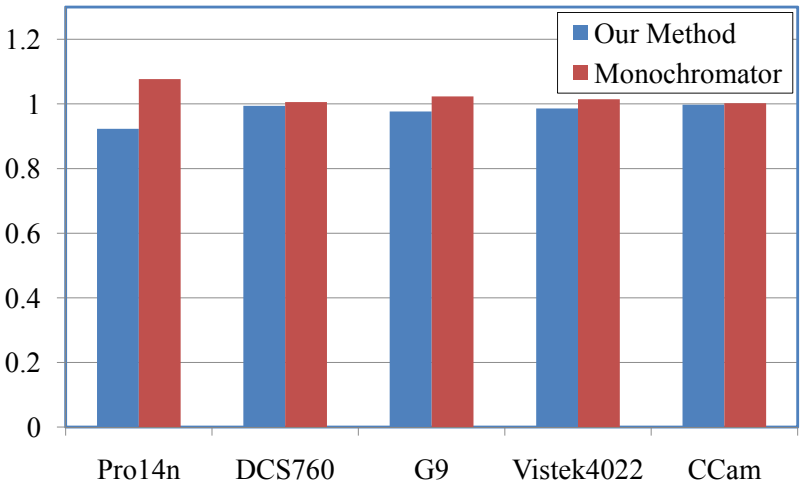
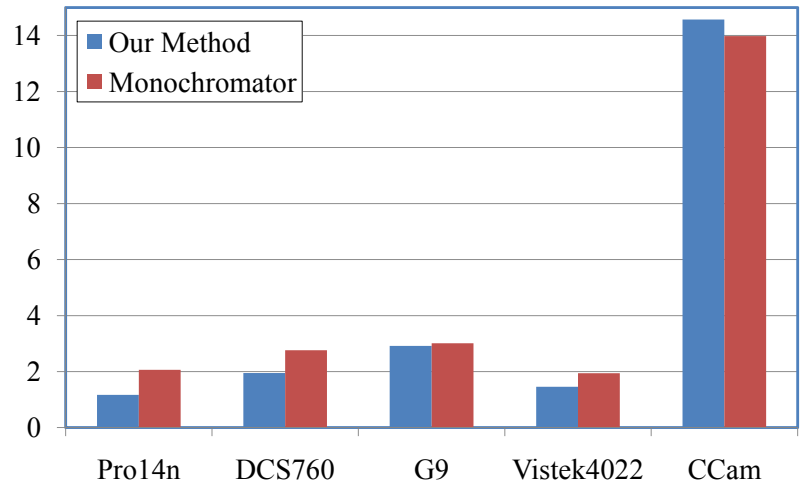


Figure 9.10: Comparing spectral responses estimated by our method (solid lines) to reference measurements acquired using the monochromator setup (dashed lines): In general spectral response curves estimated by our method show less high frequency detail compared to the references. For illumination conditions with good support across the spectral domain the shape is consistently matching the reference. For halogen lighting our methods is less accurate in the blue part of the spectrum due to the low support of the illuminant.



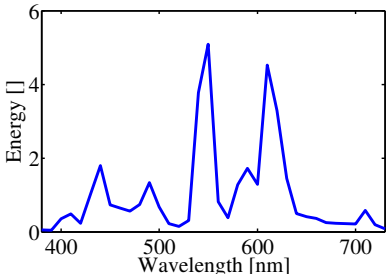
(a) L^1 distance



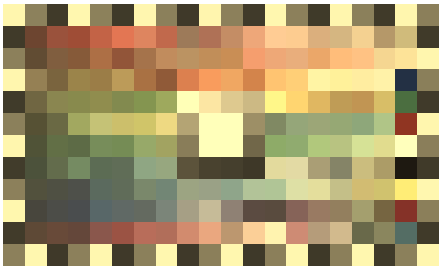
(b) ΔE^*_{94}

Figure 9.11: Cross validation: Average errors for predicting camera responses across different lighting conditions using the spectral response curves estimated by our method (blue bars) from the sunlight datasets as well as the measured references (red bars). The error was computed using L^1 distance in device RGB space (see equation (9.17)) as well as perceptual ΔE^*_{94}. Per camera, three photographs of the color chart lit by neon, HMI and halogen lamps have been used. Even though our method is much simpler, the accuracy is comparable to the technically much more involved monochromator approach.The perceptual errors for the CCam camera are much higher than for the others because it has very broad spectral filters (see Figure 9.10).

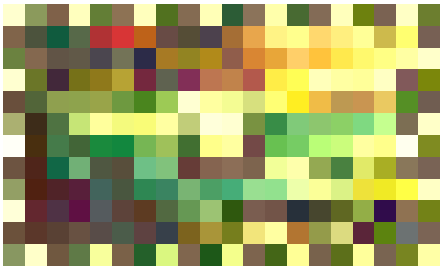
(a) Measured response



(b) Neon light spectrum

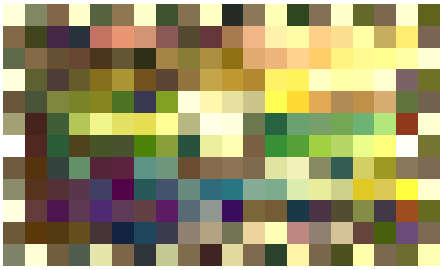


(c) Predicted response using reference response $\tilde{R}^{eff}$



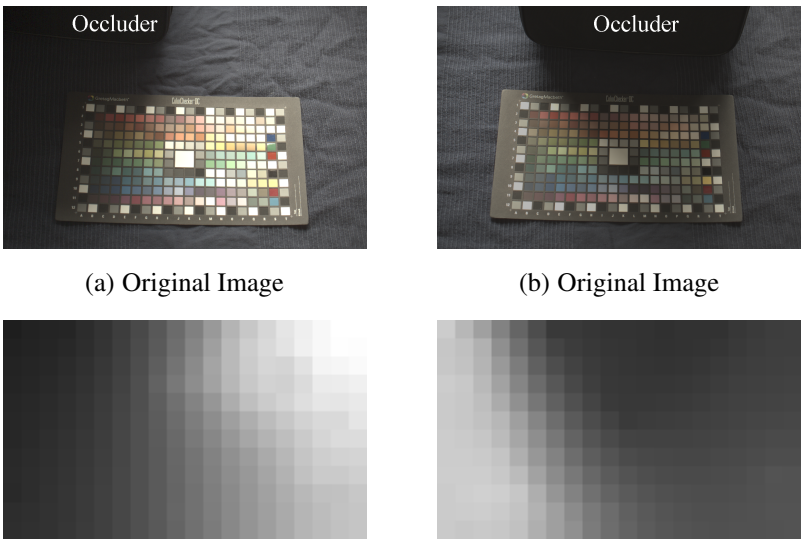
(d) Difference x10 using reference response $\tilde{R}^{eff}$, Avg: 0.040

(e) Predicted response using "our" response R^{eff}



(f) Difference x10 using "our" response R^{eff}, Avg: 0.036

Figure 9.12: Cross validation example for predicting camera response with different illumination (see also Figure 9.11) for Kodak DCS Pro14n and a neon lamp.



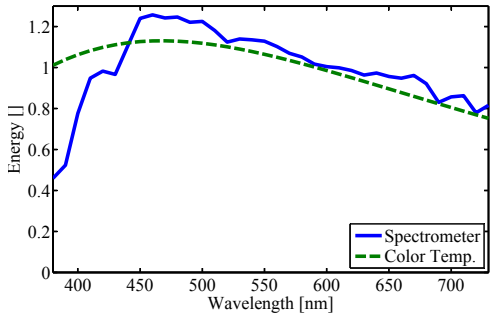
(a) Original Image (b) Original Image

(c) Scale Factor Map F (d) Scale Factor Map F

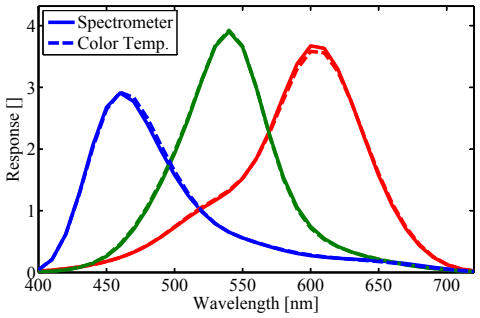
Figure 9.13: Validating shadow compensation: comparing photographs to estimated scale factors F. As can be seen the scale factors are consistent with the shadowed regions in the original images indicating that our algorithm can deal with environments strongly deviating from ideal conditions.

9.5.5 Light from Color Temperature

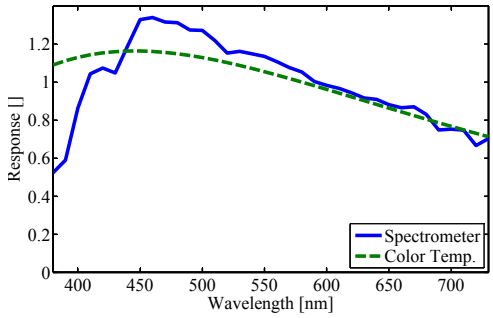
In general, our method requires an illumination spectrum to be known either by directly measuring the spectral power distribution of incident lighting using spectrophotometers or using tabulated values of well known illuminants. Interestingly, for our purpose, a variety of relevant illuminants may be well approximated by the Planck function of the respective color temperature without losing much accuracy. This includes typical daylight conditions like direct sunlight, cloudy or overcast sky. Two examples are given in Figure 9.14 for two different camera models and lighting conditions. Even though the Planckean spectrum is only roughly matching the actual illumination spectrum, with a significant difference for shorter wavelengths, the reconstructed spectral responses match surprisingly well. These results indicate that in many practical cases the spectrophotometer measurement may be replaced by much simpler devices that acquire color temperature.



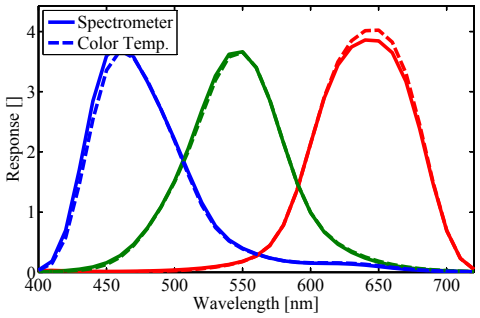
(a) Cloudy sky (6200 K)



(b) Reconstructions for Kodak14n using cloudy sky



(c) Sunlight (6700 K)

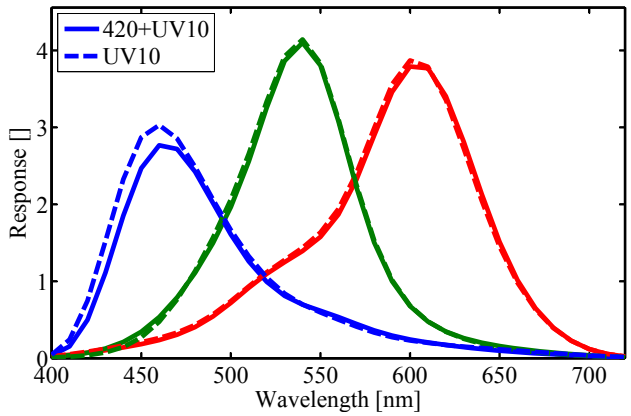


(d) Reconstructions for Vistek4022 using sunlight

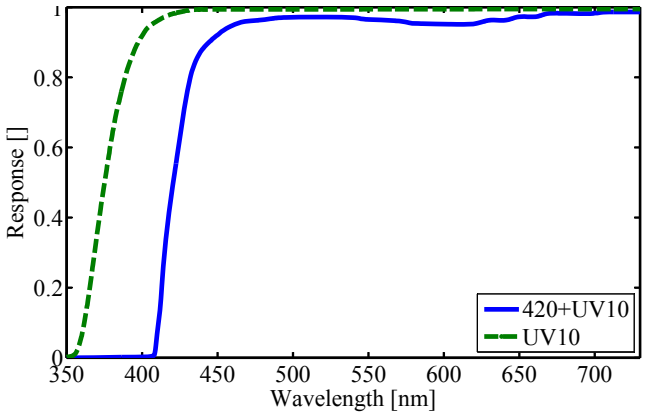
Figure 9.14: Color temperature vs measured spectral power distribution of illumination: comparing spectral responses reconstructed using cloudy sky as well as sunlight. Dashed lines indicate results based on color temperature only. For typical daylight conditions the simplified light model works extremely well.

9.5.6 Resolving Small Differences in Effective Response

To evaluate whether our method is capable of resolving subtle differences between effective responses, we used different UV cut filters with the Kodak DCS Pro 14n camera. The respective transmission spectra of both filters (a combination of two filters: 420 and UV10, as well as UV10 alone) are shown in figure 9.15(b). For both setups the effective response was estimated under sunlight conditions. The results (see figure 9.15(a)) clearly indicate that even subtle spectral differences can be identified using our method.



(a) Reconstruction of R^{eff} using two different UV filter combinations with the Kodak DCS Pro 14n camera



(b) Transmission spectra for both UV filter combinations

Figure 9.15: Resolving small differences in effective response: Our technique is able to recover even subtle differences of the effective response caused for example by different optical filters.

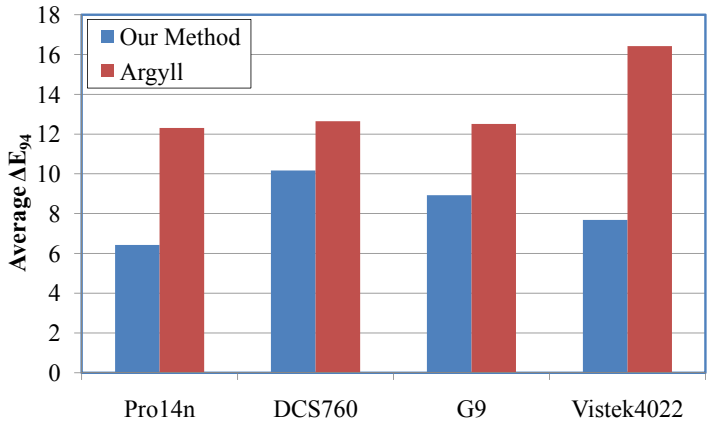


Figure 9.16: Using our method for ICC profiling. The quality of the generated ICC profiles was evaluated by measuring perceptual distances between true XYZ values of the color chart and XYZ values inferred from device RGB values using whitebalancing and device-RGB-to-XYZ transformation. The ΔE^*_{94} error (blue bars) has been computed for each of the camera models and for three different lighting conditions: neon, HMI and halogen. For comparison purposes, a state-of-the-art ICC profiling software (Argyll) was evaluated as well (red bars).

9.5.7 ICC Profiling

A very important application related to color imaging is to transform a given RGB camera space to some device independent connection space like XYZ or Lab. This is commonly achieved by applying ICC profiles that represent this color transformation, given by (multiple) 3x3 transformation matrices, as well as so called shaper curves characterizing energy response of a camera.

Normally, ICC profiles are generated by imaging known color charts for a given illuminant and then estimating the respective color transform on a trichromatic basis. Trichromatic ICC profiling suffers from the fact, that the spectrum of the illuminant used during profiling basically gets "baked" into the color space transformation.

In contrast, since we have a practical method for measuring effective spectral response R^{eff}, we can directly derive color transformation matrices to arbitrary color spaces that may be easily adapted to a new illuminant without the need for new profiling. To benchmark ICC profiles generated by taking advantage of our method we made a comparison to a state-of-the-art profiler. We have chosen Argyll CMS [Gil11], a freely available and widely recognized profiling system, because it is able to consider spectral measurements of color checker reflectances and illuminants. For this purpose, for each of the cameras an ICC profile was

generated using Argyll from color chart images of the "sunlight" dataset. Subsequently, exactly the same data was used to estimate spectral responses R^{eff}.

To recover the XYZ values of the color checker from a photograph taken under a new illuminant L with given spectrum, we have to apply both color space transformation and whitebalancing to the device RGB values. The profile generated by Argyll is first applied to the device RGB values ending up with XYZ values with wrong whitepoint. Thus, whitebalancing is necessary which can, in this case, only be performed in tristimulus space.

When using our method, we are able to compute an appropriate profile matrix M incorporating the new illuminant on the fly in the following way:

$$M = \left(R^{\mathit{eff}} L\right)^{+} M_{XYZ} \tag{9.18}$$

where A^{+} denotes the pseudo-inverse of A and M_{XYZ} is the matrix containing the XYZ primaries' spectra.

Figure 9.16 shows the resulting average ΔE_{94} errors per camera for different illuminants.

9.6 Summary

In this chapter a practical method for the spectral characterization of trichromatic cameras was presented. By simply taking images of a color chart and measuring illumination using a spectrometer, the effective spectral response was estimated. Unlike previous approaches our method does not rely on specialized and costly hardware or controlled laboratory environments. The effectiveness of the novel approach is demonstrated by comprehensive analysis including comparisons to costly reference measurements following the state-of-the-art. As shown, using our method for ICC profiling is superior in terms of accuracy and flexibility to standard colorimetric approaches. We presume that this advantage is of fundamental nature and will apply also to other applications. Obviously, the spectrophotometer is an additional piece of hardware, but it should be stated that simple handheld devices are quite inexpensive compared to a professional RGB camera and should be worth the gain provided by our method. Moreover, spectral measurements can be completely avoided if spectral camera response is estimated under daylight conditions.

CHAPTER 10

FAST RGB-SPECTRAL BTF MEASUREMENT

Having introduced the basic building blocks for efficient spectral reflectance capture during the last chapters of this thesis, we will now outline a possible setup for combined RGB-spectral BTF acquisition and show some first results.

One of the remaining central problems of *spectralization* is the geometric registration between spectral and RGB data. The spectral registration was explained in the last chapter. During tests we performed using a push-broom device, we noticed that a good alignment between a spectral scanline and a RGB image is hard to achieve, even for a fixed setup. Furthermore, it is very likely for many materials, that a fixed, spectral scanline misses important parts of the material and therefore a good spectral reconstruction will not be possible (see also Section 7.6). One way to circumvent that would be to use a freely movable line by utilizing e.g. a mirror that can be rotated around an axis. Then, two or more lines with spectral data could be acquired, potentially even based on an analysis of the RGB image content. However, the geometric registration problem would be even more complicated to solve than for a fixed setup.

As an alternative, we propose to use a CCD camera in combination with a changeable, spectral bandpass filter either in front of the camera or in front of some of the light sources in the setup. When acquiring images with a spectral camera system in addition to the RGB camera system, the geometric registration can be achieved by rather simple methods, as we will explain later on.

The central problem with this setting is the long exposure time required to capture spectral data with high signal-to-noise ratio (SNR), because of the strong reduction of photon flux by the bandpass filters. If the spectral data should be acquired with minimal additional measurement time, either only a few spectral images can be acquired or spectral images with a much lower SNR have to be used. Since many real-world materials exhibit strong angular dependence of reflectance, a reduction of spectral data points is not generally advised. In this chapter we will therefore explain, how to incorporate images with low SNR resulting in short exposure times for the spectral image capture.

To test our algorithm, we have modified an existing RGB BTF acquisition

setup by adding or spectral CCD camera from the goniometer setup described in Chapter 3.2. We will show how to calibrate such a setup to operate in low SNR range. In order to evaluate the spectral reconstruction quality that can be achieved by such a setup in combination with the *spectralization* method, we will perform simulations based on the spectral BTF database we acquired with the goniometer setup.

10.1 Reconstruction

In this section we introduce a variant of the *spectralization* method that has been introduced in Chapter 6. This variant is tailored to reconstruct full spectral images from high-quality RGB data and noisy spectral data.

In contrast to the original method we will not assume that the spectra at some sparse positions (like spots for a whisk-broom device or lines for a push-broom device) are known with high quality. Instead, as discussed above, we will use images with high spatial resolution but a very low SNR. Such images can be captured by using a monochrome camera with changeable band-pass filter running at low exposure times, which enables for fast capture. However, this configuration has the fundamental advantage that a pixel-accurate alignment between RGB and spectral images is possible. It should be clarified that this registration problem also prohibits the use of spectral snapshot techniques (see Chapter 2.3.2) due to their low resolution. One way to register the RGB and spectral images would be to use a common optical path and a beam splitter, which would allow for a one-time calibration between both images. However, if a series of spectral images has to be acquired with non-changing geometry (neither scene nor camera change), the spectral camera system can be registered by capturing one image for one spectral band with a long exposure time in beforehand. This image will have a low noise level and allows for registration against the RGB images using markers or other methods like optical flow. This second technique is especially useful for reflectance measurements, because in this case huge image series have to be captured where only changes in lighting occur.

