

Exploring the Frequency Domain for Perceptually Plausible Material Edits

Kumulative Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Dipl.-Math. Marlon Mylo

aus

München

Bonn

Dezember 2019

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

1. Gutachter: Prof. Dr. Reinhard Klein
2. Gutachter: Prof. Dr. Andreas Schilling

Tag der Promotion: 8. Juli 2020

Erscheinungsjahr: 2020

Contents

Disclaimer	xi
Abstract	xiii
Zusammenfassung	xv
1 Introduction	1
1.1 Frequency Based Similarity Detection	3
1.2 Frequency Born Attributes in Human Sensory System	4
1.3 Contribution	7
1.4 Publications	8
2 Preliminaries	11
2.1 Physically Based Rendering	12
2.1.1 Digital Image Synthesis	13
2.1.2 Radiometric Quantities	14
2.1.3 The Rendering Equation	14
2.1.4 Colour Spaces and Dynamic Range	19
2.1.5 Realism and Methodology	21

2.2	Material	22
2.2.1	Bidirectional Reflectance Distribution Functions	22
2.2.2	Bidirectional Texture Functions	26
2.2.3	Acquisition of Measured Materials	29
2.3	Material Editing	31
2.3.1	Plausible, Predictable and Meaningful Editing Operations	33
2.3.2	State of Research	33
2.4	Analysis of the Frequency Spectrum	37
2.4.1	The Fourier Transform	38
2.4.2	The Wavelet Transform	42
2.5	Systems	53
3	An Application of the Convolution Theorem	55
3.1	PushPins	57
3.2	Overview	59
3.3	Detailed Description of the Workflow	61
3.3.1	Descriptors	61
3.3.2	Classifier	61
3.3.3	Pushpins	63
3.3.4	Convolution with a Pushpin	64
3.3.5	The Lattice Detector: MSBP	65
3.3.6	Finding a Model Tile	68
3.3.7	Calculating the Optical Flow and Composing a Tile Mask	69
3.3.8	Recombination	69
3.3.9	Applying the Edits	69
3.4	Evaluation	70
3.4.1	Test Set-Up	70
3.4.2	The Results	74
3.4.3	Editing Examples	74
3.4.4	Time Requirement	76
3.5	Conclusion and Future Work	76

4	Frequency Editing of Bidirectional Texture Functions	79
4.1	From Affordances to Appearance Bending	81
4.2	Appearance Bending: Scale-Space Manipulation of Materials . .	84
4.2.1	Choice of Basis	84
4.2.2	Color Spaces and Dynamic Range	85
4.2.3	Appearance Bending on BTFs	87
4.3	Evaluation	90
4.3.1	Extreme Edits	93
4.3.2	Comparison with Image Based Editing	93
4.3.3	Performance	96
4.4	Conclusion and Physical Aspects	96
5	A Cortical Appearance Space Model	101
5.1	Cortex Models	103
5.2	Our New Cortex Model	104
5.2.1	The Computational Model of the Early Vision	108
5.2.2	An Empirically Based Model of the Visual Cortex	108
5.2.3	Parameters	109
5.3	Transferring Edits to the Model of the Visual Cortex	113
5.3.1	Learning an Operator	114
5.3.2	The Operators	115
5.4	Evaluation	116
5.4.1	Identity	116
5.4.2	Edge Aware Imaging	117
5.4.3	Affordance Editing	118
5.4.4	Spotlight Moving	121
5.4.5	Editing of High-Dimensional Material Representations . .	121
5.4.6	Time Requirement	124
5.5	Conclusion	124
6	Conclusion	129

A Notation	131
B Abbreviations	135
C Tables	137
D List of Figures	139

Danksagung

Bei meinem Betreuer Reinhard Klein bedanke ich mich für das Thema, das entgegengebrachte Vertrauen und die vielfältigen Ermutigungen. Andreas Schilling hat sich um meinen Dank verdient gemacht, indem er fachliche Fragen mit dieser unnachahmlichen Begeisterung für Detail und Tiefe mit mir erörtert hat und auch in schwierigen Momenten zu ermuntern wusste. Bei Frank Kurth bedanke ich mich für die mannigfache Unterstützung, nicht zuletzt in Fragen der Signalverarbeitung und den vielen Humor.

Dank gebührt auch den Mitautoren Matthias Hullin, Qasim Zaidi und Martin Giesel für die angenehme und entgegenkommende Zusammenarbeit.

Raban hat sich auf schwindelerregendem Niveau mit den Inhalten meiner Arbeit auseinandergesetzt, Warin hat meine langjährige geistige Abwesenheit in unvergleichlicher Nonchalance zur Kenntnis genommen, und wenn man, wie Awa, erst 8 ist, dann können einem 9 Jahre, in denen der Papa schnell noch etwas machen muss, lang vorkommen. Vielen Dank für eure Hilfe und eure Nachsicht, Kinder.

Vielen Dank, Halli, und auch Dir, Sabrina, für den Halt und die Zeit, welche Ihr mir geschenkt habt.

Schließlich möchte ich mich für fachliche Unterstützung bei meinem Freund

Stratos bedanken, bei Michael Weinmann, bei Max Hermann, bei Ralf Sarlette, bei Roland Ruiters, bei Stefan Hartmann, bei Sebastian Merzbach und bei Raoul Wessel.

Disclaimer

The content of this dissertation is based on three first author publications that have been presented at different conferences with a focus on computer graphics.

The publication with the name *Pushpins for Edit Propagation* [MK17] was published at the *International Conference on Computer Graphics (WSCG 2017)*. The paper presents a novel approach to propagating a bipartite stroke input over a near-regular-textured material. I implemented, evaluated, and wrote the paper. My co-author, Prof. Dr. Reinhard Klein, supported the work by proof-reading and discussing technical aspects.

The paper *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models* [MGZ⁺17] was published and presented at the *22nd International Symposium on Vision, Modeling and Visualization (VMV 2017)*. This paper transfers an affordance manipulation scheme from the field of visual perception, where it had been tested on photographs, to the field of computer graphics, where I could show that it may be applied to complex material representations. I implemented, evaluated, and wrote the paper. Prof. Dr. Matthias Hullin and Prof. Dr. Reinhard Klein supported the work by proof-reading and discussed technical issues with me and Dr. Qasim Zaidi and

Dr. Martin Giesel provided me with the necessary background to understand their work which is among the foundations of this publication.

The paper *Linear Subspaces of the Appearance Space* was published and presented at the *International Conference on Computer Graphics* (WSCG 2018). In this paper I introduce a model of the visual cortex to learn different material editing schemes. I implemented, evaluated, and wrote the paper. My co-author, Prof. Dr. Reinhard Klein, supported the work by proof-reading and technical discussions.

The work at hand contains text parts of the three mentioned papers that were transferred to this thesis without further modification. To make these copied sentences recognizable among the newly written parts they are highlighted in gray.

Abstract

The modeling of consistent 3D-Worlds and the rendering of photo-realistic images and films is a difficult task even for well-trained 3D-designers. This is in particular the case if the resulting media show material surfaces which do not or not yet exist. An established approach is to derive the reflectance-properties and the mesostructure of an imaginary material surface from reflectance-properties and mesostructure of existing material surfaces. In other words this problem is approached by manipulating measured, digitalized material-surfaces. According to the importance of the problem, there have been published many computer graphical research papers in this field. As the rendering of photo-realistic images is essentially done by simulation of ray-optics, most publications concentrate on manipulating the physical properties of the measured materials. But those measured optical properties often do not allow to draw inferences about the underlying physical phenomena. That is why this approach often fails. In this thesis we are trying to find editing-systems which make use of the spatial contrast between those measured reflectance models. We want to reach this target by analysing and manipulating the spatial frequency-structure of the local reflectance-models.

First we want to investigate how to make use of the Fourier-analysis to

propagate a bichromatic stroke-input robustly over a material which bears at least one near-regularly textured channel. We will show that this near-regular-structure may be used to robustly generate high-quality editing masks to separate a foreground-pattern from the background.

In the following we suggest perceptually motivated operations on the frequency spectrum. We will transfer a bandpass-filtering scheme, which has shown astonishing results in material-image editing, to bidirectional texture functions. To evaluate our approach we present many editing examples which show that this transfer of the original editing-scheme is reasonable and leads to good results.

To motivate our next step we formulate the hypothesis, that this approach is in fact a manipulation of the perception. So we will evolve our approach by doing the frequency-analysis in an empirically founded computational cortex transform model instead of a Fourier- or a wavelet-analysis. Editing operations in this model are linear operations which have to be learned by the use of editing examples. We will show by exemplary comparisons that this model is not only capable of learning the mentioned editing-scheme but can also represent optical phenomena which then may be used as editing operation.

Zusammenfassung

Die Modellierung konsistenter 3D-Welten und das Rendering fotorealistischer Bilder und Filme stellt auch geübte 3D-Designer vor eine schwer lösbare Aufgabe. Insbesondere dann, wenn in den resultierenden Medien Materialoberflächen zu sehen sein sollen, die in dieser Form nicht oder noch nicht existieren. Oft versucht man hier die Reflektanzeigenschaften und die Mesostruktur imaginärer Materialoberflächen aus denen existierender herzuleiten also gemessene, digitalisierte Materialoberflächen zu editieren. Entsprechend der Wichtigkeit des Themas weist die Computergrafik einen beachtlichen Korpus an Veröffentlichungen hierzu auf. Weil das Rendering fotorealistischer Bilder im Wesentlichen durch eine Simulation der Strahlenoptik, also der Physik, erreicht wird, konzentrieren sich die meisten Veröffentlichungen auf die Manipulation der physikalischen Eigenschaften digitalisierter Materialoberflächen. Die gemessenen optischen Eigenschaften lassen bei der Digitalisierung existierender Materialien jedoch oft wenig Rückschlüsse auf die zugrunde liegenden physikalischen Vorgänge zu, deswegen führt dieser Ansatz mitunter ins Leere. In der vorliegenden Dissertation geht es darum, Editiermethoden zu finden, die die räumlichen Kontrastverhältnisse zwischen den gemessenen lokalen Reflektanzmodellen ausnutzen. Dies soll durch Analyse und Manipulation der räumlichen Frequen-

zstruktur der lokalen Reflektanzmodelle erreicht werden.

Zunächst werden wir untersuchen, wie die Fourieranalyse genutzt werden kann, um zweifarbige Stricheingaben robust über Materialien zu propagieren, die mindestens über einen Kanal verfügen, der ein nahezu gleichmäßiges Wiederholungsmuster trägt. Wir werden zeigen, dass auf Basis der Wiederholungsstruktur mit hoher Robustheit hochqualitative Editiermasken erstellt werden können mit denen ein Muster im Vordergrund vom Materialhintergrund getrennt werden kann.

Die im Weiteren vorgeschlagenen Operationen auf dem Frequenzspektrum sind perzeptuell motiviert. Wir werden ein Bandpassfiltersystem, welches auf Materialfotos zu erstaunlichen Ergebnissen führt, auf bidirektionale Texturfunktionen übertragen. Anhand von zahlreichen Editierbeispielen werden wir belegen, dass diese Übertragung sinnvoll ist und zu guten Ergebnissen führt.

Vor dem Hintergrund der Hypothese, dass es sich hier um Manipulationen der Wahrnehmung handelt, werden wir im Anschluss die Frequenzerlegung nicht mehr mithilfe der Fourierbasis oder einer orthogonalen Waveletbasis vornehmen sondern in einem auf empirischen Daten beruhenden rechnerischen Cortex Transformationsmodell. Editieroperationen werden in diesem Modell als lineare Transformationen dargestellt, die auf Basis von Editierbeispielen erlernt werden. Wir werden anhand von exemplarischen Vergleichen zeigen, dass nicht nur die angesprochenen Editieroperationen in diesem Modell gelernt werden können, sondern auch andere optische Phänomene, die mit Hilfe des vorgestellten Ansatzes nunmehr als Materialeditieroperation zur Verfügung gestellt werden können.

CHAPTER 1

Introduction

Computer graphical progress has ever since been the figurehead of developments in computer science. As early as 1979, the astonishing results of the Whitted-raytracer gave already proof of the principal possibility to render a digital 3D-scene description to a photo-like 2D-image. But neither were computers powerful enough to cope with the vast quantities of data, necessary to describe a relatively simple all day scene, nor was the computer graphical progress far enough to allow for calculations of the light-exchange on the material surface of more complex materials.

Many things have changed since then. Today it is possible to generate pictures of highest complexity which are nearly indistinguishable from photographs. In this rendering process the material surface has a key role. If the light exchange on the material surface is not modelled neatly surfaces will not appear to be realistic.

One approach to cope with this problem is to sample the reflectance properties of an existing material surface over a discrete light- and viewing-direction grid and use those tabulated reflectance data for rendering. This very suc-

cessful technique has the significant limitation that it is extremely difficult to apply meaningful and still realistic changes to the measured data. But of course making realistic and meaningful changes to the material appearance is highly desirable for 3D-Computer artists. So manipulating measured reflectance data has been an intensively studied topic in computer graphics during the last two decades and it is also the subject of this thesis.

To be more precise we are investigating in this thesis how to make use of the spatial frequency spectrum of parameter maps of measured materials to achieve realistic, predictable and meaningful manipulations.

The oldest mathematical means to investigate the frequency domain of a given time signal f is the Fourier transform \mathbf{F} ,

$$\mathbf{F}f(\nu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i2\pi\nu t} dt. \quad (1.1)$$

The Fourier transform may be seen as an idealized frequency analyser. We call it idealized because frequencies are represented by infinite sine and cosine waves. The simplest remedy of this idealization is to limit the range of the Fourier-transform. That is, to calculate the short time Fourier transform by multiplying the signal in question with a L^2 -integrable window function (w):

$$\mathbf{F}_{w,\mu}f(\nu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\bar{w}(t - \mu)e^{-i2\pi\nu t} dt. \quad (1.2)$$

A Gaussian windowed Fourier transform is called Gabor transform.

$$\Gamma_{\mu}f(\nu) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-\pi(t-\mu)^2} e^{-i2\pi\nu t} dt. \quad (1.3)$$

The idea of the window-function is to confine the frequency analysis to a certain range in time, so the window function does not depend on the frequency-parameter ν . This differs from the scaling analysis done by the wavelet trans-

form:

$$\mathbf{W}f_{s,r} = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(t) \bar{\psi} \left(\frac{t-r}{s} \right) dt. \quad (1.4)$$

If the scaling analysis function ψ is compactly supported the support shrinks with a shrinking scaling factor s . So the support adapts to the scaling and enables localized scaling analysis. And finally we may use the Gabor-function as mother wavelet. A formula for a two dimensional Gabor mother-wavelet is given by:

$$\psi_{\kappa} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2+y^2)} \cdot \left[e^{i\kappa x} - e^{-\frac{\kappa^2}{2}} \right], \quad (1.5)$$

where κ is a function of the bandwidth of the Gaussian filter.

Those operators will be used in two different scenarios:

1. for frequency based similarity detection and
2. to model frequency born attributes in human visual system.

1.1 Frequency Based Similarity Detection

Periodical recurrence has a certain magic inherent. First asked for periodically recurring phenomena, we are inclined to think of a clock hand, of the change of day and night, and, perhaps after some reflection, of the year with its four seasons. Once we take a closer look, we see that periodicity is all over there, like in the pattern of the paved side walk, or in the transient regulation of the heating system. Material surfaces in particular bear often a periodically recurring structure, like the slings in a knit work or the imprinted pattern of structural metal.

We want to decompose such a surface pattern into an editable foreground and a non-editable background region. Periodicity supports the solution of this matting-problem in a two-fold way:

1. we may mask the frequency spectrum of the material to highlight similar regions and
2. we may calculate an average master tile to gather statistics on which part of the pattern should be considered as foreground and which as background.

Depending on the number of repetitions of the recurring structure the generation of the master tile is increasingly robust. By recombining the matted master-tile we may obtain a high-quality editing mask for a given material patch. In Chapter 3 we present such an algorithm. For our algorithm it suffices if one of the materials channels bears a near-regular pattern.

1.2 Frequency Born Attributes in Human Sensory System

A totally different approach to frequency-based material editing comes from the field of visual perception. Giesel and Zaidi could tie certain affordances to specific frequency bands in the Fourier domain of material photographs. *Affordance* is a concept from psychology; it describes a property or a quality of an object which suggests a possible use of this object. Those affordances may be weakened or strengthened by scaling of the appropriate frequency band. For realistic material editing this is a piece of good fortune as those frequency manipulations are not only realistic, predictable and meaningful but due to their linearity they may be applied to the base of compressed reflectance tensors. This makes affordance editing an ideal candidate for bidirectional texture function editing. In Chapter 4 we will investigate this approach and show that we may substitute the Fourier-transform by a member of Daubechies orthogonal wavelet transform family. Using the wavelet family brings a better localization of the edit and therefore admits editing of multi-component materials without the occurrence of strong artefacts.

Our results indicate that frequency based affordance editing is a promising approach to material editing. We can even present an example where manipulating a BTF-material brings more stable results than manipulating a photo of the material. While we leave aside the physical aspects of those frequency space manipulations we want to have a closer look at the chosen frequency analysis. The Fourier transform is a good choice to get a first impression of the effects of specific frequency manipulations. But as the base vectors of the discrete Fourier transform span the whole image the resulting manipulation is not robust against irregularities in the material. As mentioned before, the robustness may be increased by the use of a multi-scale analysis. The chosen Daubechies-wavelets have the additional advantage that they allow for a fast wavelet transform.

In Chapter 5, we want to give up the aspect of computational performance and deepen the idea of a manipulation of the visual perception. That is from neuroscience we gather the results on the filtering system of the visual striate cortex and compose an artificial Gabor-filter based filtering system which respects the neuroscientific empirical analysis. We show that such a system admits reconstruction of a given input by weighted summation of the filter responses. To transfer the results of Chapter 4, we will learn how to represent bandpass-filtering in our new cortex transform model. Additionally we will show that our cortex transform model is capable of representing other optical phenomena like edge enhancement or small movements of a light source.

The contribution of this thesis will be described in detail within the next section. We want to support the point of view that appearance manipulation is not only a matter of the underlying optical phenomena but also a matter of visual perception.

Throughout this publication we are trying to keep the same notation for the same object (see Appendix A) and we are trying to avoid ambiguities with exception of radiometric quantities, which we notate like given in [NRH⁺77], and citations from third party publications where we use the same notation as in the publication to increase readability. We identify a given direction, mostly

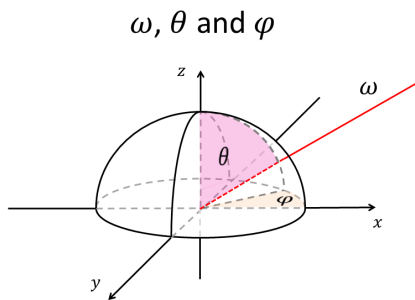


Figure 1.1: A direction and the elevation angle and the azimuth.

notated as ω , with the unit vector into this direction and we write $\omega = (\theta, \varphi)$ for the elevation angle θ and the azimuthal angle φ with implicit reference to a suitable local coordinate system. We use both, ϕ and φ , the first of which for scaling functions, the second for azimuthal angles.

1.3 Contribution

In this thesis we want to find answers to the question how the spatial frequency spectrum of parameter- or reflectance-maps of measured materials may be utilized to apply plausible editing-operations to the measured material. The term *plausible* will be refined in the second Chapter, Subsection 2.3.1. We want to contribute to the field of material editing by presenting the following results:

- we introduce pushpins, a correlation based technique to identify similar regions in near regular textured texture channels,
- those pushpins are embedded in a novel workflow to identify a foreground component of a measured material,
- we transfer an editing scheme of material images to parameter-maps of complex local reflectance models,
- we explore how this Fourier-transform based editing scheme may be transferred into a wavelet based editing scheme to circumvent undesired global effects of the Fourier-transform,
- starting with the previous findings, we develop a model of the visual cortex to represent the editing scheme in this cortical filter model,
- finally we show that our novel cortex transform model is capable of learning new plausible material edits.

1.4 Publications

This cumulative thesis is based on the following publications:

1. Marlon Mylo, and Reinhard Klein. Pushpins for edit propagation. In *International Conference on Computer Graphics, Visualization and Computer Vision*, volume 25 of *WSCG proceedings*, pages 143 – 152, [MK17].

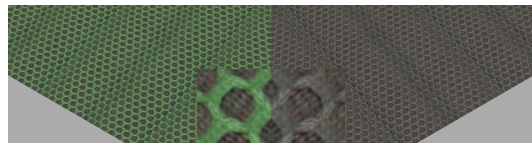


Figure 1.2: Pushpins for edit propagation.

2. Marlon Mylo, Martin Giesel, Qasim Zaidi, Matthias Hullin, and Reinhard Klein. Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models. In *Proceedings of Vision, Modeling & Visualization, The Eurographics Association*, [MGZ⁺17]

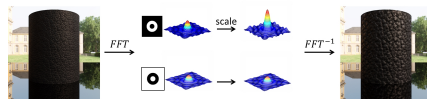


Figure 1.3: Appearance Bending.

3. Marlon Mylo, and Reinhard Klein. Linear Subspaces of the Appearance Space. In *International Conference on Computer Graphics, Visualization and Computer Vision*, volume 26 of *Journal of WSCG*, pages 95 – 103, [MK18]
4. Zaidi, Qasim and Giesel, Martin and Mylo, Marlon and Klein, Reinhard. Perception and appearance bending of material properties. In *Journal*

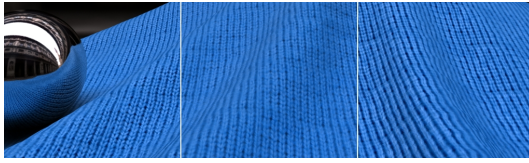


Figure 1.4: Linear Subspaces of the Appearance Space.

of Vision, volume 17(15), page 21, year 2017, [ZGMK17]

CHAPTER 2

Preliminaries

In this chapter we want to collect the fundamentals. We start by introducing the foundational ideas and definitions of digital image synthesis. In the first section we describe the tools, models and workflows, necessary to render a photorealistic image from the description of a 3D-scene.

It follows the concept of digitalized materials. We will present three different local reflectance models in form of analytical models of *bidirectional reflectance distribution functions* (**BRDF**). An important representation of digitalized material optics is the *bidirectional texture function* (**BTF**). The computer graphics department of the university of Bonn has established considerable expertise in measuring, manipulating and editing BTFs and it is a focal point of this thesis, too. BTFs will be introduced after the local reflectance models. With regard to editing schemes it is necessary to outline the state of the art of the different material measurement schemes and the resulting data structures.

It feels a little bit preposterous to denote an artificially modified digital material representation as realistic. Still realism is the currency, so we will introduce the term *plausible* for manipulated material surfaces which occur as

if they were / could be real. Evaluated is the plausibility by computer graphics experts. This definition and an overview over existing editing schemes with focus on manipulations of high-dimensional materials will be given in the fourth section of this chapter.

This work is about the usability of spatial frequency descriptions of measured material maps for the process of editing those material maps. So the first section introduces the Fourier transform. Fourier uses an ad-hoc definition of the integral whereas we assume in this thesis of course Lebesgues integral definition. A particularly useful tool is the convolution theorem. We want to emphasise its fundamental meaning for computer-graphical applications in general and stress that the convolution theorem is a core element of our first contribution, presented in the next chapter (Chapter 3). Subsequently we will introduce localized frequency analysis techniques, namely, the orthogonal Wavelet-transforms (Subsection 2.4.2) and filtering a signal by the use of Gabor-base (Subsection 2.4.2) functions. We close the chapter with a tabulated overview over the systems we used in our experiments.

2.1 Physically Based Rendering

In this section we want to gather the facts and notations from 3D-rendering necessary for the understanding of the subsequent chapters. A very good introduction to physically based rendering may be found in [PH04], which is also freely available online under <https://www.pbrt.org/>. According to [PH04], by 3D-rendering we denote the process of producing an image from the description of a 3D-scene.

Physically based rendering is a form of digital image synthesis [Gla95], which will be explained in more detail in Subsection 2.1.1. Rendering a scene with opaque material-surfaces may be done by integrating the rendering equation (Subsection 2.1.3). The last subsection is left for some words on colour-spaces and dynamic range.

2.1.1 Digital Image Synthesis

While the term may be used in a broader sense, by *digital image synthesis* we denote here the synthesis of realistic images by the use of a computer. A special form of digital image synthesis is 3D-Rendering. 3D-Rendering is the process of evaluating a scene to a digital image. This evaluation is done by simulating the light exchange on the surfaces of the objects in the scene and in the participating media.

The Scene

The greek word *skené* means *tent* or *hut*. With respect to the Greece Theatron, *skené* meant a building in the back of the stage. So the *skené* holds an assembly of the figures and objects to be presented to the audience and this is our starting point.

Here the scene describes all objects, their optical behaviour, their position and extent in a three-dimensional space and the audience of the scene in form of a camera-model. The temporal behaviour of those objects is not necessarily part of a scene and does not play a role in this thesis. A formal and computable description of a scene may be done by a scene-graph. Pharr et al. [PH04] name seven objects / phenomena which have to be part of every scene:

1. the camera,
2. a description to enable ray-object intersection,
3. light sources,
4. a visibility function which decides whether two points may be connected via a ray,
5. surface scattering, which describes the light exchange on the material surface,
6. indirect light transport and

7. rules for the ray propagation through the medium of the scene.

this list is tailored to the demands of raytracing algorithms, but it holds in comparable form for rasterizers. Here we are concerned with the fifth point: the light exchange on the material surface.

2.1.2 Radiometric Quantities

Before we describe the rendering equation and subsequently the BRDF we have to introduce the involved radiometric quantities. An overview with a short description is given in Table 2.1.

2.1.3 The Rendering Equation

To evaluate a scene to a digital image we have to calculate the light exchange in the scene. The massive increase of computational power during the 1970s led during the early 1980s to enhanced interest on modeling and calculating the light exchange in a scene [CPC84, Whi79, SH81]. In 1986 James Kajiya published an equation which summarized the mathematical formulation of those different approaches, the rendering equation [Kaj86]: With his first formulation of the rendering equation

$$I(\mathbf{x}, \mathbf{x}') = g(\mathbf{x}, \mathbf{x}') \left[\epsilon(\mathbf{x}, \mathbf{x}') + \int_{\mathfrak{E}} \rho(\mathbf{x}, \mathbf{x}', \mathbf{x}'') I(\mathbf{x}', \mathbf{x}'') d\mathbf{x}'' \right]. \quad (2.1)$$

Kajiya does not only summarize the mathematical formulation of the former approaches under one equation but also connects the geometrical description of a scene with the relevant radiometric quantities by the use of *transport quantities*: given three arbitrary points \mathbf{x}, \mathbf{x}' and $\mathbf{x}'' \in \mathbb{R}^3$ of a scene, then

$I(\mathbf{x}, \mathbf{x}')$ **unoccluded transport intensity**

is an intensity, passing from one point to the other,

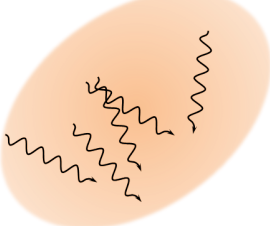
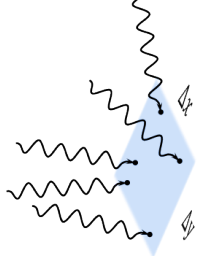
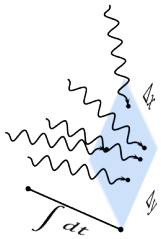
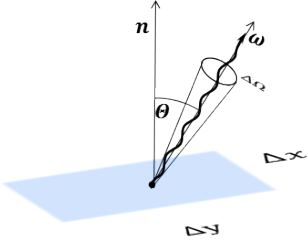
Quantity	Symbol Unit	Explanation	Sketch
Radiant Flux	Φ [W]	The <i>radiant flux</i> is the power propagated as optical electromagnetic radiation.	
Irradiance	E [W/m ²]	<i>Irradiance</i> is the flux per-area-density on a given surface-element.	
Exposure	H [J/m ²]	<i>Exposure</i> is the radiometric energy per area which has been received by a surface-element. It is the integral of the irradiance over time. CCD-sensors measure the exposure.	
Radiance	L [W/(m ² ·sr)]	<i>Radiance</i> is the flux per-area- and per-solid-angle-density on a given surface-element within a given solid-angle-element. $L = \frac{\partial \Phi}{\partial A \partial \Omega \cos \Theta}$	

Table 2.1: Radiometric quantities

$g(\mathbf{x}, \mathbf{x}')$ **mutual visibility**

describes the mutual visibility of the two points \mathbf{x} and \mathbf{x}' , and

$\rho(\mathbf{x}, \mathbf{x}', \mathbf{x}'')$ unoccluded three point transport reflectance

describes how much light is scattered in point \mathbf{x}' from point \mathbf{x}'' to point \mathbf{x} .

 $\epsilon(\mathbf{x}, \mathbf{x}')$ unoccluded two point transport emittance

describes the intensity of emitted light from \mathbf{x} to \mathbf{x}' .

 \mathfrak{S} integration range

\mathfrak{S} ranges over all surface elements in the scene.

The physical dimensions of the transport quantities differ slightly from those in the final rendering equation but we will not go into detail here.

Kajiya shows, how to derive a formulation of the rendering equation in terms of ordinary radiometric quantities [Kaj86, eq. 5-11]:

$$L_o(\mathbf{x}, \omega_o) = L_e(\mathbf{x}, \omega_o) + \int_{\mathfrak{S}} \rho(\mathbf{x}, \omega_i, \omega_o) L_i(\mathbf{x}, \omega_i) \langle \omega_i, \mathbf{n} \rangle d\omega_i. \quad (2.2)$$

Here \mathbf{x} is a surface point, ω_i is the direction of the incoming light, ω_o is the direction of reflectance, and the vector $\mathbf{n} = \mathbf{n}(\mathbf{x})$ stands normal to the surface element in \mathbf{x} .

$L_{(o,e,i)}(\mathbf{x}, \omega_{(o,e,i)})$ is a radiance. For L_o and L_e the radiance is outgoing, the radiance L_i is incoming.

$\rho(\mathbf{x}, \omega_i, \omega_o)$ is a bidirectional reflectance distribution function. The BRDF is by definition the ratio between the reflected radiance into direction ω_o , confined to reflections from direction ω_i and the irradiance confined to the incoming direction ω_i (see Figure 2.1 and Subsection 2.2.1-**BRDF**).

$$\rho(\mathbf{x}, \omega_i, \omega_o) := \frac{\partial L(\mathbf{x}, \omega_i, \omega_o)}{\partial E(\mathbf{x}, \omega_i)}. \quad (2.3)$$

The unit is $[sr^{-1}]$.

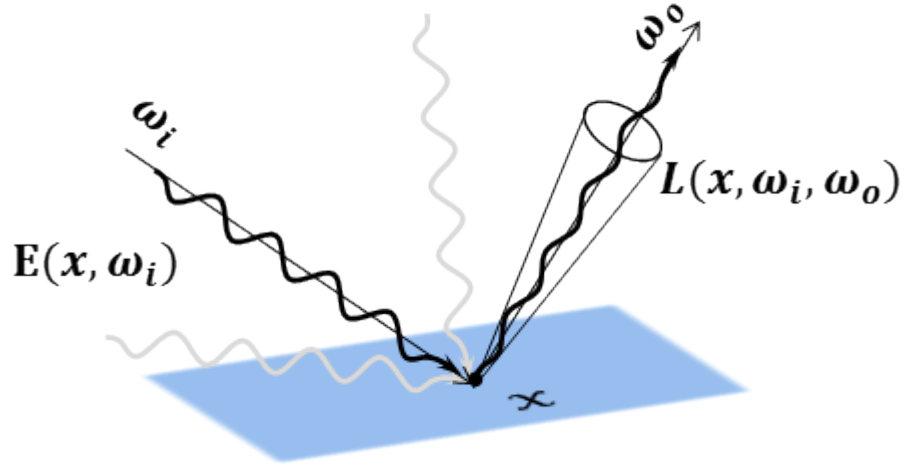


Figure 2.1: The BRDF is the quotient between the incoming irradiance E confined to a direction ω_i and the outgoing radiance L , confined to ω_o .

$\langle \omega_i, \mathbf{n} \rangle$ the cosine between the surface normal and the incoming light direction accounts for the foreshortening of the unit area perpendicular to the incoming ray.

The rendering equation makes a couple of assumptions:

1. it is based on a geometrical approximation of optics,
2. in its plain form it does not allow any treatment of diffraction,
3. the media between surfaces is not taken into consideration,
4. W.r.t the distribution ρ , Kajiya speaks explicitly about a *reflectance*-distribution, meaning that his formulation of the rendering equation confines to opaque materials.

In spite of adverse effects over mesh edges, all surfaces are assumed to have finite curvature, meaning that the tangent space, the surface normals and the

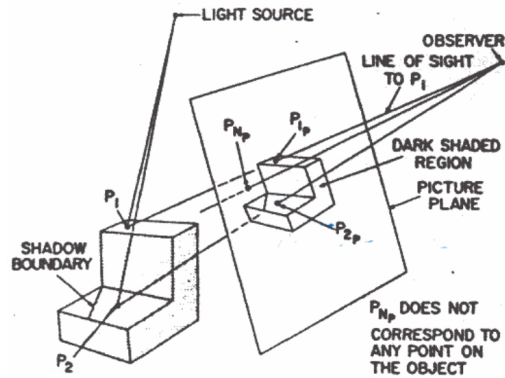
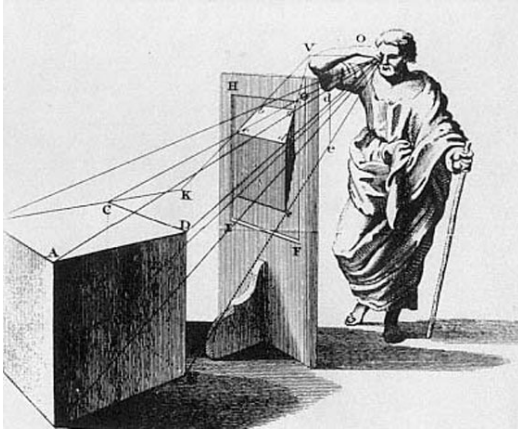


Figure 6 – Point by point shading

Figure 2.2: To explain perspective drawing, Leon Battista Alberti uses as early as 1436 visual rays for his argumentation (left image). Though visual rays lost their ontological meaning since then, they describe exactly the concept which has been used in the early raytracing algorithms from the 1960's (right image from [App68]).

upper hemisphere \mathbb{S} are well-defined. This thesis is about opaque materials, so $\omega_o, \omega_i \in \mathbb{S}$.

