



Technische Universität Berlin
Fachgruppe Angewandte Funktionalanalysis

The Complex Shearlet Transform and Applications to Image Quality Assessment

Master Thesis
by

Rafael Reisenhofer

18. Juli 2014 – 17. Oktober 2014

Primary reviewer: Prof. Dr. Gitta Kutyniok
Secondary reviewer: Dr. Wang-Q Lim
Supervisor: Prof. Dr. Gitta Kutyniok

Rafael Reisenhofer
Beusselstraße 64
10553 Berlin

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe.

Berlin, den 17. Oktober 2014

Rafael Reisenhofer

Contents

List of Figures	5
1 Introduction	7
2 Complex Shearlet Transforms	9
2.1 From Fourier to Shearlets via an Analysis of Optimality	10
2.1.1 Fourier Analysis	14
2.1.2 Wavelets	20
2.1.2.1 Wavelets in Two Dimensions	26
2.1.3 Shearlets	27
2.2 Wavelet and Shearlet Transforms	33
2.3 Complex Wavelet and Shearlet Transforms	34
3 Complex Shearlet Transforms and the Visual Cortex	41
4 Applications	45
4.1 Edge Detection	45
4.1.1 Phase Congruency	47
4.1.2 Detection of Singularities With Pairs of Symmetric and Odd-Symmetric Wavelets	51
4.1.3 Edge Detection With Pairs of Symmetric and Odd-Symmetric Shearlets	56
4.2 Image Quality Assessment	62
4.2.1 Structural Similarity Index (SSIM)	63
4.2.2 Complex Shearlet-Based Image Similarity Measure	67
4.2.3 Numerical Results	71
5 Discussion	75
A Some Definitions and Formulas	79
A.1.1 Polynomial Depth Search	81
References	83

List of Figures

2.1	A piecewise smooth function and a cartoon like image	11
2.2	Example of Fourier modes.	15
2.3	Fourier series approximation of step function	19
2.4	The Meyer wavelet	21
2.5	Wavelet approximations of step function	25
2.6	The two-dimensional Meyer wavelet	27
2.7	Decay of two-dimensional wavelet coefficients	28
2.8	Application of scaling and shear operators	30
2.9	Frequency tilings of wavelet- and shearlet-based systems	31
2.10	Decay of shearlet coefficients	32
2.11	Complex Mexican hat wavelet	37
2.12	Complex shearlets	40
3.1	Behavior of simple and complex cells	42
3.2	Gabor wavelets fitted to experimentally measured simple cell receptive fields	43
4.1	Concept of edge detection	46
4.2	Phase congruency	48
4.3	Comparison of wavelet- and shearlet-based phase congruency	51
4.4	Symmetric and odd-symmetric complex wavelets at jump discontinuity	53
4.5	Wavelet-based detection of singularities	56
4.6	Shearlet-based edge detection	59
4.7	Comparison of the shearlet-based edge measure with phase congruency and the Canny edge detector	60
4.8	The shearlet-based edge measure can discriminate between smooth and sharp transitions	61
4.9	A natural image perturbed by a block structure and Gaussian noise . .	62
4.10	Comparison of MSE and SSIM	65
4.11	The complex shearlet-based image similarity measure	70

Chapter 1

Introduction

In the course of this thesis, two image processing applications, namely a computational rule for edge detection and a full-reference image quality measure, based on so-called complex shearlet transforms will be developed. A complex shearlet transform can be seen as a complex-valued generalization of the shearlet transform, which was originally introduced by Kutyniok, Labate, Lim and Weiss in 2005 [1] and applies anisotropically scaled and sheared wavelet-based atoms for decomposing signals of two and possibly higher dimensions. While shearlet-based decompositions have been shown to provide optimally sparse approximations of a certain class of natural images [2], the complex-valued shearlet transform introduced in chapter 2 has the virtue of exhibiting a Fourier-like shift invariance in its magnitude response, a property also known to be present in the responses of certain cells in the primary visual cortex, so-called complex cells. Both these observations motivate the assumption that complex shearlet transforms might provide a useful model for the functional behavior of parts of the early stages of the human visual system, suggesting their applicability to image processing tasks like edge detection or image quality assessment.

In chapter 2, some basic results and ideas from Fourier analysis, the theory of wavelets and the theory of shearlets will be reviewed, while the construction of wavelets and shearlets will be motivated via quest of finding optimally sparse representations of piecewise smooth functions in one and two dimensions. The essential part of chapter 2 in the context of this thesis, however, is the introduction of complex shearlet transforms. After discussing some basic properties of the Hilbert transform in one and two dimensions, such transforms will be defined by considering complex-valued atoms constructed from Hilbert transform pairs of shearlet generators.¹

Chapter 3 will briefly discuss the neurophysiological findings, forming the basis of our assumption that the complex-valued transforms defined in chapter 2 might provide useful models for the functional behavior of cells in the primary visual cortex. While this chapter is not integral to the understanding of the methods developed in chapter 4, it should provide an intuition beyond the mathematical arguments laid out in chapter

¹It should be noted that such a construction was already given by Storath in [3].

2, why the application of complex-valued shearlet atoms was considered in the first place.

In chapter 4, which should be regarded as the main part of this thesis, we will motivate, develop and examine two image processing applications based on the complex shearlet transforms introduced in chapter 2. First, a complex shearlet-based edge measure is defined in section 4.1, which exhibits various desirable properties such as a high degree of contrast invariance, a surprisingly precise localization of edges mostly independent of the smoothness of the transition between two distinct geometrical regions and also provides estimates of the tangential direction of an edge at a specific location. As the construction of said measure was strongly inspired by the so-called phase congruency measure proposed by Kovesei in [4], its definition is preceded by a short review of this concept.

Finally, in section 4.2, a complex shearlet-based computational rule for full-reference image quality assessments is derived. In this task, we will consider pairs of images where one is a somehow distorted version of the other and try to determine the loss of image quality induced by the distortion as subjectively perceived by a human observer. One of the most significant applications of such quality assessments is the field of image compression, where the question as to which parts of an image can be subjected to a certain kind of compression without severely effecting the subjectively perceived image quality is of great importance. Again, the definition of the complex shearlet-based image quality metric will be preceded by the examination of two methods widely used for computing full-reference image quality assessments, namely the structural similarity index and the multiscale structural similarity index proposed by Wang et al. [5, 6].

The thesis will be closed by a short discussion of the complex shearlet transforms defined in chapter 2 and the applications developed in chapter 4.

Furthermore, all function spaces used in this thesis and some other useful definitions and formulas can be found in the appendix.

Please note that scripts for all MATLAB generated figures shown and all numerical experiments carried out in the course of this thesis can be found on the attached DVD.

Chapter 2

Complex Shearlet Transforms

In the remainder of this thesis, we will consider complex shearlet transforms to be transforms based on complex-valued functions of the form

$$\psi^c = \psi^{\text{even}} + i\psi^{\text{odd}},$$

where ψ^{even} is an even-symmetric real-valued shearlet (i.e. a function with possibly anisotropic support constructed from applying shearing, scaling and translation operators to a wavelet-like generating function) and ψ^{odd} is an odd-symmetric real-valued shearlet such that ψ^{even} and ψ^{odd} form a Hilbert transform pair. Such a construction has the benefit of bringing wavelet- or shearlet-based multiscale representation systems conceptually closer towards the classical Fourier basis in the sense that the Fourier modes

$$e^{i\xi\cdot} = \cos(\xi\cdot) + i\sin(\xi\cdot)$$

also follow this structure with the cosine being an even function, the sine being an odd function and both being 90 degree phase shifts of one another. While it can be argued that such a structure might be the cause of several desirable properties, it most notably introduces a large extent of shift invariance to the magnitude response of a complex shearlet transform, as we shall see later in this chapter.

The first section of this chapter will give a brief introduction to the basic concepts and results in Fourier analysis, the theory of wavelets and the theory of shearlets while motivating the latter two in the context of piecewise smooth functions in one and two dimensions. This will be pivotal to both understanding and formalizing the applications developed later in this thesis and provide the necessary background for the main part of this chapter, which is the discussion of the respective complex transforms.

2.1 From Fourier to Shearlets via an Analysis of Optimality

The main goal of this thesis is to develop methods for image processing applications like edge detection or image quality assessment by mimicking certain parts of the human visual system on a functional level. The central assumption we adopt here - and whose validity and origins will be discussed in detail in chapter 3 - is that at the early stages of human visual processing, some kind of time-frequency decomposition is performed and that evolutionary development in some sense optimized this decomposition for the tasks ahead. This rationale immediately raises the question as to which properties a time-frequency decomposition should have to be optimal for human visual processing, if such notions of optimality can be formalized in mathematical language and if it is possible to actually design decompositions fulfilling them.

The first step towards such an analysis of optimality is certainly to provide a precise mathematical description of the signals which are to be decomposed and what we actually mean by decomposition. For our purposes, we will consider images to be of a cartoon-like character, a notion first introduced by Donoho in [7] which can be seen as a generalization of the concept of one-dimensional piecewise smooth functions (i.e. functions which are smooth apart from a finite number of jump singularities) to higher dimensions. For two examples of such functions, see figure 2.1.