Of course, the noisy spectral images cannot be used directly. To get rid of the noise in the spectral images, filtering is necessary which leads to a loss of high frequent spatial details. We will analyse, if the low-noise and high-resolution RGB images provide enough information to recover the spectral image series. Imai et al. [IB98] also proposed to capture with different spatial resolutions, and to reconstruct the high-resolution spectral image afterwards. However, they also stick at a low dimensional subspace of the spectral space and only aim at reconstructing $L^*a^*b^*$ images. In contrast to this we would like to reconstruct a full spectral image and don't want to accept limitations on the dimensionality of the imaged

spectra. We incorporate the filtering directly into the *spectralization* method by ensuring, that the filter response to the unknown target image matches the filter response on the noisy input images. Then, the optimizer is responsible for finding a solution which respects both high frequent detail from the RGB image and low frequent spectral information from the filtered spectral images.

The new principle is implemented into the following modified energy function:

$$\begin{aligned} E_n(S) &= \alpha \sum_{j=1}^{K} \left\| \left(\sum_{i=1}^{n} F_i^j f_i S_{i,b_j} \right) - \left(\sum_{i=1}^{n} F_i^j f_i D_i^{b_j} \right) \right\|^2 \\ &+ \beta \sum_{i=1}^{n} \left\| (S_i f_i) R_{cam} - f_i V_i \right\|^2 \\ &+ \gamma \sum_{i=1}^{n} \sum_{j \in N_i} \left\| S_i - S_j \right\|^2 \end{aligned} \tag{10.1}$$

Overall, this energy function $E_n(S)$ is a more generic variant of the spectralization energy function $E(S)$ in Equation 6.9. The third term is even unchangedly taken from that. Modifications occur in the first two terms.

The first term ensures that applying a spatial filter F^j to the unknown spectral image S matches the filter response on the noisy, measured image D^{b_j} belonging to wavelength band b_j. K is the number of spatial filters applied to the data, i denotes a pixel position and n the number of all pixels. When choosing a filter F^j to be zero for all pixels i except for one, we end up with Equation 6.9.

It should be noted that the spectral-to-RGB projection matrix R_{cam} actually contains the spectral sensitivity of the camera multiplied by the spectral power distribution of the respective light source of the measurement setup. That said, S_i is not the spectrum reflected from the material surface, but it is the spectral reflectance of that material - the quantity that we want to measure.

The first and second term of E_n do now also incorporate a per-pixel factor f_i, which is computed from the RGB pixel V_i. This allows e.g. for compressing the dynamic range of the image data before the reconstruction. This is especially important since we want to reconstruct multiple images of the same material lit and viewed from different directions, at once. Shading, shadowing and BRDF effects will lead to largely different brightness, but often the color hue is similar. Therefore, a dynamic range reduction helps to still spread the spectral information between different images in a stable and meaningful way. In this work we used:

$$f_i = \frac{1}{\left(\sum_j V_i^j \right)^{p_{dyn}}} \tag{10.2}$$

with exponent p_{dyn} controlling the degree of dynamic range compression. $p_{dyn} = 0$ would mean no compression where $p_{dyn} = 1$ would normalize all pixel colors. During our experiments we noticed that values for p_{dyn} between 0.5 and 0.6 yield the best results. Nevertheless, it should be noted that the formal notation would also allow for other types of dynamic range compression which might be better suited for other material classes.

It is important to note that one needs to also scale V_i by f_i before projecting to the appearance space to determine the neighborhoods N_i. This way - and by using the neighborhood search algorithm outlined in Section 6.2 - the degree of connectivity in the neighborhood graph between different images is increased and the reconstruction stabilized.

Implementation-wise we can pre-compute $D_j = \sum_{i=1}^{n} F_i^j f_i D_i^{b_j}$, they are filtered, scaled versions of the measured images D^{b_j}. Also, we can pre-compute $f_i V_i$ and optimize for a scaled $S' = (S_i f_i)$ instead of S and later on derive the result pixels as $S_i = S'_i / f_i$. Especially pre-computing the D_j yields a large performance gain.

10.1.1 Single Spectral Image

In simple use cases, just one spectral image is to be acquired. Then, our method needs one high quality RGB image and one noisy spectral image per desired spectral band as input data. With known noise level and camera sensitivity, an appropriate filter size and shape can be chosen per spectral band. In our implementation we used a gaussian filter with limited support to speed up calculations. With known filters the D_j are calculated from the noisy spectral images. It is important to note that the filters F^j on the image should overlap largely because this also helps to recover the high-frequent spatial details. In our implementation the neighborhood sets N_p are also determined in RGB space based on the values in V. In Section 10.4.1 we show and discuss results on single images.

10.1.2 Multiple Spectral Images

For a lot of applications multiple spectral images of very similar scenes need to be taken. In the area of computer graphics, this especially includes measurements of material's reflectance. There, a series of images of a material sample is captured, where illumination and/or viewing directions vary from image to image. However, between images of similar view or light direction we expect a lot of self-similarity inside and across images in many practical cases. That means, that it is no longer necessary to capture a noisy spectral image for every spectral band and every RGB / desired output image. The regularization term will "transport" the spectral information between the images and lead to a reconstruction of the full spectrum

for each image even if there are only few or even only one noisy spectral image per RGB image. As we will demonstrate in Section 10.4.2 it is even possible to add RGB images to the reconstruction having no spectral information at all - as long as the image content is sufficiently similar to the content of the other images. In effect, this drastically reduces measurement time when capturing series of similar images. For many hardware configurations an exposure time for the spectral camera suffices that is comparable to that of the RGB camera. When taking images in parallel, this can nearly speed up spectral imaging to standard RGB imaging.

10.2 Acquisition

In this section we describe the hardware setup used during our experiments. We start with some explanations on the original RGB setup and on the additional dedicated hardware required for the spectral data. Afterwards the calibration of the setup is explained with greater detail.

10.2.1 RGB Measurement Setup

For our first practical experiments both with single images as well as with full reflectance capture we utilized an existing RGB setup [SSWK13], [SSW+14]. The setup has a hemispherical gantry covered with 198 2.5W Barthelme LEDs which can be switched on and off individually. 11 RGB industry video cameras (SVS Vistek 4022, the same model used in Chapter 9) with a resolution of 2048x2048 pixels are integrated into the gantry forming and arc on one side from the top position down to an elevation angle of 75° measured from the normal. To measure anisotropic materials as well, the sample holder is mounted onto a precision rotation stage which can orient the samples to arbitrary azimuth angles. Figure 10.1 shows a photograph of the setup.

10.2.2 Spectral Integration

To capture spectral data, we utilize the monochrome camera with the liquid crystal tuneable filter from our gonioreflectometer setup (see Chapter 3.2). This camera system is mounted at 45° elevation and with 15° azimuthal distance to the RGB camera at the same elevation. This is important since the LEDs are mounted at the gantry with the same azimuthal distance. After the sample is rotated by 15° the spectral camera captures the same content as the RGB camera before. Using this geometry a nearly pixel-correct alignment between spectral and RGB data is possible. Pixel matching with this accuracy would be very hard to achieve when

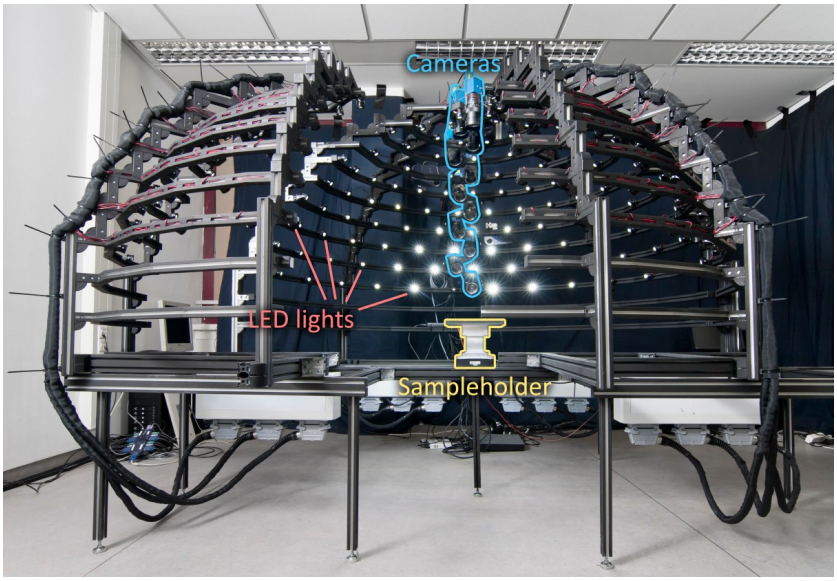


Figure 10.1: Photograph of the RGB measurement setup utilized during our experiments.

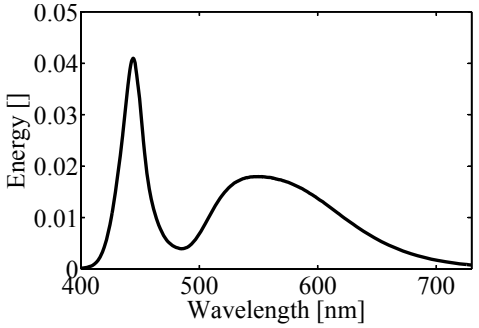
a push-broom or whisk-broom device is used. Nevertheless, our reconstruction method is quite robust against small misalignment of the data due to the filtering.

The measurement process is as follows: the sample is rotated to 24 different positions in 15° steps. For every rotation angle, an image is taken by the RGB cameras and the spectral camera using multiple LEDs switched on to have a high quality image in which the sample holder can be detected with sub-pixel accuracy to have a perfect alignment between the images of the single cameras. Afterwards the LEDs are serially switched on, images are captured, and the respective LED is switched off again. Since the RGB and the spectral camera system are operated in parallel, and since the spectral camera system is driven at short exposure times, the measurement process is nearly as fast as a traditional RGB measurement using the same setup.

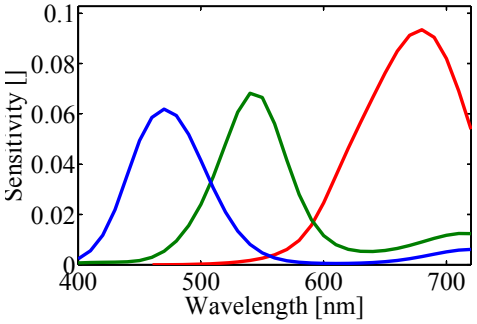
10.3 Calibration

We performed a careful calibration of the setup. Here, cameras and light sources have to be calibrated geometrically and radiometrically. Since geometric calibration of such setups is out of scope of this paper we focus on the radiometric calibration here.

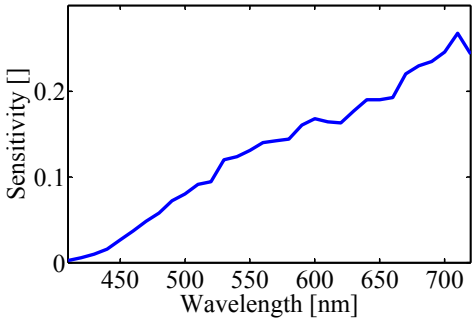
First of all we measured the spectrum of our LED light sources using an Ocean Optics USB4000 spectrophotometer. The spectrum is shown in Figure 10.2(a).



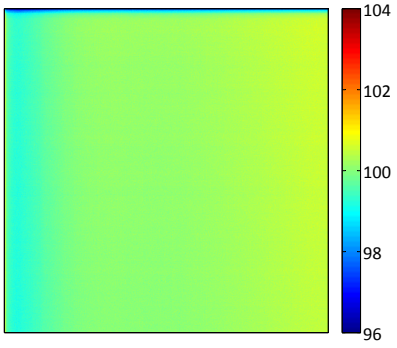
(a) Spectrum of LED



(b) Spectral sensitivity of the three different color pixels of the RGB camera



(c) Spectral sensitivity of the spectral camera system



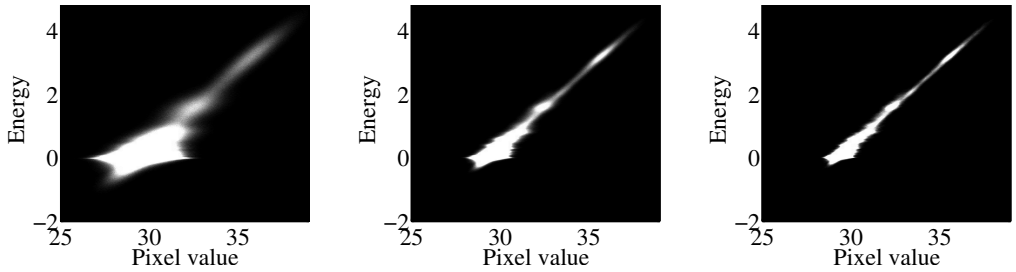
(d) Color coded dark-frame of Coolsnap camera

Figure 10.2: Calibration of the measurement setup.

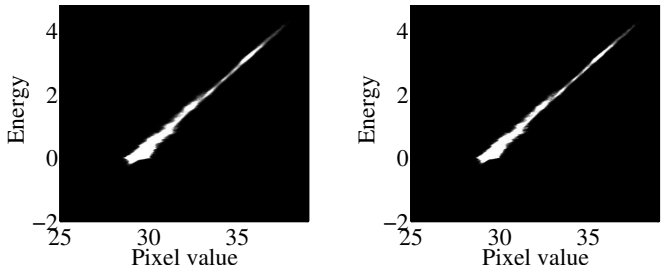
For the cameras two different kinds of calibrations are required: on the one hand recovery of the opto-electronic conversion function (*OECF*) to correct for non-linearities in the sensors response to light and on the other hand spectral calibration which aims at finding the response efficiency of the optics-camera combinations depending on the wavelength.

Linearizing the response of the RGB cameras is straightforward and was performed using the algorithm of Robertson et al. [RBS03]. To obtain the effective spectral response R_{cam} of the RGB cameras we employed our method from Chapter 9. The spectral sensitivity of the RGB cameras is shown in Figure 10.2(b).

For the spectral camera much more care has to be taken since here measurements with a very low SNR will be used. Therefore, the calibration of the camera must also hold for very low pixels values that are typically not used for measurement purposes. We, however, want to use this range as well, because we use largely underexposed spectral images as input data. We therefore require a cali-



(a) Filter size 5 (b) Filter size 13 (c) Filter size 23



(d) Filter size 31 (e) Filter size 39

Figure 10.3: Checking the stability of the output signal of the monochrome camera in the low SNR range: Two-dimensional response histograms (brighter means more values in the respective bin) on filtered, dark-subtracted (offset by 30) input images after applying gaussian filters of increasing size. As one can see, negative values after dark-subtraction need to taken into account since a clipping would bias the result largely (for that reason the offset). The cameras output in this range is statistically stable despite the fact that the output of the single pixels is shot-noise limited. Due to the stability, data in this pixel value range can be used as an input for our method.

bration which is exact for the low pixel values, too.

The first step is to obtain exact knowledge about the cameras dark current in all pixels. For this we took 200 photographs with closed shutter and repeated this process for different exposure times. The average over the dark frames is computed with floating point accuracy. One of the mean dark-frames is shown in Figure 10.2(d). One can easily spot the limited electron transfer efficiency along the CCD shifting direction.

To recover the *OECF* in the low-pixel-value range we propose to apply a modified variant of the Robertson algorithm [RBS03]. This algorithm recovers both a linearized HDR image and the inverse of the *OECF* from a series of images of a static scene captured with different exposure times. We propose to apply this algorithm to dark-subtracted and filtered images. The filtering is done using a

gaussian kernel as in the spectral recovery method.

The original energy recovery (Equation (7) in [RBS03]) can be stated as follows:

$$x_j = \frac{\sum_i w(y_{ij}) t_i f(y_{ij})}{\sum_i w(y_{ij}) t_i^2} \tag{10.3}$$

where x_j is the (unknown) energy of pixel j, y_{ij} the device counts measured in pixel j with exposure time t_i and w a weighting function for the pixel values. f is the unknown response curve, the inverse of the *OECF*, which has a discrete nature in [RBS03].

To account for noisy images with significant dark-current contribution we modify equation (10.3) to include filtering:

$$\tilde{y}_{ij} = \sum_{p=1}^{P} F_p^j [y_{ip} - D_p] \tag{10.4}$$

$$\tilde{x_j} = \frac{\sum_i w(\tilde{y}_{ij}) t_i f(\tilde{y}_{ij})}{\sum_i w(\tilde{y}_{ij}) t_i^2} \tag{10.5}$$

where F^j is a gaussian filter centered at pixel j, D_p is the dark value at pixel p and P is the number of all pixels. The same filter kernel is used as in the recovery method (see Equation 10.1). Here, f is no longer discrete. However, it turned out to be sufficient to store values for f at integral pixel values and interpolate linearly in between.

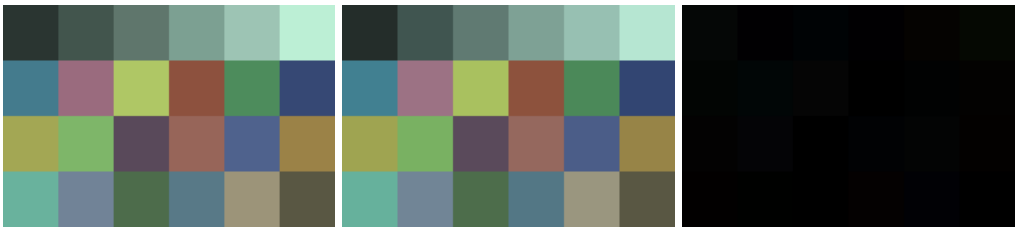
The recovery of f (Equation (11) in [RBS03]) has to be adopted to account for the fractional pixel values. Here, generation of the pixel value histogram has to include interpolation as well. Let $\delta_{ij} = \tilde{y}_{ij} - \lfloor \tilde{y}_{ij} \rfloor$, then the recovery of a integral point of f changes to:

$$L(m) = \{(i,j) | \lfloor \tilde{y}_{ij} \rfloor = m\} \tag{10.6}$$

$$R(m) = \{(i,j) | \lceil \tilde{y}_{ij} \rceil = m\} \tag{10.7}$$

$$\begin{aligned} f(m) &= \frac{1}{2 \sum_{(i,j) \in L(m)} (1 - \delta_{ij})} \sum_{(i,j) \in L(m)} (1 - \delta_{ij}) t_i \tilde{x}_{ij} \\ &+ \frac{1}{2 \sum_{(i,j) \in R(m)} \delta_{ij}} \sum_{(i,j) \in R(m)} \delta_{ij} t_i \tilde{x}_{ij} \end{aligned} \tag{10.8}$$

To show the effect of filtering on response recovery we computed *response histograms* for an exposure series having low pixel values. A *response histogram* is a two dimensional histogram where each bin is associated with a pixel value $\tilde{y}_{ij}$ and the associated energy estimate $f(\tilde{y}_{ij})/t_i$. Such a histogram shows the function f together with noise or other unwanted effects over the whole pixel value range.



(a) Real RGB response (b) Simulated RGB response (c) Difference

Figure 10.4: Cross validation of the whole spectral calibration: (a) shows the real response of the RGB cameras to color fields, (b) shows a simulation of the responses by applying the RGB filtermatrix R_{cam} to a spectral image captured using the spectral camera system (c) is the difference image showing that our calibration used to generate the image in (b) is very exact. For display purposes, gamma mapping with $\gamma = 2$ has been applied to the RGB images.

An ideal, linear camera without any noise or bias would just show a straight line of filled bins in this histogram, with all other bins being zero.

Figure 10.3 shows the *response histograms* for our camera using different filter sizes during the recovery of f. The histograms converge towards the ideal case of a single line with increasing filter size. Therefore, the response of the camera is stable under filtering even for pixel values with an SNR of 1 or smaller. Notice, that the dark-current-noise of the camera is normal distributed with standard deviation 3. The discrete nature of the photon and electron shot-noise completely cancels out when multiple pixels are combined by filtering.

The spectral calibration of the spectral camera system is straightforwardly performed by acquiring images of a white diffuse reference surface lit by the known LED illumination and then dividing the linearized camera response by the LED spectrum. The resulting sensitivity for each wavelength band can be taken from Figure 10.2(c).

Finally, we checked the complete calibration of both types of cameras by a cross-validation experiment. This is done by simulating the RGB response to a X-Rite colorchecker passport from a spectral image acquired by the spectral camera system. This way all calibration results are combined. The simulated RGB response matches the real response extremely well which is shown in Figure 10.4.