Solving the rendering equation

While Kajiya also introduces a new solver for the light-exchange integration problem, rendering techniques are no topic of this thesis, so we leave Kajiya's traces here, but not without mentioning, that Immel et al. [ICG86] independently published another approach to the rendering equation in the same year.

To produce a picture, the rendering equation is evaluated on the pixels of an image. That is the irradiance received by each pixel of the pixelated projection surface of a camera model is calculated for each wavelength-filter of the selected colour space with the desired colour-intensity-dynamics. This recursive integration of the rendering equation is mostly been done by a derivative of the raytracing-algorithm. The first raytracers may be seen as implementation of the medieval idea that the human eye feels for its environment by the use of eye-rays [App68] though the modern idea is of course not the simulation

of eye-rays but an approximation of geometrical optics.

2.1.4 Colour Spaces and Dynamic Range

In this thesis pixels occur in three different contexts: as elements of parameter maps for spatially varying BRDFs, as colour dependent reflectance values of bidirectional reflectance maps and as colour-values of photos or renderings. In this subsection we want to clarify the meaning of the pixel values of photos and reflectance maps, starting with the colour.

As humans are trichromatic it is approximately possible to express every perceptible colour in a trichromatic colour-space. The RGB-colour space is additive and derived from human colour-sensation. Additivity is important for compression. It is somewhat tricky to speak of the RGB-colourspace because every camera and every monitor has its own filtering system. Of course there exist standardizations like eci-RGB (ISO 22028).

In this thesis we used mostly the YUV-colour encoding system. The transformation between RGB and YUV is linear. YUV separates the RGB-color space into luma and chrominance. This may be exploited by the fact, that the human perception is more sensible to luma changes than to chrominance changes [Alb13]. So, for compression techniques it is profitable to distinguish between luma and chrominance because the chrominance-channel may be highly compressed with comparably low quality loss.

As we are doing physically based rendering the amount of a pixel-value in an image should always be seen as representative of a radiometric quantity. For material acquisition we use of-the-shelf cameras. The CCDs (*charge-coupled device*) of those cameras measure the exposure and map the irradiated energy ($E\Delta t$) non-linearly to pixel values. The most significant mark of this non-linearity is the saturation barrier. As we want to recover reflectance functions from the photos, we need a linear relation between the pixel-values and the irradiance. This is necessary to obtain energetically correct renderings because in the rendering equation (Equation 2.2) the reflectance values are multiplied

with a radiance value and sum up to a power density.

The process of acquiring energy linear radiance-maps (from the perspective of the scene) is described in detail by Debevec and Malik [DM97]. If we want to measure the reflectance of a material surface, we can use their setting to estimate a radiance from the material surface into the direction of the camera. In order to calculate reflectance-values from those radiance maps, we still have to estimate the irradiance onto the material surface. This can be done by a radiometric calibration with a material surface with known reflectance properties. Schwartz et al. use e.g. SphereOptics Zenith UltrawWhite[©] [SSW⁺14].

The *dynamic range* is a term from signal-processing and describes for a given quantity the ratio between the largest and the smallest possible value. Under the assumption, that the smallest value is also the smallest possible step size, the dynamic range is a measure for the amount of information which may be represented by that quantity. In computer graphics it is not unusual to speak of a *dynamic range* w.r.t. a non-linear scale. In this thesis about the amount of information are not possible.

In computer-numbers the amount of representable information is limited by the number of bits. According to Fechners law the visual system responds to changes in energy with logarithmic changes in stimulus. This makes floating-point numbers more suitable for representing radiometric values. In this thesis values are stored in three different number types:

1. Images are mostly stored as 8 bit integer RGB-values.
2. The values stored in reflectance arrays like BTFs are stored in floating point numbers with half precision (16 bit).
3. The parameters of the analytical reflectance models are stored as floating point numbers (32 bit).

2.1.5 Realism and Methodology

It is crucial to get an idea of what we understand under the term *realism*. From [PH04]: *The goal of photorealistic rendering is to create an image of a 3D scene that is indistinguishable from a photograph of the same scene. [...] it is important to understand that in this context the word indistinguishable is imprecise because it involves a human observer, and different observers may perceive the same image differently. [...] accounting for the precise characteristics of a given observer is a very difficult and largely unsolved problem. For the most part, we will be satisfied with an accurate simulation of the physics of light and its interaction with matter, relying on our understanding of display technology to present the best possible image to the viewer.* Next to the mentioned difficulties when it comes to finding precise characteristics of human observers, in this thesis we have an even worse problem: we do not want to evaluate a rendering algorithm which may as well be applied to the modeling of an existing scene, but we want to *edit* materials, meaning we want to apply changes to the reflectance properties of a given material that leads to an appearance which *is not* realistic but which *could be* realistic. In our case a real-world pendant of the manipulated material-surface does not exist. So we cannot use reality as ground-truth.

Altogether we see three possibilities to evaluate the realism of a given scene:

1. determination of the quantitative difference from a photography of a real world pendant of the rendered scene by the use of an appropriate metric,
2. by a user study
3. or by validation through other experts like it is done in the peer-review process of a conference or a journal.

Comparison by the use of a metric, may it be radiometric, photometric or perceptual (see [MRC⁺86, FKH⁺18, LMS⁺19]), necessitates that the rendered image corresponds exactly to a photography which is in most cases not avail-

able. Furthermore different metrics may contradict each other and it is not per se clear, which metric to choose.

User studies are a strong but costly instrument to evaluate a research result and have to be designed carefully to avoid misleading interpretations [KHI⁺03].

So finally we depend on the judgement of other computer-graphics experts to decide whether an image is realistic or not.

2.2 Material

By *material* we mean the digital representation of an existing or imaginary material-surface represented as a description of the light exchange in every point of the surface.¹ The influence of the material is represented in the rendering equation (Equation 2.2) by the spatially varying reflectance term $\rho(x, \cdot, \cdot) : \Omega_i \times \Omega_o \rightarrow \mathbb{R}$. We confine ourselves to opaque materials so that $\Omega_{\{i,o\}} \subseteq \mathbb{S}$. In the next section we will discuss bidirectional reflectance distribution functions (Subsection 2.2.1) and bidirectional texture functions (Subsection 2.2.2). A last subsection is dedicated to an overview over material acquisition systems (Subsection 2.2.3).

2.2.1 Bidirectional Reflectance Distribution Functions

The foundational work for the definition of the BRDF has been developed during the 1960's by Nicodemus [Nic63,Nic65]. In 1977 Nicodemus introduced the concept, the definition and the nomenclature of the BRDF as part of a standardization process for the US-government [NRH⁺77, Eq. 9]. For the definition see Equation 2.3.

Here we do not want to give an exhaustive overview over existing BRDFs as this has been done before in many variations [MU12,NDM05] but we want to highlight those techniques which are essential for the understanding of the role of the BRDF for the rendering process and we want to introduce

¹from *Pushpins for Edit Propagation*.

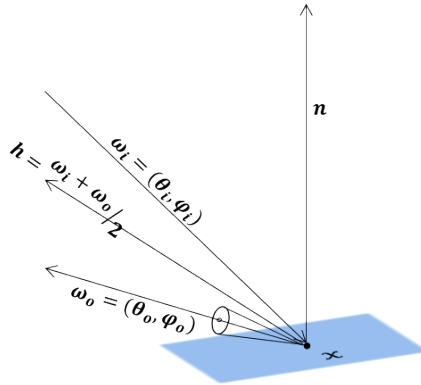


Figure 2.3: The parameters of a BRDF.

the BRDF-model which is used in this thesis, namely the Ashikhmin-Shirley-BRDF model.

BRDF A *BRDF* maps an incoming and an outgoing light direction onto a reflectance probability density. Being reflectance distributions, BRDFs are limited to the upper directional hemisphere.² As the radiant flux is part of the definition of the BRDF, it also depends on the radiation parameters frequency, phase and polarization [NRH⁺77]. The frequency, in the rendering process often represented by the colour channel, can of course not be neglected. Phase and polarization will not be subject of this thesis. And the materials will neither be fluorescing nor phosphorescent.

We concentrate on *analytical, measured* BRDFs, meaning, that an analytical reflectance model has been optimized to fit a given set of reflectance measurements. We tested our algorithms also on tabulated reflectance representations. But to describe those reflectance tables in order to make them applicable for usage with a classifier it is necessary to bring them into a comparable format like for example Rusinkiewicz-parametrization [Rus98] which makes a resampling-step necessary and

²from *Pushpins for Edit Propagation*.

to collect at least some elementary statistics.³ Investigations of this kind are beyond the scope of this thesis.

The Torrance-Sparrow reflectance model To explain the phenomenon of the off-specular peaks, Torrance and Sparrow published 1967 a reflectance model for isotropically roughened surfaces [TS67]. The authors describe those surfaces *as consisting of small randomly disposed mirror-like facets* [TS67, Page 1114].

Cook-Torrance model The microfacet-model, originally suggested by Torrance and Sparrow in 1967, has been further developed by Cook and Torrance in 1981 [CT81]. The Cook-Torrance reflectance model is strictly based on plausible physical assumptions.

$$\rho(\omega_i, \omega_o) = \frac{FDG}{\pi \langle \omega_i, \mathbf{n} \rangle \langle \omega_o, \mathbf{n} \rangle} \quad (2.4)$$

with the Beckmann-distribution D [BS63], the Fresnel-term F [SC78] and the geometric term G , which accounts for masking and shadowing.

Ward-model In his paper from 1992 [W⁺92], Ward presents an isotropic and an anisotropic BRDF-model. Here we may confine to the isotropic model, as this is the reflectance model first fitted by the linear light source reflectometer (see subsection 2.2.3-**SVBRDF acquisition**) [GTHD03].

$$\rho(\omega_i, \omega_o) = \frac{R_d}{\pi} + R_s \cdot \frac{1}{\sqrt{\cos \theta_i \cos \theta_o}} \cdot \frac{\exp\left(-\frac{\tan^2 \delta}{\alpha_s^2}\right)}{4\pi \alpha_s^2} \quad (2.5)$$

with the diffuse reflectance factor R_d , the specular reflectance factor R_s , the specular roughness α_s and the angle $\delta = \angle(\mathbf{n}, \omega_i + \omega_o)$.

Ashikhmin Shirley reflectance model The Ashikhmin-Shirley model [AS00] is a Phong-like model which additionally controls the eccentricity of the

³from *Pushpins for Edit Propagation*.

specular lobe and is given by:

$$\begin{aligned}
& \rho(\omega_i, \omega_o) \\
&= \frac{\sqrt{(n_x + 1)(n_y + 1)}}{8\pi} \frac{\langle \mathbf{n}, \mathbf{h} \rangle^{n_x \cos^2 \varphi + n_y \sin^2 \varphi}}{\langle \omega_i, \mathbf{h} \rangle \max(\langle \omega_i, \mathbf{n} \rangle, \langle \omega_o, \mathbf{n} \rangle)} \\
&\quad \cdot (R_s + (1 - R_s)(1 - \langle \omega_i, \mathbf{h} \rangle)^5) \\
&+ R_d(1 - R_s) \frac{28}{23\pi} \\
&\quad \cdot \left(1 - \left(1 - \frac{\langle \omega_i, \mathbf{n} \rangle}{2} \right)^5 \right) \left(1 - \left(1 - \frac{\langle \omega_o, \mathbf{n} \rangle}{2} \right)^5 \right) \quad (2.6)
\end{aligned}$$

for the incoming and outgoing directions ω_i and ω_o . The vector \mathbf{n} is the surface normal, $\mathbf{h} = (\omega_i + \omega_o) / \|\omega_i + \omega_o\|$ and φ is the azimuth of \mathbf{h} .

This model has four reflectance parameters: the wavelength dependent diffuse and specular reflectance shares R_d and R_s and the surface roughness along the x -axis n_x and the surface roughness along the y -axis n_y . In the following, we will refer to R_d and R_s as the *diffuse colour* and the *specular colour*. We will assume that those colours are RGB colours and the term *lightness* will refer to the HSL description of the RGB-space. We assume that the parameters are stored in rectangular maps.⁴

SVBRDF A *spatially varying BRDF (SVBRDF)* \mathcal{S} is a material where the light exchange is described by a BRDF

$$\mathcal{S} : A \times \Omega_i \times \Omega_o \rightarrow \mathbb{R}, \quad (2.7)$$

where A is the spatial domain, i.e., the extent of a material, and $\Omega_{i,o}$ are the space of all (incident) lighting and (outgoing) viewing directions, respectively. $\mathcal{S}(\mathbf{x}, \omega_i, \omega_o)$ is the amount of light scattered at point \mathbf{x} from direction ω_i into direction ω_o . The unit of the SVBRDF is $[\text{sr}^{-1}]$.

⁴from *Pushpins for Edit Propagation*.

2.2.2 Bidirectional Texture Functions

Our focus lies on advanced representations of spatially varying material appearance, and in particular the bidirectional texture function (BTF), which is capable of capturing a wide range of optical phenomena including nonlocal shading and volumetric effects⁵ produced by mesostructural material components.

The BTF is a digital representation of an opaque real-world material surface with the same domain and codomain as the SVBRDF (see equations 2.7 and 2.8). Dana sees the difference between BRDF and BTF in the scaling of the reflecting surface [DvGNK97]: where BRDFs see surfaces from a scale which is that coarse that texturing and local variations cannot be resolved by the acquisition device, the BTF is capable of describing local surface variations and texturing. That is the BTF holds a texture to each light-view-direction of a given sampling of the bidirectional domain.

So the BTF is a high-dimensional collection of sampled data and therefore not easily factorized into meaningful components like colour, reflectance, texture and mesostructure. Consequently, the development of decompositions and user interfaces for the editing of such material models is still an active field of research.⁶

BTF compression

In this subsection, we will describe a BTF-compression scheme which may be found in [KMBK03] (full matrix factorization) and in [Mül09] and is called *decorrelated full matrix factorization* (**DFMF**). We will make use of it in Chapter 4.

The BTF \mathcal{B} maps each surface position and each light and viewing direction

⁵from *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models*.

⁶from *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models*.

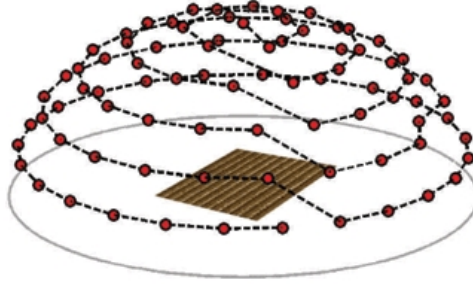


Figure 2.4: Directional sampling of the camera-dome. The light-sampling and the view-sampling differ only slightly (see Figure 2.6). From [FH09].

and each colour-channel to a reflectance value:

$$\mathcal{B} : X \times Y \times \Phi_{\Omega_i} \times \Theta_{\Omega_i} \times \Phi_{\Omega_o} \times \Theta_{\Omega_o} \times C \rightarrow R, \quad (2.8)$$

where X and Y are the spatial dimensions of the texture-function, Ω_i and Ω_o are the incoming and outgoing light-directions, C is the colour-domain with dimension K and R is the reflectance. Θ_{Ω} describes the domain of the elevation angle and Φ_{Ω} the domain of the azimuth angle. Equation 2.8 describes the same functional domain and codomain as Equation 2.7. As the support is a discrete sampling, we may interpret the BTF as a stack of textures, which we enumerate over all captured bidirections $(\varphi_i, \theta_i, \varphi_o, \theta_o)_m \in (\Phi_{\Omega_i} \times \Theta_{\Omega_i} \times \Phi_{\Omega_o} \times \Theta_{\Omega_o})$ and we may also enumerate the Pixel $(x, y)_n \in X \times Y$ and obtain:

$$\mathcal{B}_k((x, y)_m, (\varphi_i, \theta_i, \varphi_o, \theta_o)_n) =: \rho_{mnk}. \quad (2.9)$$

For a fixed colour-channel k , ρ_{mnk} is a matrix and we can apply a singular value decomposition to every channel k of it:

$$\rho_{mnk} = \sum_{i,j} U_{mik} \Sigma_{ijk} V_{jnk}^T. \quad (2.10)$$

The columns of the orthogonal base U_{mik} are the *eigen-textures*, and the rows

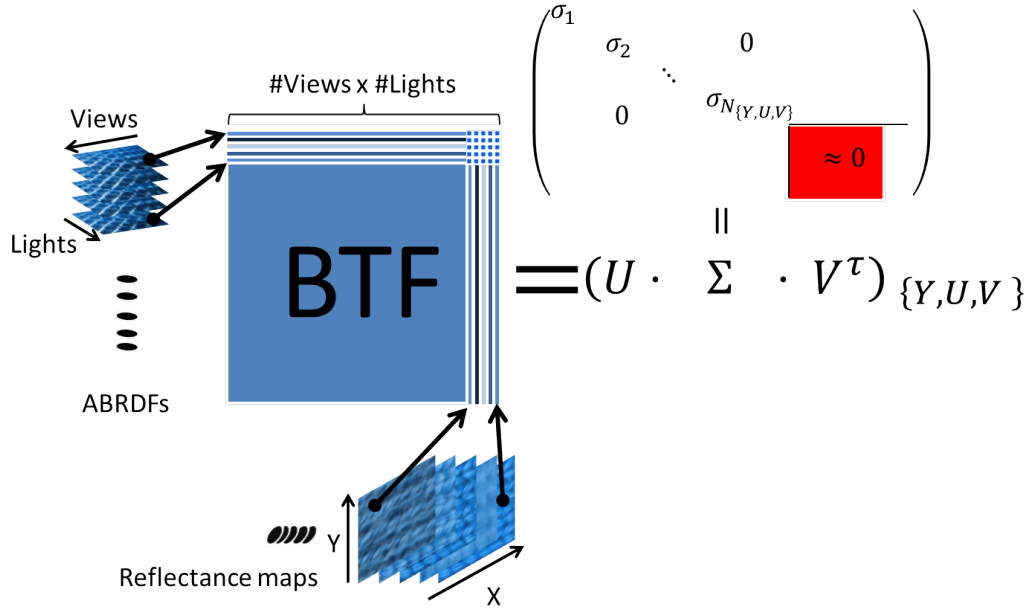


Figure 2.5: BTF-compression by the DFMF. The unwound, distorted and radiometrically calibrated images from the different light-view combinations are combined to a huge matrix. For the materials used in this thesis the directional half-sphere has been sampled equally for lights and views. The hemisphere sampling is illustrated in Figure 2.4.

of $C_{ijk}^\tau := \sum_l \Sigma_{ilk} V_{ljk}^\tau$ are called *eigen-apparent BRDFs* or *eigen-ABRDFs* (see Figure 2.5).

In the technical implementation of the BTF-acquisition device, described below, the foreshortened material images are rectified against the normal direction of the BTF.

In our setting, compression is reached by heuristically choosing a number N_k and cropping all base vectors with index greater N_k . We use the YUV-colour space and stored $N_{Y'} = 100$ and $N_{U'} = N_{V'} = 51$ components.

DFMF-compressed data may be decompressed by a simple linear combination of eigen-textures. Let us say we have a linear transformation which we want to apply to each texture, then, by linearity of the decompression step, it is possible to apply it to the eigen-textures before the decompression step,

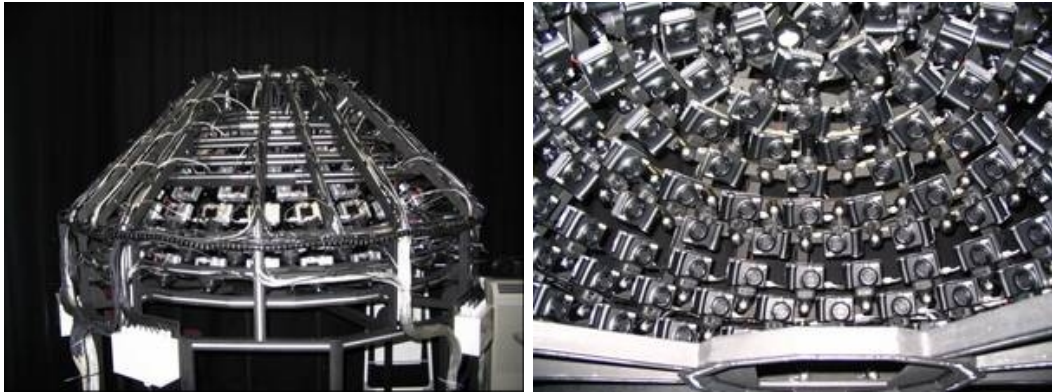


Figure 2.6: The camera dome. From [MK05].

instead. A BTF in our camera-dome-Sampling consists of about 22000 textures [MMS⁺05], whereas the base consists of 50 to 100 eigen-textures. It follows that applying the linear transformation to the eigen-base reduces the computational costs to less than half a percent of the original cost (compare Chapter 4).

2.2.3 Acquisition of Measured Materials

Here we want to describe two material acquisition devices. All materials we use in this thesis have been measured with those devices. The first device is the camera-dome from the university of Bonn to record BTFs and the second device is the linear light scanner in the version introduced by Meseth et al. [MHW⁺12] to record Ashikhmin-Shirley-SVBRDFs (see Subsection 2.2.1).

BTF acquisition

As described in the last subsection (Subsection 2.2.2) BTFs may be seen as pure texture stack. For the acquisition of this texture stack Dana et al. [DvGNK97] suggest to use a robot arm and a video camera. To allow for a faster acquisition of the BTF the university of Bonn developed an BTF-acquisition apparatus where the whole viewing hemisphere is registered in a single pass. This apparatus (Figure 2.6) has been suggested in [MMS⁺05] and

presented in [MBK05]. In this *camera-dome*, 151 cameras are assembled to a fixed sampling of the viewing hemisphere. As the light sources of this camera dome are the build-in flash lights of the cameras, the spherical view-sampling does differ only slightly from the light-sampling and there are exactly 151 light-directions. As the full BTF contains textures to all light-viewing-direction pairs, this sums up to $151 \times 151 = 22801$ textures.

All cameras are focussed onto a sample holder in the spherical centre of the dome. The post processing is described in detail in [MBK05] and in [SSW⁺14]. For a motivation of the post processing see also Subsection ??.

SVBRDF acquisition

There is a vast corpus of literature on BRDF-acquisition. We may distinguish between approaches to estimate one bidirectional reflectance distribution as done in the MERL database [MPBM03] or spatially varying bidirectional reflectance distributions. Here we want to present only the linear light scanner, presented in 2003 by Gardner et al. [GTHD03] and extended to anisotropic materials in 2012 by [MHW⁺12] et al. as the input data used in Chapter 3 have been acquired with this linear light source device. For further approaches to SVBRDF-acquisition, see Paragraph 2.3.2-**Analytical Reflectance Models**.

With his new anisotropic BRDF model, Ward suggested to use a specialized Gonioreflectometer to measure the BRDF-parameters. This approach has a couple of drawbacks:

1. The acquisition process is lengthy,
2. the equipment is relative expensive,
3. only one BRDF is measured.

Gardner et al. suggest a reflectometry apparatus made from of-the-shelf components, like LEGO[®], for about 1000,- \$ [GTHD03]: a neon bulb, the linear light source, is moved over a material probe while a fixed installed camera measures the reflected radiance. The reflectance data is fitted to the isotropic

ward model (Paragraph 2.2.1-**Ward model**, [W⁺92]). So the light scanner has to estimate the surface normal \mathbf{n} , the diffuse reflectance R_d , the specular reflectance R_s and the specular roughness α_s .

By the movement of the light source we obtain in each point of the material surface a sampling of the bidirectional domain. Additionally the linear light scanner estimates the surface displacement by the use of a laser stripe, parallel to the linear light source. The following restriction: $\angle(\omega_i, \omega_o) \approx \pi$ for this construction, so it cannot measure anisotropy in one pass. As the camera does not move, camera-pixels correspond directly with surface-elements of the measured material so BRDF-parameter estimation may be done on a pixel-wise base.

Meseth et al. improve this apparatus by placing the camera in the zenith above the material probe [MHW⁺12]. They ensure that the important angles directly under the light bulb are visible to the camera by the use of a half-silvered glass construction. Specular and diffuse reflectance components are isolated by the use of polarization filters. The system is used in two orthogonal directions in order to estimate the parameters of the anisotropic Ashikhmin-Shirley reflectance model (see Paragraph 2.2.1-**Ashikhmin-Shirley model**, [AS00]).

2.3 Material Editing

The goal of this dissertation is to find new ways of editing measured materials. Before we give an overview over former and current developments in this field, we want to introduce three criteria to assess image manipulations. Those criteria are based on our understanding of the term *realism* (see Subsection 2.1.5).

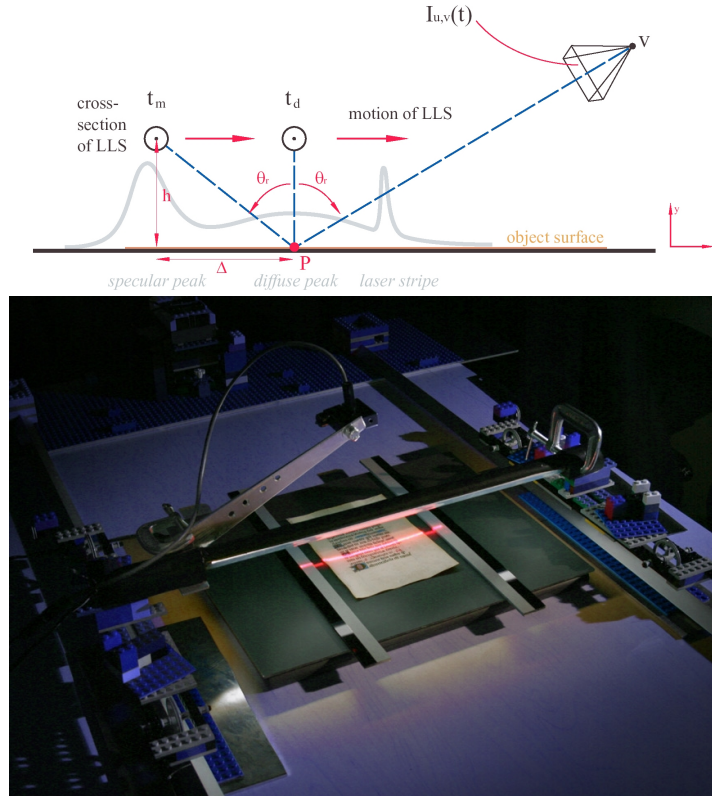


Figure 2.7: From [GTHD03]. Above the photography you can see a sketch of the linear light scanning device introduced by Gardner et al. The *linear light source (LLS)* is a white neon tube. It moves horizontally across the material probe with a fixed height h . To each pixel (u, v) there is a time 1. when the light bulb is exactly above the pixel (t_d) and 2. when the light bulb is at the mirror angle (θ_r) to the surface with respect to the camera (t_m). The grey curve shows the amount of light reflected by point P as function of the position of the LLS. The lower image shows a photo of the device while scanning an old book page. One can see the diffuse reflection and the laser-stripe to estimate the geometry.

2.3.1 Plausible, Predictable and Meaningful Editing Operations

We will call a manipulation *plausible* if we may easily produce realistic renderings from it. An edit is *predictable* if a user has an idea of the effect. In particular adequate scaling of parameters of analytical BRDFs may be considered as predictable. We will call a material-manipulation *meaningful* if it may be linked to a real world attribute.

2.3.2 State of Research

Editing of realistic material surfaces has an outstanding role in visualisation, in advertisement and in film making. This has a couple of different reasons:

1. Measuring reflectance properties in good quality is difficult and costly. So given a measured material, it is desirable to change some of the materials properties and keep others. Let us for example assume we measured a blue wool. We do not want to measure a green wool but would like to simply change the colour.
2. A director of a cine-film might wish to use materials which look plausible but simply do not exist like e.g. fantastic stones or metals.
3. A new material based on existing material-probes is planned on a computer. For example a furniture manufacturer wants to try a new leather embossment.

Here we want to summarize the state of research in general. Publications which are relevant for only one chapter will be referenced in the according chapter. In addition we give in each chapter a handful of tightly related publications to localize the research work of the chapter in the current state of the art. We want to divide the state of research into four paragraphs:

Editing on base of analytical reflectance models and general approaches to material editing. This is perhaps the area with the most previous work. Particularly Chapter 3 belongs here.

Editing of tabulated reflectance data. For our research this is the most important case because the acquisition and processing of BTFs are among the core capabilities of our research group.

Appearance Space Transformations. Recently transformations of the appearance space got much attention because neural network-based appearance transfer techniques allow for probability based descriptions of the appearance space. This terminology is misleading in so far as there does not exist a fixed definition of appearance space. In Chapter 5 we will introduce an own definition of such a space.

Edge aware imaging. Edge aware imaging has a special meaning for material editing because edges are extremely simple non-trivial features of textures. In Chapter 5 we learn an edge aware imaging operator.

Analytical reflectance models

In analytical reflectance models we have meaningful local parameters at hand like diffuse reflectance color or surface normal. The majority of prior work approximates material appearance in terms of analytical models. Those analytical models are mostly based on spatially varying BRDFs (SVBRDFs) that are associated with the geometry of the object surface. Following the pioneering work by Lensch et al. [LK01] to recover SVBRDFs in a practical way from a small number of input images, other researchers went on fitting similar models to BTF data [RK09, WDR11, WDR11, GTHD03, MHW⁺12]. There is a growing number of learning based algorithms to find SVBRDF parametrizations which represent real life materials based on a small number of photographs [AWL⁺15, AAL16]. Since SVBRDFs readily separate reflectance from geometry, they lend themselves to a number of editing techniques. Among

the approaches proposed are the transfer of reflectance functions from one measured material to another [ATDP11], the use of one-dimensional nonlinear appearance manifolds to simulate aging processes [WTL⁺06], as well as techniques using graph-based [PL07] and low-rank representations [AP08]. These are the works that we consider most representative for a larger body of prior work in SVBRDF fitting and editing. Due to the underlying model assumptions, all of them share similar problems with materials with a complex surface structure that may not adequately be represented by SVBRDFs. In 2015 Schmidt et al. [SPN⁺15] published a state of the art report for editing of Appearance, Lighting, and Material. The focus of this paper is on combined manipulation of illumination and material properties with the target to obtain a good final rendering result. ⁷

Table-based reflectance models

According to [DvGNK97], interpolated reflectance data may directly be used for rendering materials. Since the BTF describes material reflectance by its spatial and angular variation alike, it can be approached either as a collection of spatially varying angular reflectance distributions, or as a collection of textures that vary by angle.

Among the image based approaches, there is the work of Pellacini and Lawrence, who suggested, to use an k-nearest neighbour graph to construct a sparse adjacency matrix [PL07]. An and Pellacini made another step in this direction with AppProp [AP08], which has been extended to tabulated reflectance data by Xu et al. [XWT⁺09]. Müller et al. [MSK07] presented a texture synthesis approach for BTFs that guides the placement of local features using a given mesostructure constraint. Ruiters et al. present an interpolation scheme which is based on a separation of the BTF into a heightmap and a parallax compensated BTF [RSK13]. More recent publications [HLLC17] concentrate on the usability of the editing scheme. An algorithm that is particularly

⁷from *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models*.

suited for repetitive textures was introduced by Haindl and Hatka [HH05] with their BTF Roller.

Haindl and Havlíček [HH17] suggest a stochastic process based BTF-model. For the compression of BTFs (see Paragraph 2.2.2-**BTF compression**), we refer the reader to the exhaustive state-of-the-art survey by Filip and Haindl [FH09].

The perspective on the BTF as stack of ABRDFs has been taken up by Kautz et al. [KBD07] who showed that transferring operators from picture editing to the spatial or the angular domain of a BTF may bring reasonable results.⁸

Appearance Space transformations

A strong impulse into the direction of appearance transfer had been given by AppWarp, an algorithm introduced by An et al. in 2011 [ATDP11]. AppWarp is an algorithm to *retarget* measured materials, meaning to use the spatial pattern of one material and the set of reflectance distributions of another material. This is done by an appearance space warping. Here the author means by *appearance space* the span of the BRDF-parameters.

Wang et al. suggest to use of one-dimensional nonlinear appearance manifolds to simulate aging processes [WTL⁺06].

Appearance transfer has recently gathered increasing attention because neural networks allow to estimate reflectance properties of one indicated and masked object and to transfer it to another indicated and masked object in one step [LCY⁺17].

Yamada et al. investigated how properties like roughness and glossiness are represented by wood eigen-textures [YHH17], that is they tried to tie attributes like glossiness to a particular eigen-texture in the principal component analysis of a stack of wood textures. A similar but more general approach has been presented by Serrano et al. [SGM⁺16]. Serrano et al. calculated a PCA on a

⁸from *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models*.

base for MERL-formated [MPBM03] BRDFs. Based on user-studies they learn functionals from the first five components of this PCA-based representation of given materials into four different attributes. Editing may be done by solving a minimization problem to find the coefficients for a given desired attribute-value. Malpica et al. [MBG⁺17] refine this approach by using two different PCA-representations.

Zsolnai-Fehér et al. published a system which learns the user preferences for materials directly and translates these learned preferences into a parametrization of a principled shader [ZFWW18].

Edge aware imaging

Edge aware imaging has been subject of intensive study during the last years. Bilateral filters [TM98, PKT⁺09] are among the most popular tools for edge-aware image processing. One recent approach gives a linear approximation of a bilateral filter [NPC17]. He et al. suggest to improve the edge-preservation property of filters by the use of a guidance map [HST13]. Paris et al. use Laplacian Pyramids for strengthening or weakening edges in images: they argue, that edges are a jump in not only one level in the laplacian pyramid but merely in all levels [PHK11]. Laparra et al. use those insides to build a system for perceptually optimized image rendering [LBBS17]. Fattal [Fat09] detects edges by the use of second generation wavelets. Using Gabor-filters for edge detection has a long tradition, e.g. [MNR92].⁹

2.4 Analysis of the Frequency Spectrum

Exploration of the frequency domain in order to allow for plausible material manipulations is the core topic of this thesis. Here we want to introduce integral transforms and filter kernels suitable to investigate the frequency domain. Note that the term frequency domain is not unambiguous. To distinguish

⁹from *Linear Subspaces of the Appearance Space*.

between functions before and after a transformation step, we will refer to the function space as the time-domain for functions in $\mathbb{R} \rightarrow \mathbb{K}$ or the space-domain for functions in $\mathbb{R}^2 \rightarrow \mathbb{K}$ and the frequency domain is the image of all allowed functions under the given frequency or scaling transformation. Note that this distinction may become a little bit awkward for the description of processing pipelines which concatenate more than one frequency- or scaling-analysis-step without recurring into the original modelling context like for example the hierarchical analysis of signals in neural networks.