Definition 2.1.1 (Piecewise smooth functions). *Let $f \in L^2(\mathbb{R})$ with f being compactly supported in the unit interval, i.e. $f(x) = 0$ for all $x \in \mathbb{R} \setminus [0, 1]$. If the interval $[0, 1]$ can be partitioned into non-intersecting open subintervals $(t_i, t_{i+1})_{i \in \{1, \dots, K\}} \subset [0, 1]$ with $K \in \mathbb{N}$ such that the restrictions $f|_{(t_i, t_{i+1})} \in C^2(t_i, t_{i+1})$ are twice continuously differentiable and the one-sided limits $\lim_{x \downarrow t_i} f|_{(t_i, t_{i+1})}(x)$, $\lim_{x \uparrow t_{i+1}} f|_{(t_i, t_{i+1})}(x)$ are finite for all $i \in \{1, \dots, K - 1\}$, f is called piecewise smooth.*

Definition 2.1.2 (Cartoon-like image functions, as in [8]). *Let $f \in L^2(\mathbb{R}^2)$ with its support contained in the closed unit square, i.e. $f(x, y) = 0$ for all $(x, y) \in \mathbb{R}^2 \setminus [0, 1]^2$. If f can be written as*

$$f = f_0 + \mathbf{1}_B f_1,$$

where $\mathbf{1}$ denotes the characteristic function, $B \subset [0, 1]^2$ is a set with ∂B being a closed C^2 -curve with bounded curvature, $f_0, f_1 \in C^2(\mathbb{R}^2)$ have their support contained in $[0, 1]^2$ and $\|f_0\|_{C^2}, \|f_1\|_{C^2} \leq 1$, then f is said to be a cartoon-like image function.

When considering the here defined notion of a cartoon-like image as a model for natural images, we shouldn't worry too much about the fact that images typically contain more than one significant object. This restriction only was introduced to ease the mathematical analysis and there will be no danger in claiming that any time-frequency decomposition which is in some sense optimal for decomposing a function, like the one shown in figure 2.1, will also be optimal for decomposing images containing several smooth structures, where each one is enclosed by a C^2 -curve. Still, we should keep in mind that the claim brought forward by this definition of a cartoon-like image, namely

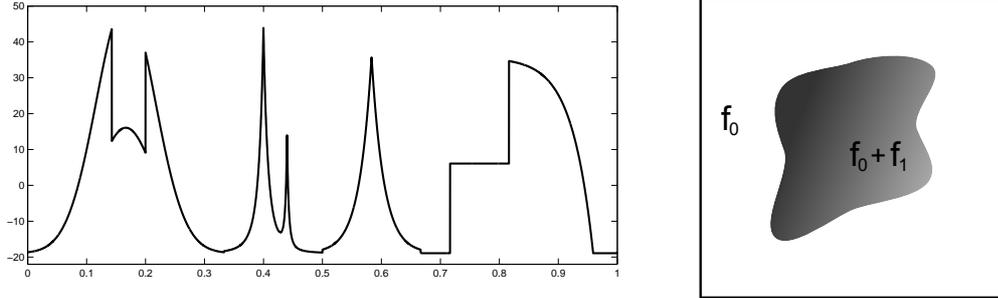


Figure 2.1: Example of a piecewise smooth function (left, computed using [9]) and a grayscale plot of a cartoon-like image function with $f_0 = 0$ (right).

that at least a vast majority of natural images consist of smooth areas, which are nicely separated from one another by (piecewise) smooth closed curves, is a very strong one. Furthermore, this claim is the basis of the mathematical analysis concerning the optimality of shearlet-based approximations in section 2.1.3 and therefore an essential premise for this whole thesis.

Having fixed the mathematical descriptions of the signals we aim to analyze, we can now turn back to the question of how we are going to decompose them. It is no coincidence that in the definitions of piecewise smooth functions and cartoon-like image functions, we chose the separable Hilbert spaces $L^2(\mathbb{R})$ and $L^2(\mathbb{R}^2)$. By doing so, we made sure to have a scalar product at our disposal, which can serve as a natural measure of the correlation of two functions inside $L^2(\mathbb{R}^d)$ (if not specified otherwise, we will from now on assume d to be either 1 or 2). Furthermore, we know from the theory of orthonormal bases that any $f \in L^2(\mathbb{R}^d)$ can be written uniquely as

$$f = \sum_{i \in I} \langle f, \phi_i \rangle \phi_i, \quad (2.1)$$

where $(\phi_i)_{i \in I} \subset L^2(\mathbb{R}^d)$ is an orthonormal basis of $L^2(\mathbb{R}^d)$, I is some index set and the L^2 -scalar product is naturally given by

$$\langle \cdot, \cdot \rangle: L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \rightarrow \mathbb{C}: (f_1, f_2) \mapsto \int_{\mathbb{R}^d} f_1(x) \overline{f_2(x)} dx.$$

Hence, at least in the setting of orthonormal bases it is clear that the function f can easily be reconstructed from the set of measurements $(\langle f, \phi_i \rangle)_{i \in I}$.

So for decomposing signals $f \in L^2(\mathbb{R}^d)$, we will try to find suitable subsets $\Phi \subset L^2(\mathbb{R}^d)$, so-called dictionaries, whose span lies dense in $L^2(\mathbb{R}^d)$, i.e. $\overline{\text{span}(\Phi)} = L^2(\mathbb{R}^d)$ and consider the set of inner products $\{\langle f, \phi \rangle : \phi \in \Phi\}$ for analyzing f . Additionally, we will ask for a tractable, ideally linear, way of reconstructing f from these coefficients. This might not sound necessary for tasks like image quality assessment or edge detection but it certainly is for others, like decomposition-based texture synthesis, where one tries to reconstruct different textures from only first and second order statistics computed on the set of inner products. Furthermore, our analysis of optimality will revolve around

the question of how few coefficients of a decomposition will be sufficient to compute good enough approximations of the original signal f , which obviously cannot be tackled without having a way of reconstructing from these coefficients.

Before formally defining this notion of optimality, we will introduce two additional conditions for candidate dictionaries $\Phi \subset L^2(\mathbb{R}^d)$, which are motivated mostly by technical considerations but will have the great benefit of putting us within the mathematical framework of frame theory, where, amongst other advantages, the existence of a simple reconstruction formula closely resembling (2.1) will come essentially for free (for a short introduction to the basic vocabulary of frame theory, please refer to [10, p. 2-12]). From now on, we will restrict ourselves to dictionaries $\Phi = (\phi_i)_{i \in I} \subset L^2(\mathbb{R}^d)$ where the index set I is countable and who fulfill the so-called frame condition

$$A\|f\|^2 \leq \sum_{i \in I} |\langle f, \phi_i \rangle|^2 \leq B\|f\|^2 \quad (2.2)$$

for all $f \in L^2(\mathbb{R}^d)$ and some fixed real constants $0 < A \leq B$. Using standard frame theoretical terminology, we can now perform the decomposition of a function $f \in L^2(\mathbb{R}^d)$ by applying the so-called analysis operator T , given by

$$T: L^2(\mathbb{R}^d) \rightarrow \ell^2(I): f \mapsto (\langle f, \phi_i \rangle)_{i \in I} \quad (2.3)$$

and reconstructions by using the adjoint operator T^* , the so-called synthesis operator,

$$T^*: \ell^2(I) \rightarrow L^2(\mathbb{R}^d): (c_i)_{i \in I} \mapsto \sum_{i \in I} c_i \phi_i, \quad (2.4)$$

via the reconstruction formula

$$f = \sum_{i \in I} \langle f, \phi_i \rangle (T^*T)^{-1} \phi_i, \quad (2.5)$$

where $T^*T: L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$, the so-called frame operator, is self-adjoint and invertible (for a simple proof of the invertibility of T^*T and formula (2.5), see [10, p. 5f.]). It is worth noting at this point that the frame condition (2.2) is a relaxation of Parseval's identity $\sum_{i \in I} |\langle f, \phi_i \rangle|^2 = \|f\|^2$, which is fulfilled for an orthonormal sequence $\Phi = (\phi_i)_{i \in I} \subset L^2(\mathbb{R}^d)$ and all $f \in L^2(\mathbb{R}^d)$ if and only if Φ is also a basis of $L^2(\mathbb{R}^d)$. Hence, frames can be seen as a generalization of the concept of bases which also allows for redundant representation systems, i.e. systems whose span is dense in the ambient space but who are not linearly independent. Considering such overcomplete systems can be very useful when trying to reconstruct signals via equation (2.5) when some of the coefficients $\langle f, \phi_i \rangle$ got lost, and - as we shall see later in this chapter - can also lead to extremely efficient representations of signals which would not be achievable with orthonormal bases.

As it was already stated, we will consider frames $\Phi \subset L^2(\mathbb{R}^d)$ to be well suited for decomposing and analyzing piecewise smooth functions or cartoon-like image functions,

if only a small number of analysis coefficients suffice to reconstruct good approximations of an original signal $f \in L^2(\mathbb{R}^d)$ i.e. if Φ provides sparse approximations. To formally compare different frames with respect to sparsity, we will look at the decay of the error of so-called N -term approximations (i.e. approximations of the form (2.5) where the sum on the right hand only consists of $N \in \mathbb{N}$ terms) when sending N to infinity. Naturally, this approach requires a mechanism for selecting these terms, which would ideally always choose the set of terms providing the best approximation possible. It can easily be shown that for the special case of an orthonormal basis (see lemma A.1.4 in the appendix), the best N -term approximation in fact coincides with the N -term approximation associated with the N largest coefficients (in magnitude). While this does not remain true when considering general frames, we will still adopt this rather simple but efficient selection approach for reasons of practicality.