10.4 Simulated Data

To evaluate, how well extrapolation of spectra across different pixels AND angles works, we performed some first simulations with the spectral BTF data captured

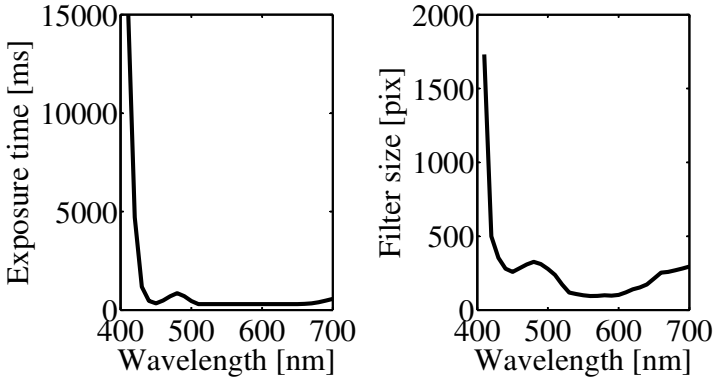


Figure 10.5: The right parameters for our setup: Our simulation helps to choose good values for the capture and reconstruction parameters in beforehand. Here we show the exposure times t_λ for the spectral camera system and the filter sizes of the gaussian filters in the reconstruction process. The standard deviation σ for the gaussians was then set to 1/5 of the filter size.

by the gonioreflectometer (see Part I).

Moreover, simulations are extremely helpful because a pixel-wise comparison between reference and reconstruction can be made. This is not easily possible when using real data since the geometries and resolutions of the setup used for the ground-truth measurements and the setup described in Section 10.2 are quite different. Furthermore, basic understanding about the method and its properties can be gained without having too much bias by real-world data.

To have an as-realistic-as-possible scenario we performed a detailed simulation of the capture process. For the image capture in the CCD cameras the calibration results from Section 10.3 are used. The following equation describes how a pixel value M is generated from an incident irradiance x and exposure time t:

$$M_{cam}(x,t) = clamp\left(\left\lfloor f_{cam}^{-1}(x\,t\,E_\lambda^{cam}) + O_{ADC} + N_\sigma \right\rfloor, 0, n-1\right) \qquad (10.9)$$

E_λ^{cam} denotes the combined quantum efficiency of optical system and sensor at wavelength λ and f_{cam}^{-1} is the *OECF* i.e. the inverse of the response function. O_{ADC}^{cam} is the ADC offset of the camera and $N_{\sigma_{cam}}$ a random value drawn from a normal distribution with standard deviation σ_{cam} mimicking the various noise sources inside of the capture process. f_{cam}, E^{cam}, O_{ADC}^{cam} and σ_{cam} are taken from the calibration described in section 10.3. We therefore end up with a detailed simulation of our measurement setup that helps us to evaluate the quality of spectral reconstruction in dependence on the various parameters.

Starting with a spectral reference image S_{ref} and the light spectrum L we first

simulate an exposure series for the RGB images by computing

$$M_{RGB}(CS_{ref}(p)diag(L), t_i) \tag{10.10}$$

for all pixels p of the reference image S_{ref} and for different exposure times t_i. Here, M_{RGB} is the function from Equation 10.9 with the parameters for the RGB camera system. The RGB-LDR images are recombined to form HDR images using Robertsons algorithm [RBS03].

For the spectral images, one single exposure per wavelength band is simulated by computing

$$M_{spectral}(S_{ref}(p)diag(L), t_\lambda) \tag{10.11}$$

In our setup, standard RGB reflectance capture is performed using four exposure times $t_i \in \{50, 186, 666, 2500\}$ for the RGB cameras. For the spectral camera, we calculated the exposure time as $t_\lambda = \frac{t_{base}}{\sqrt{E(\lambda)L(\lambda)}}$. That is, we accept more noise in those spectral bands having bad support from light source and camera system quantum efficiency. The size of the gaussian filters in the reconstruction has to be adjusted accordingly. Figure 10.5 shows both the resulting exposure times and filter sizes. It should be noted that this figure depicts the filter sizes that are required to nearly completely cancel the noise from the spectral band images. However, we found that this is often not necessary and smaller filters can be used. The reason is that the appearance smoothness term of the energy functions also acts like a filter and compensates for remaining noise in the filtered, spectral data. Smaller filters can be appreciable because they lead to a speed-up when calculating the matrix multiplication corresponding to the energy function E_n.

Since some of the exposure times t_λ are larger than the total RGB exposure time $\sum_i t_i$ we cannot obtain the spectral data at exactly the same speed as the RGB data for this setup. It is clear that we need to finish at least one spectral band exposure before switching to the next light source. However, for other RGB images we can acquire noisy spectral images for more than a single band. When using a bin-packing algorithm, the assignment between spectral channels and RGB images can be optimized. In our case we would be able to acquire the 30 spectral images during the acquisition of 8 RGB images with about 41% increase in total exposure time. However, we decided to add two more RGB images to the sequence and acquire the noisiest images two times. With this, we end up with 10 RGB images and an increase of about 61% in total exposure time. Figure 10.16 later in this thesis shows the noisy spectral images for the "Lego" dataset for reference.

It should be noted that the whole procedure itself is very general and the parameters can be easily adjusted for other measurement setups.

10.4.1 Single Spectral Image

In simple use cases, just one spectral image has to be acquired. We will investigate this case to derive first insights into the reconstruction quality of the modified *spectralization* method utilizing noisy spectral data.

In this single-image case, the input to the reconstruction algorithm is a high quality RGB image and as many noisy spectral images as we have spectral bands (in our case 30, ranging from 410 to 700nm in 10nm steps). We used one image (texture) per example material as reference image S_{ref}. After running our new algorithm, the result S is compared to S_{ref}. To evaluate the spectral accuracy of our reconstruction algorithm, we calculated histograms for the relative reconstruction error and the perceptual ΔE^*_{94} for all four example materials. See Equations 6.7 and 6.8 for details on the error computation.

Figure 10.6 show first results for two datasets. The reconstruction quality is not very satisfactory. Mainly, the reason is noise in the resulting spectral curves. We found that this is due to the fact that the quadratic function has very small gradient as soon as the first two terms of the energy function (see Equation 10.1) approach zero. As a solution, we extend the energy function by adding a spectral smoothness term:

$$\begin{aligned} E_n(S) &= \alpha \sum_{j=1}^{K} \left\| \left(\sum_{i=1}^{n} F_i^j f_i S_{i,b_j} \right) - \left(\sum_{i=1}^{n} F_i^j f_i D_i^{b_j} \right) \right\|^2 \\ &+ \beta \sum_{i=1}^{n} \left\| (S_i f_i) R_{cam} - f_i V_i \right\|^2 \\ &+ \gamma \sum_{i=1}^{n} \sum_{j \in N_i} \left\| S_i - S_j \right\|^2 \\ &+ \delta \sum_{i=1}^{n} \sum_{j=1}^{k-1} \left\| S_{i,j} - S_{i,j+1} \right\|^2 \end{aligned} \tag{10.12}$$

This is physically well supported since we are reconstructing reflectance spectra here, which are - as mentioned earlier - mostly smooth. With the introduction of the new term with the new weighting factor δ we change the normalized to the number of terms that was previously introduced in Equation 6.11. The new

variant is:

$$\begin{aligned}
\alpha &= \alpha' \frac{mk + 3n + k\sum_{i=1}^{n} |N_i| + n(k-1)}{mk} \\
\beta &= \beta' \frac{\bar{D}}{\bar{V}} \frac{mk + 3n + k\sum_{i=1}^{n} |N_i| + n(k-1)}{3n} \\
\gamma &= \gamma' \frac{mk + 3n + k\sum_{i=1}^{n} |N_i| + n(k-1)}{k\sum_{i=1}^{n} |N_i|} \\
\delta &= \delta' \frac{mk + 3n + k\sum_{i=1}^{n} |N_i| + n(k-1)}{n(k-1)}
\end{aligned} \tag{10.13}$$

The rationale of this normalization stays the same: to account for the difference in number of terms and for the different scale of D and V in the energy function. Since the new spectral smoothness term is just meant to help with final convergence, we set $\delta' = 0.1$, so much lower compared to the other three terms.

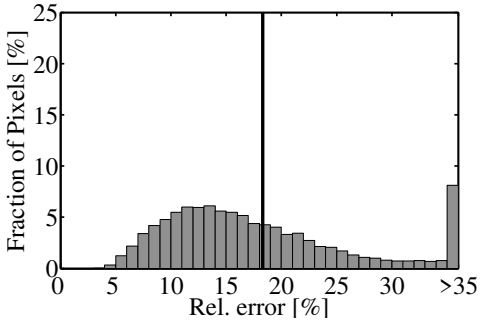
Figures 10.7 and 10.8 shows the resulting reconstruction error histograms for all four datasets when using this new energy function. The ΔE^*_{94} histograms show a high concentration of errors below one which corresponds to indistinguishable stimuli. Very few ΔE^*_{94} errors are above two for all materials showing that the reference and the reconstruction are perceptually nearly equivalent. The histograms of relative errors indicate high reconstruction accuracy as well. An exception is the "Red fabric" dataset - we will discuss this in Section 10.4.2.1.

Plots of different reflectance spectra are depicted in Figures 10.9, 10.10, 10.11, and 10.12. One can see that the spectral shapes are reproduced faithfully (again, with exception of the "Red fabric").

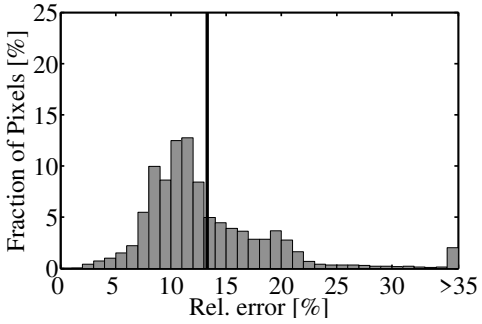
There are mainly two remaining differences:

1. Differences due to noise in the reference data: During *spectralization* the pixel-wise noise is mainly filtered out due to the prior - the appearance smoothness constraint using the RGB-based appearance neighborhoods.

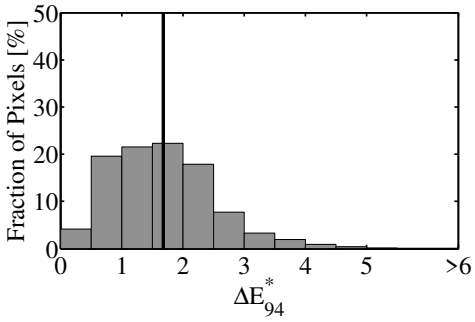
2. A bit of over-smoothing of the spectral curves near the limits of the spectral range: This is due to the fact that the capture system has rather low sensitivity in that spectral bands (see Figures 10.2(c) and 10.2(a)) and therefore the spectral smoothness term has a slightly higher ratio compared to the other wavelength bands. This could be compensated by scaling the per-band spectral smoothness terms accordingly. However, the higher prior weight reflects the higher uncertainty in the data for those bands and we would get more noise into those bands, if we would do so. We will accept this inaccuracy for the time being and refer to the fact that a capture system with more evenly distributed spectral efficiency would not suffer from this problem.



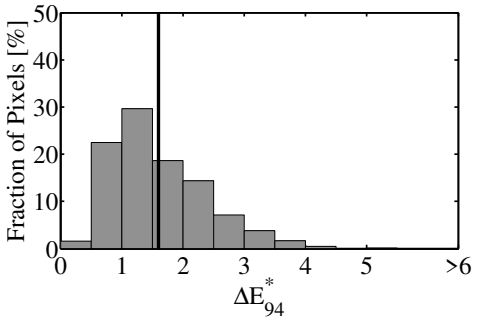
(a) Red fabric: Histogram



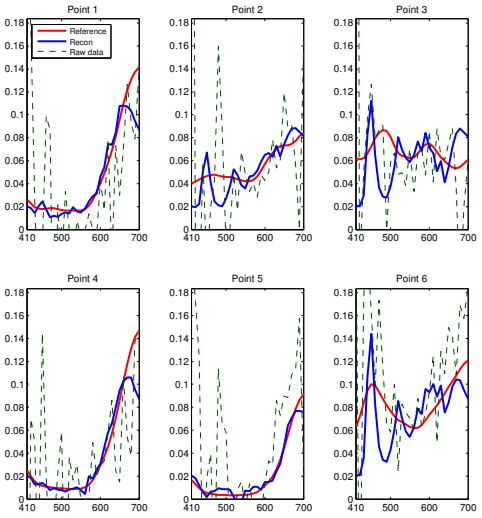
(b) Lego bricks: Histogram



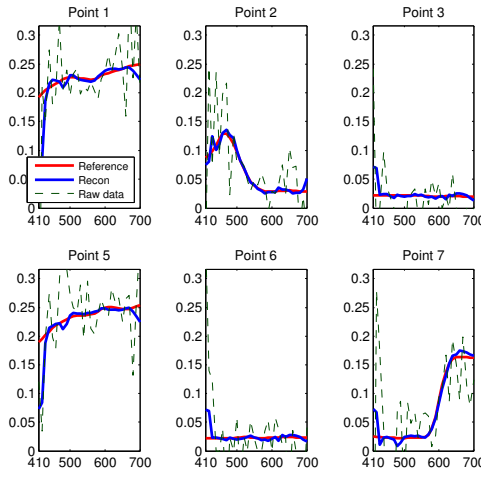
(c) Red fabric: Histogram



(d) Lego bricks: Histogram

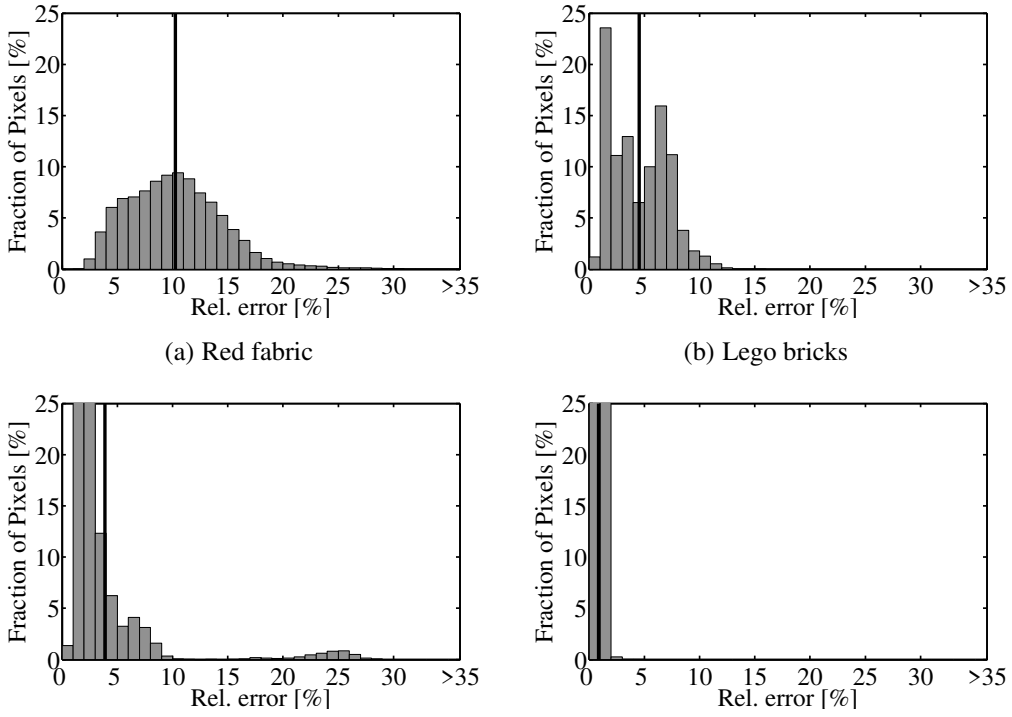


(e) Red fabric: Example pixels



(f) Lego bricks: Example pixels

Figure 10.6: Relative error for reconstructing single images using the basic energy function from Equation 10.1: histogram and pixel spectra for two example textures. The dashed curves indicate the noisy spectral input data (normalized to the LED light source spectrum for comparison). The reconstruction errors are not very satisfactory due to high noise in the reconstructed spectra.



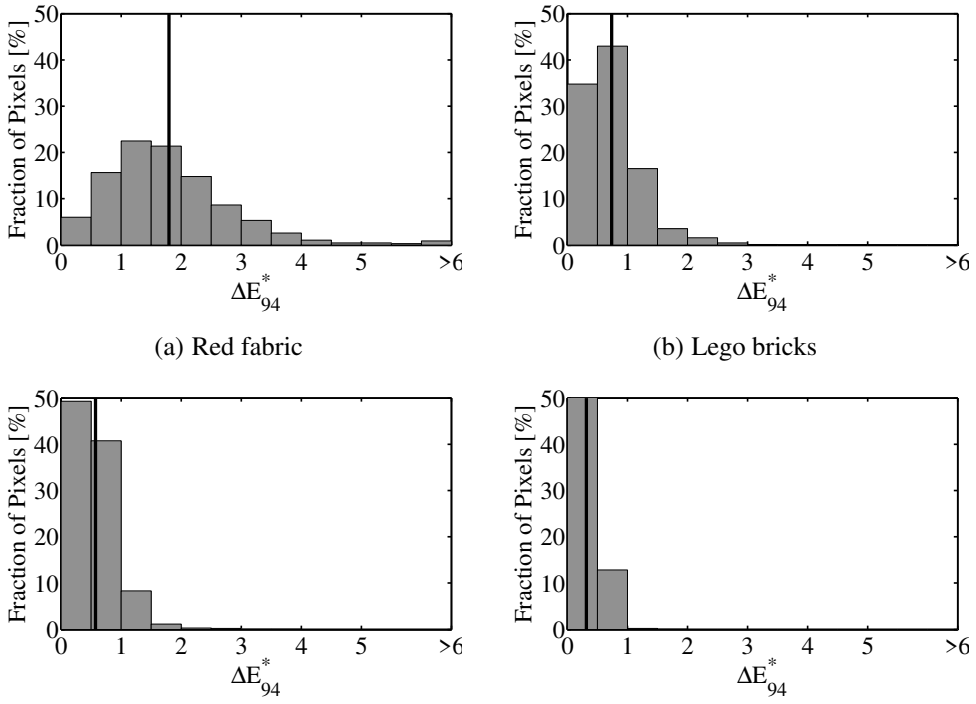
(a) Red fabric

(b) Lego bricks

(c) Color checker

(d) Wallpaper

Figure 10.7: Relative error for reconstructing single images using the energy function with spectral smoothness term (Equation 10.12): relative error histogram for all of the four example textures with the black line indicating the mean error. The high concentration at low errors shows that our method now provides high-quality spectral images. The only exception is the "Red Fabric" dataset which shows are broader error distribution with higher mean error. Please see Section 10.4.2.1 for a discussion on this dataset.



(a) Red fabric

(b) Lego bricks

(c) Color checker

(d) Wallpaper

Figure 10.8: Perceptual error for reconstructing single images using the energy function with spectral smoothness term (Equation 10.12): ΔE^*_{94} histogram for all of the four example textures with the black line indicating the mean error. The concentration of errors near or even below one show that our reconstruction algorithm produces perceptually indistinguishable results compared to the reference. The only exception is the "Red Fabric" dataset which shows are broader error distribution with higher mean error. Please see Section 10.4.2.1 for a discussion on this dataset.