2.4.1 The Fourier Transform

Joseph Fourier published the *Theorie analytique de la chaleur* after significant struggle in 1822 [Fou22]. Given a real-valued function f , we decompose f into an even and an odd part:

$$f_e(x) := (f(x) + f(-x))/2 \quad \wedge \quad f_o(x) := (f(x) - f(-x))/2. \quad (2.11)$$

Fourier delivers - with slightly different notation - the following reconstruction equations for f_e and f_o [Fou22, Paragraph 361 - Chap.IX Sect.I]:

$$f_o(x) = \frac{2}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} f(a) \sin(qa) da \sin(qx) dq. \quad (2.12)$$

$$f_e(x) = \frac{2}{\pi} \int_0^{\infty} \int_{-\infty}^{\infty} f(a) \cos(qa) da \cos(qx) dq. \quad (2.13)$$

Note that

$$\int_{-\infty}^{\infty} f(a) \sin(qa) da = 2 \int_0^{\infty} f_o(a) \sin(qa) da.$$

The Fourier transform in modern notation Today we use the Fourier transform in the following form: for

$$f \in L^1(\mathbb{R}^n) \quad (2.14)$$

$$\hat{f}(\nu) := \mathbf{F}f(\nu) := \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} f(t)e^{-i\langle \nu, t \rangle} dt \quad (2.15)$$

where the integral-operator \mathbf{F} is realized by a Lebesgue-integral.

The Fourier inversion theorem states, that if f and $\hat{f} \in L^1(\mathbb{R}^n)$, then:

$$\hat{\hat{f}}(x) = f(-x). \quad (2.16)$$

This motivates the definition of the Fourier-inverse (\mathbf{F}^{-1}):

$$\mathbf{F}^{-1}\hat{f}(\nu) = \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \hat{f}(\nu)e^{i\langle \nu, t \rangle} d\nu. \quad (2.17)$$

The convolution between f and k : The convolution between two functions $f, k : \mathbb{R}^d \rightarrow \mathbb{K}$ is defined by

$$(f * k)(x) := \int_{\mathbb{R}^n} f(\tau)k(x - \tau)d\tau. \quad (2.18)$$

and is again a function $f * k$ over \mathbb{R}^d if the integral is well-defined. The function k is often referred to as the convolution kernel.

The convolution theorem The convolution theorem states a very important connection between the convolution in the time domain and point-wise multiplication in the frequency domain:

$$\mathbf{F}(f * k)(\nu) = \mathbf{F}f(\nu) \cdot \mathbf{F}g(\nu) \quad (2.19)$$

For two matrices F and G we obtain:

$$(F * G)_{ij} = \sum_{kl} F_{kl}G_{i-k, j-l}. \quad (2.20)$$

In this thesis, higher mode tensors, like coloured images, are convolved matrix wise along the first two dimensions.

An *image channel* is a matrix. As there are also eigen-image channels with negative values, we will not restrict the value domain to a limited positive range. An *image* is an ordered set of image channels or alternatively a mode-three tensor. In the course of this thesis, the height will be the first dimension. Our use of the word *texture* corresponds to the TextureImage-class of the pbrt-project [PH04]. Operations on textures are identified with the corresponding operations on the image matrix. Depending on the context, we will refer to an image-channel as two dimensional signal or as discrete function.

The discrete Fourier transform

Trigonometric series had been known before Fourier's *Theory of Heat*. Fourier developed his integral-transform from those trigonometric series. Altogether we have four different kinds of Fourier-transformations:

1. Aperiodic continuous signal, continuous, aperiodic spectrum
2. Periodic continuous signal, discrete aperiodic spectrum
3. Aperiodic discrete signal, continuous periodic spectrum
4. Periodic discrete signal, discrete periodic spectrum

The classical Fourier-series are of the second type and the Fourier transform (Equation 2.15) is of the first type. In this thesis, we deal with discretised data, meaning that both, the time-signal as well as the frequency spectrum are discrete.

The one-dimensional discrete Fourier transform

For a given sequence $(a_k)_N \in \mathbb{C}^n$, the *discrete Fourier transform* (**DFT**) is defined as:

$$\mathbf{F}a_k = \sum_{n=0}^{N-1} a_n e^{-\frac{2\pi i n k}{N}}. \quad (2.21)$$

where we used the same symbol \mathbf{F} for the Fourier operator for the discrete case as we have used in the continuous case. The inverse transform is then

$$\mathbf{F}^{-1}a_n = \frac{1}{N} \sum_{k=0}^{N-1} a_k e^{\frac{2\pi i n k}{N}}. \quad (2.22)$$

This definition is in accordance with the DFT-implementation of Matlab[©] and the `fftw`.

The two-dimensional Fourier transform

The discrete two-dimensional Fourier transform is given by:

$$\mathbf{F}a_{k,l} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} a_{m,n} e^{-2i\pi\left(\frac{mk}{M} + \frac{nl}{N}\right)} \quad (2.23)$$

The fast Fourier transform

The fast Fourier transform is an important algorithm because it reduces the computational complexity of the discrete Fourier transform from $\mathcal{O}(n^2)$ to optimally $\mathcal{O}(n \log n)$. This is particularly interesting in combination with the convolution theorem: it is possible, to convolve two images with edge length n_e in $n_e^2 \log n_e^2$ instead of n_e^4 steps. This is possible because of the periodic character of the trigonometric functions combined with the regularity of the points of support of the discretization: the Cooley and Tukey algorithm decomposes the Fourier transform into a Fourier transform of the even indexed elements and a Fourier transform of the odd indexed elements of the sequence.

$$\mathbf{F}a_k = \sum_{n=0}^{N-1} a_n e^{-\frac{2\pi i n k}{N}} = \sum_{m=0}^{N/(2-1)} a_{2m} e^{-\frac{2\pi i m k}{N/2}} + \left(\sum_{m=0}^{N/(2-1)} a_{2m+1} e^{-\frac{2\pi i m k}{N/2}} \right) e^{-\frac{2\pi i k}{N}}. \quad (2.24)$$

By recursively applying this rule and counting steps, we obtain the complexity of $\mathcal{O}(n \log n)$.

2.4.2 The Wavelet Transform

This thesis is about image transformations. When we talk about an *edit*, we mean a set of scalable image transformations, which contains at the very least the identity. So, what we want to do is, we want to change to a frequency description of a given texture, do appropriate manipulations and change back to the spatial domain. Convolution with a Gabor-kernel produces a frequency description of a given image but we cannot recover the image from the filter response. One way to obtain a localized frequency description is, to define a wavelet transform.

Definition A function $\psi \in L^2(\mathbb{R})$ is admissible if:

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\nu)|^2}{|\nu|} d\nu < \infty. \quad (2.25)$$

It follows:

$$\hat{\psi}(0) = 0, \quad (2.26)$$

and

$$\int_{-\infty}^{\infty} \psi(\nu) d\nu = 0. \quad (2.27)$$

ψ may be interpreted as mother wavelet. From this mother wavelet we may derive a family of wavelets:

$$\psi_{s,r}(t) = \frac{1}{\sqrt{s}} \cdot \psi\left(\frac{t-r}{s}\right).$$

Define

$$\mathbf{W}f_{s,r} := \langle f, \psi_{s,r} \rangle = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \overline{\psi}\left(\frac{t-r}{s}\right) dt. \quad (2.28)$$

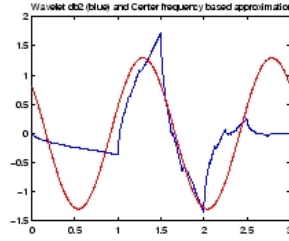


Figure 2.8: Comparison between scale and frequency, based on the result of the Matlab function `scal2frq`.

By $\bar{\psi}$ we denote the complex conjugate of ψ . If $\{\psi_{s,r} : s, r \in \mathbb{Z}\}$ is a Hilbert Base, we call ψ an orthonormal wavelet. With $\psi_{s,r}$ being a Hilbert Base the following formula holds:

$$\forall f : \quad f = \sum_{s,r} \langle f, \psi_{s,r} \rangle \psi_{s,r}. \quad (2.29)$$

Scale and Frequency

While it is not problematic from a purely heuristic point of view to speak of the frequency of a wavelet, we have to point out, that the term *frequency* is usually reserved for correlation with the trigonometric functions. The choice of the scaling-factor s has an obvious influence on the granularity of the structure which $\psi_{s,r}$ correlates with. Here we speak of the *scale* of the wavelet. In general, there is no canonical mapping between scale and frequency. But the pseudo-frequency of a wavelet may be declared as the frequency of the appropriately translated and scaled sine-wave best fitting the wavelet in the interval $[-\pi, \pi]$.

The Discrete Wavelet Transform

From equation 2.29 it follows, that it is in principle possible, to manipulate certain scales from a wavelet analysis and reconstruct the parameter map subsequently, meaning that wavelets allow in principle manipulations confined more or less exactly to a certain frequency band. Let us now assume that:

$m := a_0^{-j \in \mathbb{Z}}$ and $n := b_0 k \in \mathbb{Z}$ for some fixed values a_0 and b_0 . We then consider the wavelet at position n and scale m :

$$\frac{1}{\sqrt{a_0^j}} \psi(a_0^{-j} x - b_0 k). \quad (2.30)$$

We will discretize the wavelet transform in two different contexts:

1. For the definition of orthogonal wavelet families in order to apply the fast wavelet transform as substitute for the Fourier transform. The wavelet transform has the advantage that it allows localisation of the image manipulations. In this setting $a_0 = 2$ and $b_0 = 1$.
2. For the Gabor-wavelets in order to have a multi-scale analysis based on a filtering system with high similarities to the receptive fields profiles in the visual cortex.

In the following paragraphs we will collect the theoretical principles of those two points.

Multi Resolution Analysis

A multi resolution analysis [Mal89] is a sequence of nested, closed subspaces of $L^2(\mathbb{R})$:

$$\{0\} \cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots \subset L^2(\mathbb{R}), \quad (2.31)$$

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}), \quad (2.32)$$

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad (2.33)$$

$$f \in V_j \iff f(2^j \cdot) \in V_0 \quad \text{and} \quad (2.34)$$

$$f \in V_0 \implies f(\cdot - k) \in V_0. \quad (2.35)$$

Equation 2.34 assures, that all spaces are scaled versions of V_0 . We require, that there is a function ϕ , such that $\{\phi_{0k}\}_{k \in \mathbb{Z}}$ is a Riesz-base of V_0 :

$$V_0 = \text{span}(\{\phi_{0k}\}_{k \in \mathbb{Z}}), \quad (2.36)$$

with $\phi_{jk}(x) := \sqrt{2^{-j}}\phi(2^{-j}x - k)$. Then it follows that

$$V_j = \text{span}(\{\phi_{jk}\}_{jk \in \mathbb{Z} \times \mathbb{Z}})$$

and there are h_k , such that:

$$\phi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} h_k \phi(2x - k). \quad (2.37)$$

Given such a sequence of closed subspaces, then there exists an orthonormal wavelet basis $\{\psi_{jk}\}$ of the complement $W_j : V_{j-1} = V_j \oplus W_j, V_j \perp W_j$ such that:

$$\forall f \in L^2(\mathbb{R}) \quad P_{j-1}f = P_j f + \sum_{k \in \mathbb{Z}} \langle f, \psi_{jk} \rangle \psi_{jk}, \quad (2.38)$$

with P_j an orthogonal projection on V_j [Dau93]. W_j is called *detail space*. It follows that $\{\psi_{jk}\}$ is an orthonormal basis of $L^2(\mathbb{R})$. The wavelet may be derived from the scaling-coefficients h_k (for further information see [LMR94]):

$$g_k = (-1)^k h_{1-k}, \quad (2.39)$$

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \phi(2x - k). \quad (2.40)$$

And $\psi_{j,k}(x) := \sqrt{2^{-j}}\psi(2^{-j}x - k)$.

The Fast Wavelet Transform

From the scaling-function and the wavelet function, we derive a discrete low-

pass filter function and a discrete high pass filter function. Define:

$$d_{j,k} = \langle f, \psi_{j,k} \rangle, \quad d_j := \{d_{j,k}\}_{k \in \mathbb{Z}} \quad (2.41)$$

$$s_{j,k} = \langle f, \phi_{j,k} \rangle, \quad s_j := \{s_{j,k}\}_{k \in \mathbb{Z}} \quad (2.42)$$

then

$$d_{j,k} = \langle f, \psi_{j,k} \rangle = \sum_{i \in \mathbb{Z}} g_i \langle f, \phi_{j-1, 2k-i} \rangle = \sum_{i \in \mathbb{Z}} g_{i-2k} s_{j-1, i} \quad (2.43)$$

$$s_{j,k} = \langle f, \phi_{j,k} \rangle = \sum_{i \in \mathbb{Z}} h_i \langle f, \phi_{j-1, 2k-i} \rangle = \sum_{i \in \mathbb{Z}} h_{i-2k} s_{j-1, i} \quad (2.44)$$

Note that we do not need the function f itself, anymore, but may apply the algorithm recursively to the starting sequence $s_{0k} := \langle f, \phi_{0k} \rangle$. This interprets the discrete signal s_0 as the low pass filtered version of a signal with a higher resolution. Finally we define the linear filters $\mathbf{H}, \mathbf{G} : l^2(\mathbb{Z}) \rightarrow l^2(\mathbb{Z})$ by:

$$(\mathbf{H}s)_k := \sum_{i \in \mathbb{Z}} h_{i-2k} s_i \quad (2.45)$$

$$(\mathbf{G}s)_k := \sum_{i \in \mathbb{Z}} g_{i-2k} s_i \quad (2.46)$$

Note that for a finite signal $s \in \mathbb{R}^m$ every filtering step reduces the number of coefficients by the factor of two. This leads to the characteristic downsampling scheme (Figure 2.9).

The synthesis works exactly the opposite but with zero-padding in the up-sampling step.

Extension to images

For the two-dimensional case we may use the same argumentation as in the one-dimensional case. It is assumed that both, the scaling function Ψ and the wavelet function Φ , are separable: $\Psi(x, y) = \psi(x)\psi(y)$ and $\Phi(x, y) = \phi(x)\phi(y)$.

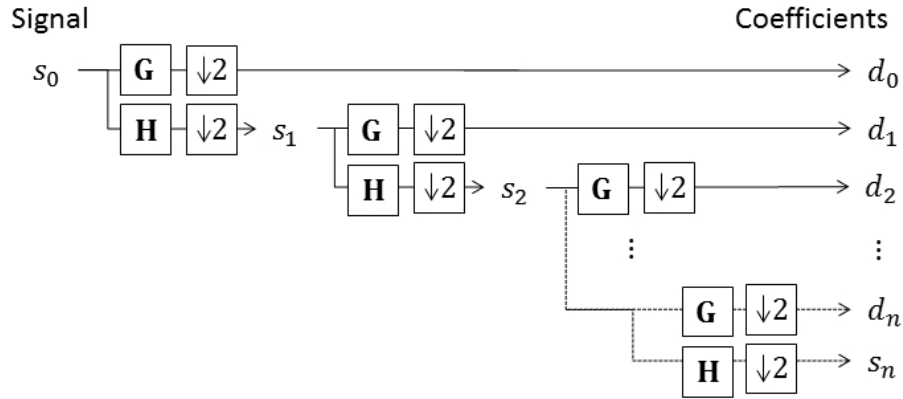


Figure 2.9: The downsampling scheme of the fast wavelet transform. Subsampling ($\downarrow 2$) is part of the definition of \mathbf{H} and \mathbf{G} .

This results in the known coefficient scheme (see Figure 2.10):

$$s_{j,(m,n)} = \langle f, \phi_{j,m} \phi_{j,n} \rangle \quad (2.47)$$

$$d_{j,(m,n)}^1 = \langle f, \psi_{j,m} \psi_{j,n} \rangle \quad (2.48)$$

$$d_{j,(m,n)}^2 = \langle f, \phi_{j,m} \psi_{j,n} \rangle \quad (2.49)$$

$$d_{j,(m,n)}^3 = \langle f, \psi_{j,m} \phi_{j,n} \rangle \quad (2.50)$$

with $(n, m) \in \mathbb{Z} \times \mathbb{Z}$.

Daubechies Wavelets

In her work, *Ten Lectures on Wavelets*, [Dau92], Ingrid Daubechies presents us with her famous orthogonal *wavelet family*. Daubechies-wavelets are characterized by vanishing moments and compact support.

The wavelet function ψ has A vanishing moments means that for every polynomial $p(x) = \sum_{n \in \mathbb{N} < A} a_n x^n$:

$$\int_{-\infty}^{\infty} p(t) \psi(t) dt = 0, \quad (2.51)$$

meaning that all polynoms with degree less than A may be completely repre-

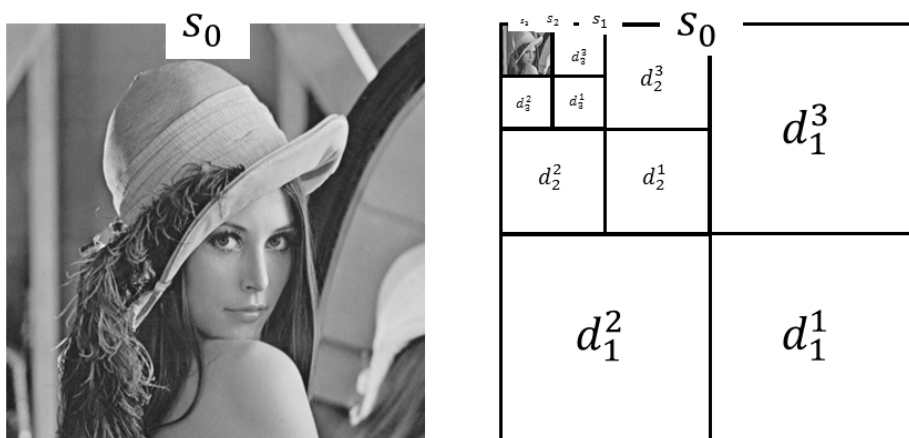


Figure 2.10: The downsampling scheme for two dimensional wavelets.

sented by the coefficients of the scaling filter.

This equation holds also for the discrete case. So with an increasing number of vanishing moments the wavelets are capable of representing polynomial signals with increasing degree. This is a nice feature for compression but we are merely interested in the frequency response. The length of the non-zero wavelet sequence is twice the number of vanishing moments. In fact the frequency response of ψ and ϕ approximates with increasing number of vanishing moments increasingly a bandpass and a low-pass filter [LMR94, Lemma 2.4.31]. So by choosing an appropriate number of vanishing moments, we may apply a band-pass filter by raising or reducing the coefficients of the corresponding wavelet-coefficient-octave (see Figure 2.10).

We obtain the two dimensional daubechies filter db6 with 6 vanishing moments, by rotating the one-dimensional filters (see Figure 2.11) around the z -axis (see Figure 2.12).

The Gabor wavelet

As explained in the last section, the *Fourier-integral considers phenomena in an infinite interval, sub specie aeternitatis*, meaning that "changing frequency" becomes a contradiction in terms. To circumvent this problem, Dennis Ga-

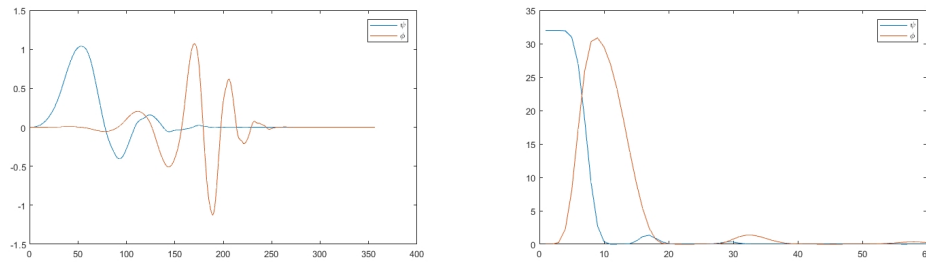


Figure 2.11: Scaling and wavelet function for Daubechies db6 wavelet. The right plot shows the frequency response, which approximates a low-pass and a band-pass filter.

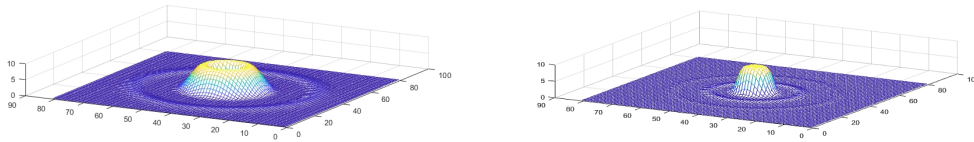


Figure 2.12: Frequency response of the rotated scaling and wavelet function for Daubechies db6 wavelet.

bor introduces Gabor functions as "*elementary signals*" which are constructed as *harmonic oscillations modulated by a "probability pulse"* in time [Gab46]. Those elementary signals have a limited life-time and allow therefore to compose signals with changing frequencies.

The uncertainty principle

Based on the results of quantum mechanics, Dennis Gabor develops Gabor signals $\gamma(t)$ to describe atomic information portions in signal-processing. To be more precise, Gabor signals are the only signals $s(t)$ for which the inequality

$$\left| \int_{-\infty}^{\infty} ts(t)dt \right| \left| \int_{-\infty}^{\infty} \nu \mathbf{F}s(\nu)d\nu \right| \geq \frac{1}{2} \|s\|_2^2 \quad (2.52)$$

is an equality. It is hardly imaginable to investigate properties of localized frequency- or scaling-analysis without mentioning the uncertainty principle [Küp24, Hei27, Nyq24], first published by Karl Küpfmüller in 1924. And there has been considerable work during the 80s and 90s on mathematical prop-

erties of filter responses in visual cortex (see e.g. [Dau85]), which also lead to a vivid academic debate. [SW90, Yan92]. While interesting from a general point of view, those properties are only in so far relevant for our work as the receptive fields recorded from neurons have similarities with two-dimensional Gabor filters. Whether the uncertainty principle plays a role in the development of those receptive fields is to our knowledge an unanswered question.

Certainly the most prominent adaptation of the Gabor functions from one dimension to two dimensions stems from John Daugman [Dau85] and has been published in 1985. Daugman introduces a two dimensional pendant of Gabor's signal atoms and shows that those two dimensional Gabor functions also fulfil an uncertainty principle. He fits those functions to empirically gained response rates of the retinal receptive fields from cats striate cortex. In other words: two dimensional Gabor functions have explicitly been developed to describe the cortical neural response to optical stimuli of the retina. This is exactly the context we use them in here.

The one dimensional Gauss-function

The probability pulse, mentioned by Gabor, is a Gauss-function:

$$G_{\sigma,\mu}(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \quad (2.53)$$

with the mean μ and the standard deviation σ .

The one-dimensional Gabor filter

According to [Gab46, Sec. 5, eq. 1.27], we declare the one dimensional Gabor base function as the product of a harmonic oscillation and a probability pulse in form of a Gaussian. The Gabor base function is the product of a wavefunction, called *carrier*, and a Gaussian *envelope*:

$$\gamma_{\sigma,\mu,\nu,\beta}(t) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} e^{i(2\pi\nu t + \beta)} \quad (2.54)$$

with the frequency ν and the phase angle β .

The two-dimensional Gaussian filter

It is convenient, to introduce the two dimensional filter-functions by starting with a transformation of the euclidean plane:

$$R_\varphi \circ T_{\mathbf{p}}(x, y) = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x - p_x \\ y - p_y \end{pmatrix} \quad (2.55)$$

with the point $\mathbf{p} := \begin{pmatrix} p_x \\ p_y \end{pmatrix}$ and the rotation angle φ . We define the local coordinates (u, v) :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} (x - p_x) \cos \varphi - (y - p_y) \sin \varphi \\ (x - p_x) \sin \varphi + (y - p_y) \cos \varphi \end{pmatrix}. \quad (2.56)$$

To get an eccentric gaussian filter, we multiply two gaussian filters with different standard deviations:

$$G_{\sigma_u, \sigma_v} = \frac{1}{2\pi\sigma_u\sigma_v} e^{-\frac{1}{2}\left(\left(\frac{u}{\sigma_u}\right)^2 + \left(\frac{v}{\sigma_v}\right)^2\right)}. \quad (2.57)$$

We will use the symbol \mathcal{G} for the discretized filter kernel.

The two-dimensional Gabor filter

Then we obtain for the two dimensional Gabor filter:

$$\gamma_{\nu, \sigma_u, \sigma_v, \beta}(u, v) = \frac{1}{2\pi\sigma_u\sigma_v} e^{-\frac{1}{2}\left(\left(\frac{u}{\sigma_u}\right)^2 + \left(\frac{v}{\sigma_v}\right)^2\right)} e^{i(2\pi\nu u + \beta)}. \quad (2.58)$$

Note that we assumed that the wave vector of the carrier is parallel to the u -axis.

Daugman derived the two-dimensional Gabor filter function by merging his own work [Dau80], where he introduced two dimensional response functions to model the perceptive field profiles, with the work of Marcelja [Mar80], who used one-dimensional Gabor filter functions to describe the perceptive field.

Two dimensional Gabor Wavelets

It is possible, to use a Gabor-filter function as mother wavelet. Tai Sing Lee presented an approach in 1996 by using the mother wavelet:

$$\psi^\gamma(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2+y^2)} \cdot \left[e^{i\kappa x} - e^{-\frac{\kappa^2}{2}} \right] \quad (2.59)$$

where the admissibility is reached by subtracting the mean $e^{-\frac{\kappa^2}{2}}$ from the wave function [Lee96]. κ is a non-linear function of the half-amplitude bandwidth. Note that rotations and translations are part of the parameter-set of his wavelet frame and are therefore no part of the mother-wavelet definition. It is yet not possible, to construct an orthogonal (equation 2.29) Gabor wavelet. But we may reconstruct an image from a Gabor-analysis, if the analysing function system forms a *tight frame*. A *frame* is a family of functions ψ_P , so that there exist two constants, the *frame bounds*, $0 < A \leq B < \infty$ and the following formula holds:

$$\forall f : A\|f\|^2 \leq \sum_{i \in I} |\langle f, \psi_i \rangle|^2 \leq B\|f\|^2. \quad (2.60)$$

If the frame bounds A and B are close to each other, then:

$$\forall f : \quad f \approx \frac{2}{A+B} \sum_{i \in I} \langle \psi_i, f \rangle \psi_i, \quad (2.61)$$

in which case we call the frame a *tight frame*.

Note that we may neither make use of the performance gain, provided by a multiscale analysis nor can we profit from the storage scheme, available from the downsampling step (see figures 2.9, 2.10). Gabor wavelets are resource-hungry.

2.5 Systems

Experiments have been done on three different architectures

ID	Name	# cpu (kernel/threads)	CPU clock rate [GH]	RAM [GB]
A	i5-2500	4/4	3.3	8
B	Pentium Dual Core E6600	4/4	3.066	8
C	i7-4770	4/8	3.4	8

CHAPTER 3

An Application of the Convolution Theorem

In this chapter we want to present an approach to segregating user indicated editable foreground patterns against the uneditable background in measured SVRBDF-materials bearing at least one near regular textured channel. The content of this chapter is based on a peer-reviewed publication

Marlon Mylo, and Reinhard Klein. Pushpins for edit propagation. In *International Conference on Computer Graphics, Visualization and Computer Vision*, volume 25 of *WSCG proceedings*, pages 143 – 152, [MK17]. [MK17]).

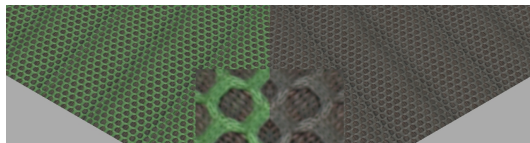


Figure 3.1: Pushpins for edit propagation.

All self-citations in the current chapter refer to this publication. Our problem is closely related to the matting problem where we seek to find an opacity

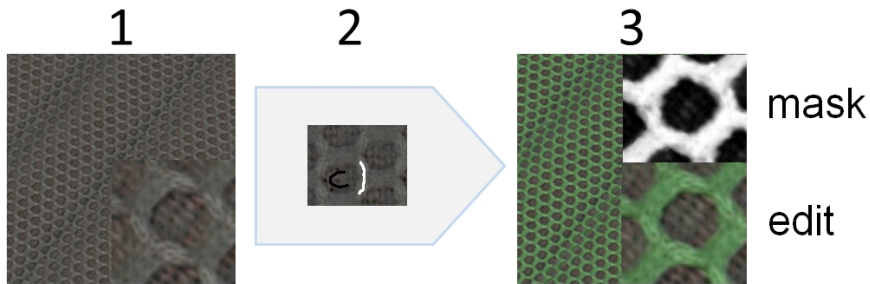


Figure 3.2: Given a near regular textured SVBRDF-material (1) and a stroke input (2). Our algorithm produces reliably an editing mask (3-mask) so that edits can be confined to the foreground material (3-edit).

value α so that the colour of each pixel C may be represented as convex-combination of a foreground colour F and a background colour B :

$$C = \alpha F + (1 - \alpha)B \quad (3.1)$$

but in our context, α should be merely interpreted as share of the texel surface area. The presented workflow contains an application of a convolution based technique, we call *pushpins*. While the whole workflow is new, our focus lies on the development of the pushpins. Those pushpins make use of the frequency spectrum in two different ways:

1. the localized frequency information is used to identify similar structure (see 2.4.2) and
2. the convolution theorem in connection with the fast Fourier transform enables fast identification of those structures (see ?? and 2.4.1).

So in this chapter we do not manipulate the spatial frequency spectrum itself but use it as means to propagate a given stroke input. The contribution of our approach to the realism of the final edit depends on an exact segregation of foreground and background. In contrast to that the following two chapters

(Chapter 4 and Chapter 5) present approaches, where the influence of manipulations of the frequency spectrum on the visual perception are explored.

The technical description of pushpins will be given in Section 3.1. This technical description is followed by an overview (Section 3.2) and a detailed description of the single steps of the presented workflow (Section 3.3). Afterwards we will describe the test setup, the competing algorithms and the results (Section 3.3.7). To obtain a deeper understanding of the mechanisms of pushpins, we will suggest as future work to investigate our technique with the focus on its applicability in the field of signal processing (Section 3.5).

In this chapter we present an approach to propagating a given bipartite stroke input. This is closely related to the propagation algorithms of An and Pellacini [PL07, AP08] and to the BTF-version of AppProp by Xu et al. [XWT⁺09]. With Deepprop Endo et al. present a deep learning based approach to edit propagation [EIKM16].

3.1 PushPins

Masking the frequency domain in order to isolate particular features is a common technique in signal processing but we did not find our approach in the computer graphics literature so we will briefly introduce it. The target is, to highlight all positions in a near regular textured image-channel \mathcal{T} which have a high similarity with a given patch $\mathcal{P} \subset \mathcal{T}$. It is important to note that we want a seedless approach in order to be robust with respect to energy barriers within the texture \mathcal{T} .

Lets first assume that we are looking for a texture \mathcal{X} with the same size as \mathcal{P} so that:

$$\mathcal{P} * \mathcal{X} = \delta$$

where δ models a spike in form of the dirac distribution. By convolution and

application of the convolution theorem, we get:

$$\mathbf{F}\mathcal{P}\mathbf{F}\mathcal{X} = \mathcal{C} \quad (3.2)$$

for a constant texture \mathcal{C} . \mathbf{F} is the fourier-transform and \mathbf{F}^{-1} its inverse. Thus, a candidate for \mathcal{X} is:

$$\mathcal{X} = \mathbf{F}^{-1} \left(\frac{\mathcal{C}}{\mathbf{F}\mathcal{P}} \right) \quad (3.3)$$

so

$$\mathcal{P} * \mathbf{F}^{-1} \left(\frac{\mathcal{C}}{\mathbf{F}\mathcal{P}} \right) = \delta \quad (3.4)$$

Here we presumed correct scaling and frequency sampling and point wise multiplication.

For numerical reasons it is advisable to suppress high frequencies. Thus we substitute the dirac spike δ in the time-domain by a gaussian filter (see subsection 2.4.2) $\mathcal{G}_{-\sigma}$ with variance σ so that we obtain $\mathcal{G}_{-1/\sigma\pi^2}$ as frequency response instead of \mathcal{C} and get:

$$\mathcal{P} * \mathbf{F}^{-1} \left(\frac{\mathcal{G}_{-1/\sigma\pi^2}}{\mathbf{F}\mathcal{P}} \right) = \mathcal{G}_{-\sigma} \quad (3.5)$$

for the variance σ . Note that equation 3.5 becomes wrong, when \mathcal{P} is not continued by zeros, but by the surrounding pixels in the material. This can be circumvented by calculating \mathcal{X} not by convolution but by deconvolution as the solution of

$$\sum_{i=0, j=0}^{n-1} \mathcal{X}(i, j) \mathcal{T}(i_0 - i, j_0 - j) = \mathcal{G}_{\sigma}(i_0, j_0) \quad (3.6)$$

$\forall i_0, j_0 \in \text{supp}(\mathcal{P})$. n is the edgelenhth of \mathcal{P} . As n is also the edge length of \mathcal{X} , we have the same number of variables and equations.

We will call the solution \mathcal{X} of equation 3.6 a *pushpin* and $\mathcal{P}(\lfloor \frac{n}{2} \rfloor, \lfloor \frac{n}{2} \rfloor)$ the *puncture* of the pushpin. By *nailhead* we mean the support of \mathcal{P} .

A pushpin, constructed in this way, does respond a bit stiff: tiles have to

be very similar to the original tile to generate a detectable spike. This may be relaxed massively by using a regularization: instead of solving the equation system 3.6, we constrain this equation system by a spatial smoothing term namely by the minimization of the discrete laplace operator (Δ). This leads to a minimization problem:

$$\mathcal{X} = \operatorname{argmin}_{\mathcal{W}} \left(\sum_{i_0=0, j_0=0}^{n-1} \left\| \sum_{i=0, j=0}^{n-1} \mathcal{W}(i, j) \mathcal{T}(i_0 - i, j_0 - j) - \mathcal{G}_\sigma(i_0, j_0) \right\| + \left\| \sum_{i=0, j=0}^{n-1} \Delta_{(k, l)(i, j)} \mathcal{W}(i, j) \right\| \right) \quad (3.7)$$

3.2 Overview

Given an SVBRDF, where at least some of the parameter channels bear a *roughly* periodic pattern in the following sense: there exists a periodic pattern which may be warped into those channel maps alongside of a small continuous flow field. Here we mean by periodic pattern an image which may be generated from a model tile and a concatenation of translations and rotations according to an appropriate wallpaper group. The geometrical aspect of this definition coincides with the definition of *near regular textures (NRT)*, given in [LLH04]. A specification of the term *small* is difficult and depends not only on the settings of the algorithm but also on the texturizing of the SVBRDF, itself.