Definition 2.1.3 (N -term approximation with N largest coefficients). *Let $\Phi = (\phi_i)_{i \in I} \subset L^2(\mathbb{R}^d)$ be a normalized frame for $L^2(\mathbb{R}^d)$, i.e. $\|\phi_i\| = 1$ for all $i \in I$, $f \in L^2(\mathbb{R}^d)$ and $N \in \mathbb{N}$, then the N -term approximation associated with the N largest coefficients is given by*

$$f_N = \sum_{i \in I_N} \langle f, \phi_i \rangle (T^*T)^{-1} \phi_i, \quad (2.6)$$

where T is the analysis operator, T^* its adjoint and the index set $I_N \subset I$ selects the N largest analysis coefficients in magnitude, i.e.

$$|I_N| = N \text{ and } |\langle f, \phi_i \rangle| \geq |\langle f, \phi_j \rangle|$$

for all $i \in I_N$ and $j \in I \setminus I_N$.

Please note that the set I_N and therefore also the approximation f_N in definition 2.1.3 is not necessarily unique. Furthermore, assuming the frame Φ to be normalized can of course be done without loss of generalization.

To measure the sparsity of the decompositions of piecewise smooth functions or cartoon-like image functions provided by a frame, we will from now on look at the decay rate of the error of the N -term approximation associated with the N largest coefficients in magnitude, i.e. we will try to find the largest exponent $\alpha \in \mathbb{R}^+$ such that

$$\|f - f_N\| \lesssim N^{-\alpha} \quad (2.7)$$

for all $f \in L^2(\mathbb{R}^d)$ within the class of piecewise smooth functions or cartoon-like images as N goes to infinity, where f_N is a N -term approximation of type (2.6). The symbol \lesssim denotes that the function on the left-hand side decays asymptotically at least as fast as the function on the right-hand side, i.e. there exists a constant $C \in \mathbb{R}$ and a number $N_0 \in \mathbb{N}$ such that $\|f - f_N\| < CN^{-\alpha}$ for all $N \geq N_0$, which is equivalent to $\|f - f_N\| \in \mathcal{O}(N^{-\alpha})$ in Landau notation.

The question remains here, whether it is possible for some classes of functions to define a theoretical upper bound for the exponent α in equation (2.7), such that any frame whose N -term approximations obey this decay rate can truly be considered to provide optimally sparse representations. It was in fact shown by Donoho in [7] that for each

dictionary $\Phi = (\phi_i)_{i \in I} \subset L^2(\mathbb{R}^2)$, there exists at least one cartoon-like image function f such that for each (countable) sequence of coefficients $(c_j)_{j \in J}$ with $J \subset I$ fulfilling $f = \sum_{j \in J} c_j \phi_j$ and generated from a practically feasible selection process (for a precise definition of feasibility in this context, see section A.1.1 in the Appendix), it holds that

$$\|(c_j)_{j \in J}\|_{\ell^p} = \left(\sum_{j \in J} |c_j|^p \right)^{1/p} = \infty, \quad (2.8)$$

for all $0 < p < \frac{2}{3}$. This means in return that when considering $(c_n^{\text{ord}})_{n \in \mathbb{N}}$ to be a rearrangement of the original sequence with decreasing order (in magnitude), the best decay we can hope for is given by $|c_n^{\text{ord}}| \lesssim n^{-\frac{3}{2}}$. As the decay of the analysis coefficients can indirectly be related to the decay of the error of a N -term approximations via the frame condition (2.2) and by using a few other methods and results from frame theory, we finally arrive at the result on optimally sparse approximations for cartoon-like images stated in definition 2.1.4. For a detailed explanation of this derivation, please refer to [2, p. 155ff.].

Definition 2.1.4 (Optimally sparse approximations of cartoon-like images, as in [2]). *Let $\Phi = (\phi_i)_{i \in I} \subset L^2(\mathbb{R}^2)$ be a normalized frame for $L^2(\mathbb{R}^2)$, i.e. $\|\phi_i\| = 1$ for all $i \in I$, then Φ is said to provide optimally sparse approximations of cartoon-like image functions (see definition 2.1.2) if the N -term approximations associated with the N largest coefficients in magnitude $f_N \in L^2(\mathbb{R}^2)$ (see definition 2.1.3) fulfill*

$$\|f - f_N\| \lesssim N^{-1}$$

for all cartoon-like image functions $f \in L^2(\mathbb{R}^2)$.

In the upcoming sections, we will now look at Fourier-, wavelet- and shearlet-based constructions of frames for $L^2(\mathbb{R})$ (respectively $L^2(a, b)$ in the Fourier case) and $L^2(\mathbb{R}^2)$, and analyze the decay rate of the error of their N -term approximations of piecewise smooth functions and cartoon-like image functions in the spirit of equation (2.7) and definition 2.1.3. In the case of cartoon-like image functions, we will be especially interested in finding out whether one of these frames actually achieves the optimal decay of N^{-1} proposed in definition 2.1.4. It will turn out that wavelets, while being nicely suited for representing one-dimensional piecewise smooth functions, fail to do so due to their incapacity of efficiently describing two-dimensional geometric features such as curve-like discontinuities occurring at edges. This shortcoming, however, will be remedied in the construction of shearlet-based frames whose N -term approximations of cartoon-like image functions actually achieve the desired error decay rate of N^{-1} .

2.1.1 Fourier Analysis

The first approach we'll examine is to decompose one-dimensional piecewise smooth functions using sine and cosine waves with varying frequencies, that is, we will consider

atomic functions of the form

$$e^{in\cdot} = \cos(n\cdot) + i \sin(n\cdot)$$

with $n \in \mathbb{Z}$, which puts us right into the realm of Fourier analysis. The origins of this branch of mathematics date back to the 18th century, when mathematicians became more and more interested in examining the idea that many real world phenomena like ocean waves could be described as the result of linear combinations of occurrences of a much more simple nature. Eventually, Jean Baptiste Joseph Fourier (1768-1830) made the claim¹ that in fact any integrable 2π -periodic function (i.e. a function $f \in L^1(\mathbb{R}/2\pi\mathbb{Z})$) could be written as

$$f(x) = \sum_{n \in \mathbb{Z}} \left(\frac{1}{2\pi} \int_0^{2\pi} f(t) \overline{e^{int}} dt \right) e^{inx}, \quad (2.9)$$

for all $x \in [0, 2\pi)$ and showed how such an expansion could be exploited to solve linear partial differential equations in his famous book *The Analytic Theory of Heat* (this little historical fact is taken from [11, p. 3]). To get a quick intuition about the applicability of sinusoidal waves in this context, one should consider that the Fourier modes $e^{in\cdot}$ are in fact eigenfunctions of the Laplace operator ∇^2 used in the heat equation $\frac{\partial f}{\partial t} = \alpha \nabla^2 f$ and that by the linearity of ∇^2 , each linear combination of solutions will again be a solution.

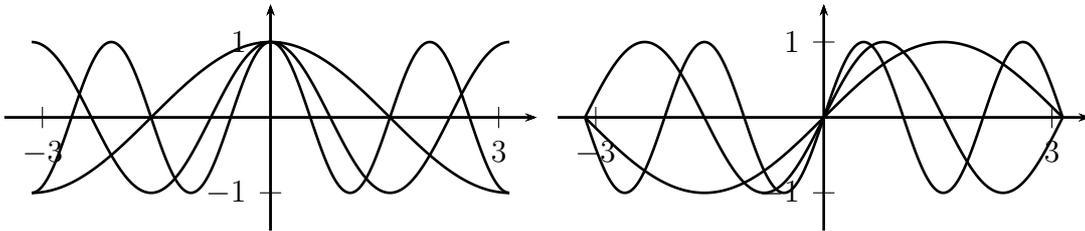


Figure 2.2: Plots of the functions $e^{ni\cdot}$ for $n \in \{1, 2, 3\}$ on the interval $[-\pi, \pi]$, with the real part $\cos(n\cdot)$ shown on the left and the imaginary part $\sin(n\cdot)$ shown on the right side.

But here, we are certainly not in the business of solving linear partial differential equations. Our program is to find representation systems for the Hilbert space $L^2(\mathbb{R})$, providing us with sparse representations of piecewise smooth functions. The first thing one should note at this point is that - due to their lack of decay - the functions $e^{in\cdot}$ don't even belong to $L^2(\mathbb{R})$. We will therefore constrain our examinations for the time being to Hilbert spaces $L^2(a, b)$ that are restricted to a closed interval $[a, b]$. While this is not a problem for the analysis of sparse approximations of elements of the class of piecewise smooth functions, whose support is by definition contained in $[0, 1]$, it is a

¹While this claim holds for many functions, it eventually turned out to be false. For a continuous counterexample, see [11, p. 67ff.]. For a more elaborate analysis of pointwise convergence, see chapter 15 in [11].

first hint that it could be useful to consider atomic functions that are actually localized on the real line.

It is now time to state the first important result about the system of functions of the form $(e^{in})_{n \in \mathbb{Z}}$, namely that they form an orthogonal basis for the space of square integrable functions on the interval $[-\pi, \pi]$, $L^2(-\pi, \pi)$, which of course can easily be normalized to form an orthonormal basis. This implies that - with respect to the frame condition (2.2) - the normalized Fourier modes constitute a frame for $L^2(-\pi, \pi)$ with constants $A = B = 1$ and a frame operator $T^*T = \text{Id}$.