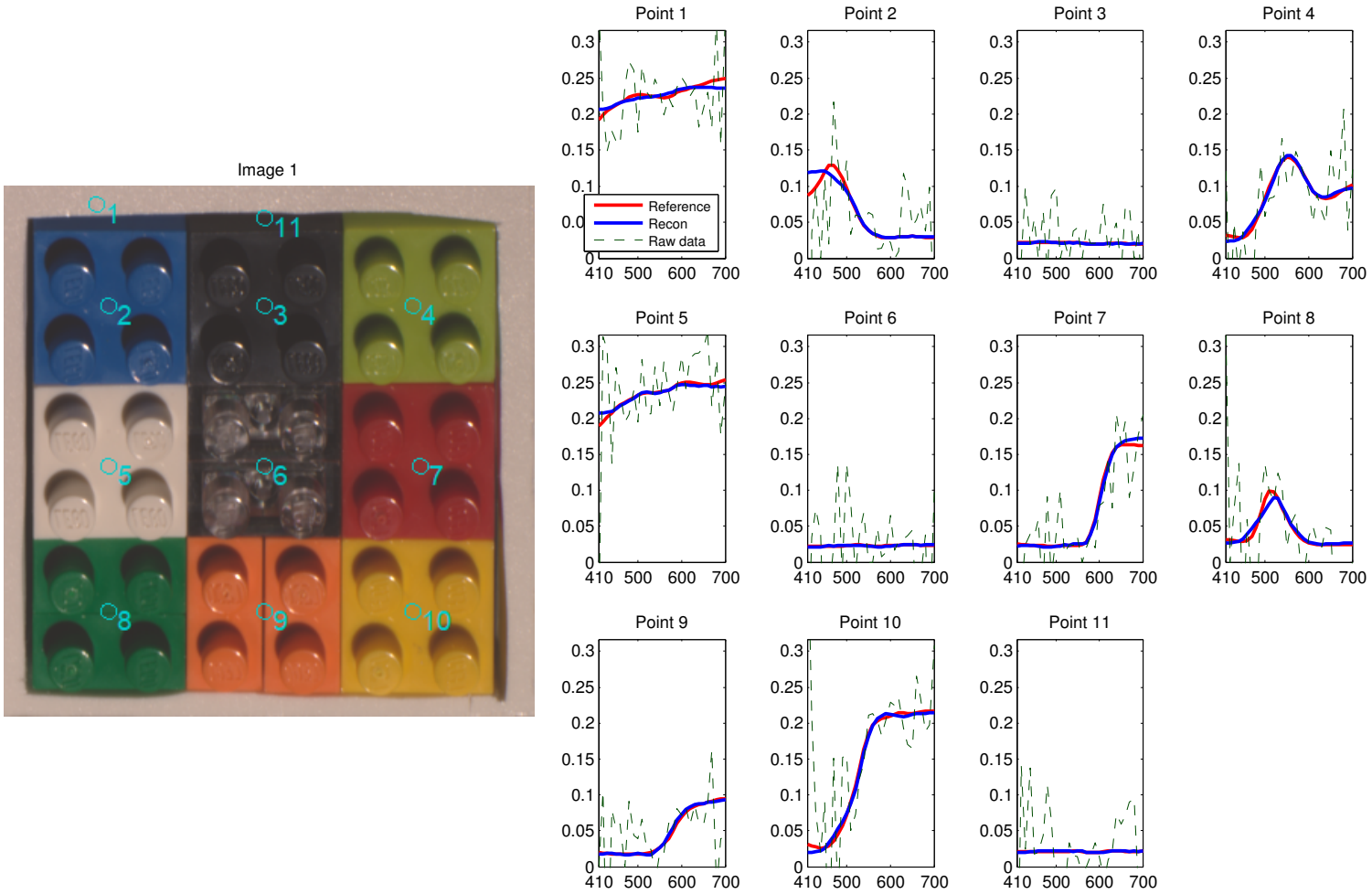


Figure 10.9: Comparison of reference and reconstructed reflectance spectra of selected pixels in the single-image "Lego" dataset. The shape of the spectra is met with high precision. The dashed curves indicate the noisy spectral input data (normalized to the LED light source spectrum for comparison).

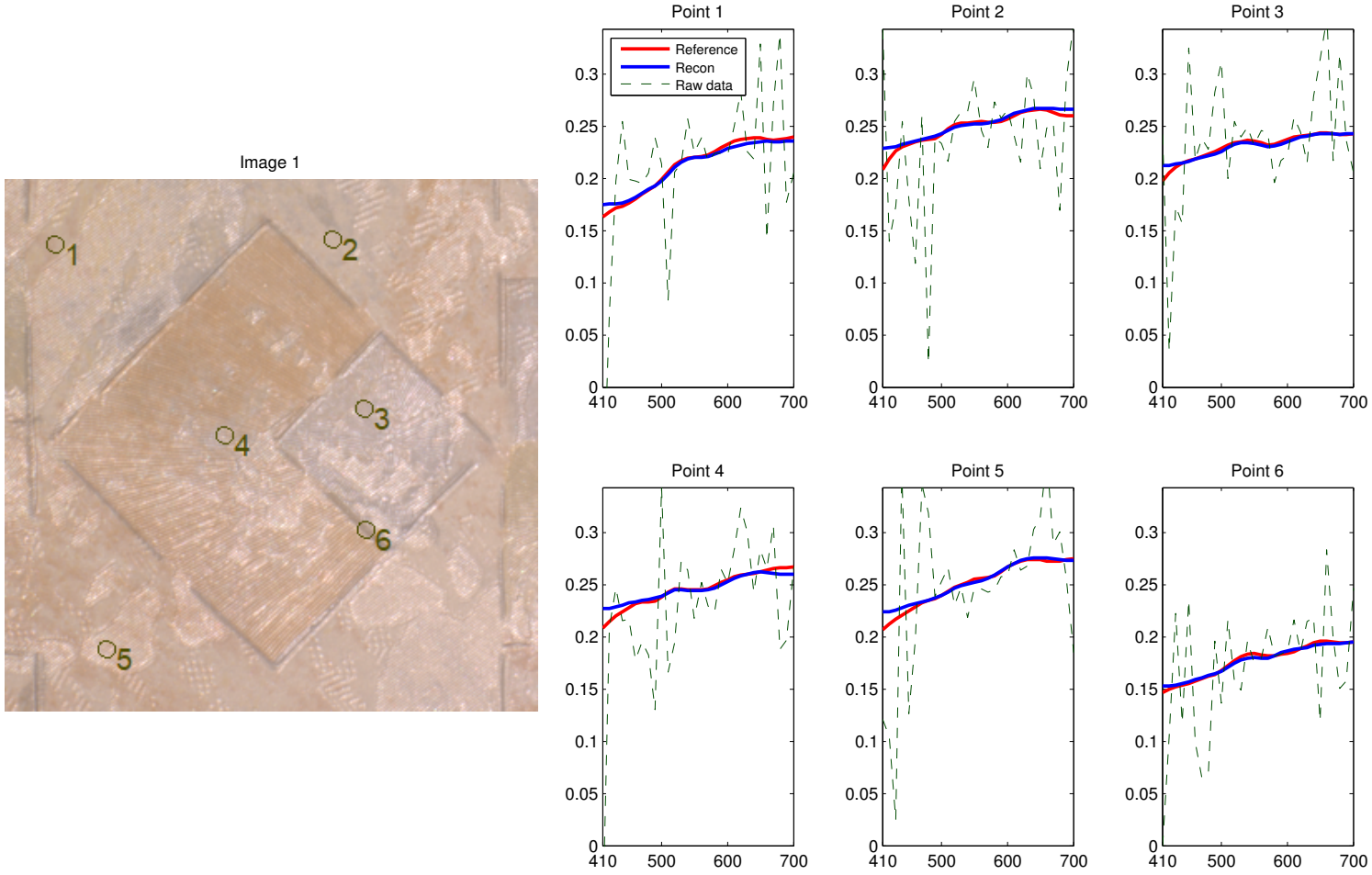


Figure 10.10: Comparison of reference and reconstructed reflectance spectra of selected pixels in the single-image "Wallpaper" dataset. The shape of the spectra is met with high precision. The dashed lines indicate the noisy spectral input data (normalized to the LED light source spectrum for comparison).

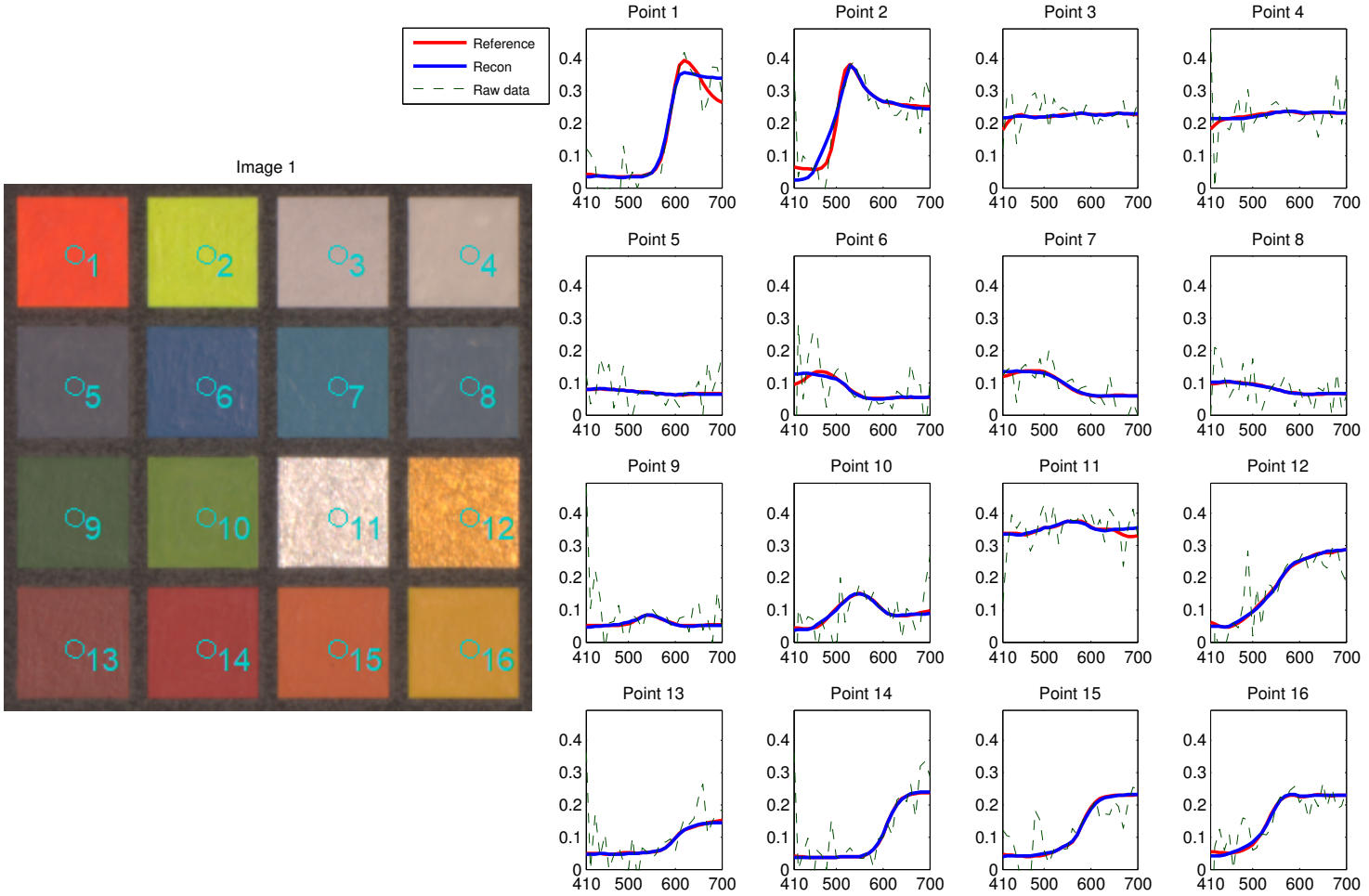


Figure 10.11: Comparison of reference and reconstructed reflectance spectra of selected pixels in the single-image "Colorchecker" dataset. The shape of the spectra is met with high precision. The dashed lines indicate the noisy spectral input data (normalized to the LED light source spectrum for comparison).

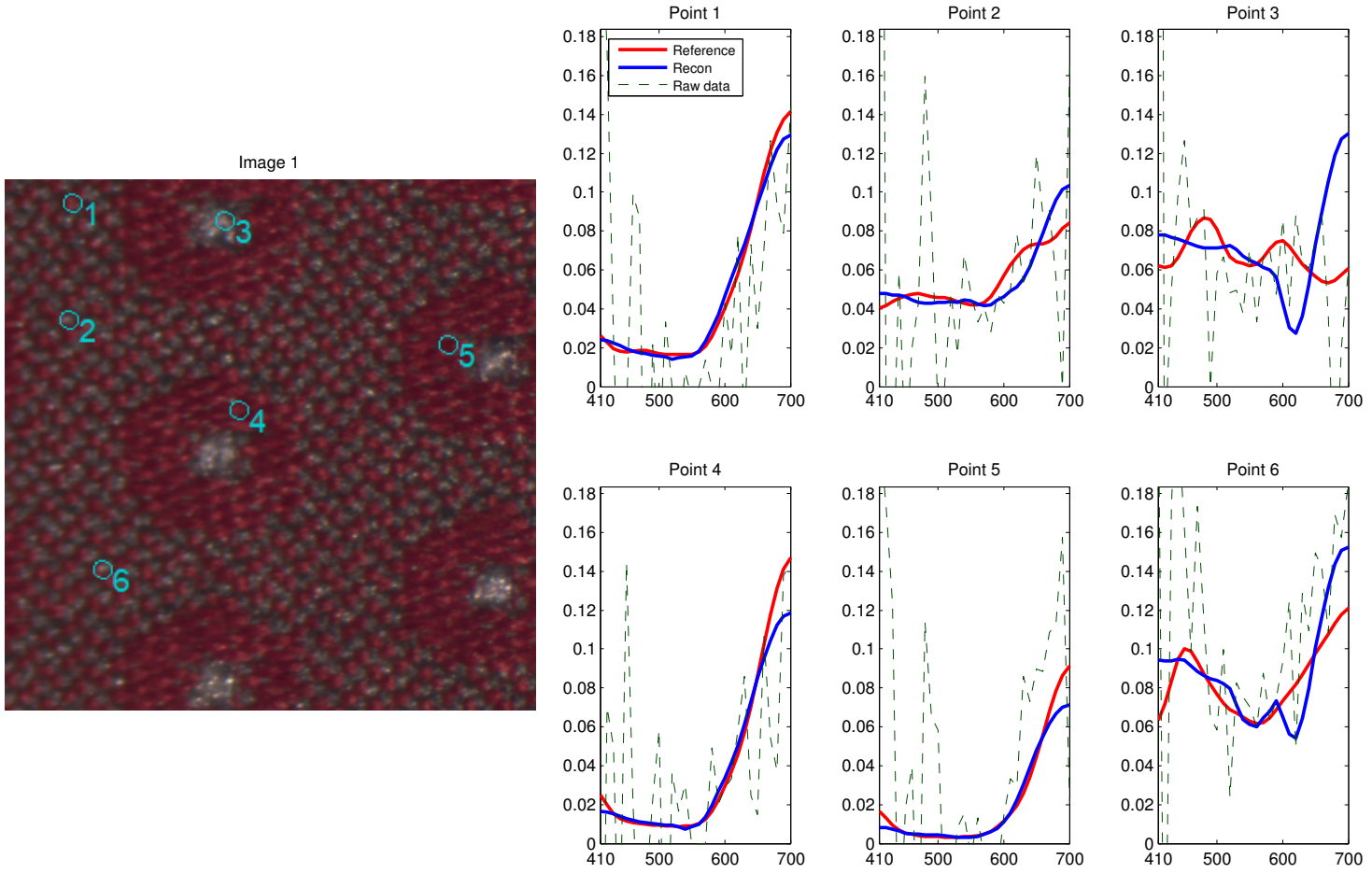


Figure 10.12: Comparison of reference and reconstructed reflectance spectra of selected pixels in the single-image "Red Fabric" dataset. The dashed lines indicate the noisy spectral input data (normalized to the LED light source spectrum for comparison). One can depict some significant inaccuracy in the reconstructed spectra. This is a failure case of our method which is in more detail discussed in Section 10.4.2.1.

10.4.2 Multiple Spectral Images

Reconstructing multiple images at once as described in Section 10.1.2 is the much more challenging case.

In the described setting, there is no single RGB image for which noisy spectral information is available for all spectral bands. We want the prior term of Equation 10.12 (see also Chapter 6.2) to "spread" the information between the images. The idea is to use this algorithm to reconstruct complete bi-angular, spectral reflectance data without significant increase in exposure time.

As described in detail at the start of Section 10.4 our measurement setup is able to acquire the necessary noisy spectral images within the time needed for 10 high quality, high dynamic range, RGB images. For that reason we use test datasets with 10 images having noisy spectral information and 10 images without any spectral information. The 10 images with spectral information are chosen to correspond to views of 45° elevation angle since the spectral camera is mounted at this angle. Exemplary, Figure 10.15 shows the 20 images for the "Lego" dataset.

Figure 10.13 depicts histograms of the reconstruction errors on the four datasets. For each data-set there are two histograms: one showing the error distribution for the first 10 images and the other one for the last 10 images with no spectral information. We can derive the following conclusions:

1. The reconstruction errors are in general rather low - our algorithm is able to reconstruct high quality spectral images.
2. There is only a slight increase of reconstruction error from the first 10 images to the last 10 images, showing as a shift in the histograms. This means that the first 10 images already contain enough spectral information to allow for a good transport via the appearance smoothness prior. The main reason is that the materials have a complex structure and thus every image contains a lot of information ranging from diffuse colors to highlights. In case of less- to non-structured materials this might be different and might potentially render our method non-applicable. However, we refer to our BRDF reconstruction in Section 7.3 as an option in these cases.
3. For the "Red fabric" dataset, the reconstruction is less accurate compared to the other datasets. We had already noticed that in Section 10.4.1 and discuss this in the following section. However, finding 2 also holds for this data-set.

10.4.2.1 Red fabric

The results depicted in Figures 10.7, 10.12, 10.13, and 10.19 indicate that the reconstruction accuracy on the "Red fabric" dataset is worse compared to the other three datasets. At a first glance this is rather astonishing since this material is also rather structured and seems to have a lot of self-similarity - within and across

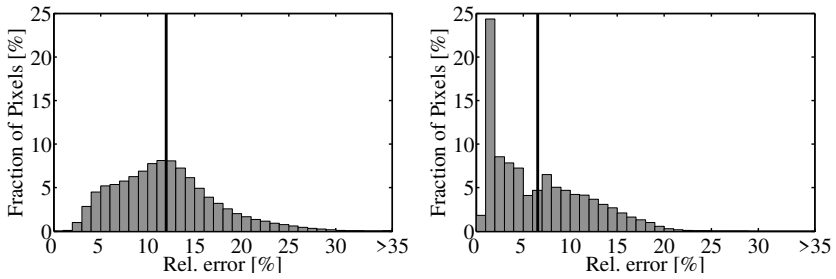
different images. However, the problem is revealed when having a more detailed look at the appearance neighborhoods N_i. It turns out that the textures of this material are full of metamers with respect to the capture system, i.e. there are different spectra in the texture that map to very similar RGB values. Figure 10.20 shows two examples of those metamers together with one example of a "good" neighborhood. In principle our method could deal with metamers, if they would show up on larger structures. However, in this case, the RGB-self-similarity is on a very small scale (the weaving pattern). For illustration, Figure 10.21 shows a comparison of raw and filtered data for the 600nm spectral band for the "Red fabric" and the "Lego" dataset. It can be seen that the fine details for the both materials are lost due to the filtering. However, for the "Lego" dataset, the larger structures (the bricks) are still clearly visible. Even two metameric bricks would therefore be reconstructed successfully if e.g. π_{app}^{RGBXY} from Equation 6.5 is used to create the appearance space. But for the "Red fabric" the appearance neighborhoods do only connect very similar positions in the same repeating small-scale structures that are all suffering from the same problem. That said, for textures with just fine-scale structures **and** metameric colors, our method as-is does not perform well.

As a possible solution we propose two improvements:

- Choose an appearance space projection with better neighborhood information. We tested even 5×5 pixel patches and found that they were not enough to distuingish between different metameric pixels. Since a 5×5 pixel patch already leads to a $5 \cdot 5 \cdot 3 = 75$ dimensional appearance space and since the nearest neighbor search gets increasingly costly with the space dimension, simply increasing the patch size is no valid option. Other descriptors might be better suited to describe the local textural structure with lower dimensional feature vectors.

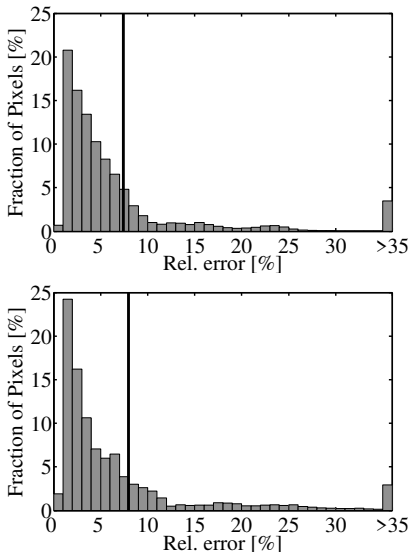
- Also capture one additional image for the whole reflectance measurement as proposed e.g. by Dong et al. (see also Section 10.4.3). From this, the spectrum of pixels could be converted into spectral constraint for the first term of the energy function $E_n(S)$ (see Equation10.12) if the appearance space distance would be below a certain threshold. Capturing the one additional image would not increase overall capture time by a lot, but would help resolve the aforementioned ambiguities.

However, we did not implement these possible improvements for this thesis and leave this to future research.

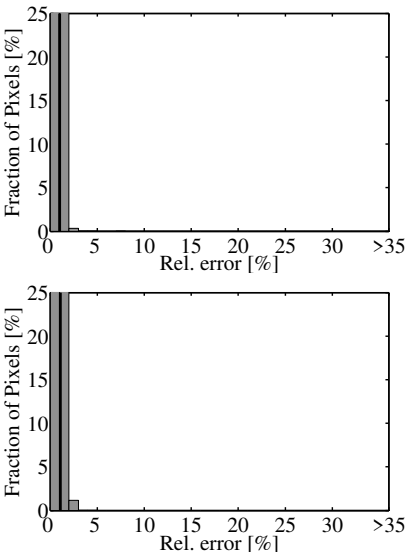


(a) Red fabric

(b) Lego bricks



(c) Color checker



(d) Wallpaper

Figure 10.13: Relative error for reconstructing multiple images at once: histograms for 20 images. The upper histograms correspond to the images from $45°$ elevation over which the spectral information was spread. The lower histograms were computed from the remaining 10 images, which had no spectral information at all and covered different elevation angles. There is only a slight increase of the errors - proving that the data acquired from $45°$ elevation already contains enough spectral information for our test cases.