Further we assume, that a user has marked a *foreground* component \mathcal{F} of the SVBRDF and a *background* component \mathcal{B} by the use of a stroke input $\mathcal{S}_{\mathcal{F}}$ for the foreground and a stroke input $\mathcal{S}_{\mathcal{B}}$ for the background stroke. Then we want to propagate this stroke input in a way that the periodic pattern is respected and a texel with the index i and the average reflectance distribution ρ_i obtains a value α_i which decomposes ρ_i into a convex combination of a

foreground BRDF $\rho^{\mathcal{F}}$ and a background BRDF $\rho^{\mathcal{B}}$

$$\rho_i = \alpha_i \rho_i^{\mathcal{F}} + (1 - \alpha_i) \rho_i^{\mathcal{B}} \quad (3.8)$$

For the classical matting problem, the parameter α is described as opacity or transparency. For our application, this interpretation is not good, as transparency leads to complicated reflectance properties. α should be merely seen as area share of the foreground reflectance distribution.

We define the *foreground*

$$\mathcal{F} = \{t_i \in \mathcal{T} | \alpha_i = 1\}, \quad (3.9)$$

the *background*

$$\mathcal{B} = \{t_i \in \mathcal{T} | \alpha_i = 0\}, \quad (3.10)$$

and the *boundary*

$$\partial = \{t_i \in \mathcal{T} | 0 < \alpha_i < 1\}. \quad (3.11)$$

Note that an important requirement for a successful classification step is:

$$\mathcal{S}_{\mathcal{F}} \cap \mathcal{B} = \mathcal{S}_{\mathcal{B}} \cap \mathcal{F} = \emptyset$$

In Figure 3.3 you can see an overview over our new algorithm. As input we take a SVBRDF together with a stroke input. Then we apply in parallel a segmentation via a support vector machine (Subsection 3.3.2) on the descriptors described in Subsection 3.3.1 and estimate a lattice on the diffuse colour (Subsection 3.3.3-3.3.5). Based on the detected lattice we extract a model tile (Subsection 3.3.6), calculate an optical flow between this model tile (Subsection 3.3.7) and all other tiles and warp the tiled SVM-classification results into

the model tile. This set of warped masks is used to compose an average tile-mask which is then warped into the original tile positions (Subsection 3.3.8). And finally the edit is applied (Subsection 3.3.9).

3.3 Detailed Description of the Workflow

We start on the left side of figure 3.3 with the classification step. Based on the stroke input, we classify in this step all texels of the material probe, without reference to the near regular structure, into foreground and background texels. We tried several different descriptors and several different classifiers:

3.3.1 Descriptors

Additionally to the 8 reflectance parameters and the 2 parameters of the surface normal provided by the Ashikhmin-Shirley model (Equation 2.6), we add the filter responses of Gabor filters (Equation 2.58). We use 8 different orientations and a wavelength of 3 texels. Gabor filters are applied to the volume-channel of the diffuse color. This strengthens the influence of line features on the classification result. We compare every texel on a patch with size 5×5 texel. So the dimension of our descriptor is altogether $(8 + 2 + 8) \times 5 \times 5 = 450$.

3.3.2 Classifier

We tried different state of the art classifiers: Support Vector Machines [CV95], Deep Belief Networks [HOT06] and Convolutional Neural Networks [LBD⁺89]. The latter have been implemented in Theano for Python, for the SVM we used the implementation by Chang [CL11].

Though we made good experiences with neural networks in the past, they failed in the current scenario. According to a rule of thumb given in [Mas93], the number of samples should be equal or more than the number of weights of

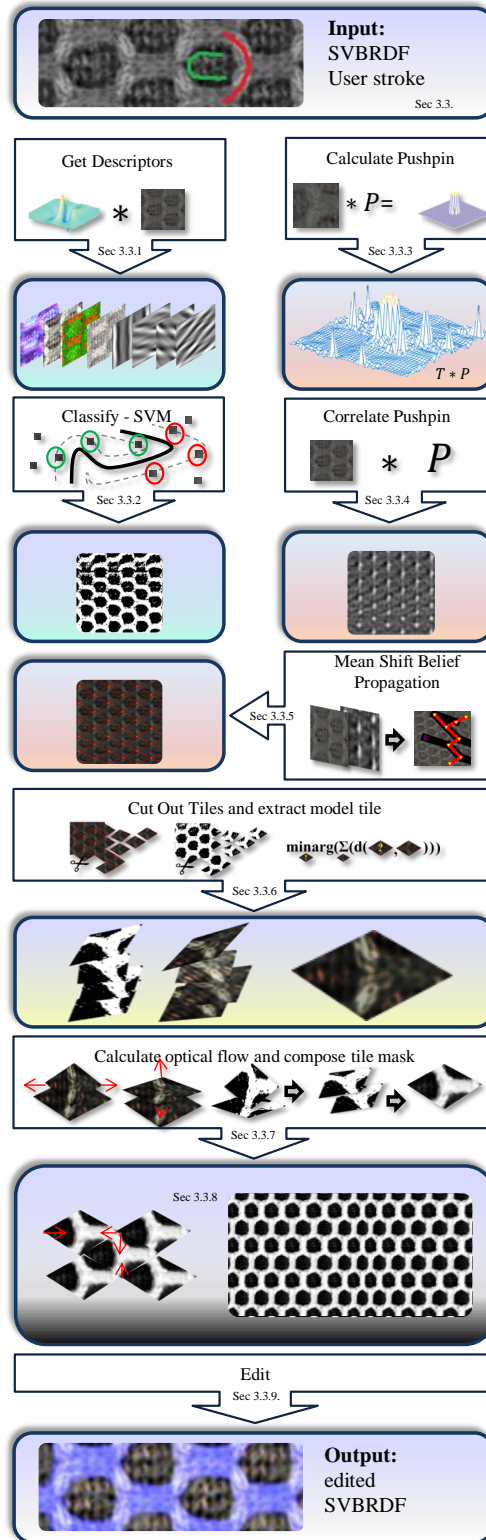


Figure 3.3: Overview.

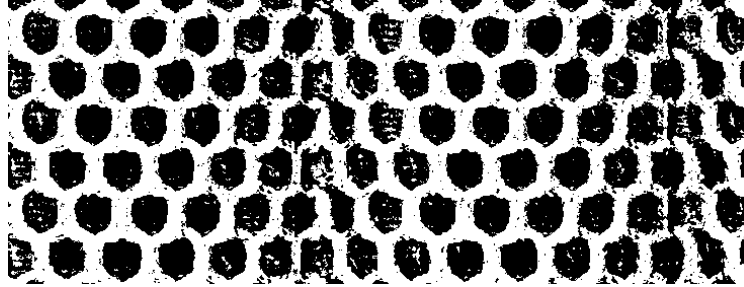


Figure 3.4: The mask resulting from the SVM classification step.

the neural network. As stated in Paragraph 3.3.1, the descriptor of a texel has the dimension of 450 which makes, dependent on the concrete topology, about 50,000 weights in a three layer neural network, whereas a stroke input provides between 100 and 500 samples. So the networks have simply not enough data for training. SVMs, in contrast, can be trained with a small amount of data and are easy to apply and quickly trained. We use radial basis functions and parameter estimation is done by grid-search and 5-fold cross-validation.

In Figure 3.4 you can see that the result of the svm classification step is already a good segmentation. Still there are some noticeable misclassifications. To handle this problem, we aim at getting an average tile. To obtain such an average tile, we need a tiling first. But before detecting the lattice, we add another channel to our material, which we obtain by the use of pushpins. The next step is found on the top of the right column of the overview image (Figure 3.3) .

3.3.3 Pushpins

To make the results of the lattice detection more stable and more predictable, we guide this following step (subsection 3.3.5) by a weaker but therefore global repetition detector. The main idea is to mask the frequency spectrum of a given material \mathcal{T} in such a way that the region with the stroke input $\mathcal{P}_{\supset \mathcal{S}_F, \supset \mathcal{S}_B} \subset \mathcal{T}$ shows a peak in the spatial domain and therefore all similar regions in the spatial domain show a peak. This may be done by convolving \mathcal{T} with \mathcal{P} ,

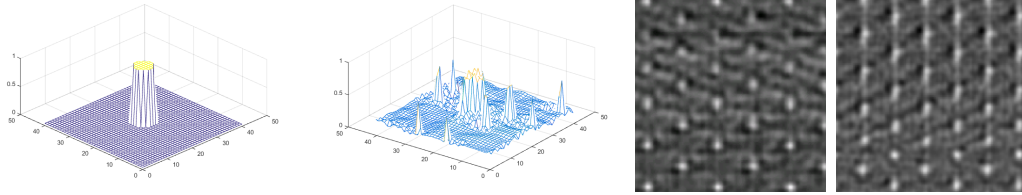


Figure 3.5: The effect of regularization to pushpins. From left: the response of the original tile to an unregularized pushpin, the response to a regularized pushpin, the response of distorted material-patch to the unregularized pushpin and the response of the same material-patch to the regularized filter.

but simple cross correlation does not bring the desired results. Instead, we construct a patch which generates a peak when convolved with \mathcal{P} . This technique is described in detail in section 3.1.

3.3.4 Convolution with a Pushpin

Pushpins can be made tolerant against noise or small distortions by adding energy terms to equationsystem 3.7. To justify the regularization step (Equation 3.7) we want to illustrate its effectiveness in Figure 3.5 by comparing the result of the convolution with a regularized pushpin against convolution with an unregularized pushpin.

And the other way round it is possible to concentrate on certain regions of the pushpin by adding weights to the corresponding equations.

To construct a pushpin, it is necessary to determine a centerpoint and a radius. We have chosen the mean of the positions of the stroke inputs \mathcal{P} as puncture and the size of the nailhead was chosen so as to cover the whole stroke.

Nested symmetry groups Pushpins generate automatically a region of dominance. This shall be demonstrated by a simple example. In figure 3.6 you may see a simple texture consisting of small squares arranged in groups to bigger squares. On the left side, you can see the response of a pushpin with a nailhead diameter in the size of a small square, on the right side we used

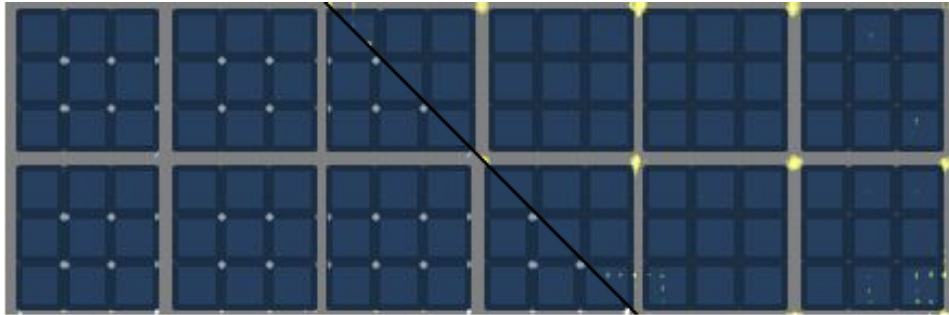


Figure 3.6: On the left side you can see in light blue the filter response of a pushpin with nailhead radius approximately equal to the size of half a small square on the right side we used a pushpin with a nailhead radius approximately equal to half a big square.

a pushpin with a nailhead diameter in the size of a big square. The clipped filter response of the smaller pushpin is held blue, the filter response of the bigger pushpin is shown in yellow. You can see that the pushpin on the left side detected the crossings between the small squares whereas the pushpin on the right side detected *exclusively* the crossings between the big squares. This means that pushpins can distinguish between nested symmetry groups. That is an improvement against plain mean shift belief propagation because MSBP simply uses the symmetry group it gets first.

Though pushpins are not limited to a certain number of channels, particularly not to 1, we confine their use to the lightness channel of R_s or R_d . Note that the use of more channels does also lead to more noise in the filter response.

Our algorithm gains its strength from the combination of lattice detection and pattern-recognition. In our tests, the most successful approach to detecting lattices was the *mean shift belief propagation (MSBP)*, published by Park et al. [PCL08].

3.3.5 The Lattice Detector: MSBP

MSBP makes the assumption that a repeating structure in an image is a slightly deformed periodic pattern. As such it is possible to find an ideal

pattern element and two linearly independent lattice base vectors to reconstruct this periodic pattern by operating via the corresponding wallpaper group [GS87, LCT04]. By clustering points of interest, MSBP estimates the base vectors for the periodic pattern and a seed point, and the algorithm extracts a characteristic tile around this seed point. The lattice base vectors define symmetry-mappings, so the seed point and all symmetry-images of this seed point may be mapped to further symmetry-images by translation along the base vectors. Those images are the vertices of the constructed lattice. As the lattice is deformed by assumption, the exact symmetry mapping has to be found by searching for a good fit for the characteristic tile in the area of the estimated new vertex position. This search is done for all new lattice-vertex candidates simultaneously, meaning that the search for two neighbouring vertices is constrained by an energy term which punishes deviation from the according base translation. Mean shift belief propagation has proven to be an extremely powerful algorithm. Still we had to struggle with two problems:

1. The results are not deterministic.
2. Regions of big distortions like the fold in the grey mesh material often stop the expansion of the lattice.

Both difficulties are illustrated in figure 3.7. The result of MSBP, reflected by the red lattice in the left image was successful: the algorithm found the smallest possible tile and the lattice covers the whole material patch. On the right image, we have an example for an abortive run of MSBP: you can see that the algorithm was not able to cross the fold in the material and the base vectors are the sum and the difference of the base vectors found in the right image. Those problems may be solved by summing the pushpin-responses into the diffuse colour-channel.

In figure 3.8 we visualize the influence of pushpins to the lattice detection process. We have made several test runs some of which had one or two nodes missing, but we obtained always the same lattices covering the whole material patch.

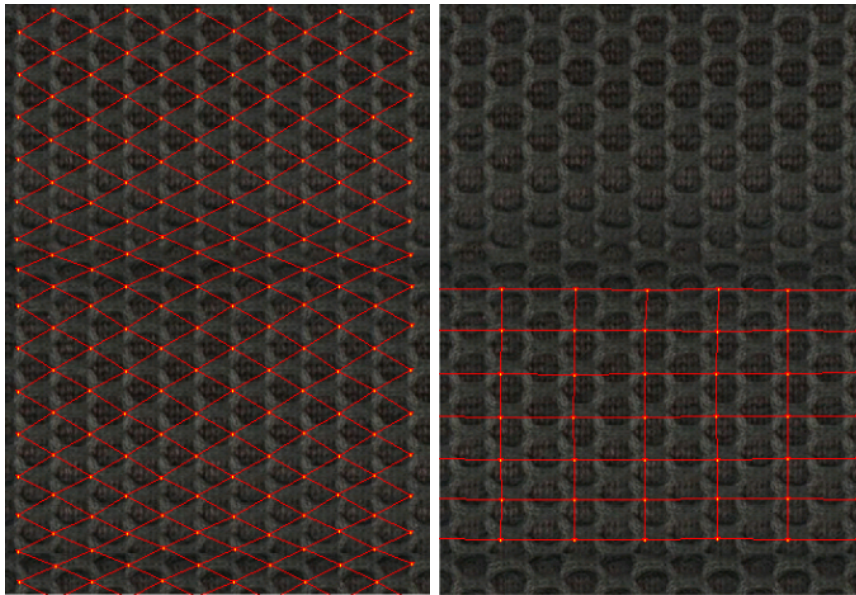


Figure 3.7: Two different runs of mean shift belief propagation on the grey mesh material.

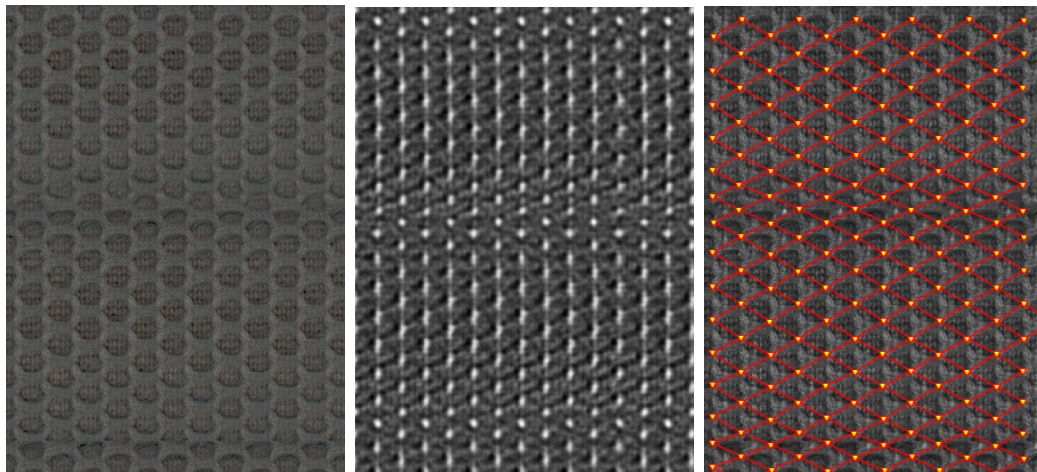


Figure 3.8: The left image shows the grey mesh material, the image in the middle depicts the filter response of a pushpin applied to the volume channel of the diffuse color of the grey mesh material and the third image shows the result of MSBP on a combined map of the filter response and the diffuse channel. Now the lattice detection is extremely stable.

In the next step we combine the results of the classification and of the lattice detection to obtain a model mask tile and a warping field to plaster the whole material patch with this model mask patch. After cutting the mask and the material into a set of tiles which we interpret as distorted version of the same model tile, we extract a model tile, we calculate an optical flow between the model tile and all other tiles and we compose a mask for the whole material probe.

3.3.6 Finding a Model Tile

To generate a single tile bearing a reliable segmentation we first choose one tile which is every bodies friend $\mathcal{P}_{\text{model}} \in \text{tiling}$:

$$\mathcal{P}_{\text{model}} = \operatorname{argmin}_{\mathcal{X}} \left(\sum_{\mathcal{W} \in \text{tiling}} d(\mathcal{X}, \mathcal{W}) \right) \quad (3.12)$$

The simplest way to define a distance d between different tiles is, to bring all tiles to the same size and apply the euclidean norm to the diffuse colour channel.

We assume that changes in the size of tiles are due to perspective distortion. Thus the best fit for an average tile should be a tile with maximum edge-length. So in the first step we resize all tiles to the maximum edge-length and store the resize-factors. As the number of tiles is small, we simply apply a brute force approach and compare all tiles pairwise. This procedure is quadratic in the number of tiles, so for big numbers, the time requirement may be optimized by using a dynamic programming approach. Note that generating a mean tile instead of searching the tile with the most friends is not advised as we want to calculate the optical flow between this model tile and all other tiles. This is more difficult with a mean tile because the algorithm has to find features, which are prone to be smoothed out by averaging.

3.3.7 Calculating the Optical Flow and Composing a Tile Mask

With the model tile at hand, we may calculate the optical flow between the model tile and all other tiles. For the estimation of the optical flow between the model tile and the test tile, we use the algorithm suggested by Sun et al. [SRB10] and apply it to the diffuse colour-channel. The resulting optical flows are applied to the resized tile masks. This and the following back-warping step are done by the use of thin-plate splines [Boo89]. We obtain a principal tile mask by averaging over all resized and warped tile masks.

3.3.8 Recombination

Now every tile mask of the original mask is substituted by the appropriately back-warped and back-sized principal tile mask.

3.3.9 Applying the Edits

Our algorithm assigns an alpha value to every texel. This value will scarcely be exactly one or zero. So we will do a segmentation by thresholding. Aside from distortions the alpha-values may be seen as voting for the background or the foreground, so 0.5 is a good threshold. The segmentation mask is of course not suitable for editing as it will obviously lead to strong artefacts. So we will substitute all texels, which have at least one corner-neighbour from the opposite component, by its alpha-value, so that the intuitive use of the word *boundary* and the definition given in Equation 3.11 coincide.

On the foreground component, editing can of course be done as e.g. described in the literature cited in Subsection 2.3.2, but on the boundary we have to be more careful. The manipulation of a reflectance value may trouble-free be done by convex-combining the old and the manipulated value. But it has to be kept in mind that not all parameters of all analytical BRDF-models superpone.

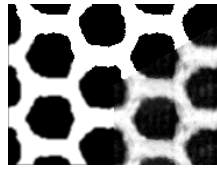


Figure 3.9: On the top the binary mask, on the right the original superposed mask and on the left the mixed mask.

3.4 Evaluation

In this section we want to show the potential of our approach. Evaluation targets into three directions:

1. we will show that our algorithm is capable of dealing with materials, which do not show the strong colour-contrasts, mostly necessary for matting and foreground-segmentation purposes.
2. the other strength of our algorithm is, that it provides us even for strongly distorted material surfaces with an editing mask, which shows almost no artefacts.
3. The coincidence between perceived structure and warped mask tile is very good.

3.4.1 Test Set-Up

To describe our test set-up we will start with a short description of the input data. Next, we will give a detailed overview over the competing algorithms to convey an idea where those algorithms run into problems. Of course the test set-up is strongly biased into the direction of our algorithm as both algorithms, AppProp and RepSnapping are by far more general. But we did not find a more fitting approach in literature.

Input data

The materials we use in this paper have been acquired with an enhanced version of the *linear light source reflectometer (LLSR)*, introduced by Gardner et al. [GTHD03]. This new system has been developed by Meseth et al. [MHW⁺12] and is capable of measuring anisotropic reflectance distributions (subsection 2.2.3-SVBRDF Acquisition). Meseth et al. use the Ashikhmin-Shirley reflectance model (Subsection 2.2.1-Ashikhmin Shirley reflectance model)

Additionally to the reflectance properties (Equation 2.6), LLSR has to estimate a surface normal \mathbf{n} . All values have been stored as 16 bit integer values. One texel represents a surface of roughly $1/4 \text{ mm}^2$.

We use two different materials for the comparison: the grey mesh material which we have used to demonstrate the single steps of the algorithm and a structured steel material (see Figure 3.10).

The grey mesh material is nearly uni coloured. It is particularly difficult to derive a near regular structure because it contains a strong bulge and the material normals do not convey much information.

While it is really simple, to derive the regular structure from the structured steel material, the only visible difference between foreground and background is a slightly less isotropic distribution of the noise. The metal material does not have a diffuse colour channel R_d so we have to use R_s , instead.

Comparison with other algorithms

Our algorithm combines techniques from the field of material manipulation with techniques from the field of repetition finding in images. Thus for comparison we have chosen one outstanding algorithm from each of those field. For the task of segmenting repetitions in images we decided for the RepSnapping algorithm [HZZ11], published in 2011 by Huang et al. And to cover the field of SVBRDF-editing we will compare against AppProp [AP08], published by An and Pellacini in 2008. Moreover, we compare those results with the segmentation of the SVM from step 3.3.2.

AppProp The authors use a low rank approximation of the full appearance adjacency matrix and minimize the following functional:

$$\sum_{i,k} w_k z_{ik} (e_i - g_k)^2 + \lambda \sum_{i,j} z_{ij} (e_i - e_j)^2$$

with

$$z_{ij} := \exp(-\|f_i - f_j\|^2/\sigma_a) \exp(-\|x_i - x_j\|^2/\sigma_s).$$

Where i and j go over all texel in the texture, k goes over all texels in the stroke input, w are weights, e is the edit and therefore the solution of the optimization problem, g is the stroke-input and therefore the right hand of the optimization problem, x is the position of the texel, λ the weight of the smoothing term and f is a texel-dependent appearance term. The resulting equation system is roughly solved by a low-rank approximation. The appearance comparison of AppProp is not limited to three dimensions or a single texel, so we can apply it to our descriptor (Section 3.3.1).

The spatial parameter σ_s is not interesting in our setting, but to find a reasonable value for σ_a is difficult for our high dimensional descriptor and has to be done in a preprocessing step for every material separately. This is not surprising because the term

$$\exp(-\|x_i - x_j\|^2/\sigma_s) = \prod_k \frac{1}{e^{(x_i^k - x_j^k)^2/\sigma_s}}$$

consists in our case of 450 factors and has therefore the inclination to explode or to collapse beyond numerical accuracy. λ controls the consistency of the edit and had not much influence. We set λ and w_k to one. Thresholding has been done manually, in order, to get the best possible segmentation.

RepSnapping RepSnapping has been published by Huang et al. in 2011 [HZZ11], and is based on the idea of co-segmentation [HS09]. It is specialized to cutting out repeated elements in natural images. The algorithm solves the

energy functional:

$$E(e) := \sum_i D_i(e_i) + \sum_{i < j} V_{i,j}(e_i, e_j) + \sum_{i,j \in H} U_{i,j \in NbH}(e_i, e_j)$$

by the use of graph cuts [KZ04]. Here D_i describes the probability that $e_i \in \mathcal{F}$ and is given as a normalized set distance to a clustering (H) of the foreground.

$$D_i(e_i = 1) = \frac{\min_{k \in \mathcal{H}(F)} \|f_i - f_k\|}{\min_{k \in \mathcal{H}(F)} \|f_i - f_k\| + \min_{k \in \mathcal{H}(B)} \|f_i - f_k\|}$$

with the appearance function f and

$$D_i(e_i = 0) = 1 - D_i(e_i = 1)$$

The term

$$V_{i,j} = \lambda |e_i - e_j| \exp(-\beta \|f_i - f_j\|^2)$$

is a smoothing term and goes over all adjacent pixel pairs. And

$$U_{j,j} = \mu |e_i - e_j| \exp(-\beta \gamma(i, j)^2)$$

assures that pixels with similar appearance are treated similar. The main idea is that the neighbourhood graph is extended by the neighbourhood-system NbH which contains edges between the pixels i and j iff $\gamma(i, j) < \epsilon$, where γ is a correlation based similarity measure, described in [HS09].

We applied RepSnapping with the parameters given in [HZZ11], namely: $\mu = 10$, $\beta = 0.1$, $\lambda = 2$ and $\epsilon = 4$. RepSnapping might easily be extended to the high-dimensional descriptor used in our algorithm but it would suffer from the same stability issues as AppProp.

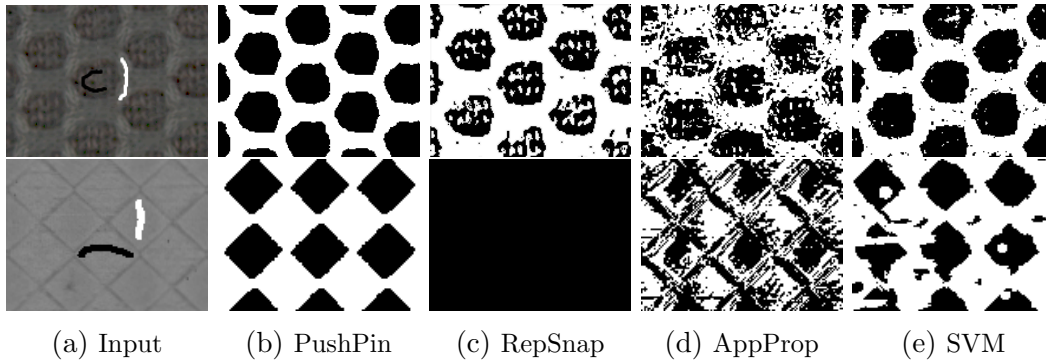


Figure 3.10: On the left a patch from the original SVBRDF-channel with the stroke input the second image shows the mask generated by RepSnapping. The third mask is the result of AppProp and the last mask is our result. The first row shows the grey mesh material the second row shows a metal.

3.4.2 The Results

In Figure 3.10 we present the comparison of the image segmentation step. You can see that our algorithm delivers artefact-free masks for both materials (3.10.b). The other three algorithms are more successful on the grey mesh material than on the metal material. An interesting result is, that the raw SVM delivers the second best results. We see the main reason in the descriptors: AppProp is numerically overcharged with the big number of descriptors, which results in this big amount of noise, and RepSnapping uses a correlation based approach to describe texel neighbourhoods. Autocorrelating the R_s -channel of the metal material reveals that the surface does not have enough structure to provide significant correlation results. Together with the fact that R_s is uni-coloured, this explains, why RepSnapping fails completely.

3.4.3 Editing Examples

In this section we want to present edits on four different materials (figure 3.11 - figure 3.14).

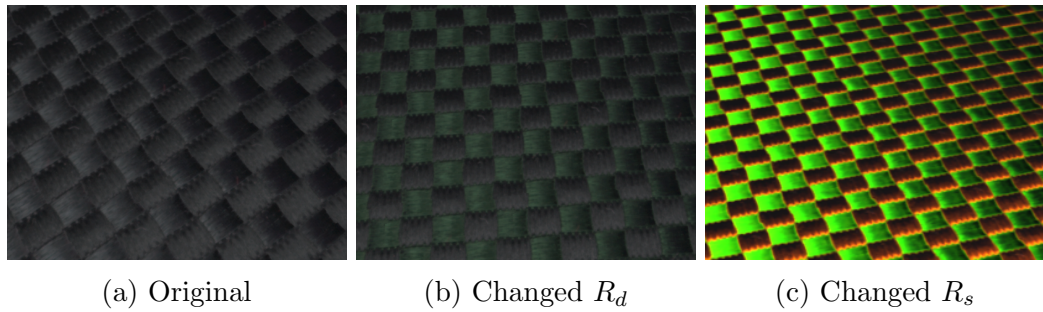


Figure 3.11: On the left the original material in the middle a rather subtle edit of R_d , on the right a more noticeable manipulation of R_s .

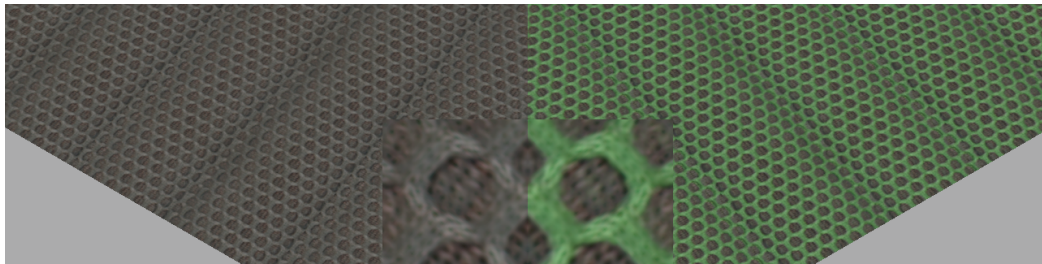


Figure 3.12: In the close-up of the edit of the grey mesh material one may see that the editing boundary coincides exactly with the perceived boundary of the foreground material.



Figure 3.13: A shiny material.

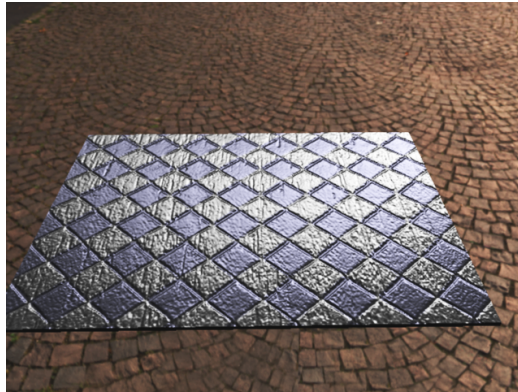


Figure 3.14: The edited metal material.

3.4.4 Time Requirement

Tests have been done on system A (see 2.5). The bottleneck of the algorithm was to calculate the optical flow on all tiles (Paragraph 3.3.7). For the grey-mesh material we had about 180 tiles. Calculating the optical flow on one tile ($\sim 80 \times 40$ texel) took about 2.2 s, which sums up to about 7 min. Depending on the number of training samples, the SVM classification step took between 10 s and 3 min (Paragraph 3.3.2). MSBP (Paragraph 3.3.5) ran for about 45 s. Warping a tile with tps took about 0.03 s. Finding a principal tile took less than a second. So the overall processing time lay between 8 and 12 minutes.

3.5 Conclusion and Future Work

In this chapter we developed an algorithm to solve the task of extracting a repeating foreground pattern from a high dimensional reflectance representation map in a way, which is even in many difficult cases robust and reliable enough, to allow for editing the foreground component and makes additional optical debugging steps unnecessary. While the task is relatively simple on suitable materials, we could show, that the competing state of the art algorithms failed for difficult material probes. Our algorithm permits high quality foreground-segregation and editing on complex materials.

We laid our focus on the exploration of the frequency domain in order to find quickly and globally repetitions of a given pattern. By this global approach, we could provide additional information to the seed based repetition finding algorithm, which made the lattice detection more reliable and robust. As recommendation for future work, we want to emphasize two aspects of the pushpins, which may be of further interest:

1. Pushpins are in the presented form scarcely applicable to classical signal-processing tasks as they are too sensitive to noise. We could show, on the other hand, that autocorrelation, which is widely used in signal-processing, is not sensitive enough to fine structural elements, to allow for a repetition finder like the presented. It would be interesting to find a more general concept which incorporates pushpins and autocorrelation as special-cases.
2. Under the condition, that the examined function is relatively smooth, pushpins provide sharp, detectable peaks. Acoustic signals, in particular pieces of music, are often relatively smooth. It would be interesting to investigate how pushpins perform in music-mining tasks.

CHAPTER 4

Frequency Editing of Bidirectional Texture Functions

In the current chapter we transfer a band-filtering scheme first published by Giesel and Zaidi [GZ13] from images to the compressed representation (Paragraph 2.2.3-**BTF compression**) of bidirectional texture functions. The content of this chapter is based on a peer-reviewed publication:

Marlon Mylo, Martin Giesel, Qasim Zaidi, Matthias Hullin, and Reinhard Klein. Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models. In *Proceedings of Vision, Modeling & Visualization, The Eurographics Association*, [MGZ⁺17]

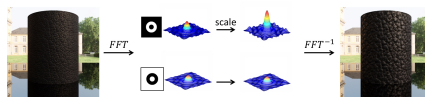


Figure 4.1: Appearance Bending.

All self-citations in the current chapter refer to *Appearance Bending: A Perceptual Editing Paradigm for Data-Driven Material Models*.

In the first year of my Doctorial-course I heard a talk of Martin Giesel on a connection between certain affordances and spatial frequencies in the images of material patches. I was intrigued from the very first moment because from this work followed that a single parameter decided in each case, whether we perceive a material as thick or shallow, as flat or undulated or as soft or rough. As this parameter did not control a physical quantity, we assume, that there has to be a process in the human brain which is comparable to the scaling of those frequency bands. This process may be identified with the associated property. On the other side the frequency manipulation gives rise to an image based material editing system which is in the sense of 2.3 plausible and predictable.

Next to the related work given in Section 2.3 the band sifting system published by Boyadzhiev et al. [BBPA15] is closely related to our frequency based material bending scheme. Katsunuma et al. [KHH17] use the results of Giesel and Zaidi to concept a texture transfer system. Their work is based on frequency and colour-channel transfer and closely related to the work presented in this chapter.

The rest of the thesis deals with this relation between frequencies and perceptual properties.

The first section (Section 4.1) gives an overview over Giesel and Zaidis work and introduces an operator-notation to make those insights usable for editing measured materials. This is followed by a detailed description of our new technique, which we call *appearance bending* (Section 4.2) and we will evaluate appearance bending by showing and discussing representative results (Section 4.3). We feel that our results permit to speak of a successful transfer from vision resarch results into the field of computer graphics but we have to mention that our edited materials are probably physically incorrect (Section 4.4).