Theorem 2.1.5 (Orthonormal Fourier basis in $L^2(-\pi, \pi)$). *The set $\{(2\pi)^{-\frac{1}{2}}e^{in} : n \in \mathbb{Z}\} \subset L^2(-\pi, \pi)$ is an orthonormal basis for $L^2(-\pi, \pi)$.*

Proof. See for example [12, p. 78] or [13, p. 72]. □

While the Fourier modes cannot form a basis for spaces of integrable or square integrable functions defined on the whole real line, one can still ask whether expressions of the form $\int_{\mathbb{R}} f(x)e^{i\xi x}$ are well defined for arbitrary functions f in $L^1(\mathbb{R})$ (respectively $L^2(\mathbb{R})$) and whether there might be a way of reconstructing f from such coefficients. As it turns out, when allowing the frequency ξ to take values on the whole real line, there is in fact a reconstruction formula for arbitrary continuous and integrable functions basically given by a continuous version of (2.9), i.e. the sum is replaced by an integral. This result is formalized in the following theorem, which defines the famous Fourier transform and its inverse.

Theorem 2.1.6 (Fourier transform and inverse Fourier transform). *Let $f \in L^1(\mathbb{R})$ then the so-called Fourier transform, given by*

$$(\mathcal{F}f)(\xi) = \hat{f}(\xi) = \int_{\mathbb{R}} f(x)e^{-ix\xi} dx \quad (2.10)$$

for $\xi \in \mathbb{R}$, is a well defined, bounded and continuous operator from $L^1(\mathbb{R})$ to $L^\infty(\mathbb{R})$. If additionally $\hat{f} \in L^1(\mathbb{R})$ (which implies the continuity of f), the inverse Fourier transform is given by

$$(\mathcal{F}^{-1}\hat{f})(x) = \check{\hat{f}}(x) = (2\pi)^{-1} \int_{\mathbb{R}} \hat{f}(\xi)e^{ix\xi} d\xi, \quad (2.11)$$

and it holds that

$$f(x) = \check{\check{f}}(x) \quad (2.12)$$

for all $x \in \mathbb{R}$.

For functions in $L^2(\mathbb{R}^n)$ with $n > 1$, the product $x\xi$ in equations (2.10) and (2.11) is replaced by the inner product $\langle x, \xi \rangle$.

Proof. For the properties of (2.10), see [12, p. 213f.], for the Fourier inversion theorem (2.12), see [12, p. 218]. □

As a matter of fact, for functions $f \in L^2(\mathbb{R}) \setminus L^1(\mathbb{R})$, the integral in equation (2.10) must not converge. This problem, however, can be overcome by defining the Fourier transform of such functions as the limit of the Fourier transforms of functions inside $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$, which lies dense in $L^2(\mathbb{R})$. So the Fourier transform is actually well defined on the Hilbert space $L^2(\mathbb{R})$. For more details on this construction, see [12, p. 221f.].

While in this thesis, we will only consider the approximations provided by discrete transforms (i.e. transforms based on countable systems), the here defined Fourier transform will be an important tool for the construction of wavelet- and shearlet-based systems as well as in the definition of the notion of phase congruency (see section 4.1.1). Furthermore, we can now elaborate on the remark made at the very beginning of this chapter, that in magnitude, the Fourier transform is shift invariant.

Remark 2.1.7 (Shift invariance in magnitude of Fourier transform). *Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then the magnitude response of the Fourier transform is invariant under translations, i.e.*

$$|\mathcal{F}f(\xi)| = |\mathcal{F}(T_\lambda f)(\xi)|, \quad (2.13)$$

for all $\xi, \lambda \in \mathbb{R}$, where T_λ is the translation operator, given by

$$(T_\lambda f)(x) = f(x - \lambda) \quad (2.14)$$

for all $x \in \mathbb{R}$.

Proof. We simply compute

$$\begin{aligned} |\mathcal{F}(T_\lambda f)(\xi)| &= \left| \int_{\mathbb{R}} f(x - \lambda) e^{-i\xi x} dx \right| \\ &= \left| \int_{\mathbb{R}} f(x) e^{-i\xi(x+\lambda)} dx \right| \\ &= \underbrace{|e^{-i\xi\lambda}|}_1 \left| \int_{\mathbb{R}} f(x) e^{i\xi x} dx \right| \\ &= |\mathcal{F}f(\xi)|. \end{aligned}$$

□

Please note that the important step in the computation above was that due to the interplay of sine and cosine, the absolute value $|e^{ix}|$ is 1 for all $x \in \mathbb{R}$. This observation already gives an intuition, how by mimicking the structure of Fourier modes, shift invariance in magnitude can also be introduced to wavelet- and shearlet-based systems.

We will conclude this short trip into the realm of Fourier analysis by returning to our original goal, namely to analyze the decay behavior of N -term approximations of piecewise smooth functions obtained from the Fourier basis defined in theorem 2.1.5. To do this, we will consider the most basic kind of a piecewise smooth function, a simple step function restricted to the interval $[-\pi, \pi]$.

Example 2.1.8 (Fourier series approximations of a step function). Let $f \in L^2(-\pi, \pi)$ be given by

$$f(x) = \begin{cases} 1 & \text{if } x \in [-\pi, 0) \\ 0 & \text{else} \end{cases},$$

i.e. $f = \mathbf{1}_{[-\pi, 0)}$ with $x \in [-\pi, \pi]$, then the error of the N -term approximations associated with the N largest coefficients in magnitude $f_N \in L^2(-\pi, \pi)$ (see definition 2.1.3) provided by the orthonormal Fourier basis $\{(2\pi)^{-\frac{1}{2}}e^{in\cdot} : n \in \mathbb{Z}\} \subset L^2(-\pi, \pi)$ obeys

$$\|f - f_N\| \asymp N^{-\frac{1}{2}},$$

where \asymp denotes that $\|f - f_N\| \lesssim N^{-\frac{1}{2}}$ and $N^{-\frac{1}{2}} \lesssim \|f - f_N\|$. By lemma A.1.4, f_N coincides with the best N -term approximation. For a visual representation, see figure 2.3.

Proof. First, we note that the coefficients are given by

$$\begin{aligned} \langle f, (2\pi)^{-\frac{1}{2}}e^{in\cdot} \rangle &= (2\pi)^{-\frac{1}{2}} \int_{-\pi}^0 e^{-inx} dx \\ &= (2\pi)^{-\frac{1}{2}} \frac{1 - e^{i\pi n}}{n} i \\ &= (2\pi)^{-\frac{1}{2}} \frac{1 - \cos(n\pi)}{n} i \\ &= \begin{cases} (2\pi)^{-\frac{1}{2}} \frac{2i}{n} & \text{if } n \text{ is odd} \\ 0 & \text{else} \end{cases}, \end{aligned}$$

for $n \in \mathbb{Z} \setminus \{0\}$. Hence, we can define the set $I_N \subset \mathbb{Z}$ of the N largest coefficients as $I_N = \{n \in \mathbb{Z} : n \text{ is odd and } |n| < N\} \cup \{0\}$ (if N is even, we simply remove the index $N - 1$ from the set) and by lemma A.1.4, we get

$$\begin{aligned} \|f - f_N\|^2 &= \sum_{n \in \mathbb{Z} \setminus I_N} \left| \langle f, (2\pi)^{-\frac{1}{2}}e^{in\cdot} \rangle \right|^2 \\ &= \sum_{n \in \mathbb{Z} \setminus I_N} \left| (2\pi)^{-\frac{1}{2}} \frac{2i}{n} \right|^2 \\ &= \frac{2}{\pi} \sum_{n \in \mathbb{Z} \setminus I_N} \frac{1}{n^2} \\ &\asymp \int_N^\infty x^{-2} dx \\ &= N^{-1}. \end{aligned}$$

Furthermore, a closed formula for the N -th approximation error is given by

$$\|f - f_N\|^2 = \frac{\pi}{2} - \frac{2}{\pi} \sum_{n \in I_N \setminus \{0\}} \frac{1}{n^2}.$$

□

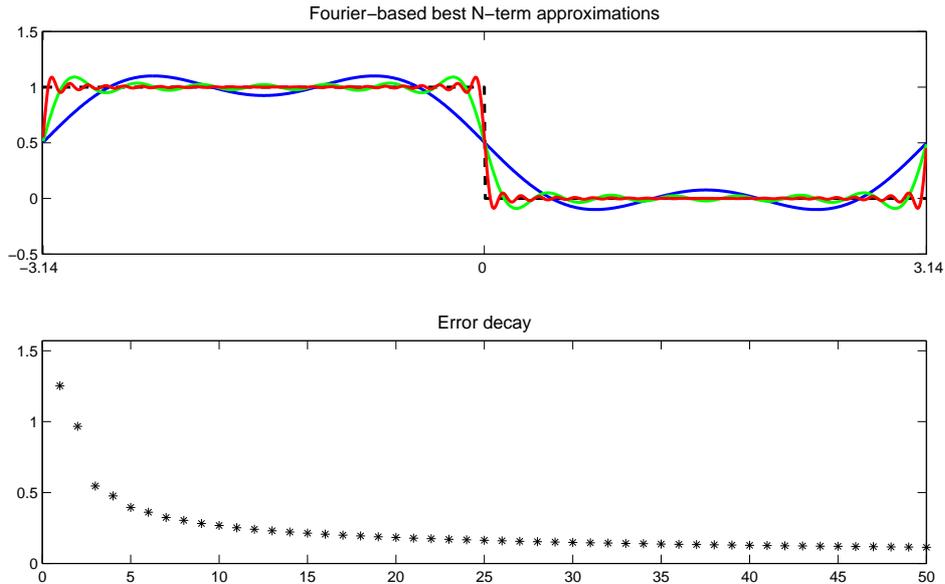


Figure 2.3: The first image shows the Fourier series approximations f_N of $f = \mathbf{1}_{[-\pi, 0)}$ on the interval $[-\pi, \pi]$ from example 2.1.8 for $N = 5$ (blue), $N = 15$ (green), and $N = 50$ (red). The second plot shows the errors $\|f - f_N\|$ for $N \in \{1, \dots, 50\}$.