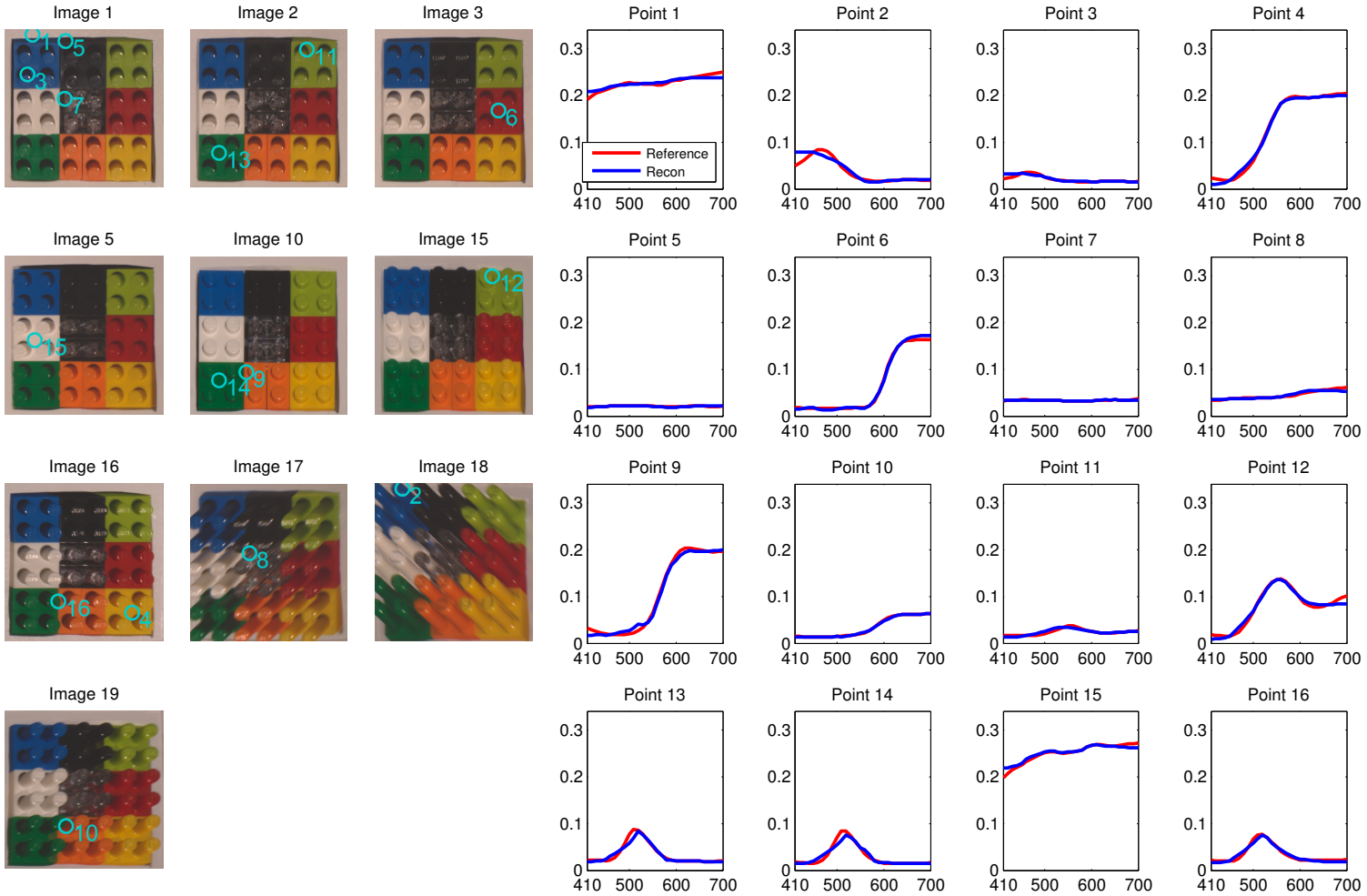


Figure 10.14: Comparison of reference and reconstructed reflectance spectra of selected pixels in the multi-image "Lego" dataset. The shapes of the spectra are met with high precision.

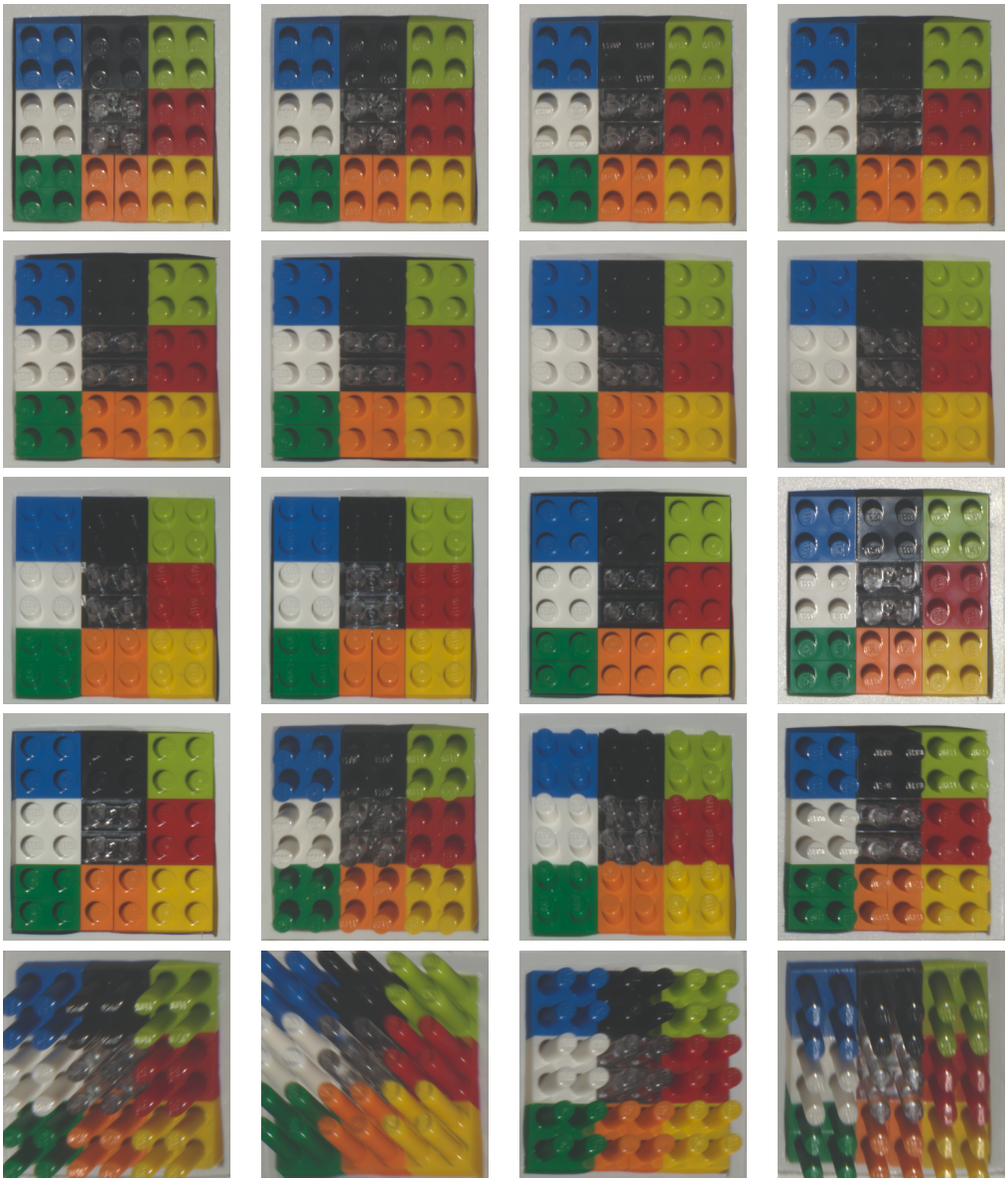


Figure 10.15: The "Lego" dataset: the 20 images - projected to sRGB color space for display purposes - used to compute simulated results. The first 10 images are from 45° elevation angle and sparse, noisy spectral information is available for them. The other 10 images have different viewing angles and no spectral information is available for them.

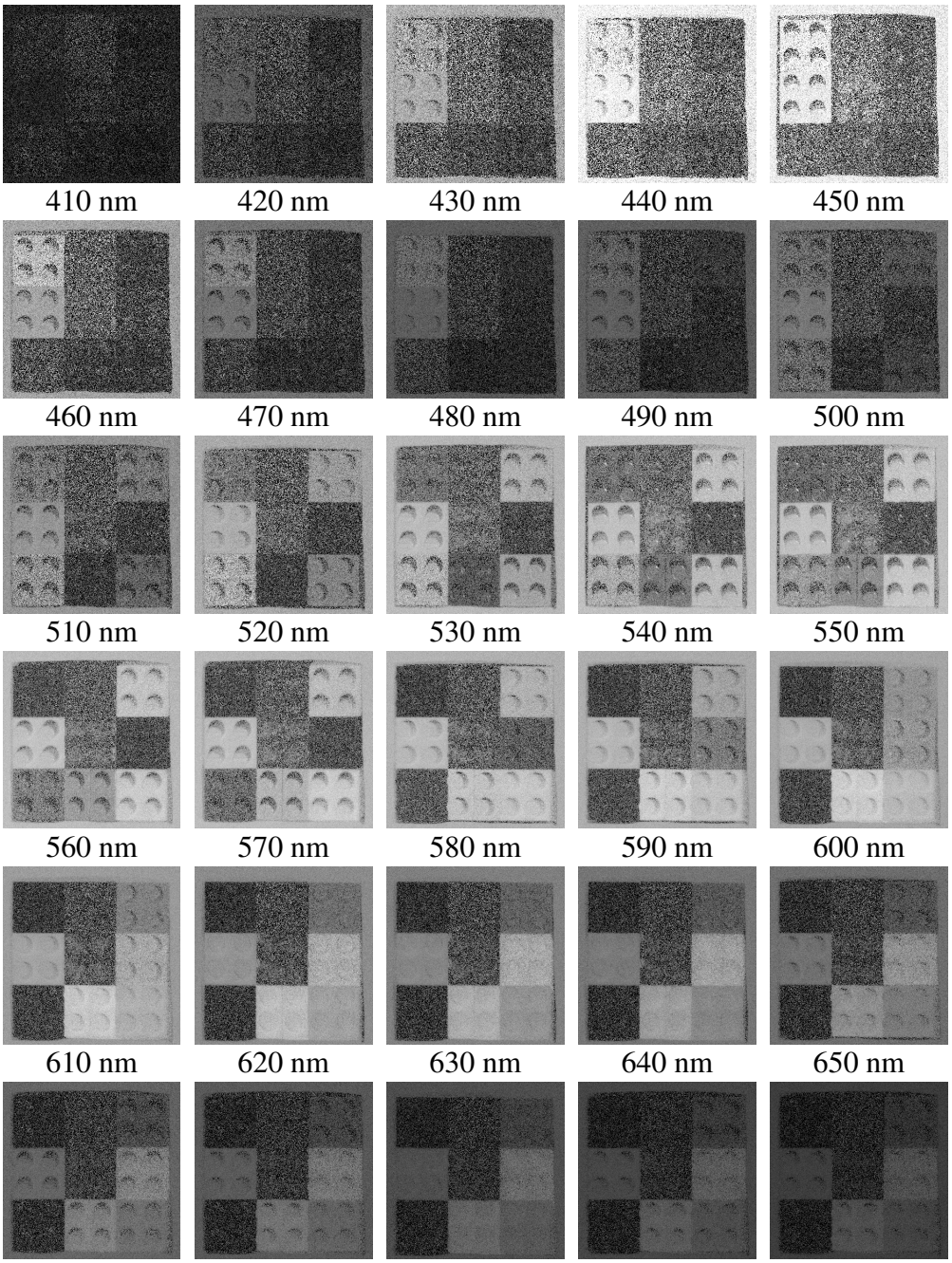


Figure 10.16: The "Lego" dataset: the noisy spectral images serving as input for our reconstruction.

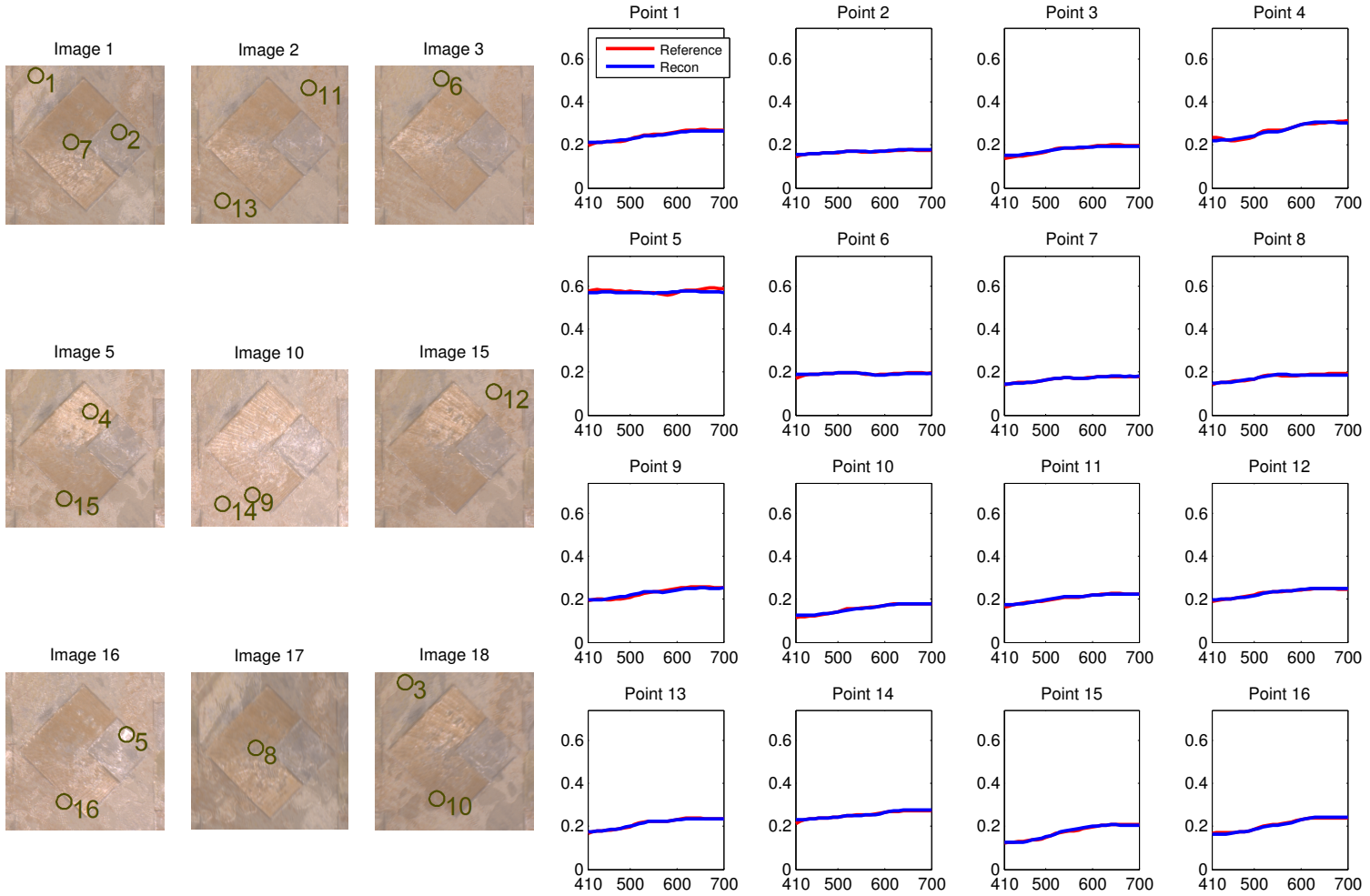


Figure 10.17: Comparison of reference and reconstructed reflectance spectra of selected pixels in the multi-image "Wallpaper" dataset. The shapes of the spectra are met with high precision. The first 10 images are those with spread spectral information, the last 10 images do not have spectral information at all. The results show that the spectral information is spread to those images by the prior term in the energy function. This is even true for a highlight (Point 5 in Image 16).

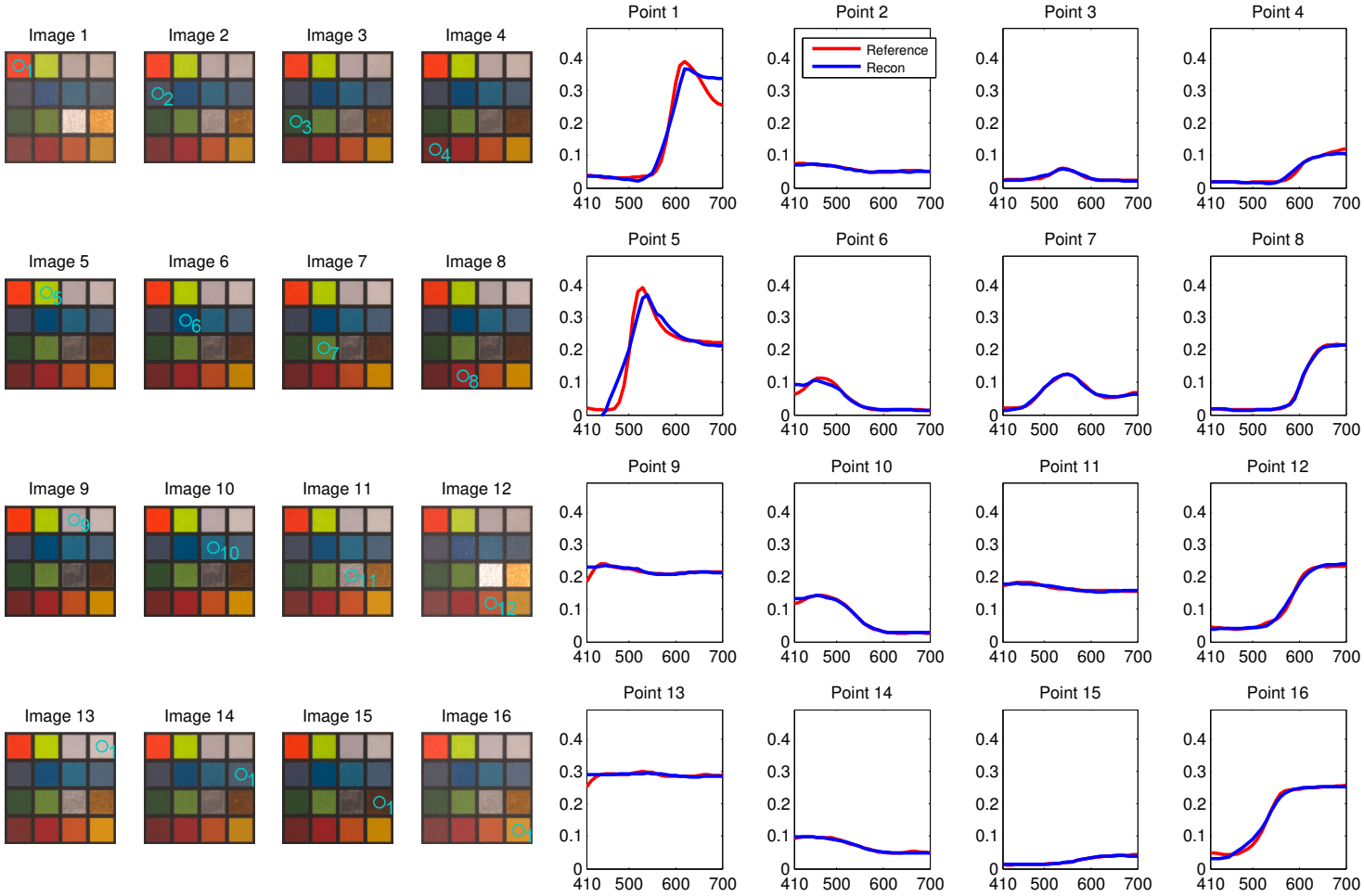


Figure 10.18: Comparison of reference and reconstructed reflectance spectra of selected pixels in the multi-image "Colorchecker" dataset. The shapes of the spectra are met with high precision. Only the very steep flanks in the spectral curves for the red and yellow fields pose a problem due to the spectral smoothness term.