In this and the following chapter, we present an approach to editing materials, which operates by scaling directly on the frequency- or scaling-domain of material reflectance maps. This stands in contrast to Chapter 3, where the editing operation itself has implicitly considered to be an arbitrary manipu-

Freq band [cpd]	Operation	Operator
0.57 -2.29	Inflated and deflated	$\mathbf{F}\mathbf{D}_U$
2.29 -4.28	Deep and flat	$\mathbf{F}\mathbf{D}_T$
6.57 -15.14	Soft and rough	$\mathbf{F}\mathbf{D}_R$
15.14 -19.42	Sparkling and dull	

Table 4.1: Affordances in cycles per degree (cpd) in Giesel and Zaidis work. In the last column, we give the symbol of the associated scaling operator. The upper \mathbf{F} indicates, that the scaling $\mathbf{F}\mathbf{D}$ operates on the Fourier-domain.

lation of reflectance parameters, changing spatially only with respect to the support of the calculated mask.

4.1 From Affordances to Appearance Bending

The concept of *affordance* has been introduced by Gibson in 1977 to psychology and it describes the possibilities of actions which may be done on a given object [Gib77]. Affordances are not only transported by the shape of the object, but also by the material it consists of [Ade01]. This is the starting point for Giesels and Zaidis work [GZ12, GZ13]. Test observers had been confronted with the task to assign images of material patches to certain properties, related to the use of a material, like whether the material was stiff, soft or water repellent. Consistently rated materials had been investigated on their frequency contents. In this way it was possible to correlate 4 properties to frequency bands, see table 4.1. The last frequency band had been omitted in the following discussions, so it will not be part of the discussion in this chapter, too.

Image based material manipulations by bandpass filtering

The connection between affordance and spatial-frequencies gives rise to a semantically founded editing scheme by simply enhancing or weakening par-



Figure 4.2: From [GZ13]: Material patches, used in Giesel and Zaidis affordance classification experiments. The researchers cropped 256 color images of fabrics to a size of 150 x 150 pixels. Images were presented on a monitor against a black background. The viewing distance was 70 cm.

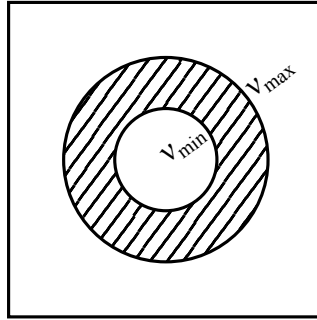


Figure 4.3: Frequency-domain support of the editing kernel $\mathbf{F}\mathbf{D}$.

ticular frequency bands [GZ13]. Because the underlying physical effects are too complicated, those effects may not be seen as the result of inverse optics [GZ13]. We want to note, that those effects may be seen as examples for complex physical operations with a linear representation in the visual cortex.

Lets give an operator based description of this editing scheme: given a texture-image \mathcal{T} . Than the researchers applied the following operators:

$$\mathcal{E} = \mathbf{F}^{-1} \circ \mathbf{F}\mathbf{D}^k \circ \mathbf{F}(\mathcal{T}) \quad (4.1)$$

The kernel $\mathbf{F}\mathbf{D}^k$ performs a pointwise multiplication in frequency domain, with a gain factor k that determines the amount of effect desired:

$$\mathbf{D}^k(\vec{\nu}) = \begin{cases} k, & \text{if } \nu_{\min} < |\nu| < \nu_{\max} \\ 1, & \text{otherwise} \end{cases} \quad (4.2)$$

The kernel's ring-shaped support in Fourier domain (Figure 4.3) needs to be derived from Table 4.1. For design purposes it may be advised to use a less abrupt transition between edited and non edited regions by applying an appropriate filter.

Although the Fourier spectrum of a texture is complex-valued, this manner of editing will not introduce an imaginary component to the signal, as long as

the kernel is symmetric about the origin ($\nu = 0$).

4.2 Appearance Bending: Scale-Space Manipulation of Materials

Before we transfer appearance bending to BTFs, we want to derive a more general formulation of equation 4.1. Mathematically speaking, our goal is to transform a given material \mathcal{T} into an edited, or “bent”, material \mathcal{E} , a process we abstract as a *bending operator* \mathbf{E} :

$$\mathcal{E} = \mathbf{E}\mathcal{T} \quad (4.3)$$

We assume the material being represented by a multi-channel image. \mathbf{E} , in turn, is composed as the chain

$$\mathbf{E} := \mathbf{R}^{-1} \circ \underbrace{\mathbf{M}^{-1} \circ \mathbf{D} \circ \mathbf{M}}_{\mathbf{B}} \circ \mathbf{R} \quad (4.4)$$

where \mathbf{R} is a range transform and \mathbf{B} is an editing operator, which consists of a multiplicative kernel \mathbf{D} acting pointwise on the coefficients in the basis \mathbf{M} . In the following, we will motivate and discuss the individual components of this chain, and how we address the differences and technical challenges that arise when marrying Giesel and Zaidi’s frequency-based operations to data-driven reflectance models.

4.2.1 Choice of Basis

When it comes to selecting the basis in which the editing shall be performed, the most straightforward choice is a Fourier transform ($\mathbf{M} = \mathbf{F}$), which was also used by Giesel and Zaidi. This means that certain frequency components will be amplified or attenuated *globally*. In Section 4.3 (Figure 4.9), we will see that Fourier editing may fail for certain multi-component materials. Thanks to the generality of appearance bending, however, it is possible just as well to operate in a wavelet basis, also enabling *local* manipulations of inhomogeneous

Associated Property	Operator	Freq band [cpd]	Octave
Undulation	\mathbf{wD}_U	0.57 – 2.29	1 - 3
Thicken	\mathbf{wD}_T	2.29 – 4.28	4 - 5
Roughen	\mathbf{wD}_R	6.57 – 15.14	5

Table 4.2: Affordances in frequency-octaves.

materials.

As it turns out, for many textures with a certain degree of randomness, the Fourier domain is a good choice. For others, however, we have become aware that the global support of the basis functions can cause objectionable artifacts in the form of ringing – textures consisting of multiple basis materials are particularly susceptible. A wavelet basis is better suited to these situations.

Contrary to the tight spacing of frequencies in Fourier domain, wavelets are organized in octaves (see Subsection **2.4.2-The Fast Wavelet Transform**). It is striking, that those manipulations cover nearly the whole frequency range of the visual cortex (Figure 4.4). In order to emulate the effect of the main frequency bands as given by Giesel and Zaidi [GZ13], we round the given cutoff frequencies upwards to the nearest octave, avoiding that the resulting bands overlap. For a 256 px×256 px image corresponding to an edge length of 3.7 cm, we come up with the scale-space octaves, listed in Table 4.2.

So the shape of the scaling operator \mathbf{wD} has to change in a way, that all coefficients of a given scaling-octave are edited. This leads to the support illustrated in Figure 4.5. For further explanations see Subsection 2.4.2.

4.2.2 Color Spaces and Dynamic Range

As in Giesel and Zaidi’s perceptual studies, we perform our manipulations on the intensity channel in YUV color space (see Subsection 2.1.4). Contrary to their example, however, our data is of much higher dynamic range, and so the same arithmetic produces significant artifacts (Figure 4.6). Therefore, we use

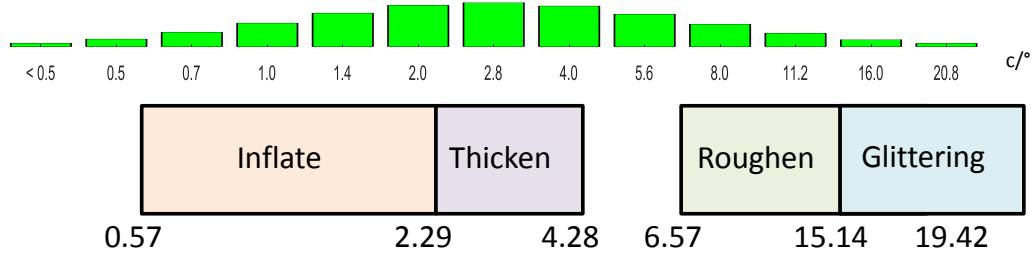


Figure 4.4: The green bars indicate the frequency distribution in [cpd] of our new computational cortex model, presented in Chapter 5. Beneath this distribution we illustrated the frequency-range of the editing operator. Note, that the frequency values on the x-axis increase exponentially, meaning that octaves have constant width.

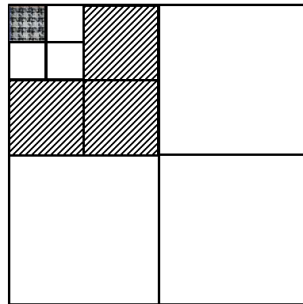


Figure 4.5: Scale space-domain support of the editing kernel \mathbf{W}_D .

the range transform operator \mathbf{R} to transform the BTF data to a logarithmic scale before the editing step, and transform the result back by its inverse, an exponential function. This is in accordance with the Weber-Fechner law according to which the perceived change of a physical stimulus is logarithmic in the strength of the stimulus [FHB66].



Figure 4.6: This leather material, represented as BTF with high dynamic range, was edited using a naïve adaptation of Giesel and Zaidi’s algorithm (left) and in log space (right). In the center is the unedited version. All patches are equally tone-mapped; negative values are marked in red. It is evident that the log-space version is robust to overshooting (it will never produce negative values) and generally less prone to implausible artifacts.

4.2.3 Appearance Bending on BTFs

Now we concretize our approach to BTFs. We may see a BTF as an image with many channels, so the operators may be applied texture by texture. Our first point concerns the editing of compressed BTF-data.

Bending compressed BTFs

Appearance bending seamlessly integrates into systems operating on compressed data, provided that the decompression operator is linear. Suitable compression schemes include full matrix factorization [KMBK03] or Müller’s decorrelated full matrix factorization [Mül09]. The tabulated data are unrolled into a two-dimensional matrix where textures form the columns and the local reflectance distributions, the $aBRDF$ s, form the rows. To compress the BTF,

the PCA is calculated and the resulting eigen-base is truncated by setting a certain threshold (see Paragraph **2.2.2-BTF Compression**)

Due to the linearity of the reconstruction step, the bending operator can be applied directly to the eigen-textures. For the lower frequencies, this yields the same results as manipulating the whole base. Manipulation of higher frequencies may suffer from the loss of such frequencies (fine-scale details) in compression.

Fourier based Appearance Bending

Now we can express the bending operator on the Fourier-domain as per Equation 4.4 as

$${}^F\mathbf{E}_{\{u,t,r\}}^k := \mathbf{Exp} \circ \mathbf{F}^{-1} \circ {}^F\mathbf{D}_{\{u,t,r\}}^k \circ \mathbf{F} \circ \mathbf{Ln} \quad (4.5)$$

While Giesel and Zaidi used textures of 150 px×150 px subtending 3.5° of visual angle, we want to make use of the full native resolution of the BTF database (346 px×346 px) for Fourier editing, and cropped versions (256 px×256 px) for “power-of-2 convenience” in wavelet editing. At a resolution of 6.92 px/mm (160 dpi), this corresponds to sample sizes of 50 mm and 37 mm, respectively, and roughly the same visual angle when viewing the real-world samples at 70 cm distance. This leads to a conversion factor of 12.211 mm/° to transform the original frequency bands to our working resolution. Table 4.3 shows the transformed frequency values and introduces the according scaling operators.

Wavelet based Appearance Bending

For the wavelet-based bending operator, we obtain the following equation:

$${}^W\mathbf{E}_{\{u,t,r\}}^k := \mathbf{Exp} \circ \mathbf{W}^{-1} \circ {}^W\mathbf{D}_{\{u,t,r\}}^k \circ \mathbf{W} \circ \mathbf{Ln} \quad (4.6)$$

Frequency [cpmm]	Associated opposing properties
0.0467– 0.1874	straight– undulated
0.1874– 0.3503	flat– thick
0.5378– 1.2392	soft– rough

Table 4.3: Affordances in cycles per millimetre ([cpmm]) with regard to the metrics of our concrete BTF-implementation.

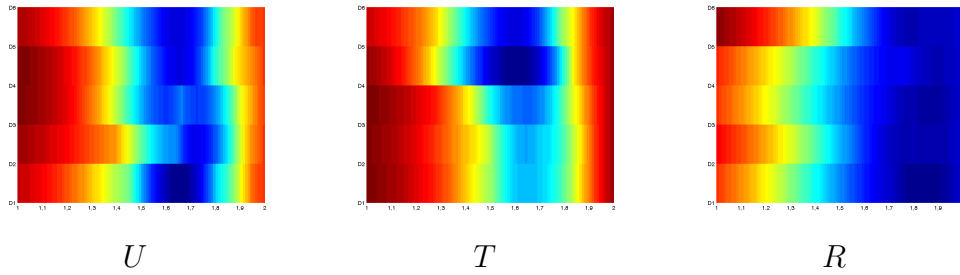


Figure 4.7: HDR-VDP-2 error plots for a reference (Fourier) gain factor of ${}^F K = 2$. The horizontal axis is the wavelet gain factor ${}^W k$, and the vertical axis the type of wavelet (D1–D6).

We compared different wavelets and obtained the best results with wavelets from the Daubechies-series (D). To obtain a scaling that most closely resembles that of Fourier-domain edits, we employed Mantiuk’s HDR-VDP-2 predictor [MKRH11] to formulate a minimization problem:

$$k_{\text{opt},\{u,t,r\}}^s = \arg \min_k \| {}^W \mathbf{E}_{\{u,t,r\}} k - {}^F \mathbf{E}_{\{u,t,r\}} \|_{\text{HDR-VDP-2}} \quad (4.7)$$

We sampled the range of k for Fourier editing at four values (0.5, 0.66, 1.5 and 2.0), and found the scalings given in Table 4.4 to be most appropriate.

Linear regression reveals the following approximate relations between the

$^F k_{\{u,t,r\}}$	0.5	0.66	1.5	2
$^W k_u$	0.7	0.85	1.4	1.7
$^W k_t$	0.65	0.75	1.4	1.8
$^W k_r$	0.85	0.9	1.2	1.3

Table 4.4: Scaling of the wavelet-edit strength by edit type and fourier-edit strength.

Fourier and the wavelet gain factors:

$$^W k_u - 1 \approx 0.66 (^F k_u - 1)$$

$$^W k_t - 1 \approx 0.77 (^F k_t - 1)$$

$$^W k_r - 1 \approx 0.31 (^F k_r - 1)$$

Figure 4.7 shows plots of the HDR-VDP-2 error as a function of gain factor and wavelet type. Our experiments suggest to use wavelets with a number of vanishing moments greater or equal to 6 for the “roughen” operation; for “thicken” and “undulate”, vanishing moments of 3 and 2, respectively are sufficient. Further explanations on the relation between vanishing moments and bandpass filtering may be found in Subsection 2.4.2.

We note that higher vanishing moments increase the spatial support, so we use the lowest-order wavelet that produces no visible artifacts.

As expected, wavelets can act as a good replacement for the Fourier basis (Figure 4.8), and their use pays off particularly for materials with strong spatial variation (Figure 4.9).

4.3 Evaluation

In this section, we present a variety of results that were obtained using appearance bending. Although we process the textures purely in a signal processing framework, the edits often appear to correspond to a semantic meaning. Figure 4.10 shows a wool material processed with the “roughen” operator at

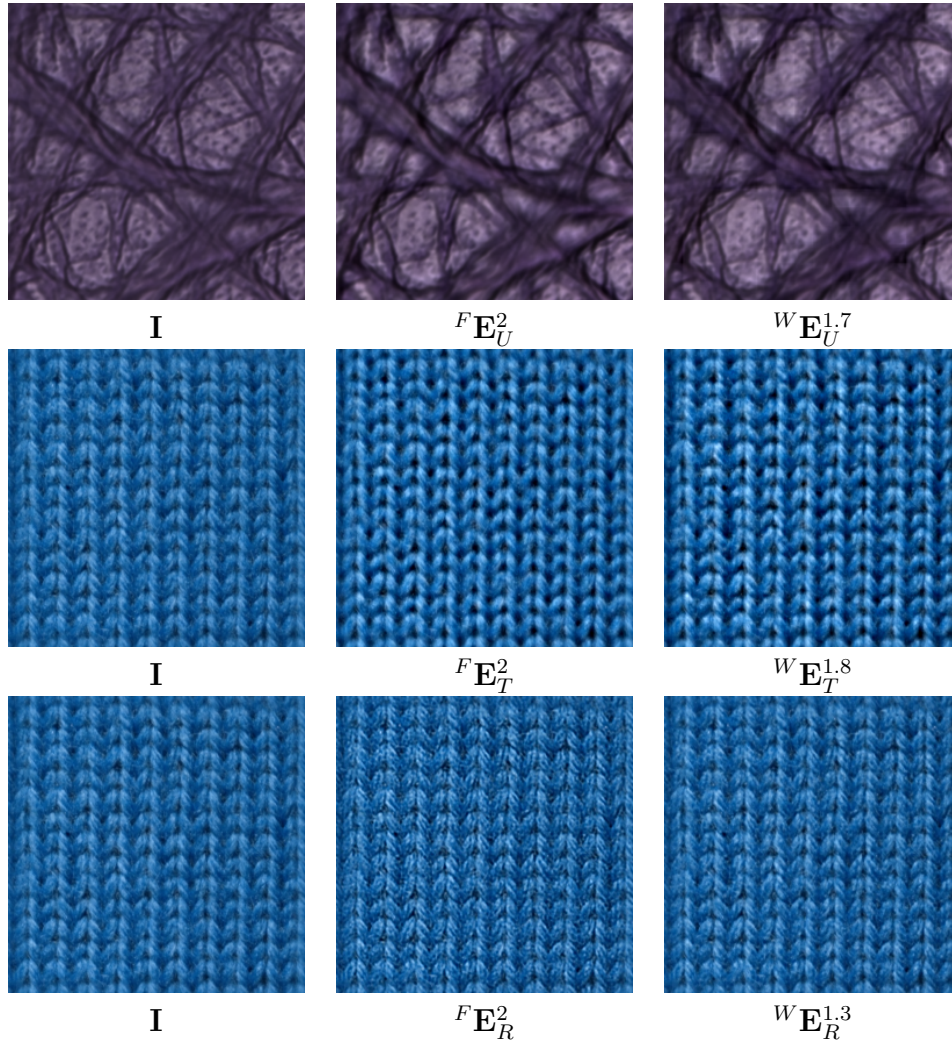


Figure 4.8: A comparison of Fourier- and wavelet-edited materials. Note that the strength of the editing operator, k , has been adapted to obtain the best possible match according to HDR-VDP-2 [MKRH11].

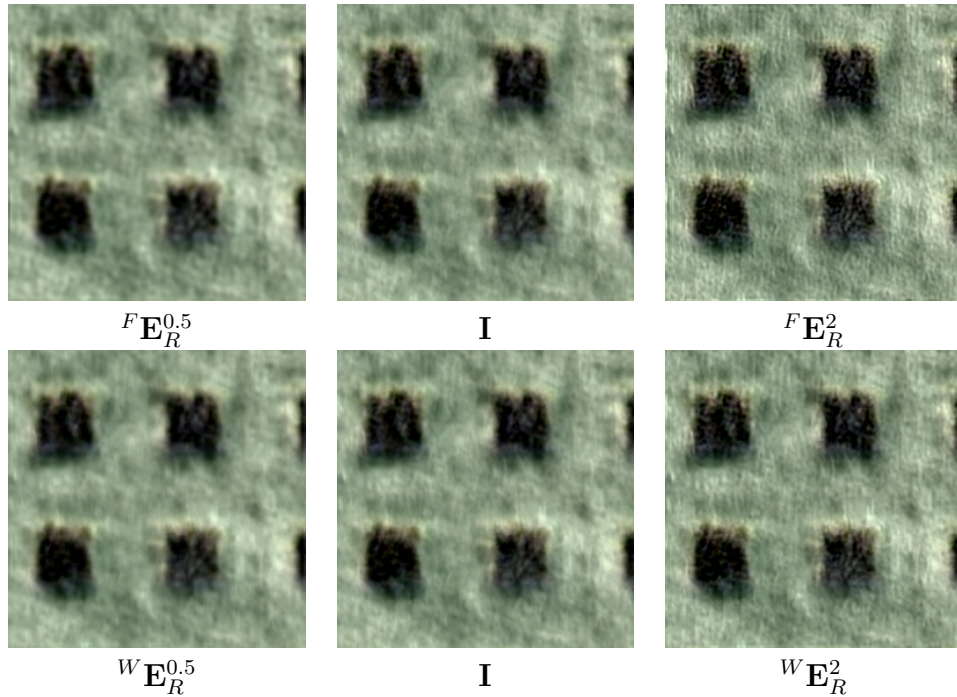


Figure 4.9: Comparison of Fourier (top) and wavelet (bottom) editing for the “roughen” operator. From left to right: $k = 0.5$, $k = 1$ (identity), $k = 2$. Note the absence of regularly patterned artifacts in the wavelet results, even for a gain factor of $k = 2$.

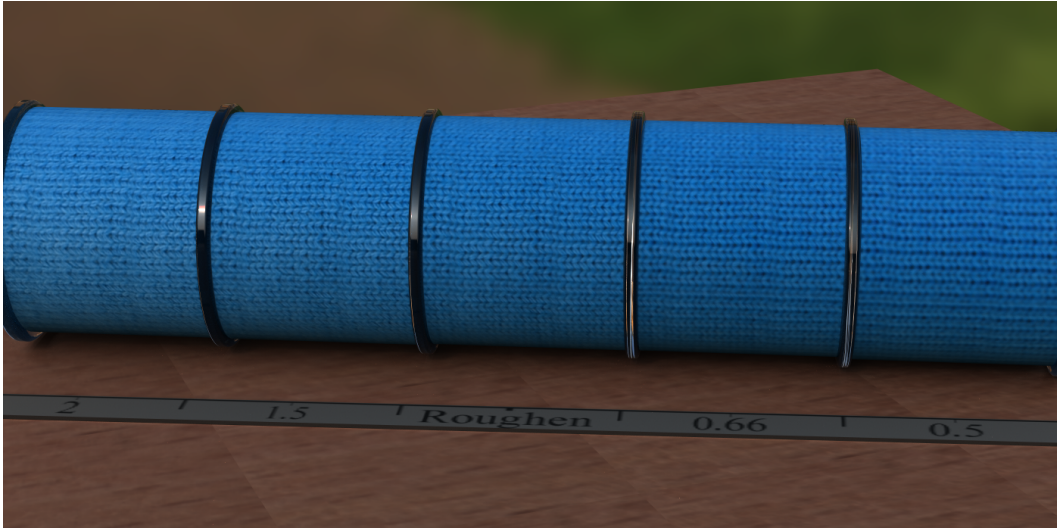


Figure 4.10: \mathbf{E}_R applied to a wool material.

various gain levels. The edit amplifies the structure of the knitware on the finest scale, conveying an impression of fiber-level detail.

Figures 4.13-4.15 show various operators applied to a selection of material classes ranging from grainy leather over textiles to stone and wallpaper. It becomes evident that not every material responds to each operator in the exact same way – however, within a class of materials, the results are consistent (Figure 4.15). We note that uncompressed materials contain more high-frequency details and are therefore better bending candidates - at an increased computational cost.

4.3.1 Extreme Edits

In Figure 4.11, we see the editing result of a material before and after editing in two different presentations: as a flat texture and on a 3D material with environment lighting. The 3D version is more forgiving, producing a plausible impression of a material even for extreme edits. We attribute this to the generally increased realism of the scene, but in particular to the low-pass nature of the environment illumination.

4.3.2 Comparison with Image Based Editing

In this section, we want to compare the result of editing materials to directly editing on the rendered image. Of course the bending operators might have been applied directly to the rendered images. But then the result of the operations depended on the distance of the material surface from the camera and on the angular between the material surface and the image plane. This effect shall be demonstrated for the undulation operator.

In the left column of figure 4.12, we rendered the material probe first and applied the undulation operation afterwards. In the right column we applied the operator to the material base vectors. Shifting of the frequency window has been done appropriately as you can see in the yellow framed image sections: the results are nearly indistinguishable. But the green framed image sections

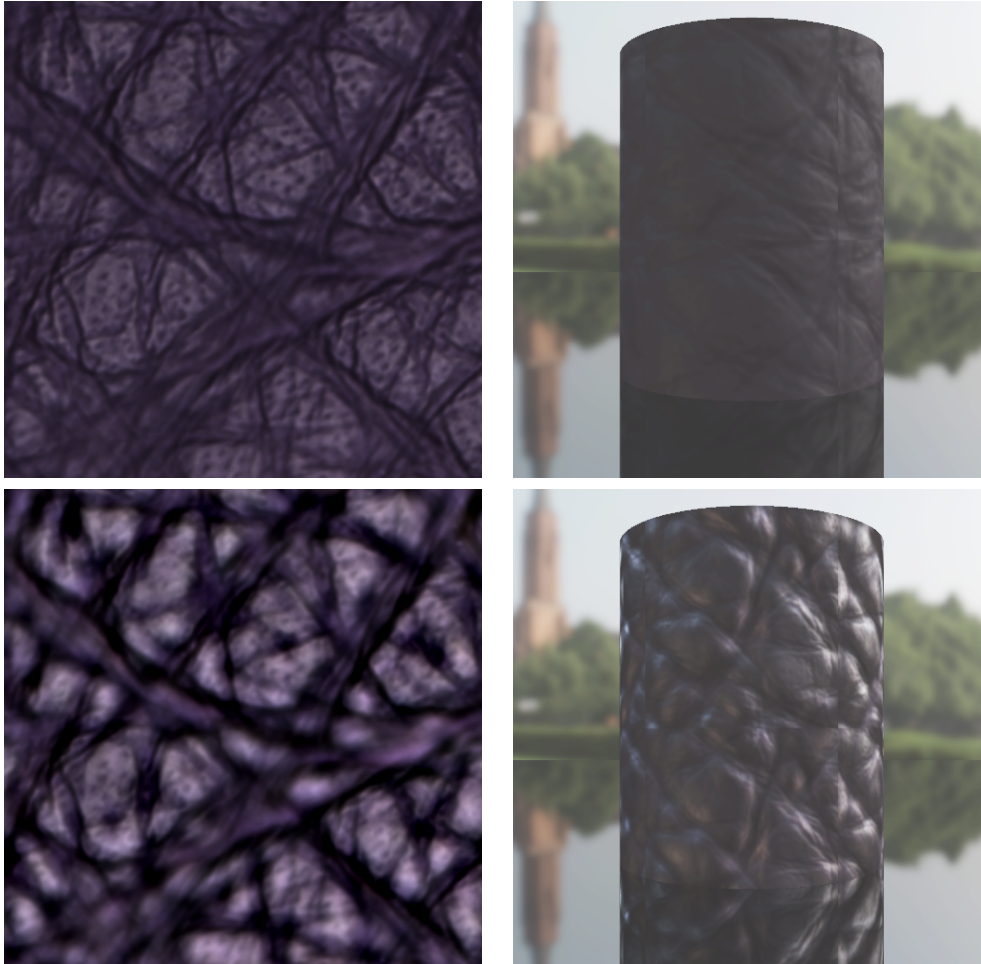


Figure 4.11: Flat texture (left) vs. 3D presentation with environment light (right). The top left image shows the input texture. The top right image shows the rendered material. On the bottom left you can see the influence of a strong undulation edit and on the right you can see the effect of this edit on the rendered material.

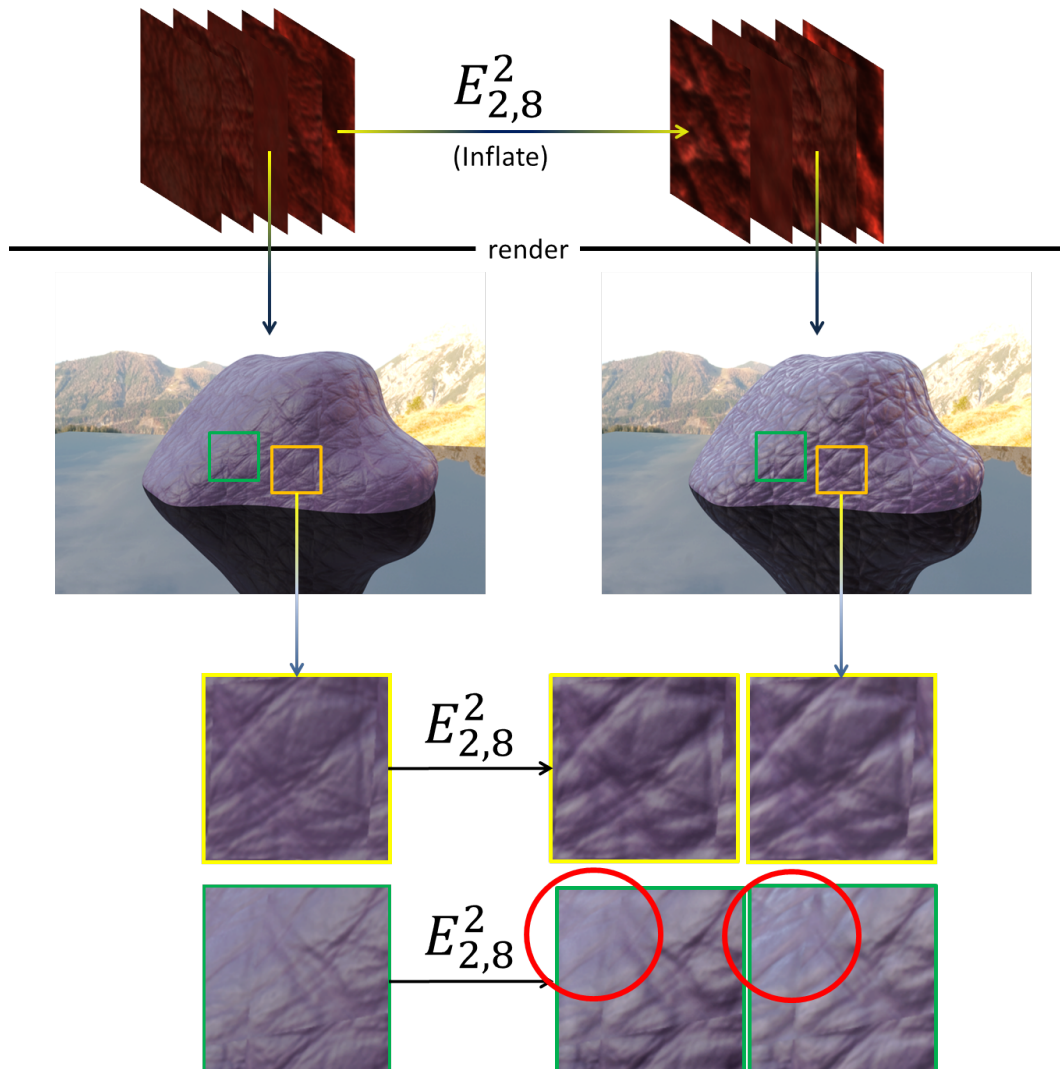


Figure 4.12: Comparison between material bending and image editing. In Flat regions there is no difference between the bending operation applied to the material base and the appropriately scaled bending operation applied to the rendered image. But the image operator is not capable of editing rendered surfaces under changed perspective as may be seen by comparing the image sections with the green boundary. Particularly in the red surrounded region the image operator fails.

show that pure manipulations of the image spectrum may not account for perspective distortion.

4.3.3 Performance

On compressed BTF data with 100 (Y), 50 (U) and 50 (V) principal components, data processing took no more than 0.5 s (Fourier), 1.2 s (D1 wavelet) and 4.8 s (D6 wavelet). For uncompressed data, the time scales up linearly with the number of textures, resulting in processing times of about 5 minutes for applying the a Fourier-based operator and 12 to 45 minutes for wavelet-based edits.

4.4 Conclusion and Physical Aspects

In this chapter we have transferred the image based affordance-manipulation scheme by Giesel and Zaidi to a BTF-affordance-editing system. The edits performed by appearance bending do, of course, not change the surface geometry. When moving an object that is textured with a strongly edited BTF, the parallax effect (or lack thereof) may result in a lack of realism. Above that the changes in singular textures are not necessarily consistent with changes in the other textures of a given BTF. To stay physically correct the reflectance-values over any BRDF have to integrate to a number smaller than 1. It can not be guaranteed, that the ABRDFs of an edited material are still correct with regard to this law.

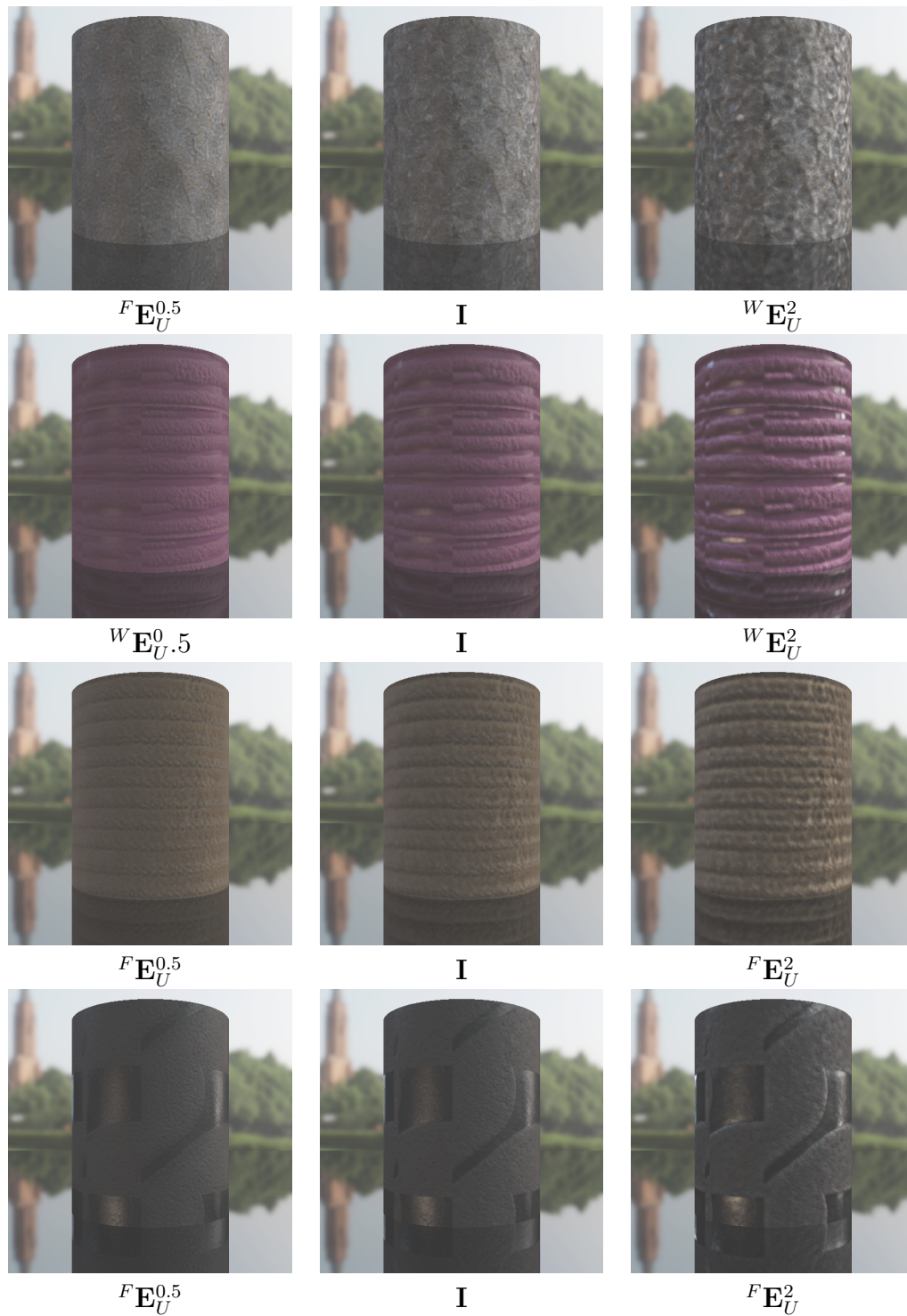
In the current version, appearance bending can only strengthen or weaken such qualities that are present in the original material. The transfer of frequencies from one material to the other would be an interesting prospective and will be the subject of further investigation.