When examining the N -term approximations plotted in figure 2.3, we notice that away from the discontinuity, the partial sums of the Fourier series quickly converge to the constant values 1 (respectively 0), while in a small neighborhood around 0, the approximations increasingly overshoot (respectively undershoot) these values - a very undesirable feature which is known in the literature as the so-called Gibbs phenomenon². The main reason why Fourier series approximations have difficulties with representing jump discontinuities is that periodic sine and cosine waves are not at all localized on the real line. Simply put, the position of a jump discontinuity on the real line matters to a sine or cosine wave only modulo the interval $[0, P]$, where P denotes the period of the wave. Another way of looking at this issue is that in the presence of a jump discontinuity, the analysis coefficients of the orthonormal Fourier basis only decay through increasing the frequency of the Fourier modes, an action that is completely independent of the localization of the jump discontinuity on the real line.

One rather simple fix to localize the functions $e^{in\cdot}$ on the real line would be to introduce a window function $g \in L^2(\mathbb{R})$ and consider atomic functions of the form

$$\phi_{a,b}(\cdot) = g(\cdot - a)e^{ib\cdot}, \quad (2.15)$$

where the frequency b and the translation parameter a can be chosen from some discrete subset of $\mathbb{R} \times \mathbb{R}$. Such functions were first proposed by Nobel prize winner Dennis Gabor in 1946 [14] and can, in contrast to the Fourier bases considered so far, in fact be used

²Named after the American scientist Josiah Willard Gibbs (1839 - 1903).

to construct orthonormal bases for $L^2(\mathbb{R})$. To give an example, the set

$$\Phi = \left\{ \phi_{m,n}(\cdot) = \mathbf{1}_{[0,1]}(\cdot - n)e^{2\pi im\cdot} : (m, n) \in \mathbb{Z} \times \mathbb{Z} \right\} \subset L^2(\mathbb{R})$$

indeed forms an orthonormal basis for $L^2(\mathbb{R})$ (see example 3.5.3 in [10]). While this approach resolves the issue of finding discrete orthonormal bases of $L^2(\mathbb{R})$, the question remains, how the window function g should be chosen to efficiently decompose certain classes of signals. The obvious problem arising with this question is that the optimal size and shape of the window g again depends on the local behavior of a signal f . It is possible that one function $f \in L^2(\mathbb{R})$ exhibits a strong oscillatory behavior on the interval $[0, 1]$, suggesting a very tight window, and is almost constant on the interval $[5, 6]$, suggesting a rather wide window. Hence, one would require a construction combining functions of different frequencies (as in the standard Fourier series), different locations on the real line (as introduced via the translation parameter in systems based on functions of the form (2.15)) and different levels of localization inside one discrete dictionary. This line of thought now directly leads us into the field of multiresolution analysis and wavelet theory.

2.1.2 Wavelets

As it was discussed at the end of the previous section, we now aim to construct discrete dictionaries for $L^2(\mathbb{R})$ containing atomic elements of varying frequency, location and localization. For this purpose, we will consider functions $\psi \in L^2(\mathbb{R})$, fulfilling the so-called wavelet admissibility condition

$$\int_{\mathbb{R}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty, \quad (2.16)$$

and sets of functions

$$\Psi_{a,b} = \left\{ \psi_{j,m}(\cdot) = a^{-\frac{j}{2}} \psi \left(\frac{\cdot - ma^j b}{a^j} \right) : (j, m) \in \mathbb{Z} \times \mathbb{Z} \right\} \subset L^2(\mathbb{R}), \quad (2.17)$$

with a fixed scaling density parameter $a > 1$ and a fixed translation density parameter $b > 0$. The convergence of the integral in equation (2.16) is a necessary condition for the existence of an inversion formula of the continuous wavelet transform, in which the scaling parameter, written as a^j in equation (2.17), and the translation parameter, written as $ma^j b$ in equation (2.17), take all possible values on the real line, with the exception of 0 in the case of the scaling parameter (see [15, p. 24f.]). Please note that in the case that $\psi \in L^1(\mathbb{R})$ (implying the continuity of $\hat{\psi}$), the admissibility condition (2.16) can only be satisfied if ψ has a vanishing mean value, that is

$$\hat{\psi}(0) = \int_{\mathbb{R}} \psi(x) dx = 0. \quad (2.18)$$

Hence, the admissibility condition enforces the function ψ - which here essentially plays the role of $e^{i\cdot}$ in the definition of the orthonormal Fourier basis (see theorem 2.1.5) - to

exhibit at least some extent of wavelike behavior. For an example of such a function, see figure 2.4.

While in the Fourier case, the dictionary was constructed by simple frequency modulations, the scaling parameter in (2.17) not only changes the frequency³ of ψ but also its degree of localization. To be precise, the atomic functions $\psi_{j,m}$ covering the lower frequencies will be badly localized on the real line while their high frequency counterparts will be highly localized. Naturally, using localized functions also requires the possibility of considering different translates, which is provided in (2.17) by the translation parameter m .

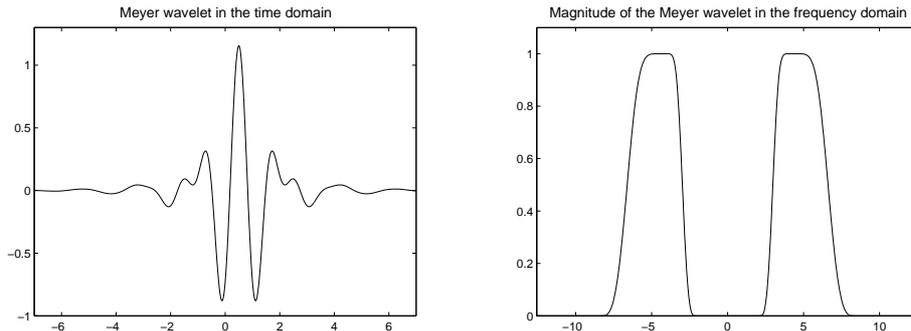


Figure 2.4: The Meyer wavelet (left) and its Fourier transform in magnitude (right). For a definition of the Meyer wavelet, see A.1.5. (computed with [9])

Analogous to the Fourier case, we will now try to construct dictionaries $\Psi_{a,b}$ based on a wavelet ψ such that $\Psi_{a,b}$ forms an orthonormal basis of $L^2(\mathbb{R})$. While Stromberg in 1982 [16] and Meyer in 1985 [17] published first specific examples of such orthonormal wavelet bases, we will consider a much more general approach developed by Meyer and Mallat in 1986 termed multiresolution analysis [18, 19]. The basic idea behind this technique is to construct a sequence of nested subspaces of $L^2(\mathbb{R})$ such that via orthogonal projections, each of these subspaces provides approximations of signals in $L^2(\mathbb{R})$ at different resolutions. Furthermore, it will be required that the union of these subspaces lies dense in $L^2(\mathbb{R})$ while their intersection is just the zero function (i.e. when going up the ladder of subspaces, the approximations go from arbitrarily bad to arbitrarily good) and that each subspace has an orthonormal basis given by the translates of a so-called scaling function ϕ considered at different dyadic scales. This means that the whole system of subspaces can be constructed by scaling and translating a single generator function ϕ - a structure closely resembling the definition of a discrete wavelet system $\Psi_{a,b}$ in equation (2.17). Assuming the existence of such a multiresolution analysis, we can then look at the difference between two neighboring subspaces described by the orthogonal complement of the smaller one inside its larger neighbor. The direct sum of this thereby constructed sequence of increments will then be equivalent to $L^2(\mathbb{R})$ and we will show that it is possible to define a generator function ψ , fulfilling the wavelet admissibility condition, such that each increment space has

³Note that $\mathcal{F}f(a\cdot)(\xi) = \frac{1}{|a|} \hat{f}\left(\frac{\xi}{a}\right)$.

an orthonormal basis constructed from translates of ψ considered at different scales. Eventually, this process will lead to a simple way of constructing functions ψ such that the sets $\Phi_{a,b}$ with $a = 2$ and $b = 1$ defined in (2.17) form an orthonormal basis of $L^2(\mathbb{R})$.

Definition 2.1.9 (Multiresolution Analysis, as in [15]). *A sequence $(V_j)_{j \in \mathbb{Z}}$ of closed subspaces with $V_j \subset L^2(\mathbb{R})$ for all $j \in \mathbb{Z}$ constitutes a multiresolution analysis if the following properties are satisfied*

$$\forall j \in \mathbb{Z}: V_{j+1} \subset V_j, \quad (2.19)$$

$$\forall j \in \mathbb{Z}: f(\cdot) \in V_j \Leftrightarrow f\left(\frac{\cdot}{2}\right) \in V_{j+1}, \quad (2.20)$$

$$\forall j, m \in \mathbb{Z}: f(\cdot) \in V_j \Leftrightarrow f(\cdot - 2^j m) \in V_j, \quad (2.21)$$

$$\bigcap_{j=-\infty}^{\infty} V_j = \{0\}, \quad (2.22)$$

$$\bigcup_{j=-\infty}^{\infty} V_j = L^2(\mathbb{R}), \quad (2.23)$$

$$\exists \phi \in L^2(\mathbb{R}): \{\phi(\cdot - m) : m \in \mathbb{Z}\} \text{ is an orthonormal basis of } V_0. \quad (2.24)$$

The function ϕ in condition (2.24) is often called father wavelet or scaling function, while the spaces V_j are typically termed scaling spaces. Please note that (2.20) tells us that each space V_j is just a scaled version of V_0 , eventually justifying the term multiresolution analysis.