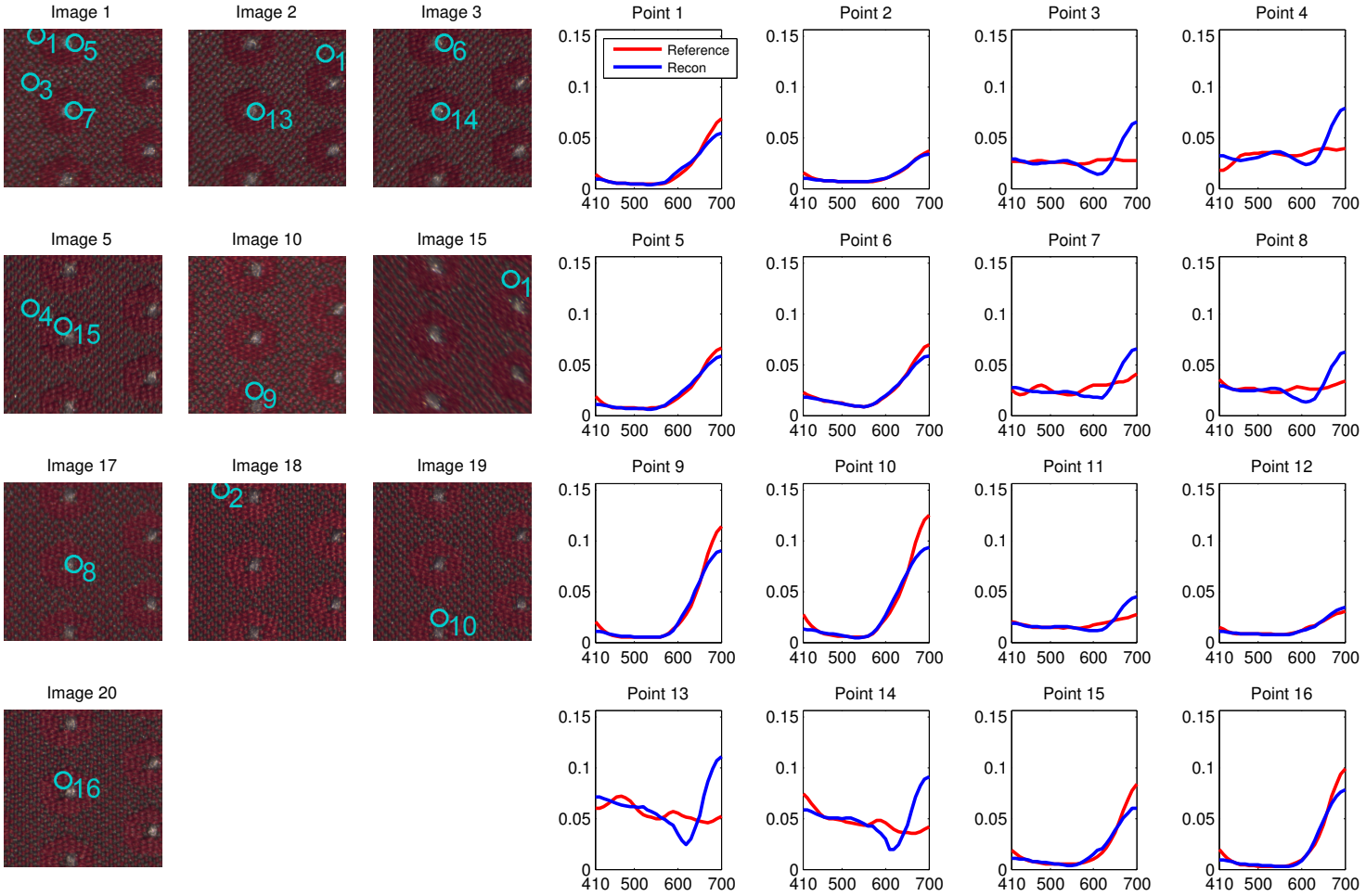


Figure 10.19: Comparison of reference and reconstructed reflectance spectra of selected pixels in the multi-image "Red Fabric" dataset. Some spectra are not reconstructed with sufficient accuracy. This failure case of our method is discussed in Section 10.4.2.1.

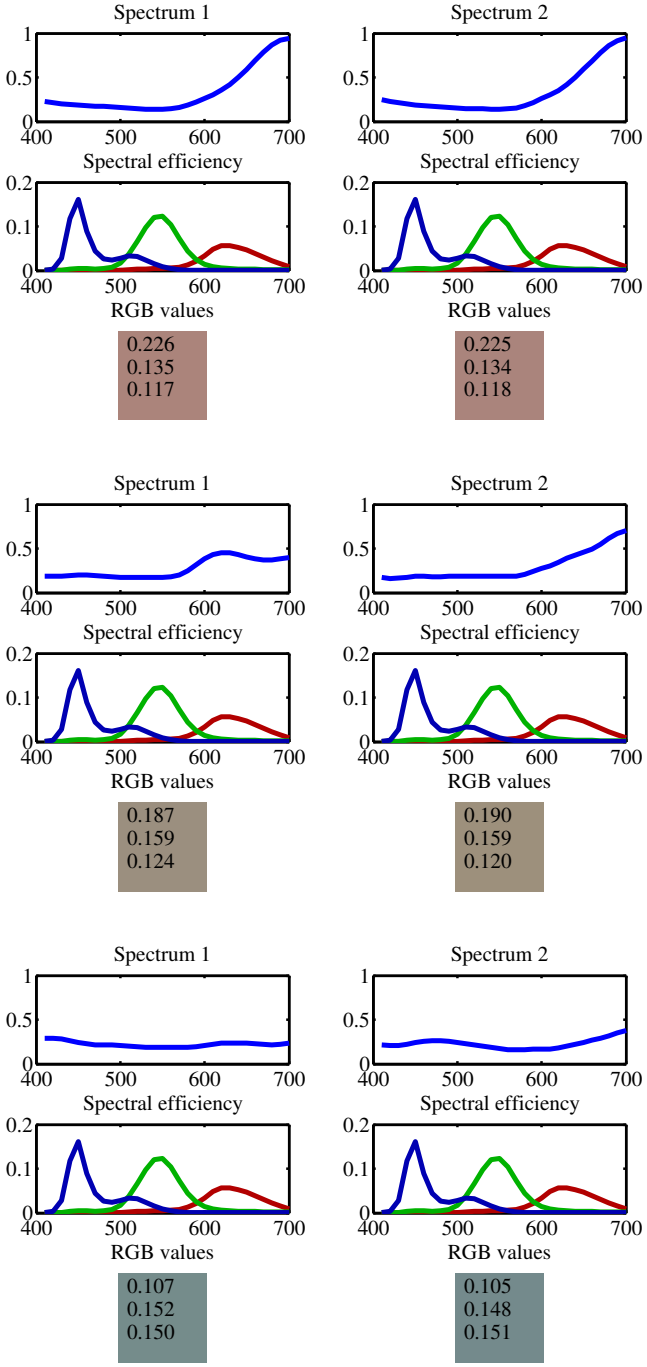
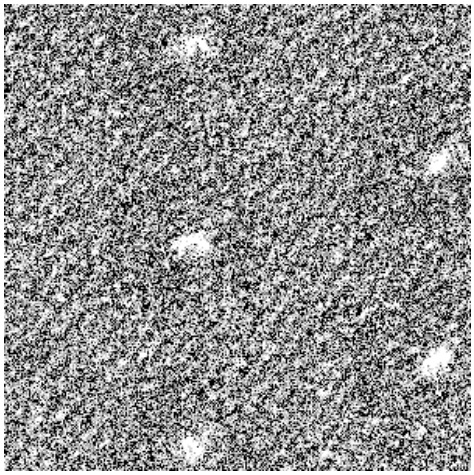
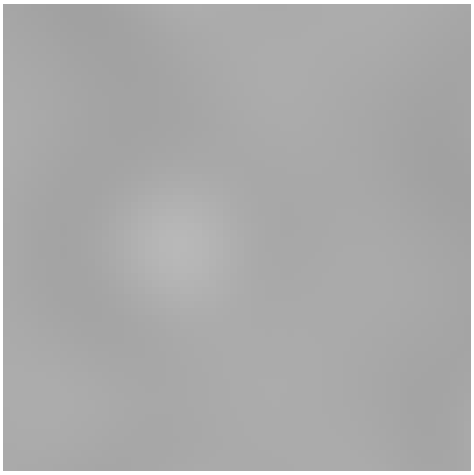


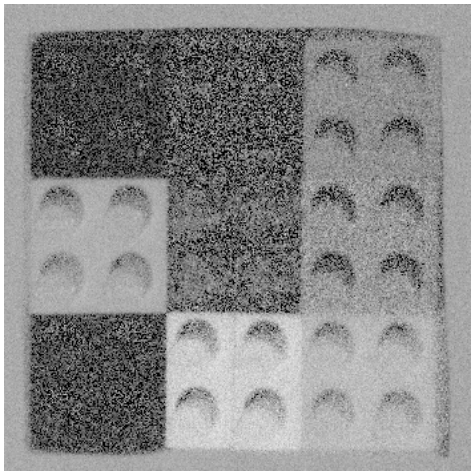
Figure 10.20: RGB neighborhoods in the "Red fabric" dataset: only for the first pair of pixels, the RGB value similarity does also imply spectral similarity. This - system induced - metamerism in combination with the fine spatial structure that is filtered (see Figure 10.21) causes the reconstruction inaccuracy for this dataset. RGB values were scaled for visualization purposes.



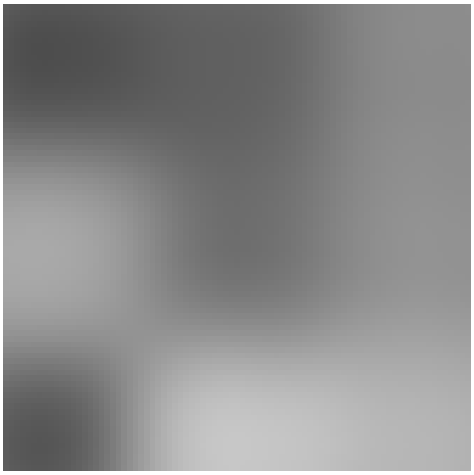
(a) Red fabric: raw



(b) Red fabric: filtered



(c) Lego: raw



(d) Lego: filtered

Figure 10.21: Comparing filtered spectral data: the images show raw and filtered data from the 600nm band for two datasets. The high frequency detail of all threads is lost for the "Red fabric" dataset. For the "Lego" dataset, the different bricks are still distuingishable. Please note that the filtered images do not cover the full image portion since the filter kernels cannot extend past the image borders. See also Figure 10.20 and Section 10.4.2.1 for a discussion.

10.4.3 Comparison to Dong et al. [DSD$^+$16]:

We compared the accuracy of the *spectralization* method to the NMF-based method from [DSD$^+$16]. Since they have a pretty similar idea to combine a full RGB BTF with sparse spectral data, their algorithm is straitforwardly applicable to our problem as well. Their method requires one high quality spectral image acquired using spherical illumination. This is of course not ideally achievable with our device, but a high quality approximation to that can be acquired by illuminating the sample with all LEDs at once. Then the incident irradiance on the sample is that high, that even with short exposure times a nearly noise-free spectral image can be acquired with the spectral camera system. In Figure 10.22 the reconstruction of the Lego dataset using their method is shown. It can be seen that their method is producing rather good results on many of the pixels. However, it fails to reconstruct some of the brick colors with high accuracy. The main reason is that the dimensionality of this dataset if too high to allow for a decent reconstruction using a three-dimensional basis. The error histograms in Figure 10.23 also confirm this finding. Here, we also noticed that the Dong method performs better on the "Red fabric" dataset, despite the fact that it is limited by the dimensionality (see also Section 10.4.2.1).

It should also be noted that using a spherical illumination, the diffuse reflectance of the material is well captured. However, such an image would not contain much information about the specular reflection or the gloss.

It should be noted, though, that their method is computationally magnitudes faster compared to our optimization approach. That means, whenever one is sure that the materials to be captured are exhibiting a low spectral dimensionality, and when an additional method is utilized to reconstruct specular highlights, their method might be preferable over ours. In a general setting, however, their method lacks reconstruction accuracy.

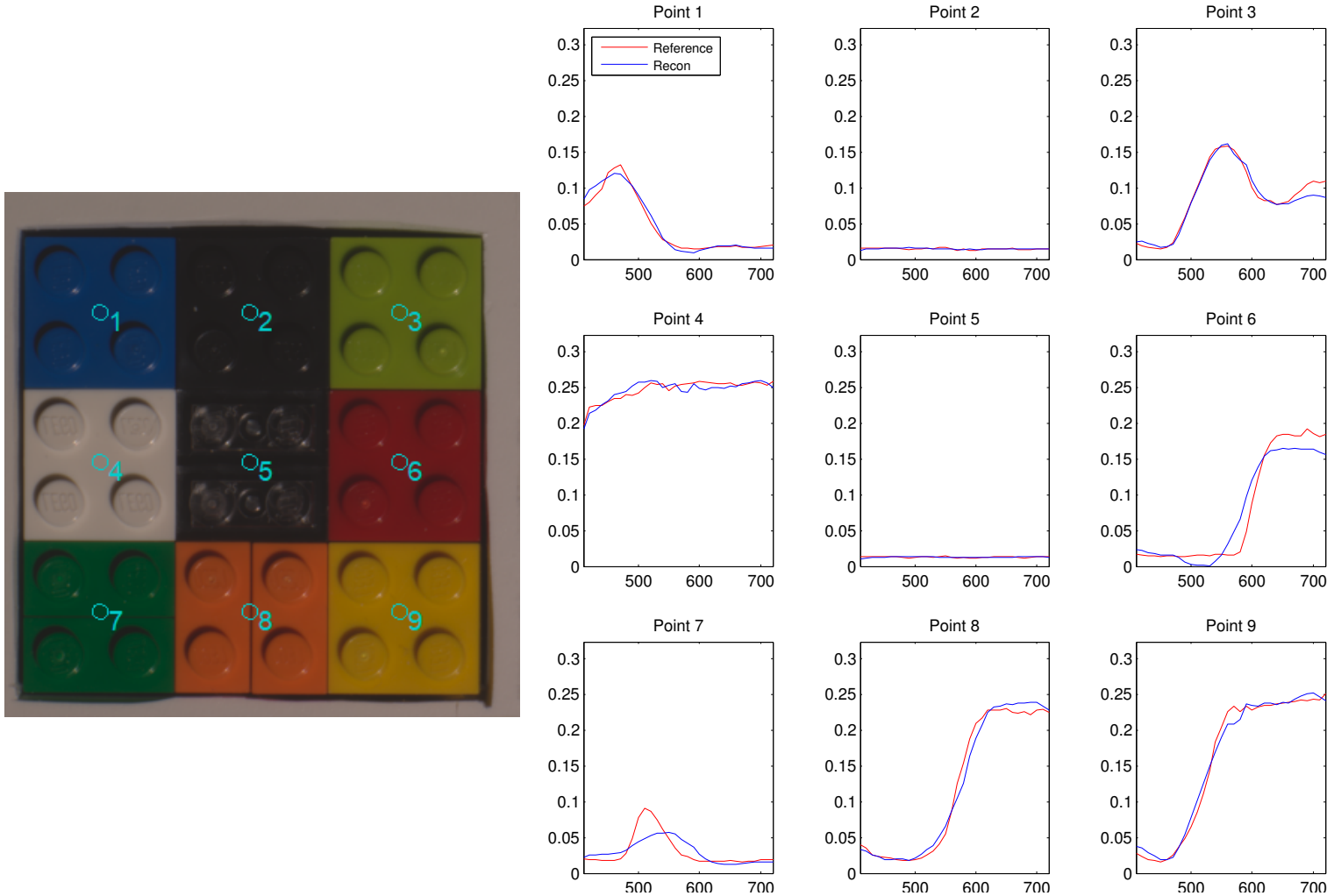
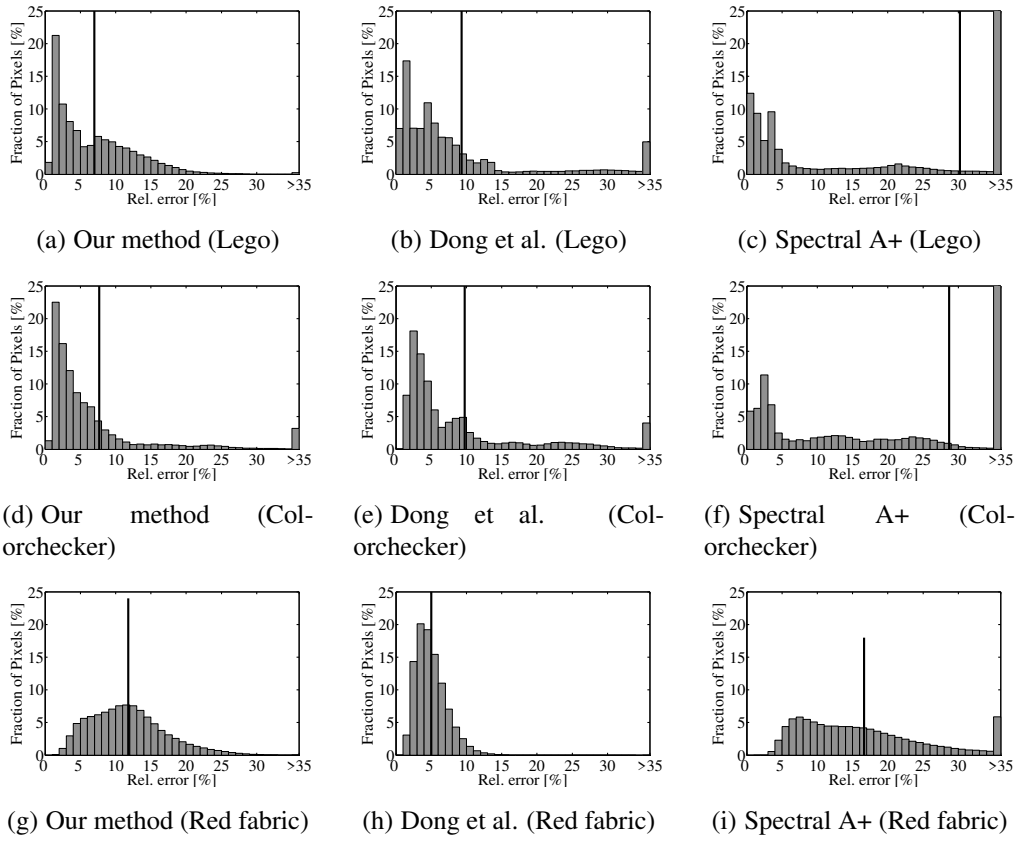


Figure 10.22: Comparison of reference and reconstructed reflectance spectra of selected pixels in the single-image "Lego" dataset using the NMF-based algorithm from Dong et al. [DSD+16]. Due to the limitation to a three-dimensional basis, not all spectra can be reconstructed using high precision (esp. the green field at Point 7.)



(a) Our method (Lego) (b) Dong et al. (Lego) (c) Spectral A+ (Lego)

(d) Our method (Colorchecker) (e) Dong et al. (Colorchecker) (f) Spectral A+ (Colorchecker)

(g) Our method (Red fabric) (h) Dong et al. (Red fabric) (i) Spectral A+ (Red fabric)

Figure 10.23: Comparison to Dong et al. and Spectral A+: Relative error histogram for the "Lego", "Colorchecker" and "Red fabric" datasets. The vertical black lines indicate the mean error. One can see that our method performs always better than Spectral A+ - and we are not sure why this happens. For Dong, our method performs better on "Lego" and "Colorchecker" but worse on the "Red fabric" dataset. The main reason for the better performance on the first two is that the dimensionality of this dataset if too high to allow for a decent reconstruction using a three-dimensional basis. This is also true for the "Red fabric", but here our method fails to deliver high quality results (see Section 10.4.2.1 for a discussion).

10.4.4 Comparison to Arad et al. [ABS16] and Aeschbacher et al. [AWT17]

The method from Arad et al. [ABS16] and Aeschbacher et al. [AWT17] is directly suitable for reconstructing full BTFs as well. To apply Spectral A+ in this setting, we construct the sparse dictionary from the sphere-lit image, acquired as for the Dong method in the previous section. We quickly noticed, that the method delivered very bad reconstruction results with the standard parameters sparsity $k = 8$, regularization $\lambda = 0.1$. However, when setting λ to very tiny values, the results improved a lot. We found that $\lambda = 0.00001$ yielded the best results on our datasets. Error histograms for two examples can be seen in Figure 10.23. We can see that - even with optimized parameters - the reconstruction accuracy is still very limited compared to the Dong method or to our method. This confirms our findings from Sections 7.1 and 7.2. There, we supposed that the 250 spectral datapoints might be insufficient to create a descent, overcomplete basis. However, we now input the whole sphere-lit, spectral image as training data. That means, that the pure number of training samples cannot be responsible for the bad performance. We did not perform any further analysis.