Still Appearance bending, or the editing of materials by scaling bands of spatial frequencies, has proven to be a versatile tool. The transfer from images to complex material representations such as the BTF yields results that match

the original data in realism.

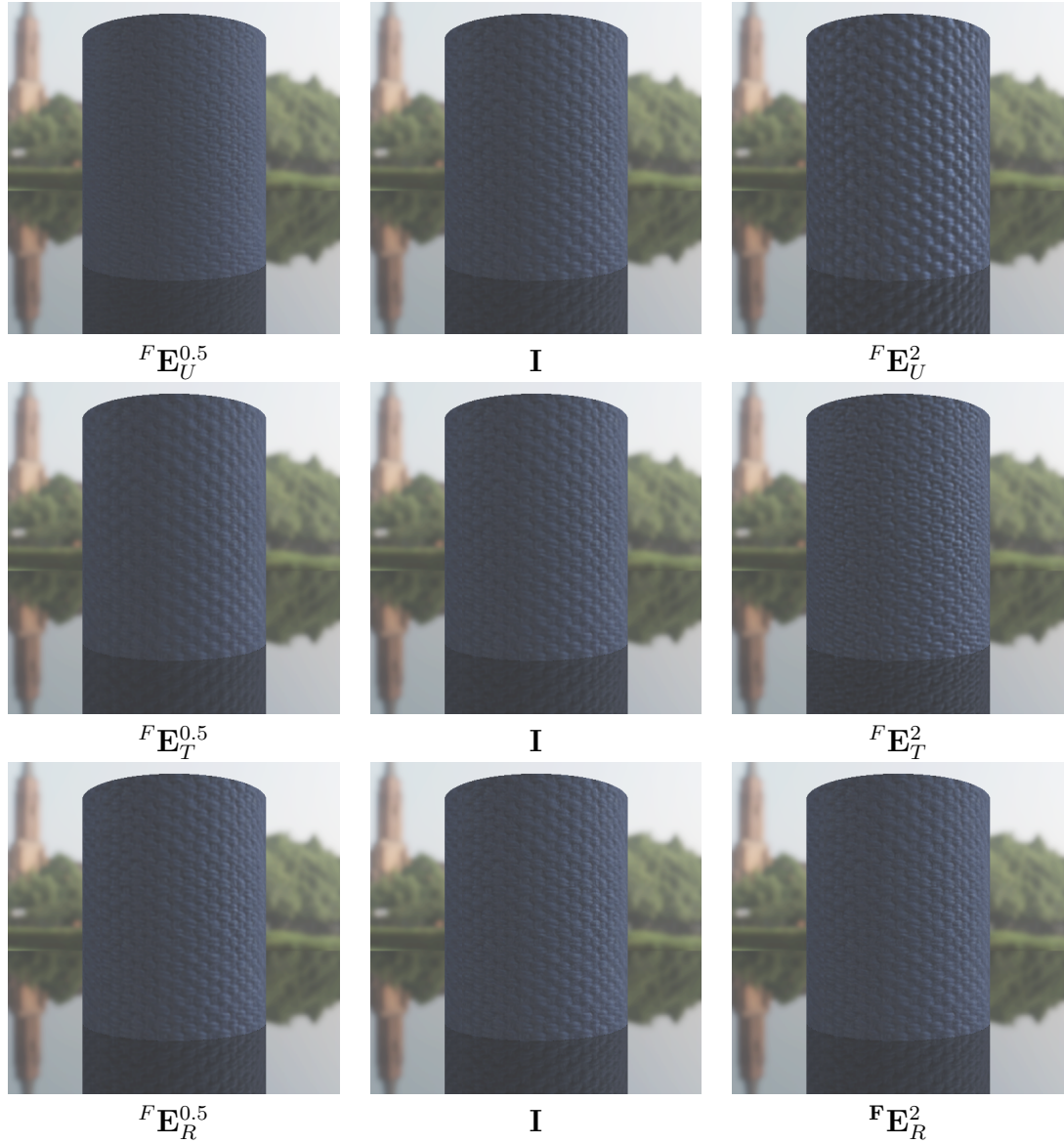
We are hopeful that further developments in psycho-visual perception will further contribute to advance the state of the art in semantically meaningful and predictable editing processes for appearance.

In the following chapter we want to deepen our understanding of the frequency band manipulations by adapting the frequency filtering system to current models of the visual cortex. It remains to be seen to which extent a deeper understanding of human perception can be used to further increase the meaningfulness and realism of editing interfaces for appearance.



The undulation operator applied to different materials.

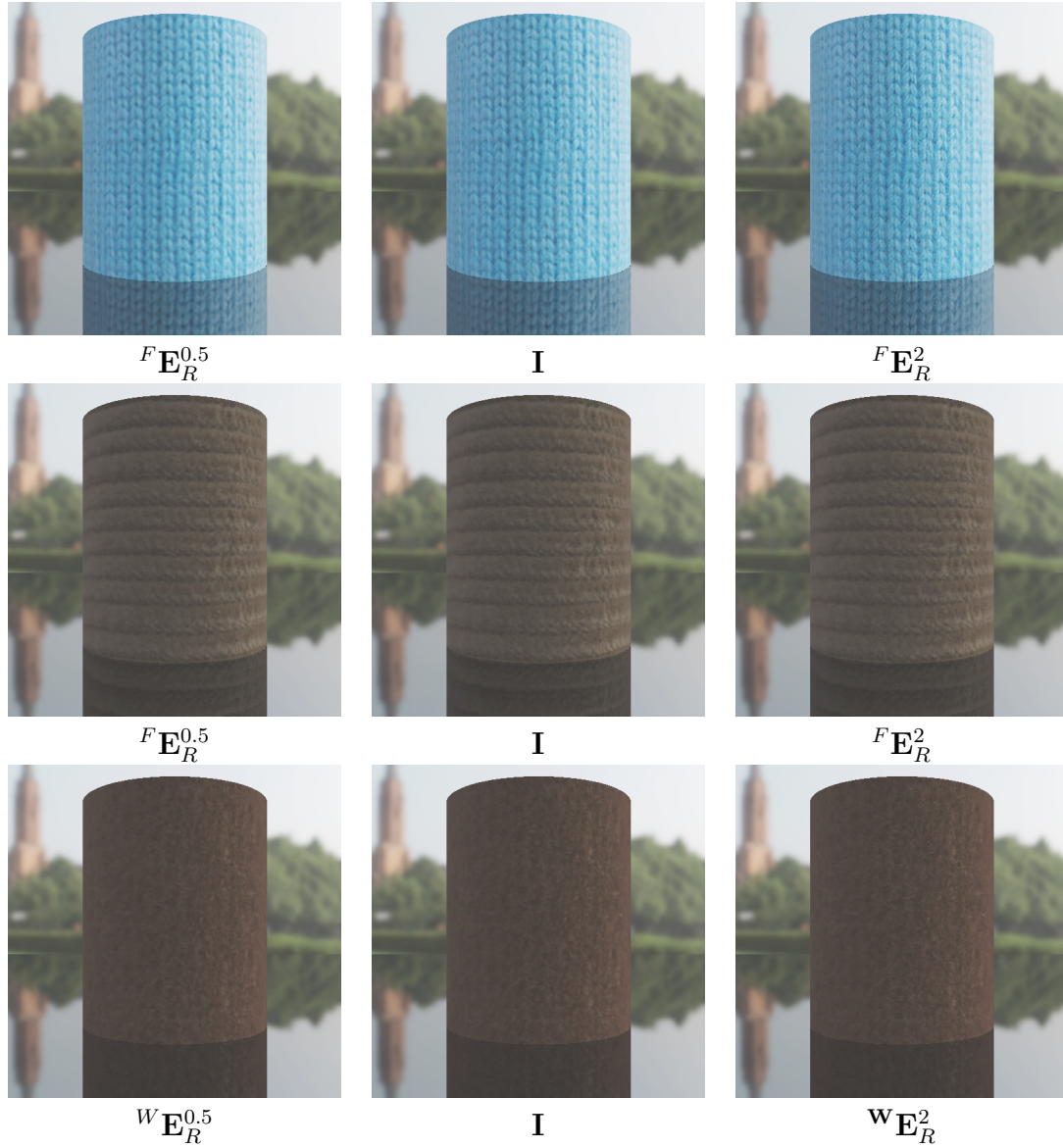
Figure 4.13: Result renderings. For a discussion, see section 4.3.



Different edits on the same material (blue wallpaper).

From top to bottom: undulate, thicken, roughen.

Figure 4.14: Result renderings. For a discussion, see section 4.3.



Effect of the “roughen” operator on three different materials.
 From top to bottom: Wool, Corduroy, Carpet.

Figure 4.15: Result renderings. For a discussion, see Section 4.3.

CHAPTER 5

A Cortical Appearance Space Model

In the current chapter we suggest to use a model of the primates striate cortex to learn linear material editing operations. The content of this chapter is based on a peer-reviewed publication:

Marlon Mylo, and Reinhard Klein. Linear Subspaces of the Appearance Space. In *International Conference on Computer Graphics, Visualization and Computer Vision*, volume 26 of *Journal of WSCG*, pages 95 – 103, [MK18]

All self-citations in the current chapter refer to this publication.

The American cognitive scientist Roger Newland Shepard opens his publication *Perceptual-cognitive universals as reflections of the world* [She01] with the words: *The universality, invariance, and elegance of principles governing the universe may be reflected in principles of the minds that have evolved in that universe-provided that the mental principles are formulated with respect to the abstract spaces appropriate for the representation of biologically significant*

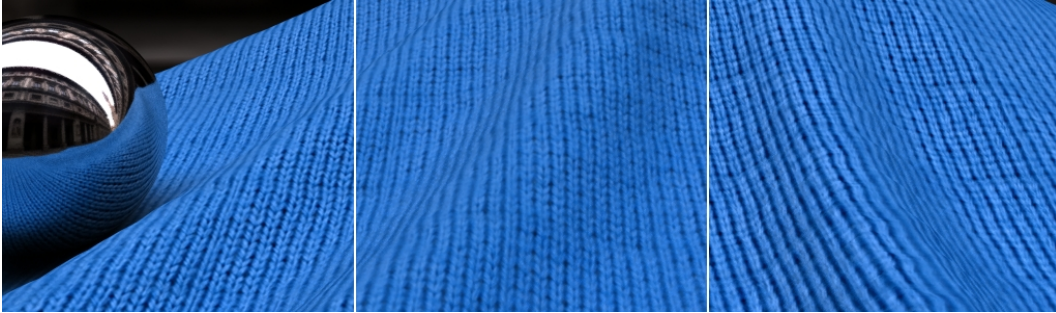


Figure 5.1: Rendering of an edited wool-BTF. The left part of the image shows the result of a combination of the edge aware operator and the thickening operator (sections 5.4.2 and 5.4.3), the right image shows the result of the corresponding band-pass filter, according to [MGZ⁺17] and in the middle, we show the original material.

objects and their properties.

We want to follow this train of thought. The work of Marcelja and Daugman [Mar80,Dau80] prepared the ground for Gabor-filter based models of the simple cells of the primary visual cortex (V1 or striate cortex). Bandpass filters (see Chapter 4) may be realized by combining Gabor-filters (see Figure 5.4). The depth or the roughness of a surface is without doubt a *biologically significant* property. So if we accept the Fourier transform as a very rough computational model of the striate cortex, we may consider the band-scaling as a *principle of the mind* which reflects physical properties like thickness and undulation. This is the motivation for composing an empirically born computational model of the V1-region of the visual cortex. The idea is to find physical phenomena which have a simple representation in our new model. We will show that our new cortex model allows for learning linear representations of complex physical phenomena. Such a linear representation enables plausible, predictable and meaningful manipulations and is therefore predestined as editing operation.

According to Poggio and Serre [PS13], *models of the visual cortex provide the much-needed framework for summarizing and integrating existing data and for planning, coordinating, and interpreting new experiments.* Poggio and Serre concentrate on computational models of vision. They state, that biological

plausibility is elusive. There exist several different computational models of the visual cortex, mainly as feedforward-model. We assume a feedforward hierarchical model of the visual cortex and we use empirically data to estimate parameters like the distribution of the frequency resolution of the receptive fields.

Using models of the visual cortex has a long tradition in computer vision for pattern recognition tasks and for the description of perceptual image-metrics but it is not yet an integral component of computer graphics. We will introduce the term *appearance space* which has mostly been used implicitly [LH06,HB95]. We argue, that the set of all possibly occurring neuronal states in the visual cortex may be seen as this appearance space. So given a computational cortex transform model, we may define a *computational* appearance space as part of it. Seeing computer graphics from the perspective of human physiology is fruitful: Bayer filter in digital image sensors follow the cone distribution in the retina, retinal displays mimic the cone-density and photo-sensors filter and cumulate the incoming energy comparable to photo receptor cells. Our approach is a kind of frequency editing. Frequency editing is a very old technique. Blurring e.g. has already been used hundred years ago in silent film. But it has mostly not been seen as an operation in the visual cortex but merely as a given visual effect.

This work contributes to the field of material editing by presenting a system to transfer image manipulations into a model of the visual cortex which in many cases brings better results than the original editing scheme and we will provide a novel realistic material manipulation, namely frequency based moving of a light source.

5.1 Cortex Models

Our understanding of the structure and the modes of action of the animals visual cortex goes back to the work of Hubel and Wiesel during the late 50. and 60. of the last century [HW59,HW62,HW68]. Twenty years later Daug-

man fitted Gaussian and Gabor-filters to the cortical responses measured by Hubel and Wiesel [Dau80, Dau88]. It is noticeable that neural networks develop Gabor-filter like structures by their own, when trained with random input [San90]. Olshausen and Field found that optimizing a vector base for sparse linear coding of images leads to a set of Gabor-like base vectors which is in spatial frequency and orientation coverage comparable to the filtering system in the visual cortex [OF96, OF97]. A publication which concentrates on mathematical aspects of the Gabor-filter-systems compatible with the neural responses of the V1-cortex is the work of Lee [Lee96]. Lee gives explicit parameters for his filter systems and calculates the tightness of the Gabor-frames. A good overview over publications on cortical parameter measurements may be derived from [MR02, Table 1]. Recently Huth et al. published a python-toolbox for simulations of early vision [HMA17].

The presented approach stands in the tradition of the pyramid-based texture analysis and synthesis published by Heeger and Bergen in 1995 [HB95]. Heeger and Bergen use steerable pyramids to model the behaviour of the visual cortex. Gutman and Hyvärinen derive a probabilistic model of image statistics by modeling two cortical layers of simple and complex cells [GH13]. This publication may also be consulted for further references to Bayesian perception. In her dissertation Diana Turcsány [Tur16] uses a convolutional neural network to model the deeper levels of the visual cortex for image editing.

Glassner gives a good computer graphics oriented introduction into the visual system [Gla95].

5.2 Our New Cortex Model

The insight that light is not coloured but that the energy in a light beam provokes a sensation of colour goes back to Newton [New04]. In their trend-setting paper *Pyramid-Based Texture Analysis/Synthesis* Heeger and Bergen used the word *appearance* to bridge the gap between the sensation of a texture and the physical phenomena on the surface of the texture [HB95]. We can

locate the term appearance between the sole occurrence of physical phenomena and the only inter-subjectively examinable set of sensations by identifying the *appearance space* with the set of all neural response-states in the visual cortex. In the ventral stream of the human visual system, the primary visual cortex follows after the *lateral geniculate nucleus (LGN)*. Mathematically the receptive field is described by the *receptive field profile (rfp)*, a bivariate real-valued function which simulates the neuronal response to a given stimulus [Dau80]. As receptive fields have directly been measured while exposing the macaque retina to visual stimuli, the influence of the LGN is an implicit part of the model but does not have to be modelled explicitly.

As frame for our (computational) appearance space we will use a *cortex transform model* [W⁺87] which we will derive from empirical data (section 5.2.1).

Our formal scaffold consists of

1. a *model of the space of retinal responses*,
2. a *model of the neural responses of simple cells in V1*,
3. a *model of the visual stream* from the retina to the neural response and
4. an *interpretation model for the retinal responses*.

The space of retinal responses describes the entrance of pictorial data into the visual system. We will use RGB-images with an edge-length of 256 pixels. Decorrelating the color space as in [HB95, Sec. 3.5], lead to strong artefacts. Confining the manipulations to the value-channel of the HSV-color-space brought good results. So we define $\mathbb{I} := [0, 1]^{256 \times 256}$ as retina model. The images, we use for testing, correspond to real-world patches with an edge length of approximately 5 cm. If a patch of this size has a distance of 57 cm from the observer, its retinal image approximately covers the fovea.

Our model of the visual stream is limited to the early ventral stream up to V1. While there have been suggested different filters for modeling V1-receptive fields [Lin13], we use Gabor-filters [Dau80], (see Section 5.2.1). Our whole V1-model consists of a filter bank of 517 filters $\{\gamma_\xi\}_{\xi \in \Xi}$ (see section 5.2.3) and we define the operators $\Gamma_x := \gamma_x \star$.

The space of neural responses will be modelled as a stack of matrices $\mathbb{G} := \mathbb{R}^{256 \times 256 \times 517}$. We do not limit the amplitude of neural responses. It is not self-telling, that the spatial dimension of the neural responses (256×256) equals the dimension of \mathbb{I} (see Paragraph 5.2.3) but it enables a direct comparison between the input and the result of the V1-transform.

The interpretation space is a set of mappings $\mathbf{G} : \mathbb{G} \rightarrow \mathbb{I}$ with $\mathbf{G} := \sum_\xi a_\xi \{\Gamma_\xi \mathcal{T}\}$.

Now we define the *appearance space* \mathbb{A} as the image of the interpretation space. This leads to the following diagram, modeling the relations, depicted in Figure 5.2:

$$\begin{array}{ccccc}
 \mathcal{T}_{\in \mathbb{I}} & \xrightarrow{\{\Gamma_\xi\}} & \{\Gamma_\xi \mathcal{T}\}_{\in \mathbb{G}} & \xrightarrow{\mathbf{G}_\mathbb{I}} & \mathbf{G}_\mathbb{I} \mathcal{T}_{\in \mathbb{A}} \\
 \mathfrak{P} \downarrow & & & \searrow \mathbf{G}_\mathfrak{P} & \downarrow \mathbf{E}_\mathfrak{P} \\
 \mathfrak{P}(\mathcal{T})_{\in \mathbb{I}} & \xrightarrow{\{\Gamma_\xi\}} & \{\Gamma_\xi \mathfrak{P}(\mathcal{T})\}_{\in \mathbb{G}} & \xrightarrow{\mathbf{G}_\mathbb{I}} & \mathbf{G}_\mathbb{I} \mathfrak{P}(\mathcal{T})_{\in \mathbb{A}}
 \end{array} \tag{5.1}$$

The filter bank $\{\Gamma_\xi\}$ maps the texture \mathcal{T} to the neural response space \mathbb{G} . By the definition of \mathbb{G} and \mathbb{A} , we may identify \mathbb{I} and \mathbb{A} . Neural responses are recombined to a texture $\mathcal{T}_\mathcal{X} := \mathbf{G}_\mathbf{X} \mathcal{T}$ in the appearance space. A physical phenomenon \mathfrak{P} induces a mapping from the appearance space $\mathbf{E}_\mathfrak{P} : \mathbb{A} \rightarrow \mathbb{A}$ to itself (compare with Figure 3.3). If we identify $\mathbf{G}_\mathbf{X}$ and $\mathbf{G}_\mathbf{X} \circ \{\Gamma_\xi\}$, the operator $\mathbf{G}_\mathfrak{P}$ may be constructed as linear approximation of $\mathbf{E}_\mathfrak{P}$.

$$\mathbf{G}_\mathfrak{P} \approx \mathbf{E}_\mathfrak{P} \tag{5.2}$$

Note that $\mathbf{G}_\mathbb{I} \approx \mathbf{I}$ is an approximation of the identity on \mathbb{A} . $\mathbf{G}_\mathfrak{P}$ is the operator,

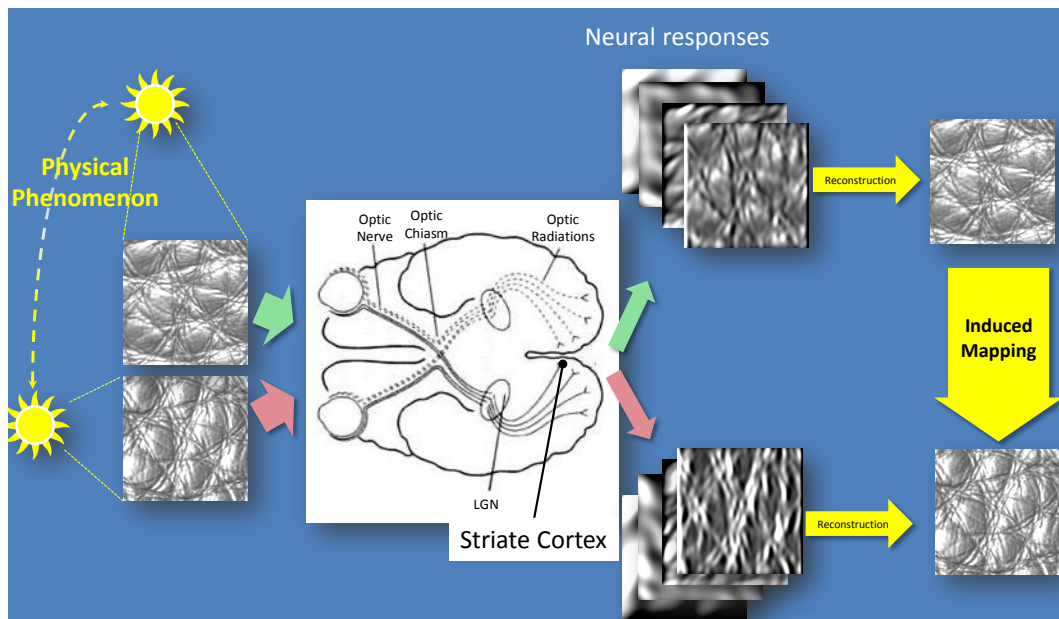


Figure 5.2: Visual path of a material patch, seen under different optical conditions. The physical phenomenon induces a mapping in the space of cortical responses. Brain drawing taken from <http://universe-review.ca/I10-85-opticpath.jpg>.

we want to learn. For a full clarification of the symbols, see the following section.

5.2.1 The Computational Model of the Early Vision

In this section we will introduce the V1-model. The concept that the neural response of a simple cell in V1 cortex is linear in the intensity of the incoming optical stimulus is essential not only for the model of the visual pathway [Dau80, AP79] but also for all measuring methods of the receptive fields like subspace reverse correlation [RSS97]. The function describing the weighted contribution from each position of the receptive field to the response of this cell is called *weighting function* and may be modelled by a linear filter [W⁺87].

This means that if we may reconstruct a texture by summing up the filter responses according to the receptive field system of the visual cortex, we may influence the neural response in the V1 directly.

5.2.2 An Empirically Based Model of the Visual Cortex

There exist many publications on the frequency distribution in Macaque V1-area [FGNP85, Sel16]. We used empirical data, measured and fitted by De Valois et al. [DVAT82]. We use two dimensional Gabor-base functions for spatial frequency filtering [Dau80].

Our filter function will be the 2D-Gabor filter from Equation 2.58. Here we confine to the real part and obtain:

$$\gamma_{\nu, \sigma_u, \sigma_v, \beta}(u, v) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\left(\frac{u}{\sigma_u}\right)^2 + \left(\frac{v}{\sigma_v}\right)^2\right)} \cos(2\pi\nu u + \beta) \quad (5.3)$$

The preimage of the directional standard deviation of the Gaussian envelope forms an ellipse. The semi-minor axis, here the u -axis, of this ellipse is according to [Lee96] and [JP87, Fig. 8A] parallel to the wave-vector of the carrier. We confine to a real plane-wave (see 5.2.3). So, with $\xi := \{\nu, \sigma_u, \sigma_v, \beta, \varphi\}$, we may define:

$$a\Gamma_{\xi, \mathbf{p}} := a\gamma_{\nu, \sigma_u, \sigma_v, \beta} \circ R_\varphi \circ T_{\mathbf{p}}(x, y) \quad (5.4)$$

R_φ and $T_{\mathbf{p}}$ are the rotation and the translation as given in equation 2.55.

5.2.3 Parameters

To compose the Gabor filter bank, we have to specify the parameters. We distinguish between the parameters, which we set up according to given publications in the field of neuro-science (the parameter set ξ , Paragraph 5.2.3), the position of the filter center \mathbf{p} (Paragraph 5.2.3) and the amplitude a (Paragraph 5.2.3), which we will use for the definition of the editing operator \mathbf{G} .

The parameter-set ξ

The parameter-set ξ contains all parameters which have to be distributed according to measurements in the macaques or in the cats striate cortex.

The spatial frequency $2\pi\nu$

In the visual cortex, frequency sensitivity occurs not in exact but in rough steps of 0.3 to 0.5 octaves. As we drew the spatial frequency according to [DVAT82, Fig. 6.], we were limited to the bin width in this figure, which is 0.5 octaves.

Differences between human and macaque visual system The monkey visual system as model for the human visual system has been validated under several different aspects [RSD15]. While the human visual system is from an anatomical and physiological perspective extremely similar to the macaque visual system, it has a slightly higher retinal magnification factor (about 0.291/0.223), which hints to a higher angular resolution [MK90]. Therefore we add another frequency bin at 20.8c/° and so we have to extrapolate to a plausible histogram of the human frequency distribution.

The distribution given by de Valois In [DVAT82, Fig. 6.] De Valois et al. describe their measurements of the spatial frequency distribution of the

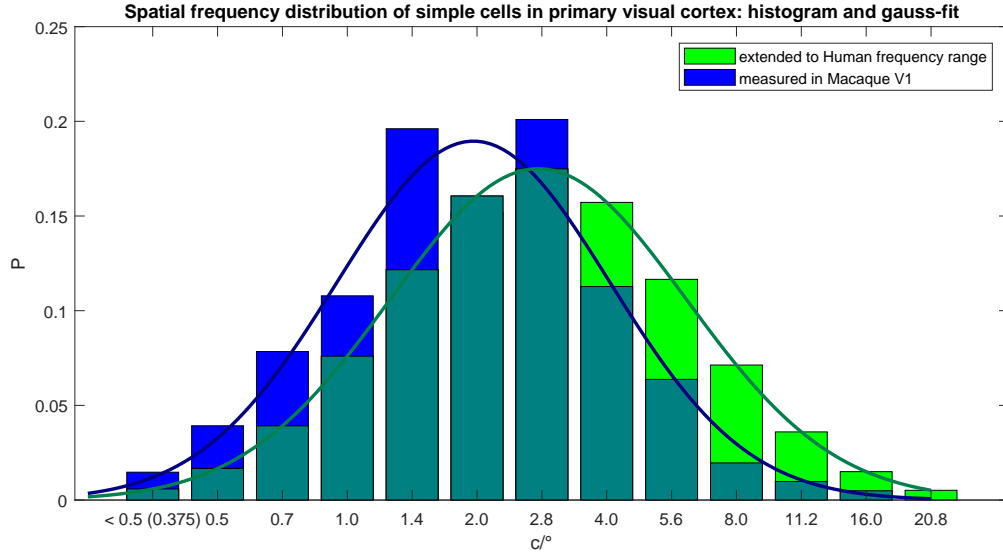


Figure 5.3: The blue part of the bars shows the histogram given in [DVAT82]. The green part describes the extrapolation results and has been added to account for the slightly wider frequency range of human vision [MK90].

receptive fields of simple cells in macaques primary visual cortex. They distinguish between the cells with receptive fields in the fovea and in the parafovea region of the retina. We assume that our texture covers a visual angle of 5° . As we cannot expect observers to concentrate on a texture without any eye-movement, we merged the distributions for the fovea and the parafovea by normalized summation. The blue (including green-blue) bars in Figure 5.3 belong to the merged histogram from [DVAT82]. To extend this histogram to the slightly bigger frequency range of the human vision, we fitted a gaussian by an iterative Least Mean Square algorithm, moved the mean of the gaussian to the logarithmic middle of the new frequency distribution range and stretched the standard deviation proportional to the ratio of logarithmic ranges.

The following table shows the number of filters we have in every frequency bin. The absolute number of 517 filters has been chosen in order to have a

good fitting to the histogram and still stay comparable with the reconstruction scheme for tight frames (Section 5.4.1).

$c/^\circ$	0.4	0.5	0.7	1.0	1.4	2.0	2.8	4.0	5.6	8.0	11.2	16.0	20.8
# filters	3	8	20	39	64	84	92	82	61	37	18	7	2

The standard deviation in direction of the wave-vector σ_u

σ_u and ν are connected via the bandwidth. As Gaussian kernels have infinite support, the bandwidth is defined as *half amplitude bandwidth*. Bandwidths have been drawn on base of [DVAT82, Fig. 7]. In this diagram, De Valois et al. visualized the bandwidth with standard deviation as a function of the spatial frequency. As spatial frequencies were known, bandwidth-samples could be drawn under the assumption of normal-distribution within the same frequency range. Given the bandwidth B and the spatial frequency $2\pi\nu$, we may calculate:

$$\sigma_u = \frac{\sqrt{2 \ln 2} ((2^B + 1)/(2^B - 1))}{2\pi\nu} \quad (5.5)$$

The standard deviation orthogonal to the wave-vector σ_v

According to [Rin02, FIG. 4.], there is a relation between $\nu\sigma_u$ and $\nu\sigma_v$. This relation may be interpreted as functional graph with a small deviation. To make use of this relation, we fitted a cubic spline to the data and used this spline as function graph.

The Phase angle β

To draw the phase parameter β , we used the histograms given in [Rin02, FIG. 7A/B].

The orientation φ

By definition of $u = (x - p_x) \cos \varphi + (y - p_y) \sin \varphi$, φ is the angle between the u and the x -axis. We drew the orientation equally distributed from $\{i \frac{\pi}{8}\}_{i \in \{1, \dots, 16\}}$. Where possible, directions have been drawn in orthogonal pairs.

All random experiments have been done in several passes and brought comparable results. The set of all parameter-sets ξ in the Gabor-filter bank, will be denoted by Ξ .

The position \mathbf{p}

Every neural measurement provides us just one sample of the domain of neural responses. Be γ the Gabor filter, best fitting the receptive field of a given neuron with filter center \mathbf{p} : now the neural response is modelled as $a \langle \gamma_{\mathbf{p}}, \mathcal{T} \rangle$.

$\langle \rangle$ is the standard inner product in the image domain. γ has to be appropriately sized and evaluated on the spatial grid of the image and \mathcal{T} has to be zero-padded, where necessary. The filter-centers are often chosen to be elements of the spatial image grid ($\mathbf{p} \in \{1, \dots, 256\}^2$) [Dau88], sometimes with the constant stride ($h := p_{z_{i+1}} - p_{z_i}$) between consecutive grid points increasing with an increasing wavelength and/or starting with a stride smaller than one (e.g. [Lee96]). In order to make use of the convolution theorem and to avoid a resampling step we will assume the parameter set ξ to be constant over the whole grid and set the stride $h = 1$ to one and keep the image-grid. Nevertheless we have to emphasize that our approach might distract the statistics: as the statistics of DeValois et al. [DVAT82] are based on the measurements of individual cells, a higher spatial resolution goes to the cost of the angular resolution and the variety of the phase values. Particularly in the case of low frequencies, the spatial domain is probably oversampled. The results of the undulation experiment 5.4.3 might indicate this problem (see figure 5.11).

To locate the neural response, we multiply it by the canonical base matrix $\mathbf{e}_{\mathbf{p}} \in \mathbb{R}^{256 \times 256}$ at position \mathbf{p} and sum those matrices up $\sum_{\mathbf{p}} \langle \gamma_{\mathbf{p}}, \mathcal{T} \rangle \mathbf{e}_{\mathbf{p}}$. As we confined to real-valued Gabor-base functions (Equation 2.58), meaning $\gamma^* = \gamma$,

we may write that summation-formula as cross-correlation \star :

$$a \sum_{\mathbf{p}} \langle \gamma_{\mathbf{p}}, \mathcal{T} \rangle e_{\mathbf{p}} = a\gamma \star \mathcal{T} = a\Gamma\mathcal{T} \quad (5.6)$$

Where we appoint the amplitude a to be fixed for a changing position \mathbf{p} . We should hold the following point of criticism in mind: for deeper frequencies a sparser sampling of the image-plane is probable.

The amplitude a

We use the amplitude to combine filter responses to operators. Be $\mathbf{E} : \mathbb{A} \rightarrow \mathbb{A}$ an operator, then we want to find $a_{\xi}^{\mathbf{E}}$ to approximate \mathbf{E} (see section 5.3.1):

$$\mathbf{E} \approx \sum_{\xi \in \Xi} a_{\xi}^{\mathbf{E}} \Gamma_{\xi} \quad (5.7)$$

This mapping operates via cross-correlation, it may be visualized by applying it to the discrete dirac $\delta \in \mathbb{A}$.

Gabor filtering systems are often described by a log-polar overview of the frequency responses in the euclidean plane (see Figure 5.4).