Given the existence of a multiresolution analysis $(V_j)_{j \in \mathbb{Z}}$, we can - following the program already sketched in the previous paragraph - define a sequence of so-called wavelet spaces $(W_j)_{j \in \mathbb{Z}}$, describing the difference between two neighboring scaling spaces, that is

$$W_j = V_j^\perp \subset V_{j-1}. \quad (2.25)$$

This implies that

$$V_{j-1} = W_j \oplus V_j$$

for all $j \in \mathbb{Z}$ and due to (2.19),

$$W_j \perp W_{j'}$$

for all $j, j' \in \mathbb{Z}$ with $j \neq j'$. Furthermore, we conclude from (2.22) and (2.23) that

$$\bigoplus_{j \in \mathbb{Z}} W_j = L^2(\mathbb{R}).$$

So the sequence $(W_j)_{j \in \mathbb{Z}}$ consists of mutually orthogonal spaces whose sum equals $L^2(\mathbb{R})$. If we were able to find a function ψ generating the spaces W_j in the same way as the scaling function ϕ generates the scaling spaces V_j , that is, for each space W_j , the set $\{\psi_{j,m}(\cdot) = 2^{-\frac{j}{2}} \psi(2^{-j} \cdot - m) : m \in \mathbb{Z}\}$ is an orthonormal basis, then we would have also constructed an orthonormal basis for $L^2(\mathbb{R})$. The following theorem shows that this is indeed possible.

Theorem 2.1.10 (Orthonormal wavelet basis for $L^2(\mathbb{R})$). *Let $(V_j)_{j \in \mathbb{Z}}$ be a multiresolution analysis (see definition 2.1.9), $\phi \in L^2(\mathbb{R})$ be the associated scaling function and let $h \in \ell^2(\mathbb{Z})$ be the sequence given by $h_n = \sqrt{2} \langle \phi(2 \cdot -n), \phi \rangle$, that is*

$$\phi(\cdot) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \phi(2 \cdot -n).$$

Then with $\hat{h}(\xi) = \sum_{n \in \mathbb{Z}} h_n e^{-in\xi}$,

$$\hat{\psi}(\xi) = 2^{-\frac{1}{2}} e^{-i\frac{\xi}{2}} \overline{\hat{h}\left(\frac{\xi}{2} + \pi\right)} \hat{\phi}\left(\frac{\xi}{2}\right) \quad (2.26)$$

defines the Fourier transform of a wavelet ψ such that for each $j \in \mathbb{Z}$,

$$\left\{ \psi_{j,m}(\cdot) = 2^{-\frac{j}{2}} \psi(2^{-j} \cdot -m) : m \in \mathbb{Z} \right\}$$

is an orthonormal basis for the space $W_j \subset L^2(\mathbb{R})$ (see equation (2.25)), and the set

$$\left\{ \psi_{j,m}(\cdot) = 2^{-\frac{j}{2}} \psi(2^{-j} \cdot -m) : (j, m) \in \mathbb{Z} \times \mathbb{Z} \right\}$$

is an orthonormal basis for $L^2(\mathbb{R})$.

Proof. See Daubechies [15, p. 129ff.] or Mallat [13, p. 278ff.]. □

The question remains, however, how we can construct a sequence of scaling spaces satisfying the conditions (2.19)-(2.24). One should note in this regard that the only free variable appearing in the definition of a multiresolution analysis is in fact the scaling function ϕ . So the right question might be which properties a function ϕ must have, such that its translates considered on all dyadic scales construct scaling spaces satisfying the six axioms from definition 2.1.9. While this can in fact be answered (see for instance [15, p.140ff.]), we will go even one step further and reconsider the so-called scaling relation

$$\phi(\cdot) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \phi(2 \cdot -n), \quad (2.27)$$

which already appeared in theorem 2.1.10. Applying the Fourier transform to both sides gives

$$\begin{aligned} \hat{\phi}(\xi) &= \frac{1}{\sqrt{2}} \hat{h}\left(\frac{\xi}{2}\right) \hat{\phi}\left(\frac{\xi}{2}\right) \\ &= \frac{1}{2} \hat{h}\left(\frac{\xi}{2}\right) \hat{h}\left(\frac{\xi}{4}\right) \hat{\phi}\left(\frac{\xi}{4}\right) \end{aligned}$$

and assuming the continuity of $\hat{\phi}$ in 0, we can go to the limit and get

$$\hat{\phi}(\xi) = \hat{\phi}(0) \prod_{j=1}^{\infty} 2^{-\frac{1}{2}} \hat{h}(2^{-j}\xi). \quad (2.28)$$

Hence, it seems to be possible to define the scaling function ϕ , and thereby implicitly a sequence of scaling spaces constituting a multiresolution analysis, by a discrete sequence $h \in \ell^2(\mathbb{Z})$. The exact conditions such a sequence h has to satisfy such that its associated scaling function is the generator of a multiresolution analysis are given in the theorem below.

Theorem 2.1.11 (Multiresolution analysis from scaling filter h). *Let $h \in \ell^2(\mathbb{Z})$ be a sequence satisfying*

$$\begin{aligned} \hat{h}(0) &= \sqrt{2}, \\ \inf_{\xi \in [-\pi/2, \pi/2]} |\hat{h}(\xi)| &> 0, \\ \forall \xi \in \mathbb{R}: \quad |h(\hat{\xi})|^2 + |\hat{h}(\xi + \pi)|^2 &= 2 \end{aligned}$$

with \hat{h} being continuously differentiable in a neighborhood of 0, then

$$\hat{\phi}(\xi) = \prod_{j=1}^{\infty} 2^{-\frac{j}{2}} \hat{h}(2^{-j}\xi). \quad (2.29)$$

is the Fourier transform of a scaling function $\phi \in L^2(\mathbb{R})$. That is, the spaces

$$V_j = \overline{\text{span} \left\{ \phi_{j,m}(\cdot) = 2^{-\frac{j}{2}} \phi(2^{-j} \cdot -m) : m \in \mathbb{Z} \right\}}$$

constitute a multiresolution analysis for $L^2(\mathbb{R})$ (see definition 2.1.9).

Proof. See Mallat [13, p. 270ff.]. □

Theorems 2.1.11 and 2.1.10 provide us with an efficient recipe for constructing wavelet orthonormal bases for the Hilbert space $L^2(\mathbb{R})$. While we will not examine specific constructions of such bases here, several examples are given in [13, p. 289ff.].

We have now introduced a way of constructing orthonormal bases for $L^2(\mathbb{R})$ whose atoms not only have varying frequencies but also varying levels of localization on the real line. Hence, we would expect that N -term approximations provided by these bases outperform the Fourier basis defined in 2.1.5 at heavily localized features such as jump discontinuities. A first hint that this is indeed the case is given in figure 2.5, where we used a wavelet multiresolution analysis to compute best N -term approximations of the step function already considered in figure 2.3. Furthermore, we will state a much more general result for piecewise Lipschitz continuous functions on a closed interval taken from [13, p. 456f.].

Theorem 2.1.12 (Error decay of wavelet approximations of piecewise regular functions). *Let $\psi \in L^2(\mathbb{R})$ be a wavelet associated with a multiresolution analysis⁴ for $L^2(-\pi, \pi)$ with $q \in \mathbb{N}$ vanishing moments, that is*

$$\int_{-\pi}^{\pi} x^k \psi(x) dx = 0 \quad (2.30)$$

⁴For the construction of multiresolutions on closed intervals, see chapter 7.5.3 in [13].

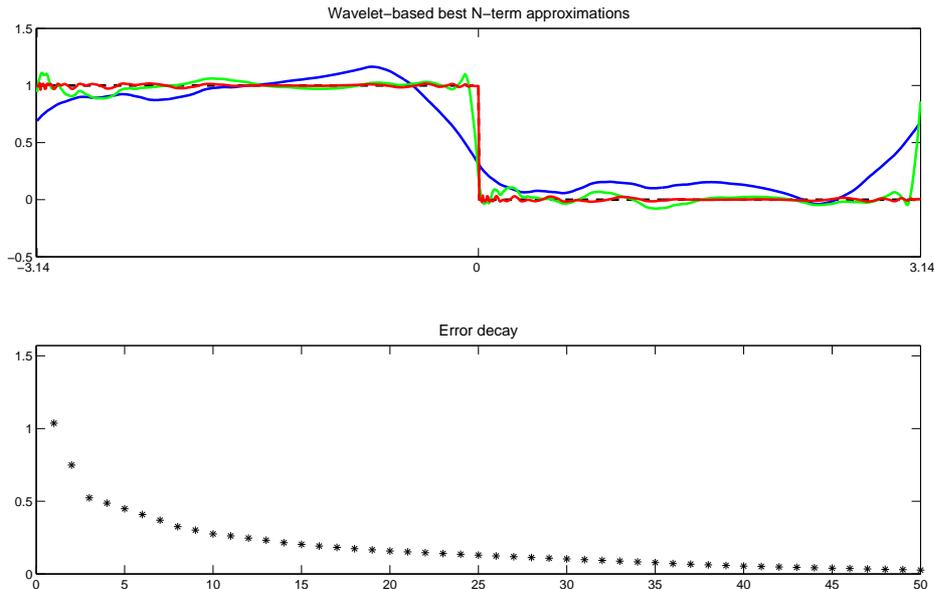


Figure 2.5: The first image shows the best N -term approximations f_N of $f = \mathbf{1}_{[-\pi,0]}$ on the interval $[-\pi, \pi]$ obtained from a wavelet multiresolution analysis for $N = 5$ (blue), $N = 15$ (green), and $N = 50$ (red). The second plot shows the approximated errors $\|f - f_N\|$ for $N \in \{1, \dots, 50\}$. In contrast to the Fourier series approximations shown in figure 2.3, the Gibbs oscillations have already almost disappeared for $N = 50$ and the error decays much faster (computed using [9]).