10.5 Real Data

To test our algorithm on real data we captured reference data for a X-Rite ColorChecker Passport Photo[1] with our spectral camera system using a total exposure time of 226212ms. Afterwards we captured RGB images and noisy spectral images with a total exposure time of 7463ms and performed a reconstruction using our novel algorithm. Figure 10.24 shows a comparison between reference and reconstructed spectra for one pixel per color field. Additionally, we used the noisy spectral images directly and added the respective pixels to the plot for comparison purposes. The figure shows, that our algorithm achieves nearly the same results as the reference capture while requiring exposure times that are about 1.5 orders of magnitude smaller even serial capture of RGB and spectral images is assumed. When both are captured in parallel, the ratio is even better.

This first test with the real setup shows that our modification of the spectralization method introduced in Section 10.1 is well suited for practical applications and delivers precisely reconstructed spectral images.

[1] https://www.xrite.com/categories/calibration-profiling/colorchecker-classic-family/colorchecker-passport-photo-2

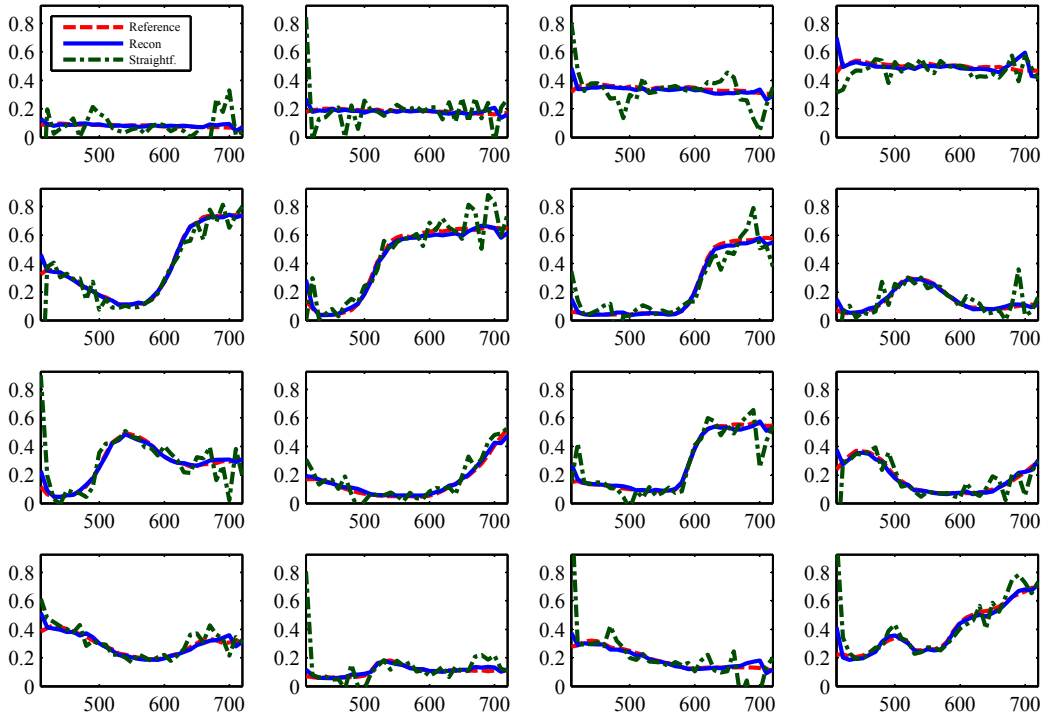


Figure 10.24: Spectral reconstruction using real data: pixel spectra from different color fields of a X-Rite ColorChecker Passport are shown. The reference data (dashed red) was obtained using long exposure times and required a total exposure time of 226212ms. The reconstruction by our algorithm (blue) was based on camera images with a total exposure time of 7463ms. The green spectra show the result obtained by using the noisy spectral input data directly. The result indicates the robustness of our reconstruction against high noise levels. Spectral imaging is accelerated by about 1.5 orders of magnitude.

10.6 Summary

In this chapter we presented one possible, practical way to utilize the *spectralization* technique to combine RGB and spectral technology to capture image series with self-similar content. In the area of computer graphics this is relevant for capturing bidirectional reflectance data like BTFs or SVBRDFs, for instance. Due to the flexibility of the *spectralization* we can integrate a lot of different kinds of spectral data into the energy optimization. In this case we used noisy spectral images, which allow for low exposure times for the spectral imaging system. The noisy data is integrated by adding a filtering into the energy function. The RGB data is the key to restore the high resolution spectral images from the incomplete

input data. Using this set-up we can also solve the problem of geometric registration of RGB data with respect to spectral data.

For this, we have first introduced a new variant of the energy function in Equation 10.1 that supersedes the original function from Equation 6.9.. We have also shown how to calibrate spectral camera system to make use of low-signal, high noise data. Using a detailed simulation of our RGB-spectral measurement setup we could analyze the reconstruction performance of our approach in detail in Sections 10.4.1 and 10.4.2 and propose an improved version of the energy function in Equation 10.12. Beyond simulated data we have also shown that the proposed algorithms work for real-world data, as well.

We have also documented a partial failure case of our method in Section 10.4.2.1 and given outlook to possible improvements.

By comparing our results to the methods from Dong et al. [DSD+16] and to Spectral A+ [AWT17], which are both applicable in the same setting, we have set our method into the context of prior work. We have shown that our method performs always better than Spectral A+ and also better than Dong et al. except for the documented failure case.

Part IV

Closure

CHAPTER 11

CONCLUSIONS

11.1 Summary

Material reflectance data with dense sampling of the spectral domain is necessary in all applications where images should be created from virtual scenes by the means of rendering methods and where exact color reproduction is a required key feature.

In this thesis we presented a novel solution for reconstructing spectral images or spectral BRDFs from sparse spectral and dense RGB data, enabling to capture complex reflectance functions like Bidirectional Texture Functions (BTF) or SV-BRDFs with high spectral resolution.

Traditionally, mostly trichromatic - e.g. RGB - solutions are available for this due to the low cost and ease of use of RGB cameras. Straightforwardly extending existing setups to allow for reflectance acquisition with higher spectral resolution is associated with large efforts in terms of hardware costs, acquisition time, and storage cost. Our method helps to reduce these costs and the acquisition time significantly compared to straightforward approaches.

Our work was divided in three steps:

1. Extend an existing RGB gonioreflectometer with dedicated hardware to enable for the capture of spectral BTF ground-truth data.
2. Develop a novel method for the reconstruction of dense spectral data from dense trichromatic and sparse spectral data, which allows for re-use of existing RGB capture setups with minimal additional hardware.
3. Propose methods to apply the novel acquisition scheme in practice.

In the first step we extended an existing RGB gonioreflectometer [SSK03] with a high-quality CCD camera having a liquid crystal tunable filter in the optical path. This system allowed for capture of spectral images with average bandwidth of 10nm in the range of 400-720nm. Using this modified system, we captured a

small database of spectral BTFs and textures to serve as a ground-truth data for our latter experiments.

In the second step we developed a novel method for reconstructing dense spectral data based on optimization. This method supports a wide variety of input devices since all possible integrals over the spectral data-cube can be incorporated. One additional contribution inside of this method is a novel prior on spectral images which predicts similarity of spectra based on similarity in an appearance space based on the RGB data.

The last part of the thesis deals with practical implementation of measurements using the proposed method from the second step. Here, a robust method for spectral calibration of RGB cameras is introduced, which allows for calibration under even largely uncontrolled conditions and for increased accuracy under laboratory conditions. Furthermore, a possible extension of an existing RGB BTF capture setup with dedicated spectral hardware is discussed to allow for efficient capture of the additional spectral data. First results for spectral reconstruction are presented.

11.2 Discussion and Future Work

Compared to other spectral image reconstruction methods, our *spectralization* method delivers results of superior quality, especially in cases where the image content is complex. All methods, that do not incorporate spectral knowledge of the image at all, produce unacceptably high reconstruction errors. Those using this additional knowledge clearly improve. However, the restriction of the methods of Hardeberg et al. [HSB+99b] and Dong et al. [DSD+16] to a three dimensional subspace of spectra and the restriction of the Murakami method [MIYO07] to images having normal distribution of pixel values leads to less precise reconstruction compared to our method using the appearance-space smoothness prior.

Furthermore, our method allows for flexible integration of nearly every source of data. All photometric capture devices provide some integral over the spectral data-cube and the respective integral can be directly modeled in our linear-least-squares energy function. Especially the Murakami method - while being the most accurate among the previous methods on images - is very limited with respect to the kind of input data it accepts and it also works on images only. In contrast to this, our method can be applied to even fully unstructured data-sets and also to various kinds of spectral input data.

The first results obtained using the real setup described in Chapter 10 indicate that the algorithm is also stable when using real data.

The main drawback of *spectralization* is the high computational demand, resulting in long reconstruction times. While the Hardeberg method needs less than

a second for any image and the Murakami and Dong methods (if implemented in an optimized fashion) only a few seconds, our method will easily need 5-10 minutes depending on the image size, the amount of additional spectral data and the neighborhood sizes in the prior term. For practical applications, especially for reconstructing whole Bidirectional Texture Functions or alike, this is too slow, of course.

It is notable that the quite simple Hardeberg method and also the Dong method already provide good results as soon as the space of spectra in the image has not significantly more than three dimensions. One possible way to accelerate spectralization would therefore be to employ a hierarchical reconstruction, where the optimization would be responsible for dealing with the global image content without any restrictions and the Hardeberg or Dong method would be used on a local level to reconstruct the spectra with low computational cost. A hierarchy might be constructed by clustering in one of the appearance spaces introduced in Chapter 6.2.

In Section 10.4.2.1 we have also proposed two possible improvements to our method while we documented a failure case: First, one could use more efficient textural descriptors for the appearance space projection to help deal with metamers in fine structures. Various image feature descriptors exist that could help improve the results in the case documented there but also in many other cases. Second, it could be advantageous to add measurement data corresponding to more than two kinds of integrals over the spectral data-cube. In this thesis we have always just used RGB images and combined them with one kind of spectral data. However, the linear-least-squares optimization would straightforwardly allow to incorporate different kinds of spectral data at the same time.

Furthermore, it would be very interesting to determine good, material class dependent angular sampling strategies for efficiently acquiring the spectral data necessary for a high-quality reconstruction of reflectance function data. This is of high practical relevance since cost efficient reflectance capture setups - tailored to specific material classes - can be designed this way. For some material classes, such knowledge is already present e.g. for effect paints there exists an ASTM standard [AST12]. However, for most material classes, especially for compound materials consisting of more than one different basic materials, this will require a thorough analysis.

An additional direction of future research would be to use the algorithm described in Chapter 10 for spectral video capture. In the context of computer graphics, spectral video could e.g. be used for capturing space-dependent spectral environment maps by imaging a reflecting sphere and moving around the apparatus in space. The camera system could be designed by using a common optical system and a beam splitter to project the image onto a RGB sensor and a monochrome sensor with a changeable filter in front of the monochrome sensor. For every video

frame, the filter is changed to the next spectral band, resulting in one RGB image and a one-band spectral image per video frame. Since adjacent video frames are typically very similar, the algorithm might be used to reconstruct a full spectral video from this input data.

Finally, an extension of the spectralization method to fluorescent materials could be interesting. The high practical relevance of fluorescence - even for everyday materials - was accounted for by Hullin et al. [HHA+10] who presented a method to capture the bi-spectral BRDF of homogeneous, fluorescent materials. By incorporating RGB images acquired using different light sources (e.g. by adding two different UV LEDs) into the spectralization method, the amount of data to be captured could be diminished even further compared to their method. This could render the acquisition of spatially varying reflectance of - perhaps partially - fluorescent materials practical.

BIBLIOGRAPHY

[ABS16] Boaz Arad and Ohad Ben-Shahar. Sparse recovery of hyperspectral signal from natural RGB images. In *European Conference on Computer Vision*, pages 19–34. Springer, 2016.

[ABST18] Boaz Arad, Ohad Ben-Shahar, and Radu Timofte. NTIRE 2018 challenge on spectral reconstruction from RGB images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 929–938, 2018.

[AEB+06] Michal Aharon, Michael Elad, Alfred Bruckstein, et al. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing*, 54(11):4311, 2006.

[AFOR04] G. Antonioli, F. Fermi, C. Oleari, and R. Reverberi. Spectrophotometric scanner for imaging of paintings and other works of art. In *Proceedings of CGIV 2004*, pages 219–224, 2004.

[AP08] Xiaobo An and Fabio Pellacini. AppProp: all-pairs appearance-space edit propagation. *ACM Transactions on Graphics (TOG)*, 27(3):1–9, 2008.

[AST12] ASTM. Standard E2194, 2012, "Practice for Multiangle Color Measurement of Metal Flake Pigmented Materials", DOI: 10.1520/e2194-12, http://www.astm.org , 2012.

[AWT17] Jonas Aeschbacher, Jiqing Wu, and Radu Timofte. In defense of shallow learned spectral reconstruction from RGB images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 471–479, 2017.

[BEWW+08] Moshe Ben-Ezra, Jiaping Wang, Bennett Wilburn, Xiaoyang Li, and Le Ma. An LED-only BRDF measurement device. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008*, pages 1–8. IEEE, 2008.

[Bli77] James F. Blinn. Models of light reflection for computer synthesized pictures. In *Proceedings of the 4th annual conference on Computer graphics and interactive techniques*, pages 192–198, 1977.

[BN76] James F. Blinn and Martin E. Newell. Texture and reflection in computer generated images. *Communications of the ACM*, 19(10):542–547, 1976.

[BS70] Willard S. Boyle and George E. Smith. Charge coupled semiconductor devices. *The Bell System Technical Journal*, 49(4):587–593, 1970.

[BTF] BTF Database Bonn. http://cg.cs.uni-bonn.de/en/projekte/btfdbb/.

[CDT+11] Xun Cao, Hao Du, Xin Tong, Qionghai Dai, and Stephen Lin. A prism-mask system for multispectral video acquisition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2423–2435, 2011.

[CLH+05] Vien Cheung, Changjun Li, Jon Y. Hardeberg, David Connah, and Stephen Westland. Characterization of trichromatic color cameras by using a new multispectral imaging technique. *Journal of the Optical Society of America A*, 22(7):1231–1240, Jul 2005.

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

[CTNM06] Kunlaya Cherdhirunkorn, Norimichi Tsumura, Toshiya Nakaguchi, and Yoichi Miyake. Spectral based color correction technique compatible with standard RGB system. *Optical review*, 13(3):138–145, 2006.

[CV01] Tony F. Chan and Luminita A. Vese. Active contours without edges. *IEEE transactions on Image processing*, 10(2):266–277, 2001.

[CZ11] Ayan Chakrabarti and Todd Zickler. Statistics of real-world hyperspectral images. In *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 193–200. IEEE Computer Society, 2011.

[DD95] Michael Descour and Eustace Dereniak. Computed-tomography imaging spectrometer: experimental calibration and reconstruction results. *Applied Optics*, 34(22):4817–4826, Aug 1995.

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

[DSD+16] Wei Dong, Hui-Liang Shen, Xin Du, Si-Jie Shao, and John H Xin. Spectral bidirectional texture function reconstruction by fusing multiple-color and spectral images. *Applied optics*, 55(36):10400–10408, 2016.

[DTCL09] Hao Du, Xin Tong, Xun Cao, and S. Lin. A prism-based system for multispectral video acquisition. In *IEEE 12th International Conference on Computer Vision, 2009*, pages 175–182, Sept 2009.

[DvGNK97] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real-world surfaces. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 151–157, 1997.

[DvGNK99] Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[DWT+10] Yue Dong, Jiaping Wang, Xin Tong, John Snyder, Yanxiang Lan, Moshe Ben-Ezra, and Baining Guo. Manifold bootstrapping for SVBRDF capture. In *ACM SIGGRAPH 2010 Papers*, pages 98:1–98:10, New York, NY, USA, 2010. ACM.

[Ebn07] Marc Ebner. Estimating the spectral sensitivity of a digital sensor using calibration targets. In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, GECCO '07, pages 642–649, New York, NY, USA, 2007. ACM.

[EMV10] EMVA Standard 1288. http://www.emva.org/cms/index.php?idcat=26, 2010.

[FANA04] David H. Foster, Kinjiro Amano, Sérgio M. C. Nascimento, and Kinjiro Amano. Information limits on neural identification of colored surfaces in natural scenes. *Visual Neuroscience*, 21(3):331–336, May 2004.

[FANF06] David H. Foster, Kinjiro Amano, Sérgio M. C. Nascimento, and Michael J. Foster. Frequency of metamerism in natural scenes.

Journal of the Optical Society of America A, 23(10):2359–2372, Oct 2006.

[FH09] J. Filip and M. Haindl. Bidirectional texture function modeling: A state of the art survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):1921–1940, November 2009.

[FNW07] Mário A. T. Figueiredo, Robert D. Nowak, and Stephen J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):586–597, 2007.

[Foo97] Sing Choong Foo. A gonioreflectometer for measuring the bidirectional reflectance of materials for use in illumination computations. Master's thesis, Cornell University, 1997.

[FOP08] Joyce Farrell, Michael Okincha, and Manu Parmar. Sensor calibration and simulation. *Proc. SPIE 6817, Digital Photography IV*, 6817:68170R, 2008.

[GAHO07] Abhijeet Ghosh, Shruthi Achutha, Wolfgang Heidrich, and Matthew O'Toole. BRDF acquisition with basis illumination. In *IEEE 11th International Conference on Computer Vision, 2007*, pages 1–8, Oct. 2007.

[Gat00] Nahum Gat. Imaging spectroscopy using tunable filters: a review. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4056, pages 50–64, April 2000.

[Gil11] Graeme Gill. Argyll CMS, http://www.argyllcms.com/, 2011.

[GJB+07] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*, 15(21):14013–14027, Oct 2007.

[Gra53] Hermann Günther Graßmann. Zur Theorie der Farbmischung. *Poggendorfs Annalen der Physik und Chemie*, 89:59–84, 1853.

[GTHD03] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. *ACM Transactions on Graphics (TOG)*, 22(3):749–758, 2003.

[HBG+00] Andrew R. Harvey, John E. Beale, Alain H. Greenaway, Tracy J. Hanlon, and John W. Williams. Technology options for imaging spectrometry. In *Society of Photo-Optical Instrumentation*

Engineers (SPIE) Conference Series, volume 4132, pages 13–24, November 2000.

[HBS92] Peter J.B. Hancock, Roland J. Baddeley, and Leslie S. Smith. The principal components of natural images. *Network: computation in neural systems*, 3(1):61–70, 1992.

[Hei06] Gunther Heidemann. The principal components of natural images revisited. *IEEE transactions on pattern analysis and machine intelligence*, 28(5):822–826, 2006.

[Her87] Heinrich Hertz. Ueber den Einfluss des ultravioletten Lichtes auf die electrische Entladung. *Annalen der Physik*, 267(8):983–1000, 1887.

[HHA+10] Matthias B. Hullin, Johannes Hanika, Boris Ajdin, Hans-Peter Seidel, Jan Kautz, and Hendrik P. A. Lensch. Acquisition and analysis of bispectral bidirectional reflectance and reradiation distribution functions. *ACM Transactions on Graphics (TOG)*, 29(4):1–7, 2010.

[HKW12] Ralf Habel, Michael Kudenov, and Michael Wimmer. Practical spectral photography. *Computer Graphics Forum*, 31(2):449–458, 2012.

[HL07] Gerald C. Holst and Terrence S. Lomheim. *CMOS/CCD Sensors and Camera Systems*, volume 408 of *SPIE PM*. JCD Publishing, 2007.

[HP03] Jefferson Y. Han and Ken Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Transactions on Graphics (TOG)*, 22(3):741–748, 2003.

[HP11] Robert William Gainer Hunt and Michael R. Pointer. *Measuring colour, Fourth edition*. John Wiley & Sons, 2011.

[HS52] Magnus R. Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6), 1952.

[HSB99a] Jon Y. Hardeberg, Francis Schmitt, and Hans Brettel. Multispectral image capture using a tunable filter. In *Electronic Imaging*, pages 77–88. International Society for Optics and Photonics, 1999.

[HSB+99b] Jon Y. Hardeberg, Francis Schmitt, Hans Brettel, Jean-Pierre Crettez, and Henri Maitre. Multispectral image acquisition and simulation of illuminant changes. In *Color Imaging: Vision and Technology*, pages 145–164, 1999.