5.3 Transferring Edits to the Model of the Visual Cortex

Now, that we have introduced our model of the visual cortex, we want to introduce the operators. First we will discuss the editing scheme itself and how to transform into it. Than we will present the editing paradigms to transfer.

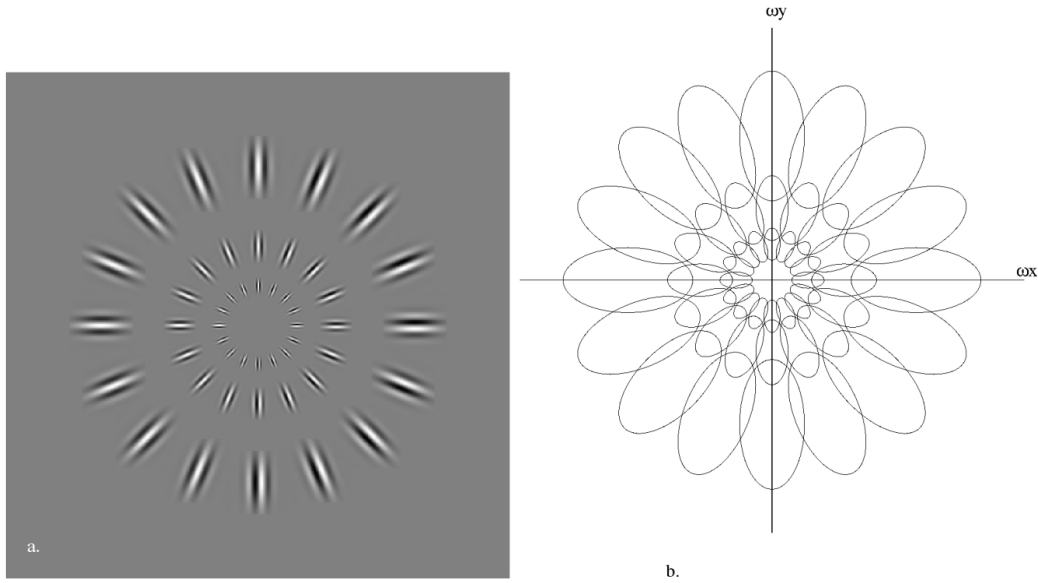


Figure 5.4: From [Lee96]: an ensemble of Gabor wavelets in the frequency domain.

5.3.1 Learning an Operator

To transfer a given edit, we take a collection of test-textures $\mathcal{T}_{i \in \{1, \dots, m\}}$ and solve

$$\mathbf{a}_i^{\mathbf{E}} = \min \arg_{\mathbf{c} \in \mathbb{R}^n} \left\| \mathbf{E} \mathcal{T}_i - \sum_{\xi \in \Xi} c_{\xi} \Gamma_{\xi} \mathcal{T}_i \right\|_2 \quad (5.8)$$

for each texture \mathcal{T}_i . We could define

$$\mathbf{a}^{\mathbf{E}} := \min \arg_{\mathbf{c} \in \mathbb{R}^n} \sum_i \left\| \mathbf{E} \mathcal{T}_i - \sum_{\xi} c_{\xi} \Gamma_{\xi} \mathcal{T}_i \right\|_2 \quad (5.9)$$

but this definition lead to undesired activities in higher frequency-bands. Instead, we apply a singular value decomposition to $\mathbf{A} := a_{i\xi}$:

$$\mathbf{A} = \mathbf{U} \Sigma \mathbf{V}' \quad (5.10)$$

and use the base vector $\mathbf{a}^{\mathbf{E}} := (a_{\xi}^{\mathbf{E}})_{\xi \in \Xi} := (V_{\xi 1})_{\xi \in \Xi}$. So we may declare our new editing operator

$$\mathbf{G}_{\mathbf{E}} := \sum_{\xi \in \Xi} a_{\xi}^{\mathbf{E}} \Gamma_{\xi} \quad (5.11)$$

5.3.2 The Operators

We will explore four different operators: the identity, linear edge enhancement, bandpass filters and spotlight moving.

The identity

The first operator maps the image to itself. This is a reconstruction. There is no canonical reconstruction scheme for Gabor-Wavelets as they are overcomplete. There has been many efforts to produce models of the visual cortex which had good mathematical properties [SA00, Lee96]. Lee introduces a reconstruction scheme which relies on the tightness of the frame [Lee96, DS52]. There are many approaches, to adjust the filter responses of a Gabor-filter bank to a partition of unity in the frequency domain [W⁺87].

Edge aware imaging

Edge aware imaging is a very important editing operation. An overview over existing approaches can be found in Paragraph 2.3.2-**Edge Aware Imaging**.

Bandpass filtering

Here we want to learn the editing scheme presented in Chapter 4.

Moving spotlight

Given a directionally illuminated texture patch. We will show, that it is possible in our model to learn and reproduce small movements of the light source.

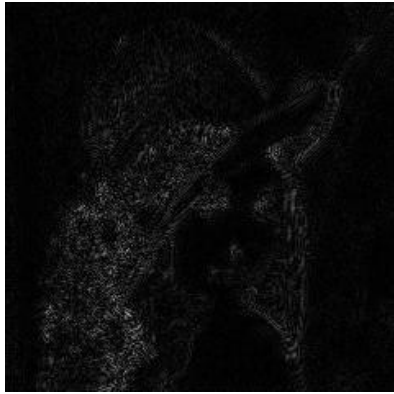
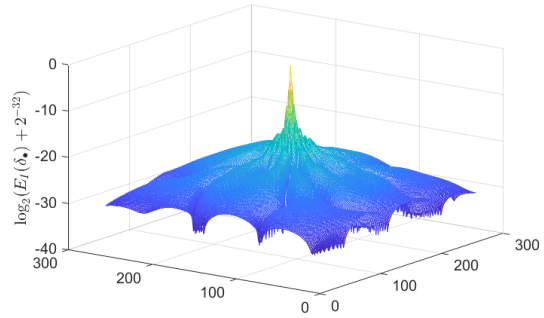
(a) $500 \cdot |\mathcal{L} - \mathbf{E}_I(\mathcal{L})|$ (b) $\log(\mathbf{E}_I(\delta) + 2^{-32})$

Figure 5.5: 500 times amplified reconstruction error of Lena image. Boundary cut off in a distance of 10 pixel. (a). To get a better impression of the delta-spike, we added 2^{-32} and applied the binary logarithm (b).

5.4 Evaluation

In this section we want to show and discuss some results. The presented results have been calculated on colour or reflectance maps. The colour-values had been stored as half-precision floating point-values (16 bit) and the reflectance-values had been stored as full-precision floating points (32 bit). All operations on the HDR-images had been performed in log-space. As training samples, we used textures from the USC-SIPI Image Database from the University of Southern California and the describable texture dataset [CMK⁺14].

5.4.1 Identity

Figure 5.5 visualizes the learned identity operator. The difference between the reconstructed and the original image is with bare eyes intractable. The maximum pixel intensity difference between the Lena image (\mathcal{L}) and the reconstruction of it was $\max_{ij} |\mathcal{L}_{ij} - \mathbf{G}_I(\mathcal{L})_{ij}| = 6.9 \cdot 10^{-3}$ which corresponds to 2 steps in an 8 bit grey-scale image. While such a small deviation will not stand-out when affecting the intensity channel, sensible people might perceive

colour aberrations if the operator was applied channel-wise to an RGB-image.

As it is not possible, to reconstruct an Gabor-filtered image perfectly, we will compare against the approximative reconstruction scheme, presented by Lee [Lee96]: a frame $\{\psi_I^\gamma\}$ (for a definition and constraints on the parameter set I , see [DS52, Lee96]) is tight when the following equation holds for a given constant c and a small positive number ε :

$$\forall \mathcal{T} : \quad c \|\mathcal{T}\|^2 \leq \sum_I |\langle \psi_I^\gamma, \mathcal{T} \rangle| \leq (c + \varepsilon) \|\mathcal{T}\|^2 \quad (5.12)$$

Lee investigated for which parameter sets I this frame becomes a tight frame ($\varepsilon \searrow 0$). Note that Lee uses complex-valued Gabor-base functions, which does not make sense in our setting as we do not apply filters to filtered values and have therefore no complex multiplications. In his definition of the Gabor-base functions, the amplitude a is part of the definition of ψ^γ and the position \mathbf{p} is an element of the parameter set I and he uses a pyramid sampling scheme. For a tight frame the following reconstruction formula may be applied:

$$\mathcal{T} \approx \frac{2}{2c + \varepsilon} \sum_I \langle \mathcal{T}, \psi_{\gamma I} \rangle \psi_{\gamma I} \quad (5.13)$$

To compare against [Lee96], we sample over 16 directions θ , made three steps per octave and set the stride h to 0.5. This yields a value for ε of approximately 0.0001, the number of base-vectors was 864.

Figure 5.6 visualizes that even with this very tight frame the quality of this reconstruction scheme is not high enough to allow for applications in computer graphics. We cannot expect the frame to be tight enough [SA00].

5.4.2 Edge Aware Imaging

For this edit, we learned randomly linear edge aware filtering kernels: we used Gabor and Sobel-filters $((1, 2, 1)' \otimes (1, 0, -1))$, we will write: \mathbf{SX}, \mathbf{SY}). Note that the parameters of the filter kernels and the intensity of the filters had been



Figure 5.6: The upper right of the image shows the Lena image reconstructed by the formula 5.13. The wavelet family forms a frame with $\delta = 0.0001$. The lower left shows our reconstruction.

drawn randomly and so they were in general not in the set γ_{Ξ} . Intensities were always enhanced. We used 1000 editing samples of varying photos for learning. The resulting filter (Figure 5.8) may be seen as the average of all projected filter-kernels. It is a good approximation of the sign-inverted discrete Laplace operator with weights on the diagonals. In comparison with other state of the art edge aware imaging operators (Figure 5.7), it is noteworthy, that the learned operator enhances very fine structure and the material still looks realistic. A physical effect, bound to this appearance, is a higher fibrousness.

5.4.3 Affordance Editing

In this section we will compare our results against pure frequency band scaling. While there is evidence, that the frequency-bands are subject to a recognition step [GZ13] and consecutively to a scaling step in the visual cortex, according to the original perceptual studies, an edge length of a material patch should cover a viewing angle of 3.5° . With an edge length of 5 cm on the monitor,

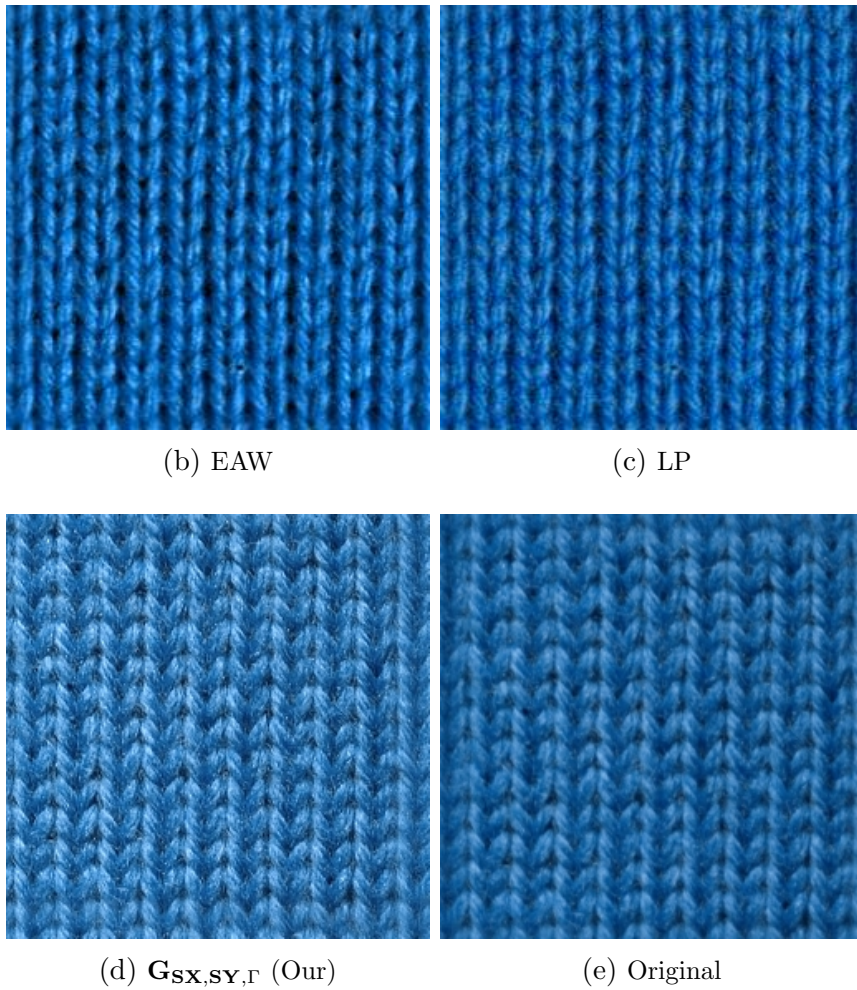
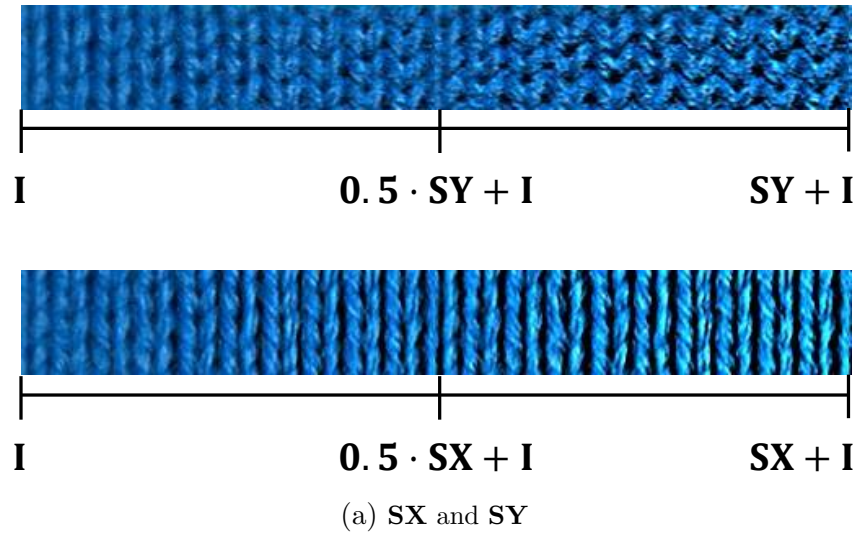


Figure 5.7: In the top rows you can see the increasing influence of the Sobel filters. Beneath we compare the results of two non-linear edge filters: the edge avoiding wavelets of Fattal et al. (5.7b, [Fat09]) with an exponent of 1.15 (slightly enhancing fine details, see publication) and the local Laplacian filters of Paris et al. (5.7c, [PHK11]) with $\sigma_{\text{publ}} = 0.2$ and $\alpha_{\text{publ}} = 0.2$. The bottom row shows the result of our algorithm (5.7d) and the original material (5.7e).

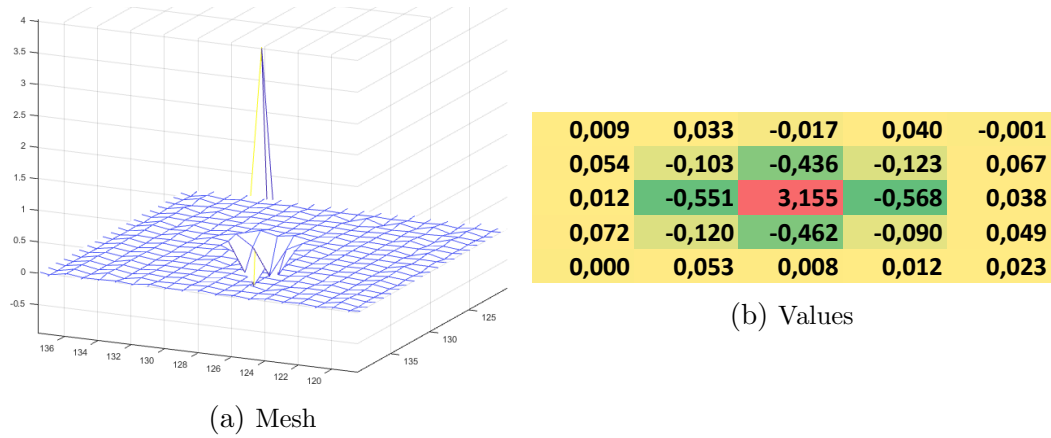


Figure 5.8: The learned filter.

this corresponds to an observer-distance of approximately 82 cm. We will confine to the roughen and the undulation operation. The thicken and the glitter-operator will be compared on bidirectional texture functions (Section 5.4.5).

The comparison edit The influence of the absolute value for the strength of the edit is not directly comparable. The learned edits were mostly weaker than the originals. To compensate for that, we made a relaxation step based on the HDR VDP 2.2-metric as published by Mantiuk et al. in 2015 [NMDSL15], by scaling the edit with a positive number s with

$$s := \operatorname{argmin}_{r>0} |d(\mathcal{T}, \mathbf{E}\mathcal{T})_{VDP} - d(\mathcal{T}, r\mathbf{G}_E\mathcal{T})_{VDP}| \quad (5.14)$$

to minimize the visual difference between \mathbf{E} and \mathbf{G}_E . Of course $r\mathbf{G}_E := r \cdot (\mathbf{G}_E - \mathbf{G}_I) + \mathbf{G}_I$

Roughening

Roughening seems to work comparably good in the Fourier-domain (operator \mathbf{F}) and in the Cortex-filter-bank. For stronger edits (Images 5.9c and 5.9d)

our operator shows less artefacts.

Inflation

For relatively small structures, the undulation-operation works slightly better in the cortical filter bank (Image 5.10a) than in the fourier domain (Image 5.10b, operator \mathbf{F}). For bigger structures, the manipulation in the cortical filter bank is not capable of reproducing the results of the bandpass filtering in the Fourier-domain (Figure 5.11).

5.4.4 Spotlight Moving

To make the moving spotlight experiment, we used the BTF-measurements of the UBO14-database of the university of Bonn [WGK14]. For learning, we used the leather materials with the numbers 1-3 and 5-12. The testing results will be presented on the leather4 material. The camera position had been in the zenith above the material. Material-patches which were illuminated from a polar angle of approximately 30° against the zenith and from an azimuthal angle of 0° were considered as unedited material samples. We interpreted material patches, taken under the same conditions but illuminated from an azimuthal angle of 15° or -15° respectively as the edited versions of the original material patch and used those patches for learning the motion of the spotlight. The results are presented in figure 5.12 and in a short movie in the additional material.

We can see that small moves of a spotlight can be represented and learned in the cortical domain.

5.4.5 Editing of High-Dimensional Material Representations

For an introduction to BTFs see Subsection 2.2.2. Energy preservation and other expressions of physical phenomena are lost after the editing step. Instead

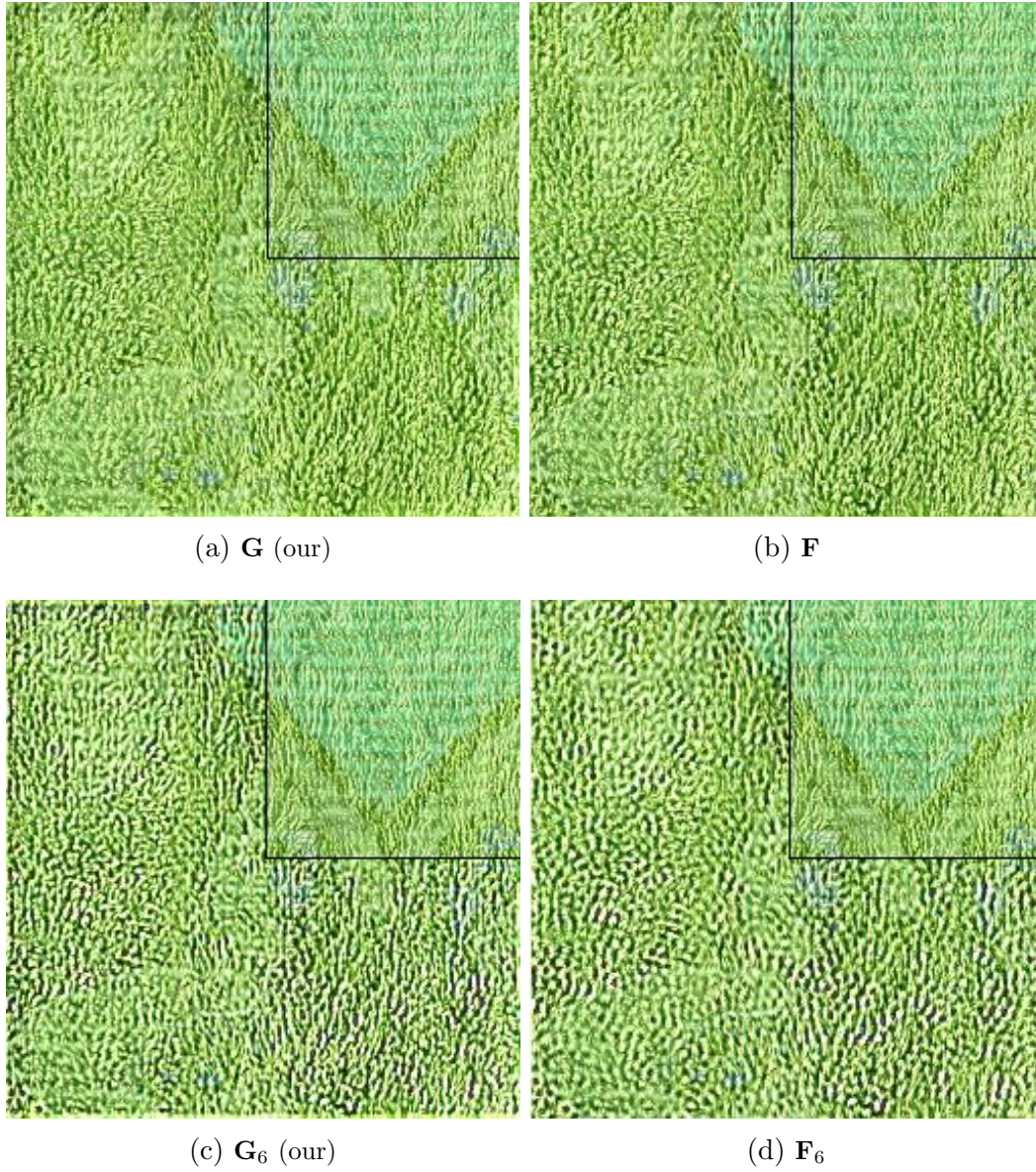


Figure 5.9: Comparison of the roughening filter. The original material in the top right corner.

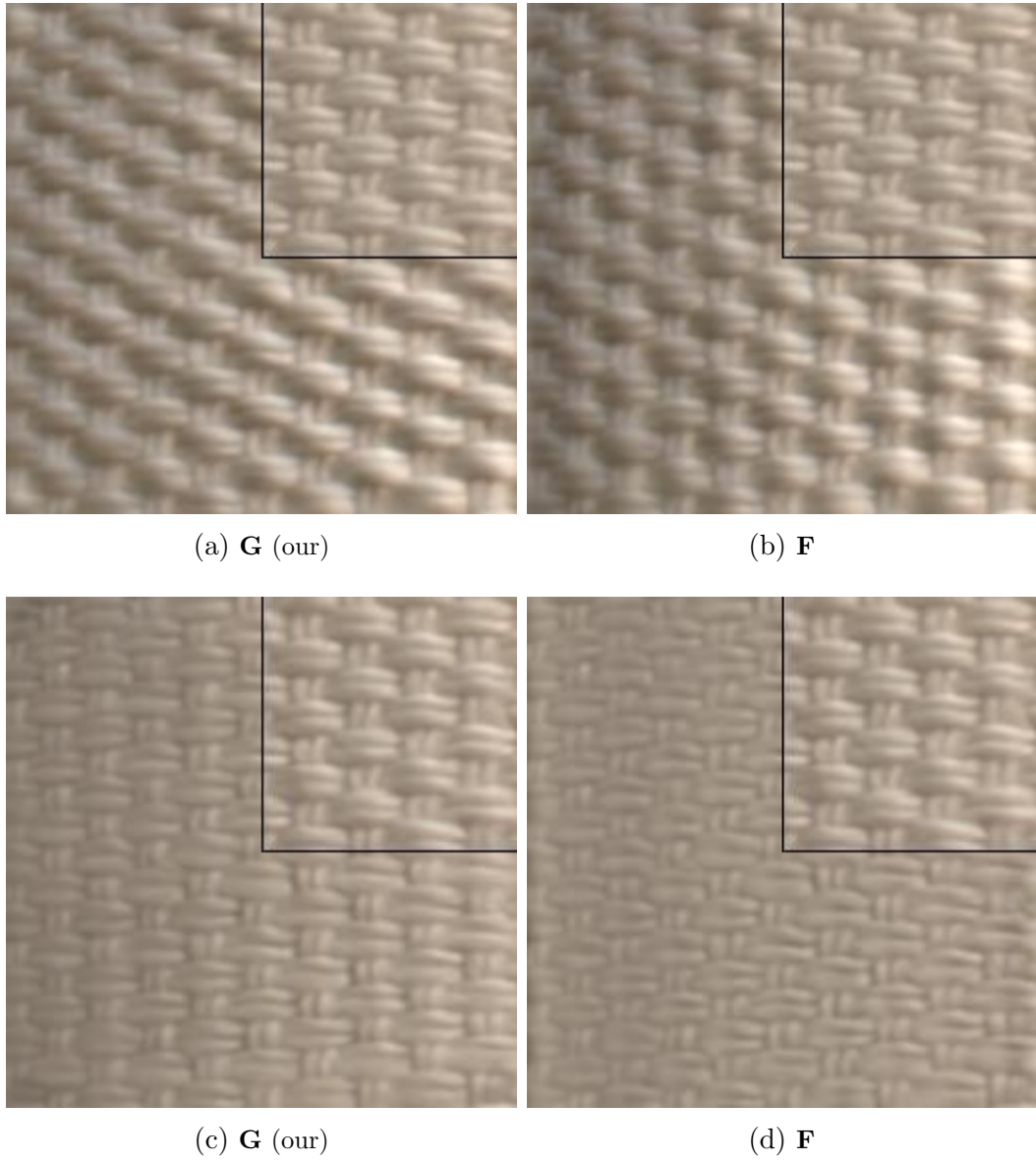


Figure 5.10: Inflating a material. The original material in the top right corner.

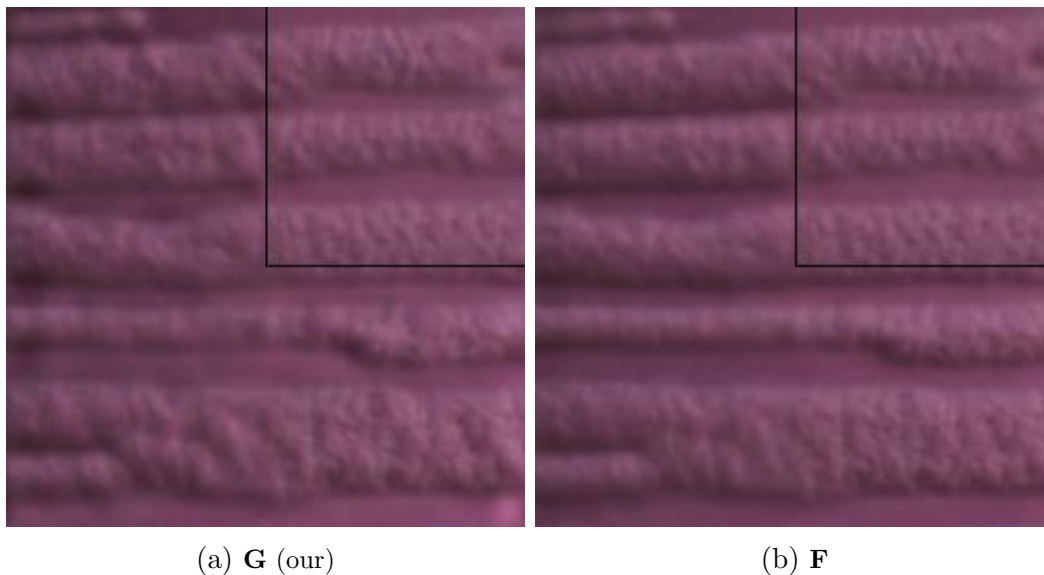


Figure 5.11: Inflating a material. The original material in the top right corner.

of using the suggested thicken or the undulation-operator, one may estimate the surface structure and operate on the new geometry. For higher frequencies, the inverse optics are highly complicated.

5.4.6 Time Requirement

Experiments have been done on an *i7-4770 CPU @ 3.4 GHz* with 8 GB RAM. Learning took between 51'' (spotlight moving, 5.4.4) and 90' (edge aware imaging, 5.4.2). The editing step took about 0.4' for a texture and 40' for a BTF.

5.5 Conclusion

In this chapter we have presented different linear editing operations based on a model of the V1-region of the visual cortex. We could show, that it is possible to reconstruct material patches in an appropriate quality by simple summation of the filter responses of the suggested Gabor-filter bank (section 5.4.1). We learned band-pass filtering which shows in many cases less artefacts when

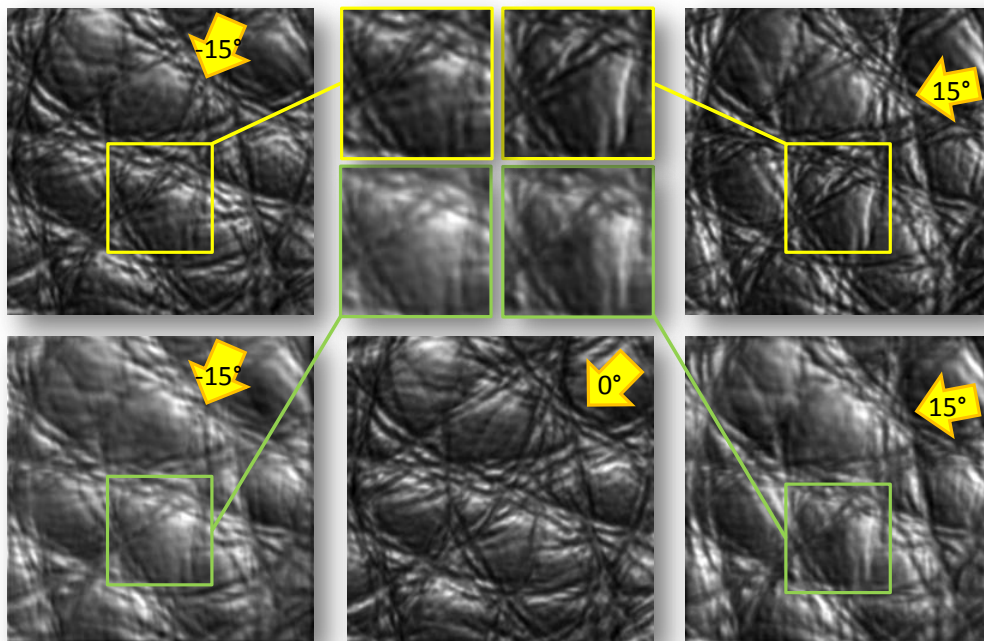


Figure 5.12: In the middle of the bottom row, you can see the original material test-patch. To compare against the real physical operation, we show in the top row a photography of the same patch, illuminated under an azimuthal angle of -15° (left) and illuminated under an azimuthal angle of 15° (right). We compare those results against the application of the spotlight moving-operator (bottom row, left (-15°) and right (15°)). Here we show only the value channel of the material patch.

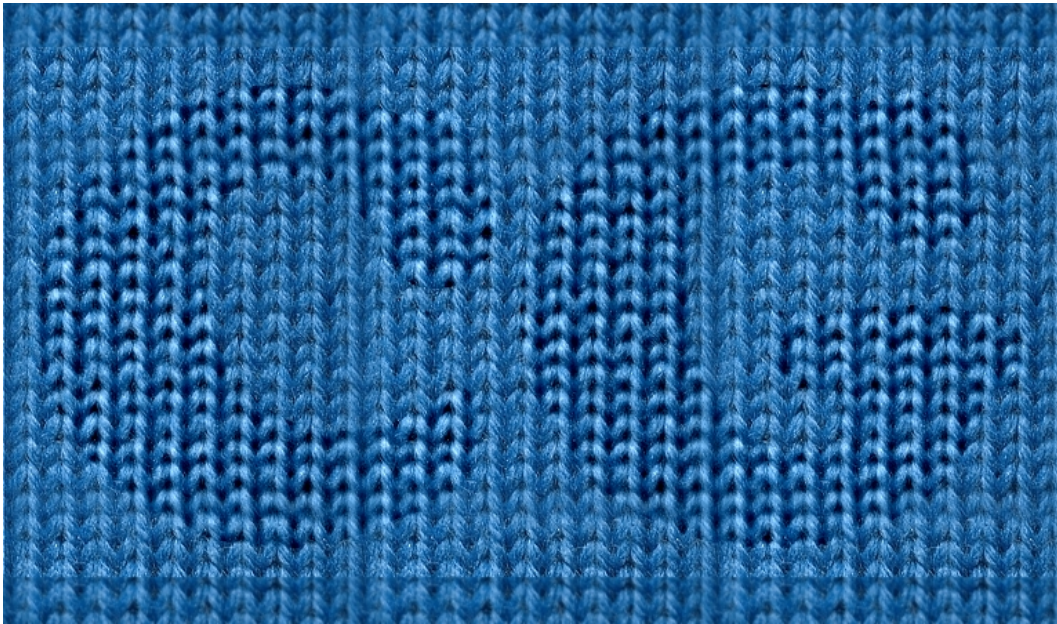


Figure 5.13: A wool material edited with our new cortex based editing system. The bars above and below the letters show the original material. The background is manipulated by our new edge aware image filter. The letters are made deeper and more sparkling.

the corresponding band-pass filter in the fourier domain (section 5.4.3). This effect might be due to the immanent pre-filtering but we have to pronounce that prefiltering the bandpass in the fourier-domain is difficult because it varies between destroying the effect and producing strong sidelobes. Applying those learned operators to a BTF brought notably better results than the corresponding band-pass filters (section 5.4.5). Above that, we could also learn a physical effect (section 5.4.4).

An important subset of the appearance space is the set of all realistic appearances, meaning appearances which are inter-subjectively considered as pictorial representations of a real environment like e.g. photos. We have shown, that starting with a valid element of the space of realistic appearances, the presented operators define an affine linear subspace with limited diameter.