for all integers $0 \leq k < q$ and let $f \in L^2(-\pi, \pi)$ be uniformly Lipschitz α continuous (see definition A.1.6) with $1/2 < \alpha < q$ between a finite number of discontinuities $K \in \mathbb{N}$. Then the error of the N -term approximations associated with the N largest coefficients in magnitude of f (see definition 2.1.3) decays as

$$\|f - f_N\| \lesssim N^{-\alpha}. \quad (2.31)$$

Proof. See Mallat [13, p. 456f]. □

Theorem 2.1.12 tells us that the error decay rate for approximations of the function $f = \mathbf{1}_{[-\pi,0]}$ associated with a wavelet basis is only bounded by the number of vanishing moments of the corresponding wavelet ψ , which is a remarkable result and a great improvement to the error decay rate of $N^{-\frac{1}{2}}$ of the Fourier approximations.

At this point, we will consider our quest for constructing orthonormal bases for the decomposition of one-dimensional signals to be finished and turn to our original goal of analyzing frames for $L^2(\mathbb{R}^2)$ with respect to the class of cartoon-like images (see definition 2.1.2). While the orthonormal wavelet bases derived from a multiresolution analysis can easily be generalized to two dimensions, it will turn out that such a construction will indeed have difficulties when dealing with certain features newly arising in the two-dimensional setting, such as curve-like discontinuities.

2.1.2.1 Wavelets in Two Dimensions

One very simple but useful way of constructing an orthonormal wavelet basis for $L^2(\mathbb{R}^2)$ is to start with a one-dimensional multiresolution analysis and consider tensor products of the scaling function ϕ and the associated wavelet ψ , that is, $\phi \otimes \psi$, $\psi \otimes \phi$ and $\psi \otimes \psi$. As it turns out, these functions will be the generators of a separable orthonormal wavelet basis of $L^2(\mathbb{R}^2)$, a result which is formalized in the following theorem.

Theorem 2.1.13 (Orthonormal wavelet basis for $L^2(\mathbb{R}^2)$). *Let $\phi \in L^2(\mathbb{R})$ be the scaling function and $\psi \in L^2(\mathbb{R})$ be the wavelet of a multiresolution analysis for $L^2(\mathbb{R})$ (see definition 2.1.9) then the wavelets given by*

$$\begin{aligned}\psi^{(1)}(x, y) &= (\phi \otimes \psi)(x, y) = \phi(x)\psi(y), \\ \psi^{(2)}(x, y) &= (\psi \otimes \phi)(x, y) = \psi(x)\phi(y), \\ \psi^{(3)}(x, y) &= (\psi \otimes \psi)(x, y) = \psi(x)\psi(y)\end{aligned}$$

are the generators of an orthonormal basis of $L^2(\mathbb{R}^2)$, that is, the set

$$\left\{ \psi_{j,m}^{(k)}(\cdot) = 2^{-j} \psi^{(k)}(A_j \cdot -m) : j \in \mathbb{Z}, m \in \mathbb{Z}^2 \right\}, \quad (2.32)$$

with the scaling matrix $A = \begin{pmatrix} 2^{-j} & 0 \\ 0 & 2^{-j} \end{pmatrix}$ is an orthonormal basis of $L^2(\mathbb{R}^2)$.

Proof. See Mallat [13, p. 341f]. □

When also including the scaling function $(\phi \otimes \phi)(x, y) = \phi(x)\phi(y)$, we get a separable two-dimensional multiresolution analysis defined by the spaces $V_j^{2d} = V_j \otimes V_j$ for $j \in \mathbb{Z}$. For a visualization of the two-dimensional Meyer wavelets and the associated scaling function, see figure 2.6.

We have already discussed that in the one-dimensional case, the major problem with approximations based on the Fourier modes was that the functions $e^{i\xi \cdot}$ are not localized and therefore have difficulties representing highly localized features such as jump discontinuities - which serve as separators of two smooth parts in the definition of piecewise smooth function (see definition 2.1.1). Now, in two dimensions and having the cartoon-like image model in our mind (see definition 2.1.2), the discontinuities separating two smooth regions of a signal are not localized at a single point but along a (smooth) curve. This raises the question, whether the way the wavelets from set (2.32) are localized in the two-dimensional plane makes them also well suited for representing such anisotropic features (i.e. features that are directionally dependent). As it turns out, the isotropic nature of the scaling matrix A defined in theorem 2.1.13 leads to a non-optimal decay behavior of the wavelet coefficients along such curve-like discontinuities (for a visual illustration, see figure 2.7). This eventually translates into a non-optimal (compare definition 2.1.4) decay rate of the error of the best N -term approximations of cartoon-like image functions obtained from a two-dimensional orthonormal wavelet basis. A formal result was given by Kutyniok, Lim and Lemvig in [2, p. 159].

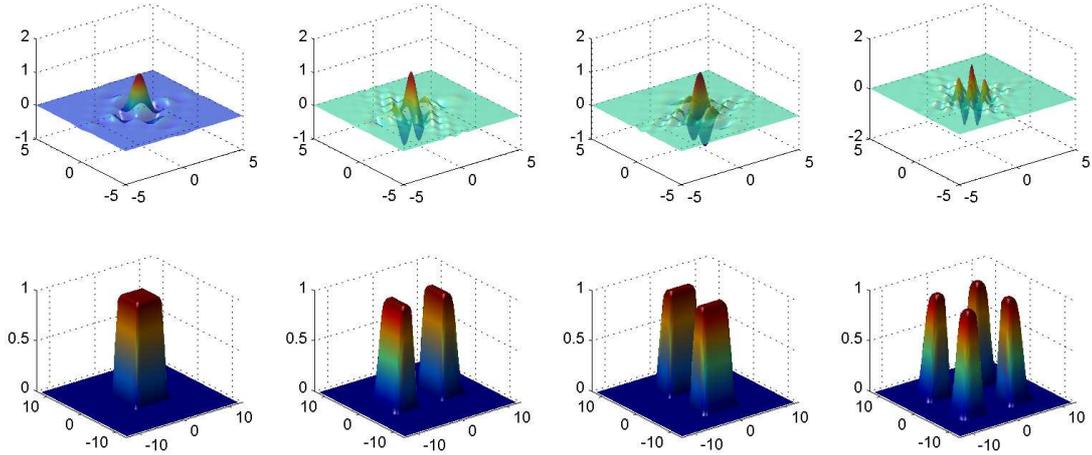


Figure 2.6: The functions $\phi \otimes \phi$, $\phi \otimes \psi$, $\psi \otimes \phi$ and $\psi \otimes \psi$ (upper row) and their respective Fourier transforms in magnitude (lower row), where ψ is the Meyer wavelet (see definition A.1.5) and ϕ the associated scaling function (computed using [9]).

Example 2.1.14 (Non-optimally sparse wavelet approximations of cartoon-like images). Let $\Phi \subset L^2(\mathbb{R}^2)$ be a wavelet basis of $L^2(\mathbb{R}^2)$ and $f = \mathbf{1}_B$, where B is a ball contained in $[0, 1]^2$, then

$$\|f - f_N\| \asymp N^{-\frac{1}{2}}$$

where f_N is the best N -term approximation obtained from Φ .

Proof. See [2, p. 159]. □

So the error decay of the N -term approximations of a cartoon-like image obtained from an orthonormal wavelet basis is far off the optimal rate, which would be N^{-1} (see definition 2.1.4). The essential reason for this is that, similar as for the Fourier basis in the one-dimensional setting, the localization of the atoms in a two-dimensional wavelet basis is not well suited for fitting the now occurring anisotropic structures, such as curve-like discontinuities. This shortcoming will be remedied by the construction of shearlet-based frames for $L^2(\mathbb{R}^2)$, presented in the upcoming section.