[HSF94] Paul M. Hubel, Doron Sherman, and Joyce E. Farrell. A comparison of methods of sensor spectral sensitivity estimation. In *Second Color Imaging Conference: Color Science, Systems, and Applications*, pages 45–48. Society for Imaging Science and Technology, Nov 1994.

[IB98] Francisco H. Imai and Roy S. Berns. High-resolution multi-spectral image archives - a hybrid approach. In *Proceedings of IS&T and SID's 6th Color Imaging Conference: Color Science, Systems and Applications*, pages 224–227. Society for Imaging Science and Technology, 1998.

[IB99] Francisco H. Imai and Roy S. Berns. Spectral estimation using trichromatic digital cameras. In *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pages 42–49. Wiley, 1999.

[ICOL05] Revital Irony, Daniel Cohen-Or, and Dani Lischinski. Colorization by example. In *Proceedings of the Sixteenth Eurographics Conference on Rendering Techniques*, EGSR'05, pages 201–210. Eurographics Association, 2005.

[Ima98] Francisco H. Imai. Multi-spectral image acquisition and spectral reconstruction using a trichromatic digital camera system associated with absorption filters. *Munsell Color Science Laboratory Technical Report*, 1998.

[JZG+17] Yan Jia, Yinqiang Zheng, Lin Gu, Art Subpa-Asa, Antony Lam, Yoichi Sato, and Imari Sato. From RGB to spectrum for natural scenes via manifold-based mapping. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4705–4713, 2017.

[Kaj86] James T. Kajiya. The rendering equation. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '86, pages 143–150, New York, NY, USA, 1986. ACM.

[KFY+10] Shota Kanamori, Kazuya Fujiwara, Takahiro Yoshinobu, Bisser Raytchev, Toru Tamaki, and Kazufumi Kaneda. Physically based rendering of rainbows under various atmospheric conditions. In *Proceedings of the 2010 18th Pacific Conference on Computer Graphics and Applications*, pages 39–45, Washington, DC, USA, 2010. IEEE Computer Society.

[KKP+08] Duck Bong Kim, Kang Yeon Kim, Kang Su Park, Myoung Kook Seo, and Kwan H. Lee. A fast and accurate image-based measuring system for isotropic reflection materials. In *Reflection, Scattering, and Diffraction from Surfaces*, volume 7065, page 70650I. International Society for Optics and Photonics, 2008.

[KL97] Nandakishore Kambhatla and Todd K. Leen. Dimension reduction by local principal component analysis. *Neural Computation*, 9(7):1493–1516, October 1997.

[KMBK03] Mellisa L. Koudelka, Sebastian Magda, Peter N. Belhumeur, and David J. Kriegman. Acquisition, compression and synthesis of bidirectional texture functions. In *3rd International Workshop on Texture Analysis and Synthesis (Texture 2003)*, pages 59–64, October 2003.

[KN07] Kiyosumi Kidono and Yoshiki Ninomiya. Visibility estimation under night-time conditions using a multiband camera. In *IEEE Intelligent Vehicles Symposium*, pages 1013–1018, june 2007.

[Koe10] Jan J. Koenderink. The prior statistics of object colors. *Journal of the Optical Society of America A*, 27(2):206–217, Feb 2010.

[KP08] Hyunsoo Kim and Haesun Park. Nonnegative matrix factorization based on alternating nonnegativity constrained least squares and active set method. *SIAM journal on matrix analysis and applications*, 30(2):713–730, 2008.

[KSKL10] Duck Bong Kim, Myoung Kook Seo, Kang Yeon Kim, and Kwan H. Lee. Acquisition and representation of pearlescent paints using an image-based goniospectrophotometer. *Optical Engineering*, 49(4):043604, 2010.

[LFTW06] Hongsong Li, Sing-Choong Foo, Kenneth E. Torrance, and Stephen H. Westin. Automated three-axis gonioreflectometer for computer graphics applications. *Optical Engineering*, 45(4):043605, April 2006.

[LGZ08] Xin Li, Bahadir Gunturk, and Lei Zhang. Image demosaicing: a systematic survey. In *Visual Communications and Image Processing 2008*, volume 6822, page 68221J. International Society for Optics and Photonics, 2008.

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

[LLW04] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH '04, pages 689–694, New York, NY, USA, 2004. ACM.

[Lut27] Robert Luther. Aus dem Gebiete der Farbreizmetrik. *Zeitschrift für technische Physik*, 8:540–558, 1927.

[Max63] James Clerk Maxwell. A dynamical theory of the electromagnetic field. *Proceedings of the Royal Society of London*, 13:531–536, 1863.

[MBK05] Gero Müller, Gerhard H. Bendels, and Reinhard Klein. Rapid synchronous acquisition of geometry and BTF for cultural heritage artefacts. In *The 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST)*, pages 13–20. Eurographics Association, November 2005.

[McA02] David K. McAllister. *A generalized surface appearance representation for computer graphics*. PhD thesis, The University of North Carolina at Chapel Hill, 2002. AAI3061704.

[MIYO07] Yuri Murakami, Kunihiko Ietomi, Masahiro Yamaguchi, and Nagaaki Ohyama. Maximum a posteriori estimation of spectral reflectance from color image and multipoint spectral measurements. *Applied Optics*, 46(28):7068–7082, Oct 2007.

[ML09] Marius Muja and David G. Lowe. Fast approximate nearest neighbors with automatic algorithm configuration. In *International Conference on Computer Vision Theory and Applications (VISAPP)*, pages 331–340, 2009.

[MMS+05] G. Müller, J. Meseth, M. Sattler, R. Sarlette, and R. Klein. Acquisition, synthesis and rendering of bidirectional texture functions. *Computer Graphics Forum*, 24(1):83–109, March 2005.

[MPBM03a] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (TOG)*, 22(3):759–769, 2003.

[MPBM03b] Wojciech Matusik, Hanspeter Pfister, Matthew Brand, and Leonard McMillan. Efficient isotropic BRDF measurement. In *Proceedings of the 14th Eurographics workshop on Rendering*, volume 44 of *ACM International Conference Proceeding Series*, pages 241–247, Aire-la-Ville, Switzerland, 2003. Eurographics Association.

[MPN+02] Wojciech Matusik, Hanspeter Pfister, Addy Ngan, Paul Beardsley, Remo Ziegler, and Leonard McMillan. Image-based 3d photography using opacity hulls. In *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '02, pages 427–437, New York, NY, USA, 2002. ACM.

[MSI94] Yoshitsugu Manabe, K Sato, and Seiji Inokuchi. An object recognition through continuous spectral images. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, volume 1, pages 858–860. IEEE, October 1994.

[MSY07] Yasuhiro Mukaigawa, Kohei Sumino, and Yasushi Yagi. Multiplexed illumination for measuring brdf using an ellipsoidal mirror and a projector. In *Proceedings of the 8th Asian conference on Computer Vision - Volume Part II*, pages 246–257, 2007.

[Mül09] Gero Müller. *Data-driven methods for compression and editing of spatially varying appearance*. PhD thesis, Universität Bonn, 2009.

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

[MWRK17] Sebastian Merzbach, Michael Weinmann, Martin Rump, and Reinhard Klein. Fast capture of spectral image series. In *12th International Conference on Computer Graphics Theory and Applications (VISIGRAPP)*, 2017.

[Nas01] Kurt Nassau. *The physics and chemistry of color: the fifteen causes of color*. Wiley series in pure and applied optics. Wiley, 2001.

[NEBS04] Heinz Niedrig, Hans-Joachim Eichler, Ludwig Bergmann, and Clemens Schaefer. *Lehrbuch der Experimentalphysik; Bd. 3: Optik*. Walter de Gruyter, Berlin, New York, 2004.

[New72] Isaac Newton. A new theory about light and colors. *Philosophical Transactions of the Royal Society of London*, 80:3075–3087, Feb 1672.

[NFF02] Sérgio M. C. Nascimento, Flávio P. Ferreira, and David H. Foster. Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A*, 19(8):1484–1490, Aug 2002.

[NGHJ18] Jan Novák, Iliyan Georgiev, Johannes Hanika, and Wojciech Jarosz. Monte Carlo methods for volumetric light transport simulation. *Computer Graphics Forum*, 37(2):551–576, 2018.

[NHE+00] Stephan Nickell, Marcus Hermann, Matthias Essenpreis, Thomas J. Farrell, Uwe Krämer, and Michael S. Patterson. Anisotropy of light propagation in human skin. *Physics in Medicine and Biology*, 45(10):2873–2886, October 2000.

[NPB14] Rang M. H. Nguyen, Dilip K. Prasad, and Michael S. Brown. Training-based spectral reconstruction from a single RGB image. In *European Conference on Computer Vision*, pages 186–201. Springer, 2014.

[NRH+77] Fred E. Nicodemus, Joseph C. Richmond, Jack J. Hsia, Irving W. Ginsberg, and Thomas Limperis. *Geometrical Considerations and Nomenclature for Reflectance*. US Dept. of Commerce, National Bureau of Standards, 1977.

[OKG+10] Se Baek Oh, Sriram Kashyap, Rohit Garg, Sharat Chandran, and Ramesh Raskar. Rendering wave effects with augmented light field. *Computer Graphics Forum*, 29:507–516, May 2010.

[ON94] Michael Oren and Shree K. Nayar. Generalization of lambert's reflectance model. In *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 239–246, New York, NY, USA, 1994. ACM.

[PB96] James E. Proctor and P. Yvonne Barnes. NIST high accuracy reference reflectometer-spectrophotometer. *Journal of Research of the National Institute of Standards and Technology*, 101(5):619–627, 1996.

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

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

[PL07] Fabio Pellacini and Jason Lawrence. AppWand: editing measured materials using appearance-driven optimization. *ACM Transactions on Graphics (TOG)*, 26(3):54, 2007.

[Pla01] Max Planck. Ueber das Gesetz der Energieverteilung im Normalspectrum. *Annalen der Physik*, 309(3):553–563, 1901.

[PRK93] Yagyensh C. Pati, Ramin Rezaiifar, and Perinkulam S. Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th Asilomar conference on signals, systems and computers*, pages 40–44. IEEE, 1993.

[PS82] Christopher C. Paige and Michael A. Saunders. LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM Transactions on Mathematical Software*, 8(1):43–71, March 1982.

[RBdB+99] Klaus Reif, Klaus Bagschik, Klaas S. de Boer, Juergen Schmoll, Philipp Mueller, Henning Poschmann, Guenther Klink, Ralf Kohley, Uli Heber, and Ulrich Mebold. BUSCA: a telescope instrumentation for simultaneous imaging in four optical bands. In *Sensors, Cameras, and Systems for Scientific/Industrial Applications*, volume 3649, pages 109–120. International Society for Optics and Photonics, 1999.

[RBS03] Mark A. Robertson, Sean Borman, and Robert L. Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12(2):219–228, April 2003.

[RK10] Martin Rump and Reinhard Klein. Spectralization: Reconstructing spectra from sparse data. In *SR '10 Rendering Techniques*, pages 1347–1354. Eurographics Association, June 2010.

[RSK10] Martin Rump, Ralf Sarlette, and Reinhard Klein. Groundtruth data for multispectral bidirectional texture functions. In *CGIV 2010*,

pages 326–330. Society for Imaging Science and Technology, June 2010.

[Rus98] Szymon M. Rusinkiewicz. A new change of variables for efficient BRDF representation. In *Rendering Techniques '98*, pages 11–22, New York, NY, 1998. Springer Wien.

[RWS+11] Peiran Ren, Jiaping Wang, John Snyder, Xin Tong, and Baining Guo. Pocket reflectometry. *ACM Transactions on Graphics (TOG)*, 30(4):1–10, 2011.

[RZK11] Martin Rump, Arno Zinke, and Reinhard Klein. Practical spectral characterization of trichromatic cameras. *ACM Transactions on Graphics (Proceedings SIGGRAPH Asia 2011)*, 30(6):170, December 2011.

[SALK05] Martin Solli, Mattias Andersson, Reiner Lenz, and Björn Kruse. Color measurements with a consumer digital camera using spectral estimation techniques. In *Image Analysis: 14th Scandinavian Conference, SCIA 2005*, pages 105–114, 2005.

[SBBC98] G. Serrot, M. Bodilis, X. Briottet, and H. Cosnefroy. Presentation of a new BRDF measurement device. In *Atmospheric Propagation, Adaptive Systems, and Lidar Techniques for Remote Sensing II*, volume 3494, pages 34–40. International Society for Optics and Photonics, 1998.

[Sch07] János Schanda. *Colorimetry: Understanding the CIE system*. John Wiley & Sons, 2007.

[SCX+18] Zhan Shi, Chang Chen, Zhiwei Xiong, Dong Liu, and Feng Wu. HSCNN+: Advanced CNN-based hyperspectral recovery from RGB images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 939–947, 2018.

[SH10] Raju Shrestha and Jon Y. Hardeberg. Multispectral image capture using two RGB cameras. In *Proceedings of the 18th European Signal Processing Conference*, pages 1801–1805, 2010.

[Smi99] Brian Smits. An RGB-to-spectrum conversion for reflectances. *Journal of Graphics Tools*, 4(4):11–22, 1999.

[SNB03] Yoav Y. Schechner, Shree K. Nayar, and Peter N. Belhumeur. A theory of multiplexed illumination. In *Proceedings of the Ninth IEEE International Conference on Computer Vision - Volume 2*, page 808, 2003.

[SSK03] Mirko Sattler, Ralf Sarlette, and Reinhard Klein. Efficient and realistic visualization of cloth. *Proceedings of the Eurographics Symposium on Rendering 2003*, pages 167–177, 2003.

[SSW+14] Christopher Schwartz, Ralf Sarlette, Michael Weinmann, Martin Rump, and Reinhard Klein. Design and implementation of practical bidirectional texture function measurement devices focusing on the developments at the University of Bonn. *Sensors*, 14(5):7753–7819, April 2014.

[SSWK13] Christopher Schwartz, Ralf Sarlette, Michael Weinmann, and Reinhard Klein. DOME II: A parallelized BTF acquisition system. In *Workshop on Material Appearance Modeling*. Eurographics Association, June 2013.

[SWRK11] Christopher Schwartz, Michael Weinmann, Roland Ruiters, and Reinhard Klein. Integrated high-quality acquisition of geometry and appearance for cultural heritage. In *The 12th International Symposium on Virtual Reality, Archeology and Cultural Heritage (VAST)*, pages 25–32. Eurographics Association, October 2011.

[SX04a] Hui-Liang Shen and John H. Xin. Colorimetric and spectral characterization of a color scanner using local statistics. *Journal of Imaging Science and Technology*, 48(4):342–346, 2004.

[SX04b] Hui-Liang Shen and John H. Xin. Spectral characterization of a color scanner by adaptive estimation. *Journal of the Optical Society of America A*, 21(7):1125–1130, Jul 2004.

[SX06] Hui-Liang Shen and John H. Xin. Spectral characterization of a color scanner based on optimized adaptive estimation. *Journal of the Optical Society of America A*, 23(7):1566–1569, Jul 2006.

[TAN+05] Masaru Tsuchida, Hiroyuki Arai, Masami Nishiko, Yoshiyuki Sakaguchi, T. Uchiyama, Masahiro Yamaguchi, Hideaki Haneishi, and Nagaaki Ohyama. Development of BRDF and BTF measurement and computer-aided design systems based on multispectral imaging. In *Proc. 10th Congress of the International Colour Association*, pages 129–132, 2005.

[TDSVG14] Radu Timofte, Vincent De Smet, and Luc Van Gool. A+: Adjusted anchored neighborhood regression for fast super-resolution. In *Asian conference on computer vision*, pages 111–126. Springer, 2014.

[TG17] Antoine Toisoul and Abhijeet Ghosh. Practical acquisition and rendering of diffraction effects in surface reflectance. *ACM Transactions on Graphics (TOG)*, 36(5):1–16, 2017.

[TSA+05] Masaru Tsuchida, Yoshiyuki Sakaguchi, Hiroyuki Arai, Masami Nishiko, Norihito Fujikawa, Masahiro Yamaguchi, Hideaki Haneishi, and Nagaaki Ohyama. High-fidelity color reproduction based on multi-channel BTF/BRDF acquisition, rendering and display. In *ACM SIGGRAPH 2005 Sketches*, page 42, 2005.

[VFTB97] Poorvi L. Vora, Joyce E. Farrell, Jerome D. Tietz, and David H. Brainard. Digital color cameras - 2 - spectral response, 1997.

[VT99] Michael J. Vrhel and H. Joel Trussell. Color device calibration: A mathematical formulation. *IEEE Transactions on Image Processing*, 8(12):1796–1806, December 1999.

[War92] Gregory J. Ward. Measuring and modeling anisotropic reflection. In *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '92, pages 265–272, New York, NY, USA, 1992. ACM.

[WGT+05] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins, and Paul Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics (TOG)*, 24(3):756–764, 2005.

[WH87] Victor R. Weidner and Jack J. Hsia. NBS measurement services: Spectral reflectance. Technical report, National Bureau of Standards, July 1987.

[WHON97] Tien-Tsin Wong, Pheng-Ann Heng, Siu-Hang Or, and Wai-Yin Ng. Image-based rendering with controllable illumination. In *Proceedings of the Eurographics Workshop on Rendering Techniques '97*, pages 13–22, London, UK, 1997.

[WMP+06] Tim Weyrich, Wojciech Matusik, Hanspeter Pfister, Bernd Bickel, Craig Donner, Chien Tu, Janet McAndless, Jinho Lee, Addy Ngan, Henrik Wann Jensen, and Markus Gross. Analysis of human faces

using a measurement-based skin reflectance model. *ACM Transactions on Graphics (TOG)*, 25(3):1013–1024, 2006.

[WPSB08] Ashwin A. Wagadarikar, Nikos P. Pitsianis, Xiaobai Sun, and David J. Brady. Spectral image estimation for coded aperture snapshot spectral imagers. In *Image Reconstruction from Incomplete Data V*, volume 7076, pages 9 – 23. International Society for Optics and Photonics, SPIE, 2008.

[WPSB09] Ashwin A. Wagadarikar, Nikos P. Pitsianis, Xiaobai Sun, and David J. Brady. Video rate spectral imaging using a coded aperture snapshot spectral imager. *Optical Express*, 17(8):6368–6388, Apr 2009.

[WS00] Günter Wyszecki and Walter S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae, 2nd Edition*. Wiley, July 2000.

[Wys63] Günter Wyszecki. Proposal for a new color-difference formula. *Journal of the Optical Society of America*, 53(11):1318–1319, November 1963.

[ZF14] Zhong Zhao and Guocan Feng. A dictionary-based algorithm for dimensionality reduction and data reconstruction. In *Proceedings of the 22nd International Conference on Pattern Recognition*, pages 1556–1561, 2014.

[Zha00] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, November 2000.

List of Publications

S. Merzbach, M. Hermann, M. Rump, and R. Klein. Learned Fitting of Spatially Varying BRDFs. In: *Computer Graphics Forum*, 38:4, July 2019.

S. Merzbach, M. Weinmann, M. Rump, and R. Klein. Fast Capture of Spectral Image Series. In: *Proceedings of 12th International Conference on Computer Graphics Theory and Applications (VISIGRAPP)*, pages 148-159, 2017.

C. Schwartz, R. Sarlette, M. Weinmann, M. Rump, and R. Klein. Design and Implementation of Practical Bidirectional Texture Function Measurement Devices focusing on the Developments at the University of Bonn. In: *Sensors*, 14:5, 2014.

M. Rump, A. Zinke, and R. Klein. Practical Spectral Characterization of Trichromatic Cameras. In: *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)*, 30:6, 2011.

M. Rump, R. Sarlette, and R. Klein. Groundtruth Data for Multispectral Bidirectional Texture Functions. In proceedings of *CGIV 2010*, pages 326-330, June 2010.

M. Rump and R. Klein. Spectralization: Reconstructing spectra from sparse data. In proceedings of *SR '10 Rendering Techniques*, pages 1347-1354, 2010.

A. Zinke, T. Lay Herrera, A. Andriyenko, M. Rump, A. Weber, and R. Klein. A Practical Approach for Photometric Acquisition of Hair Color. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia)*, 28:5(165), 2009.

M. Rump, R. Sarlette, and R. Klein. Efficient Resampling, Compression and Rendering of Metallic and Pearlescent Paint. In proceedings of *Vision, Modeling, and Visualization*, pages 11-18, 2009.

R. Ruiters, M. Rump, and R. Klein. Parallelized Matrix Factorization for fast BTF Compression. In proceedings of *Eurographics Symposium on Parallel Graphics and Visualization*, pages 25-32, 2009.

M. RUMP, G. MÜLLER, R. SARLETTE, D. KOCH, AND R. KLEIN. Photorealistic rendering of metallic car paint from image-based measurements. In *Computer Graphics Forum*, 27(2), pages 527-536, 2008.