CHAPTER 6

Conclusion

In this thesis we explored several techniques to make use of the spatial frequency spectrum of parameter- or reflectance-maps of measured materials for editing purposes. Frequency or scaling-analysis may be done by linear operators like the Fourier or the Wavelet analysis. Those two operators have the additional advantage that they allow for fast analysis and synthesis. In this thesis we have shown that we can use the Fourier-transform for fast global similarity detection. This global similarity detector has been used to support a seed based near regular pattern detector. The combination of the two algorithms enabled a robust and statistically well founded edit propagation algorithm for near regular textures.

We presented a more heuristically based approach to image editing by evolving an image based affordance-editing system to a material based editing algorithm. The most important advantage of the affordance editing system, which stems originally from visual perception research, is, that it enables plausible, predictable and meaningful edits. The affordance-editing system is based on the heuristic observation that enhancing or weakening specific frequency

bands leads to meaningful effects in the material appearance. We observed that those bands nearly gaplessly cover the whole frequency range of human V1-cortex filters. This motivated the question whether there are other filtering constellations in the human visual system that represent physical phenomena. Consequently the new editing scheme has been driven forward to a cortex-simulation based material editing system. Our new cortex-transform model has been parametrized according to measured results from neuroscientific research. We could show, that it is possible to use this new model not only to represent the known manipulations from the affordance editing system but we could also learn an edge-enhancing operator and the physical phenomenon of a slightly moving light-source.

This gives rise to the question whether there exist other optical phenomena which have linear representations in the visual cortex. As we could pose the same question with some justification for e.g. the auditory cortex, we want to formulate our question more general: are there more examples of complex physical phenomena which have a linear representation in the human brain?

APPENDIX A

Notation

\langle, \rangle Standard inner product:

$$\langle f, g \rangle_{L^2([a,b])} := \int_a^b f \bar{g}$$

$\| \cdot \|$ $\|f\| := \sqrt{\langle f, f \rangle}$

$-$ Complex conjugate: $\overline{x + iy} := x - iy$

$*$ Convolution, for the discrete case in 2D:

$$(F * G)_{ij} = \sum_{kl} F_{kl} G_{i-k, j-l}$$

\star Correlation, for the discrete case in 2D:

$$(F \star G)_{ij} = \sum_{kl} F_{kl} \overline{G_{k-i, l-j}}$$

$L^1(\mathbb{R}^n)$ Space of Lebesgue-integrable functions

$L^2(\mathbb{R}^n)$ Space of Lebesgue-square-integrable functions

\mathbb{A}	Appearance space
\mathbb{G}	Model space of neural responses
\mathbb{I}	Retinal model space
\mathbb{S}	The upper hemisphere:

$$\mathbb{S} = \{\mathbf{x} | x_3 \geq 0 \wedge |\mathbf{x}| = 1\}$$

\mathcal{C}	Constant texture
\mathcal{G}	Discretized Gauss-kernel texture



\mathcal{L}	
\mathcal{P}	Patch of a texture
\mathcal{X}, \mathcal{W}	Test-textures
\mathbf{E}	Variable edit operation
\mathbf{E}_Γ	Gabor manipulation
$\mathbf{E}_\mathbf{F}$	Fourier edit
$\mathbf{E}_\mathbf{W}$	Wavelet edit
\mathbf{F}	Fourier-transform
$\mathbf{F}_\mathbf{C}$	Cosine transform
$\mathbf{F}_\mathbf{S}$	Sine transform
\mathbf{M}	Linear texture transform
\mathbf{P}	Push-pin
\mathbf{W}	Wavelet-transform
Δ	Laplace operator
Γ	Gabor-filter
∂	Boundary operator
γ	Gabor kernel
G	Gauss kernel
α	Probability that a pixel belongs to the foreground
β	Phase angle
δ	Dirac distribution
θ	Elevation angle
μ	Mean

ν	Frequency
ξ	Parameter-set of Gabor-filter
Ξ	All parameter-sets of the Gabor-filter in a given database
ρ	Reflectance distribution
σ	Standard deviation
ϕ	Scaling function
φ	Azimuth angle, orientation of a Gabor filter
ψ	Mother wavelet
ω	Direction
A	Amplitude
$R_{s,d}$	Reflectance-parameter: specular s and diffuse d
\mathbf{p}	Position
$\mathbf{e}_{\mathbf{p}}$	$\frac{\mathbf{p}}{ \mathbf{p} }$
\mathbf{h}	Half vector
\mathbf{n}	Normal vector
$\mathbf{U}\Sigma\mathbf{V}$	Svd. See [GK65]
\mathfrak{P}	Phenomenon; $\mathfrak{P} : \mathbb{I} \rightarrow \mathbb{I}$

APPENDIX B

Abbreviations

BRDF:	<i>bidirectional reflectance distribution functions</i>	page 11
BTF:	<i>bidirectional texture function</i>	page 11
SVBRDF:	<i>spatially varying BRDF</i>	page 25
DFMF:	<i>decorrelated full matrix factorization</i>	page 26
LLS:	<i>linear light source</i>	page 32
STFT:	<i>short time Fourier transform</i>	page ??
DFT:	<i>discrete Fourier transform</i>	page 40
NRT:	<i>near regular textures</i>	page 59
LGN:	<i>lateral geniculate nucleus</i>	page 105
rfp:	<i>receptive field profile</i>	page 105

APPENDIX C

Tables

Table 4.1:	Affordances in cycles per degree (cpd) in Giesel and Zaidis work. In the last column, we give the symbol of the associated scaling operator. The upper \mathbf{F} indicates, that the scaling ${}^{\mathbf{F}}\mathbf{D}$ operates on the Fourier-domain.	page 81
Table 4.2:	Affordances in frequency-octaves.	page 85
Table 4.3:	Affordances in cycles per millimetre ([cpmm]) with regard to the metrics of our concrete BTF-implementation.	page 89
Table 4.7:	HDR-VDP-2 error plots for a reference (Fourier) gain factor of ${}^{\mathbf{F}}K = 2$. The horizontal axis is the wavelet gain factor ${}^{\mathbf{W}}k$, and the vertical axis the type of wavelet (D1–D6).	page 89

Table 4.4: Scaling of the wavelet-edit strength by edit page 90
 type and fourier-edit strength.

APPENDIX D

List of Figures

- | | | |
|-------------|--|---------|
| Figure 1.1: | A direction and the elevation angle and the azimuth. | page 6 |
| Figure 3.1: | Pushpins for edit propagation. | page 55 |
| Figure 4.1: | Appearance Bending. | page 79 |
| Figure 1.4: | Linear Subspaces of the Appearance Space. | page 9 |
| Figure 2.1: | The BRDF is the quotient between the incoming irradiance E confined to a direction ω_i and the outgoing radiance L , confined to ω_i . | page 17 |

-
- Figure 2.2: To explain perspective drawing, Leon Battista Alberti uses as early as 1436 visual rays for his argumentation (left image). Though visual rays lost their ontological meaning since then, they describe exactly the concept which has been used in the early raytracing algorithms from the 1960's (right image from [App68]). page 18
- Figure 2.3: The parameters of a BRDF. page 23
- Figure 2.4: The camera dome. From [FH09]. page 27
- Figure 2.5: BTF-compression by the DFMF. page 28
- Figure 2.6: The camera dome. 151 consumer-class, off-the-shelf cameras with an image resolution of 2048 x 1536 pixel have been assembled to cover a hemisphere with a radius of 1 meter. The cameras flashlights are used as directional light sources. Figure taken from [MK05]. page 29
- Figure 2.7: Photo of the linear light scanning device. From [GTHD03]. page 32
- Figure 2.8: Comparison between scale and frequency, based on the result of the Matlab function *scal2frq*. page 43
- Figure 2.9: The downsampling scheme of the fast wavelet transform. Subsampling ($\downarrow 2$) is part of the definition of \mathbf{H} and \mathbf{G} . page 47
- Figure 2.10: The downsampling scheme for two dimensional wavelets. page 48

Figure 2.11:	Scaling and wavelet function for Daubechies db6 wavelet. The right plot shows the frequency response, which approximates a low-pass and a band-pass filter.	page 49
Figure 2.12:	Frequency response of the rotated scaling and wavelet function for Daubechies db6 wavelet.	page 49
Figure 3.1:	Pushpins for edit propagation.	page 55
Figure 3.2:	Teaser for the PushPins algorithm.	page 56
Figure 3.3:	Overview over the pushpin based foreground segregation algorithm	page 62
Figure 3.4:	The mask resulting from the SVM classification step.	page 63
Figure 3.5:	The effect of regularization to pushpins. From left: the response of the original tile to an unregularized pushpin, the response to a regularized pushpin, the response of distorted material-patch to the unregularized pushpin and the response of the same material-patch to the regularized filter.	page 64
Figure 3.6:	On the left side you can see in light blue the filter response of a pushpin with nailhead radius approximately equal to the size of half a small square on the right side we used a pushpin with a nailhead radius approximately equal to half a big square.	page 65

- Figure 3.7: Two different runs of mean shift belief propagation on the grey mesh material. page 67
- Figure 3.8: The left image shows the grey mesh material, the image in the middle depicts the filter response of a pushpin applied to the volume channel of the diffuse color of the grey mesh material and the third image shows the result of MSBP on a combined map of the filter response and the diffuse channel. Now the lattice detection is extremely stable. page 67
- Figure 3.9: On the top the binary mask, on the right the original superposed mask and on the left the mixed mask. page 70
- Figure 3.10: On the left a patch from the original SVBRDF-channel with the stroke input the second image shows the mask generated by RepSnapping. The third mask is the result of AppProp and the last mask is our result. The first row shows the grey mesh material the second row shows a metal. page 74
- Figure 3.11: On the left the original material in the middle a rather subtle edit of R_d , on the right a more noticeable manipulation of R_s . page 75
- Figure 3.12: In the close-up of the edit of the grey mesh material one may see that the editing boundary coincides exactly with the perceived boundary of the foreground material. page 75

-
- Figure 3.13: A shiny material. page 75
- Figure 3.14: The edited metal material. page 76
- Figure 4.1: Appearance Bending. page 79
- Figure 4.2: From [GZ13]: Material patches, used in Giesel and Zaidis affordance classification experiments. The researchers cropped 256 color images of fabrics to a size of 150 x 150 pixels. Images were presented on a monitor against a black background. The viewing distance was 70 cm. page 82
- Figure 4.4: The green bars indicate the frequency distribution in [cpd] of our new computational cortex model, presented in Chapter 5. Beneath this distribution we illustrated the frequency-range of the editing operator. Note, that the frequency values on the x-axis increase exponentially, meaning that octaves have constant width. page 86
- Figure 4.8: A comparison of Fourier- and wavelet-edited materials. Note that the strength of the editing operator, k , has been adapted to obtain the best possible match according to HDR-VDP-2 [MKRH11]. page 91

- Figure 4.9: Comparison of Fourier (top) and wavelet (bottom) editing for the “roughen” operator. From left to right: $k = 0.5$, $k = 1$ (identity), $k = 2$. Note the absence of regularly patterned artifacts in the wavelet results, even for a gain factor of $k = 2$. page 92
- Figure 4.10: \mathbf{E}_R applied to a wool material. page 92
- Figure 5.3: The histogram of macaque frequency distribution [DVAT82], extended to human vision. page 110
- Figure 5.4: From [Lee96]: an ensemble of Gabor wavelets in the frequency domain. page 114

Bibliography

- [AAL16] Miika Aittala, Timo Aila, and Jaakko Lehtinen. Reflectance modeling by neural texture synthesis. *ACM Transactions on Graphics (TOG)*, 35(4):65, 2016.
- [Ade01] Edward H. Adelson. Proceedings of the spie: Human vision and electronic imaging vi. pages 1–12, 2001.
- [Alb13] Josef Albers. *Interaction of color*. Yale University Press, 2013.
- [AP79] BW Andrews and DA Pollen. Relationship between spatial frequency selectivity and receptive field profile of simple cells. *The Journal of physiology*, 287(1):163–176, 1979.
- [AP08] X. An and F. Pellacini. Appprop: all-pairs appearance-space edit propagation. *ACM Transactions on Graphics*, 27(3):1–9, 2008.
- [App68] Arthur Appel. Some techniques for shading machine renderings of solids. In *Proceedings of the April 30–May 2, 1968, spring joint computer conference*, pages 37–45. ACM, 1968.

- [AS00] Michael Ashikhmin and Peter Shirley. An anisotropic phong brdf model. *Journal of graphics tools*, 5(2):25–32, 2000.
- [ATDP11] Xiaobo An, Xin Tong, Jonathan D Denning, and Fabio Pellacini. Appwarp: retargeting measured materials by appearance-space warping. In *ACM Transactions on Graphics (TOG)*, volume 30, page 147. ACM, 2011.
- [AWL⁺15] Miika Aittala, Tim Weyrich, Jaakko Lehtinen, et al. Two-shot svbrdf capture for stationary materials. *ACM Trans. Graph.*, 34(4):110–1, 2015.
- [BBPA15] Ivaylo Boyadzhiev, Kavita Bala, Sylvain Paris, and Edward Adelson. Band-sifting decomposition for image-based material editing. *ACM Transactions on Graphics (TOG)*, 34(5):163, 2015.
- [Boo89] Fred L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence*, 11(6):567–585, 1989.
- [BS63] P Beckmann and A Spizzichino. The scattering of electromagnetic waves from rough surface. 1963.
- [CL11] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [CMK⁺14] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, , and A. Vedaldi. Describing textures in the wild. In *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2014.

- [CPC84] Robert L Cook, Thomas Porter, and Loren Carpenter. Distributed ray tracing. In *ACM SIGGRAPH computer graphics*, volume 18, pages 137–145. ACM, 1984.
- [CT81] Robert L Cook and Kenneth E Torrance. A reflectance model for computer graphics. In *ACM Siggraph Computer Graphics*, volume 15, pages 307–316. ACM, 1981.
- [CV95] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [Dau80] John G Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision research*, 20(10):847–856, 1980.
- [Dau85] John G Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *JOSA A*, 2(7):1160–1169, 1985.
- [Dau88] John G Daugman. Complete discrete 2-d gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169–1179, 1988.
- [Dau92] Ingrid Daubechies. Ten lectures on wavelets, cbms-nsf regional conf. *Series in Appl. Math*, 61, 1992.
- [Dau93] Ingrid Daubechies. Wavelet transforms and orthonormal wavelet bases. *Different Perspectives on Wavelets: American Mathematical Society Short Course, January 11-12, 1993, San Antonio, Texas*, 47:1, 1993.
- [DM97] Paul E Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *Computer graphics proceedings, annual conference series*, pages 369–378. Association for Computing Machinery SIGGRAPH, 1997.

- [DS52] Richard J Duffin and Albert C Schaeffer. A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, 72(2):341–366, 1952.
- [DVAT82] Russell L De Valois, Duane G Albrecht, and Lisa G Thorell. Spatial frequency selectivity of cells in macaque visual cortex. *Vision research*, 22(5):545–559, 1982.
- [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.
- [EIKM16] Yuki Endo, Satoshi Iizuka, Yoshihiro Kanamori, and Jun Mitani. Deepprop: Extracting deep features from a single image for edit propagation. In *Computer Graphics Forum*, volume 35, pages 189–201. Wiley Online Library, 2016.
- [Fat09] Raanan Fattal. Edge-avoiding wavelets and their applications. *ACM Transactions on Graphics (TOG)*, 28(3):22, 2009.
- [FGNP85] KH Foster, James P Gaska, M Nagler, and DA Pollen. Spatial and temporal frequency selectivity of neurones in visual cortical areas v1 and v2 of the macaque monkey. *The Journal of physiology*, 365(1):331–363, 1985.
- [FH09] J. Filip and M. Haindl. Bidirectional texture function modelling: A state of the art survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(11):1921–1939, 2009.
- [FHB66] Gustav Theodor Fechner, Davis H Howes, and Edwin Garrigues Boring. *Elements of psychophysics*, volume 1. Holt, Rinehart and Winston New York, 1966.

- [FKH⁺18] Jiří Filip, Martina Kolafová, Michal Havlíček, Radomír Vávra, Michal Haindl, and Holly Rushmeier. Evaluating physical and rendered material appearance. *The Visual Computer*, pages 1–12, 2018.
- [Fou22] Joseph Fourier. *Theorie analytique de la chaleur, par M. Fourier*. Chez Firmin Didot, père et fils, 1822.
- [Gab46] Dennis Gabor. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26):429–441, 1946.
- [GH13] Michael U Gutmann and Aapo Hyvärinen. A three-layer model of natural image statistics. *Journal of Physiology-Paris*, 107(5):369–398, 2013.
- [Gib77] James J Gibson. Perceiving, acting, and knowing: Toward an ecological psychology. *The Theory of Affordances*, pages 67–82, 1977.
- [GK65] Gene Golub and William Kahan. Calculating the singular values and pseudo-inverse of a matrix. *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, 2(2):205–224, 1965.
- [Gla95] Andrew S Glassner. *Principles of digital image synthesis: Vol. 1*, volume 1. Elsevier, 1995.
- [GS87] Branko Grünbaum and Geoffrey Colin Shephard. *Tilings and patterns*. Freeman, 1987.
- [GTHD03] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. *ACM Transactions on Graphics (TOG)*, 22(3):749–758, 2003.

- [GZ12] Martin Giesel and Qasim Zaidi. Adaptation reveals frequency band based inferences of material properties. volume (In press), 2012.
- [GZ13] Martin Giesel and Qasim Zaidi. Frequency-based heuristics for material perception. *Journal of vision*, 13(14):7–7, 2013.
- [HB95] David J Heeger and James R Bergen. Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 229–238. ACM, 1995.
- [Hei27] Werner Heisenberg. Über den anschaulichen inhalt der quanten theoretischen kinematik und mechanik. *Zeit. für Phys.*, 4:172–198, 1927.
- [HH05] M. Haindl and M. Hatka. Btf roller. In M. Chantler and O. Drbohlav, editors, *Texture 2005: Proceedings of 4th International Workshop on Texture Analysis and Synthesis*, pages 89–94, Edinburgh, October 2005. Heriot-Watt University.
- [HH17] Michal Haindl and Michal Havlíček. A compound moving average bidirectional texture function model. In *Multimedia and Network Information Systems*, pages 89–98. Springer, 2017.
- [HLLC17] Tsung-Shian Huang, Jen-Hao Liao, Wen-Chieh Lin, and Jung-Hong Chuang. An artist friendly material design system. *Journal of Information Science & Engineering*, 33(2), 2017.
- [HMA17] Jacob Huth, Timothee Masquelier, and Angelo Arleo. Convis: A toolbox to fit and simulate filter-based models of early visual processing. *bioRxiv*, page 169284, 2017.

- [HOT06] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.
- [HS09] Dorit S Hochbaum and Vikas Singh. An efficient algorithm for co-segmentation. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 269–276. IEEE, 2009.
- [HST13] Kaiming He, Jian Sun, and Xiaoou Tang. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 35(6):1397–1409, 2013.
- [HW59] David H Hubel and Torsten N Wiesel. Receptive fields of single neurones in the cat’s striate cortex. *The Journal of physiology*, 148(3):574–591, 1959.
- [HW62] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of physiology*, 160(1):106–154, 1962.
- [HW68] David H Hubel and Torsten N Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1):215–243, 1968.
- [HZZ11] Hua Huang, Lei Zhang, and Hong-Chao Zhang. Reprsnapping: efficient image cutout for repeated scene elements. In *Computer Graphics Forum*, volume 30, pages 2059–2066. Wiley Online Library, 2011.
- [ICG86] David S Immel, Michael F Cohen, and Donald P Greenberg. A radiosity method for non-diffuse environments. In *Acm Siggraph Computer Graphics*, volume 20, pages 133–142. ACM, 1986.

- [JP87] Judson P Jones and Larry A Palmer. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of neurophysiology*, 58(6):1187–1211, 1987.
- [Kaj86] James T Kajiya. The rendering equation. In *ACM Siggraph Computer Graphics*, volume 20, pages 143–150. ACM, 1986.
- [KBD07] Jan Kautz, Solomon Boulos, and Frédo Durand. Interactive editing and modeling of bidirectional texture functions. In *ACM Transactions on Graphics (TOG)*, volume 26, page 53. ACM, 2007.
- [KHH17] Takafumi Katsunuma, Keita Hirai, and Takahiko Horiuchi. Fabric appearance control system for example-based interactive texture and color design. *ACM Transactions on Applied Perception (TAP)*, 14(3):16, 2017.
- [KHI+03] Robert Kosara, Christopher G Healey, Victoria Interrante, David H Laidlaw, and Colin Ware. User studies: Why, how, and when? *IEEE Computer Graphics and Applications*, (4):20–25, 2003.
- [KMBK03] M. Koudelka, S. Magda, P. Belhumeur, and D. Kriegman. Acquisition, compression, and synthesis of bidirectional texture functions. *Proc. Third Int’l Workshop Texture Analysis and Synthesis*, pages pp. 47–52, Oct. 2003.
- [Küp24] Karl Küpfmüller. Transient phenomena in wave filters. *Elektr. Nachrichtentechn.*, 1(5):141–152, 1924.
- [KZ04] Vladimir Kolmogorov and Ramin Zabini. What energy functions can be minimized via graph cuts? *IEEE transactions on pattern analysis and machine intelligence*, 26(2):147–159, 2004.

- [LBBS17] Valero Laparra, Alex Bernardino, Johannes Ballé, and Eero P Simoncelli. Perceptually optimized image rendering. *arXiv preprint arXiv:1701.06641*, 2017.
- [LBD⁺89] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard, Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551, 1989.
- [LCT04] Yanxi Liu, Robert Collins, and Yanghai Tsin. A computational model for periodic pattern perception based on frieze and wallpaper groups. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):354 – 371, March 2004.
- [LCY⁺17] Guilin Liu, Duygu Ceylan, Ersin Yumer, Jimei Yang, and Jyh-Ming Lien. Material editing using a physically based rendering network. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2280–2288. IEEE, 2017.
- [Lee96] Tai Sing Lee. Image representation using 2d gabor wavelets. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18:959–971, 1996.
- [LH06] Sylvain Lefebvre and Hugues Hoppe. Appearance-space texture synthesis. *ACM Transactions on Graphics (TOG)*, 25(3):541–548, 2006.
- [Lin13] Tony Lindeberg. A computational theory of visual receptive fields. *Biological cybernetics*, 107(6):589–635, 2013.
- [LK01] H. Lensch and J. Kautz. Image based reconstruction of spatially varying materials. *Rendering Techniques '01 (Proc. of Eurographics Workshop on Rendering)*, 2001.

- [LLH04] Yanxi Liu, Wen-Chieh Lin, and James Hays. Near-regular texture analysis and manipulation. *ACM Transactions on Graphics (TOG)*, 23(3):368–376, 2004.
- [LMR94] Alfred K Louis, Peter Maaß, and Andreas Rieder. *Wavelets: Theorie und Anwendungen*. Springer-Verlag, 1994.
- [LMS⁺19] Manuel Lagunas, Sandra Malpica, Ana Serrano, Elena Garces, Diego Gutierrez, and Belen Masia. A similarity measure for material appearance. *arXiv preprint arXiv:1905.01562*, 2019.
- [Mal89] Stephane G Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (7):674–693, 1989.
- [Mar80] S. Marčelja. Mathematical description of the responses of simple cortical cells*. *J. Opt. Soc. Am.*, 70(11):1297–1300, Nov 1980.
- [Mas93] Timothy Masters. *Practical neural network recipes in C++*. Morgan Kaufmann, 1993.
- [MBG⁺17] Sandra Malpica, Miguel Barrio, Diego Gutiérrez, Ana Serrano, and Belén Masia. Improved intuitive appearance editing based on soft pca. In *Congreso Español de Informática Gráfica (27^o. 2017. Sevilla)(2017)*, p 97-106, pages 97–106, 2017.
- [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, Eurographics Association, November 2005.
- [MGZ⁺17] Marlon Mylo, Martin Giesel, Qasim Zaidi, Matthias Hullin, and Reinhard Klein. Appearance bending: A perceptual editing

- paradigm for data-driven material models. In *Vision, Modeling & Visualization*, 2017.
- [MHW⁺12] Jan Meseth, Shawn Hempel, Andrea Weidlich, Lynn Fyffe, Graham Fyffe, Craig Miller, Paul Carroll, and Paul Debevec. Improved linear light source material reflectance scanning. In *ACM SIGGRAPH 2012 Posters*, page 42. ACM, 2012.
- [MK90] William H Merigan and Laurence M Katz. Spatial resolution across the macaque retina. *Vision research*, 30(7):985–991, 1990.
- [MK05] Meseth J. Sattler M. Sarlette R. Müller, G. and R. Klein. Acquisition, synthesis, and rendering of bidirectional texture functions. *Computer Graphics Forum*, 24(1 (Mar.)):83–110, 2005.
- [MK17] Marlon Mylo and Reinhard Klein. Pushpins for edit propagation. 2017.
- [MK18] Marlon Mylo and Reinhard Klein. Linear subspaces of the appearance space. *Journal of WSCG*, pages 95–103, 2018.
- [MKRH11] Rafat Mantiuk, Kil Joong Kim, Allan G Rempel, and Wolfgang Heidrich. Hdr-vdp-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions. In *ACM Transactions on graphics (TOG)*, volume 30, page 40. ACM, 2011.
- [MMS⁺05] Gero Müller, Jan Meseth, Mirko Sattler, Ralf Sarlette, and Reinhard Klein. Acquisition, synthesis and rendering of bidirectional texture functions. *Computer Graphics Forum*, 24(1):83–109, March 2005.

- [MNR92] Rajiv Mehrotra, Kameswara Rao Namuduri, and Nagarajan Ranganathan. Gabor filter-based edge detection. *Pattern recognition*, 25(12):1479–1494, 1992.
- [MPBM03] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics*, 22(3):759–769, July 2003.
- [MR02] Ferenc Mechler and Dario L Ringach. On the classification of simple and complex cells. *Vision research*, 42(8):1017–1033, 2002.
- [MRC⁺86] Gary W Meyer, Holly E Rushmeier, Michael F Cohen, Donald P Greenberg, and Kenneth E Torrance. An experimental evaluation of computer graphics imagery. *ACM Transactions on Graphics (TOG)*, 5(1):30–50, 1986.
- [MSK07] Gero Müller, Ralf Sarlette, and Reinhard Klein. Procedural editing of bidirectional texture functions. In J. Kautz and S. Pattanaik, editors, *Eurographics Symposium on Rendering 2007*. The Eurographics Association, June 2007.
- [MU12] Rosana Montes and Carlos Ureña. An overview of brdf models. *University of Grenada, Technical Report LSI-2012-001*, 2012.
- [Mül09] Gero Müller. *Data-Driven Methods for Compression and Editing of Spatially Varying Appearance*. Dissertation, Universität Bonn, December 2009.
- [NDM05] Addy Ngan, Frédo Durand, and Wojciech Matusik. Experimental analysis of brdf models. *Rendering Techniques*, 2005(16th):2, 2005.
- [New04] Isaac Newton. *Opticks: Or a treatise of the reflexions, refractions, inflexions and colours of light*. 1704.

- [Nic63] Fred E Nicodemus. Radiance. *American Journal of Physics*, 31(5):368–377, 1963.
- [Nic65] Fred E Nicodemus. Directional reflectance and emissivity of an opaque surface. *Applied optics*, 4(7):767–775, 1965.
- [NMDSL15] Manish Narwaria, Rafal K Mantiuk, Mattheiu Perreira Da Silva, and Patrick Le Callet. Hdr-vdp-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images. *Journal of Electronic Imaging*, 24(1):010501–010501, 2015.
- [NPC17] Pravin Nair, Anmol Popli, and Kunal N Chaudhury. A fast approximation of the bilateral filter using the discrete fourier transform. *Image Processing On Line*, 7:115–130, 2017.
- [NRH⁺77] FE Nicodemus, JC Richmond, JJ Hsia, IW Ginsberg, and T Limperis. Geometrical considerations and nomenclature for reflectance. *Final Report National Bureau of Standards, Washington, DC. Inst. for Basic Standards.*, 1977.
- [Nyq24] Harry Nyquist. Certain factors affecting telegraph speed. *Transactions of the American Institute of Electrical Engineers*, 43:412–422, 1924.
- [OF96] Bruno A Olshausen and David J Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607, 1996.
- [OF97] Bruno A Olshausen and David J Field. Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision research*, 37(23):3311–3325, 1997.

- [PCL08] Minwoo Park, Robert Collins, and Yanxi Liu. Deformed lattice detection via mean-shift belief propagation. In *European Conference on Computer Vision (ECCV)*, October 2008.
- [PH04] Matt Pharr and Greg Humphreys. *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.
- [PHK11] Sylvain Paris, Samuel W Hasinoff, and Jan Kautz. Local laplacian filters: Edge-aware image processing with a laplacian pyramid. *ACM Trans. Graph.*, 30(4):68–1, 2011.
- [PKT⁺09] Sylvain Paris, Pierre Kornprobst, Jack Tumblin, Frédo Durand, et al. Bilateral filtering: Theory and applications. *Foundations and Trends[®] in Computer Graphics and Vision*, 4(1):1–73, 2009.
- [PL07] Fabio Pellacini and Jason Lawrence. Appwand: editing measured materials using appearance-driven optimization. In *ACM SIGGRAPH 2007 papers*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.
- [PS13] Tomaso Poggio and Thomas Serre. Models of visual cortex. *Scholarpedia*, 8(4):3516, 2013.
- [Rin02] Dario L Ringach. Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of neurophysiology*, 88(1):455–463, 2002.
- [RK09] Roland Ruiters and Reinhard Klein. Heightfield and spatially varying brdf reconstruction for materials with interreflections. *Computer Graphics Forum*, 28(2):513–522, April 2009.

- [RSD15] Rishi Rajalingham, Kailyn Schmidt, and James J DiCarlo. Comparison of object recognition behavior in human and monkey. *Journal of Neuroscience*, 35(35):12127–12136, 2015.
- [RSK13] Roland Ruiters, Christopher Schwartz, and Reinhard Klein. Example-based interpolation and synthesis of bidirectional texture functions. In *Computer Graphics Forum*, volume 32, pages 361–370. Wiley Online Library, 2013.
- [RSS97] Dario L Ringach, Guillermo Sapiro, and Robert Shapley. A subspace reverse-correlation technique for the study of visual neurons. *Vision research*, 37(17):2455–2464, 1997.
- [Rus98] Szymon M Rusinkiewicz. A new change of variables for efficient brdf representation. In *Rendering techniques 98*, pages 11–22. Springer, 1998.
- [SA00] Emilio Salinas and LF Abbott. Do simple cells in primary visual cortex form a tight frame? *Neural computation*, 12(2):313–335, 2000.
- [San90] Terence D Sanger. Analysis of the two-dimensional receptive fields learned by the generalized hebbian algorithm in response to random input. *Biological cybernetics*, 63(3):221–228, 1990.
- [SC78] Ephraim M Sparrow and RD Cess. Radiation heat transfer. *Series in Thermal and Fluids Engineering, New York: McGraw-Hill, 1978, Augmented ed.*, 1978.
- [Sel16] Benjamin Selby. Development of an integrated model of primary visual cortex. Master’s thesis, University of Waterloo, 2016.
- [SGM⁺16] Ana Serrano, Diego Gutierrez, Karol Myszkowski, Hans-Peter Seidel, and Belen Masia. An intuitive control space for material

- appearance. *ACM Transactions on Graphics (TOG)*, 35(6):186, 2016.
- [SH81] R Siegel and J.R. Howell. *Thermal Radiation Heat Transfer*. 1981.
- [She01] Roger N Shepard. Perceptual-cognitive universals as reflections of the world. *Behavioral and brain sciences*, 24(4):581–601, 2001.
- [SPN⁺15] Thorsten-Walther Schmidt, Fabio Pellacini, Derek Nowrouzezahrai, Wojciech Jarosz, and Carsten Dachsbacher. State of the art in artistic editing of appearance, lighting and material. In *Computer Graphics Forum*. Wiley Online Library, 2015.
- [SRB10] Deqing Sun, Stefan Roth, and Michael J Black. Secrets of optical flow estimation and their principles. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2432–2439. IEEE, 2010.
- [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, 2014.
- [SW90] David G. Stork and Hugh R. Wilson. Do gabor functions provide appropriate descriptions of visual cortical receptive fields? In *J. Opt. Soc. Am. A* 7, pages 1362–1373, 1990.
- [TM98] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on*, pages 839–846. IEEE, 1998.

- [TS67] Kenneth E Torrance and Ephraim M Sparrow. Theory for off-specular reflection from roughened surfaces. *Josa*, 57(9):1105–1114, 1967.
- [Tur16] Diána Turcsány. *Deep learning models of biological visual information processing*. PhD thesis, University of Nottingham, 2016.
- [W⁺87] Andrew B. Watson et al. The cortex transform- rapid computation of simulated neural images. *Computer vision, graphics, and image processing*, 39(3):311–327, 1987.
- [W⁺92] Gregory J Ward et al. Measuring and modeling anisotropic reflection. *Computer Graphics*, 26(2):265–272, 1992.
- [WDR11] H. Wu, J. Dorsey, and H. Rushmeier. A sparse parametric mixture model for btf compression, editing and rendering. *Eurographics*, 30(2):1921–1939, 2011.
- [WGK14] Michael Weinmann, Juergen Gall, and Reinhard Klein. Material classification based on training data synthesized using a btf database. In *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part III*, pages 156–171. Springer International Publishing, 2014.
- [Whi79] Turner Whitted. An improved illumination model for shaded display. In *ACM SIGGRAPH Computer Graphics*, volume 13, page 14. ACM, 1979.
- [WTL⁺06] Jiaping Wang, Xin Tong, Stephen Lin, Minghao Pan, Chao Wang, Hujun Bao, Baining Guo, and Heung-Yeung Shum. Appearance manifolds for modeling time-variant appearance of materials. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 754–761, New York, NY, USA, 2006. ACM.

-
- [XWT⁺09] K. Xu, J. Wang, X. Tong, S. Hu, and B. Guo. Edit propagation on bidirectional texture functions. *Pacific Graphics*, 28:7, 2009.
- [Yan92] Jian Yang. Do gabor functions provide appropriate descriptions of visual cortical receptive fields?: comment. *JOSA A*, 9(2):334–336, 1992.
- [YHH17] Yoshimitsu Yamada, Keita Hirai, and Takahiko Horiuchi. Correlation analysis between wood eigen textures and perceptual qualities. In *Color and Imaging Conference*, volume 2017, pages 290–295. Society for Imaging Science and Technology, 2017.
- [ZFWW18] Károly Zsolnai-Fehér, Peter Wonka, and Michael Wimmer. Gaussian material synthesis. *arXiv preprint arXiv:1804.08369*, 2018.
- [ZGMK17] Qasim Zaidi, Martin Giesel, Marlon Mylo, and Reinhard Klein. Perception and appearance bending of material properties. *Journal of Vision*, 17(15):21–21, 2017.