2.1.3 Shearlets

To have a chance of achieving optimally sparse approximations of cartoon-like images, we will require a dictionary based on atoms which are not only localized but are also well suited for fitting anisotropic features, which are very common in the two-dimensional setting. Assuming we want to construct these dictionaries in a similar fashion as in the Fourier and wavelet case, that is, we want to choose a generator function (like $e^{i \cdot}$ or a function $\psi \in L^2(\mathbb{R})$ satisfying the wavelet admissibility condition) and build the dictionary from applying certain operations to this generator (like frequency modulation, scaling or translation), one simple way of introducing a higher degree of directional

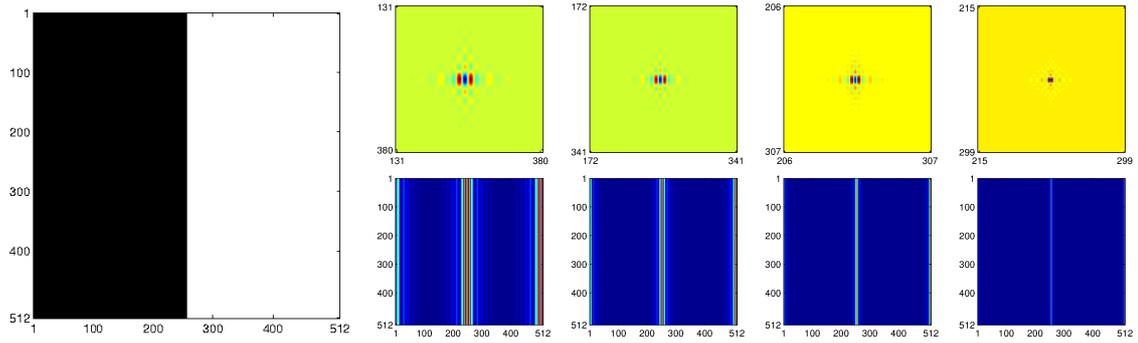


Figure 2.7: The lower right row shows the magnitudes of the wavelet coefficients corresponding to the discrete two-dimensional wavelets displayed above. The coefficients were obtained from a discrete stationary wavelet transform of the two-dimensional step function shown left (using the MATLAB Wavelet Toolbox). As the values in the diagonal of the scaling matrix $A_j = \text{diag}(2^{-j}, 2^{-j})$ (see theorem 2.1.13) are the same, the scaling is of an isotropic nature, that is, the degree of scaling is the same for both the x - and the y -direction. Hence, while still being directionally sensitive, the wavelets $\psi \otimes \phi$ on higher scales are not perfectly adapted to the anisotropic nature of the here considered discontinuity.

sensitivity would be to apply anisotropic scaling. That is, instead of using a scaling matrix $A_j = \text{diag}(2^{-j}, 2^{-j})$ (see theorem 2.1.13) where the amount of scaling is uniform in both coordinates, one can consider a scaling matrix with two different diagonal entries. Using such an anisotropic scaling matrix will result in atomic functions strongly stretched in x - or y -direction on higher scales. Hence, we will also need the ability of changing the preferred direction via a rotation-like operation. Note that this necessity of adding another degree of freedom again resembles the transition from Fourier bases to wavelets, where the use of localized generator functions required us to introduce a second parameter, namely the translation parameter (see equation (2.17)).

The first constructions of frames for $L^2(\mathbb{R}^2)$ applying anisotropic scaling were given by Candès and Donoho in 2004 under the name of curvelets [20, 21], quickly followed by a filterbank-based approach by Do and Vetterli in 2005 termed contourlets [22]. Both curvelet- and contourlet-based systems use rotations to change the preferred direction of the anisotropically scaled atoms, which leads to problems when entering the digital realm, as rotations typically don't preserve the integer grid⁵. Furthermore, these systems are not based on a single, or even a finite number of generator functions, but use different generators at each scale. Both these issues were resolved in a construction based on so-called shearlets, introduced by Kutyniok, Labate, Lim and Weiss in 2005 [1], where the designation 'shearlets' emphasizes the fact that grid-preserving shearings are applied instead of rotations.

Similar to the wavelet case, a function $\psi \in L^2(\mathbb{R}^2)$ is said to be an admissible shearlet

⁵In fact, only rotations by multiples of 90 degrees leave the integer grid intact.

if it satisfies

$$\int_{\mathbb{R}^2} \frac{|\hat{\psi}(\xi_1, \xi_2)|^2}{\xi_1^2} d\xi_2 d\xi_1 < \infty, \quad (2.33)$$

where $\hat{\psi}$ of course denotes the two-dimensional Fourier transform of ψ . So basically any function in $L^2(\mathbb{R}^2)$ whose Fourier transform is compactly supported away from the origin is an admissible shearlet and just as in the wavelet case, the convergence of the integral (2.33) implies the existence of an inverse continuous shearlet transform (see [8, p. 19f.]). One specific construction of an admissible shearlet, the so-called classical shearlet, which combines a one-dimensional wavelet and a one-dimensional bump function, is given below.

Definition 2.1.15 (Classical shearlet, as in [8]). *Let $\psi \in L^2(\mathbb{R}^2)$ be defined by*

$$\hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1) \hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right),$$

where $\psi_1 \in L^2(\mathbb{R})$ is a discrete wavelet in the sense that it satisfies the discrete Calderón condition, given by

$$\sum_{j \in \mathbb{Z}} \left| \hat{\psi}_1(2^{-j}\xi) \right|^2 = 1$$

for almost every $\xi \in \mathbb{R}$, with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and $\text{supp}(\hat{\psi}_1) \subset [-\frac{1}{2}, -\frac{1}{16}] \cup [\frac{1}{16}, \frac{1}{2}]$, and $\psi_2 \in L^2(\mathbb{R}^2)$ is a bump function in the sense that

$$\sum_{k=-1}^1 \left| \hat{\psi}_2(\xi + k) \right|^2 = 1$$

for almost every $\xi \in [-1, 1]$, satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and $\text{supp}(\hat{\psi}_2) \subset [-1, 1]$. Then ψ is called a classical shearlet.

A function $\psi \in L^2(\mathbb{R}^2)$ satisfying equation (2.33) can now be used as a generator for a shearlet-based system by applying scaling, shear and translation operators, that is, we will consider functions

$$\psi_{j,k,m}(\cdot) = \psi(S_k A_j \cdot -m),$$

where $m \in \mathbb{Z}^2$ is again a translation parameter and scaling matrices A_j and shear matrices S_k are given by

$$A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}. \quad (2.34)$$

The effects of applying a shear or scaling operator to a generator function in both the time and the frequency domain are depicted in figure 2.8. When considering classical shearlets, the generator will by definition be vertically aligned in the time domain and the functions $\psi_{j,k,m}$ even more so when increasing the parameter j . Of course, by also

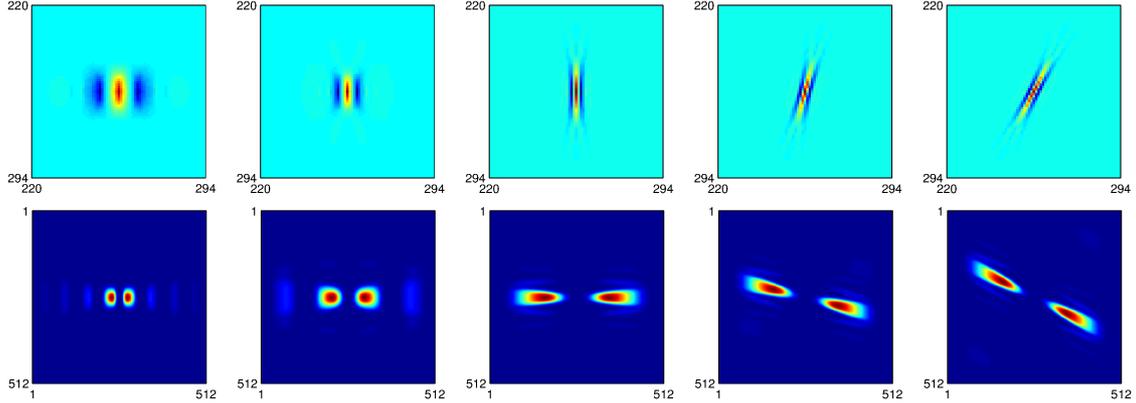


Figure 2.8: A two-dimensional digital shearlet filter is scaled twice and sheared twice in clockwise direction. The lower row shows the respective magnitude responses in the frequency domain (computed using ShearLab 3D [23]).

increasing the shear parameter k , the shearlets will progressively be sheared towards a horizontal alignment, but only at the cost of gradually stretching them out⁶. Furthermore, a perfect horizontal alignment could only be reached by sending k to infinity, which not only would lead to functions with an extremely thin support in y -direction but would also be a huge obstacle in any practical implementation of a shearlet-based transform. One possible solution to this problem is to split the frequency plane into four different cones (denoted C1 to C4 in figure 2.9) and to consider a generator function ψ having its Fourier transform essentially supported in $C1 \cup C3$ associated with the vertical directions and a second generator $\tilde{\psi}$ with the essential support of its Fourier transform contained in $C2 \cup C4$ associated with the horizontal directions. Assuming the function ψ to be a classical shearlet, i.e. $\hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2\left(\frac{\xi_2}{\xi_1}\right)$, where ψ_1 is a discrete wavelet and ψ_2 a bump function, $\tilde{\psi}$ can then simply be chosen as $\tilde{\psi}(\xi_1, \xi_2) = \psi(\xi_2, \xi_1)$. To nicely separate these cones at the lower frequencies, we will - roughly speaking - cut out a square centered around the origin in the frequency plane, that is we will restrict the scaling parameter $j \geq 0$ and add translates of a scaling function instead, thereby also covering the lower frequencies. These considerations lead to the formal definition of a discrete cone-adapted shearlet system.

Definition 2.1.16 (Cone-adapted discrete shearlet system, as in [8]). *Let $\psi \in L^2(\mathbb{R}^2)$ be an admissible shearlet (see equation (2.33)) associated with the horizontal frequency cones (see figure 2.9) and $\tilde{\psi} \in L^2(\mathbb{R}^2)$ its counterpart associated with the vertical frequency cones. Let furthermore $\phi \in L^2(\mathbb{R}^2)$ be a scaling function and scaling and shear matrices given by*

$$A_j = \begin{pmatrix} 2^j & 0 \\ 0 & 2^{j/2} \end{pmatrix}, S_k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}, \tilde{A}_j = \begin{pmatrix} 2^{j/2} & 0 \\ 0 & 2^j \end{pmatrix}. \quad (2.35)$$

⁶This is the price we have to pay for using shearings instead of rotations.

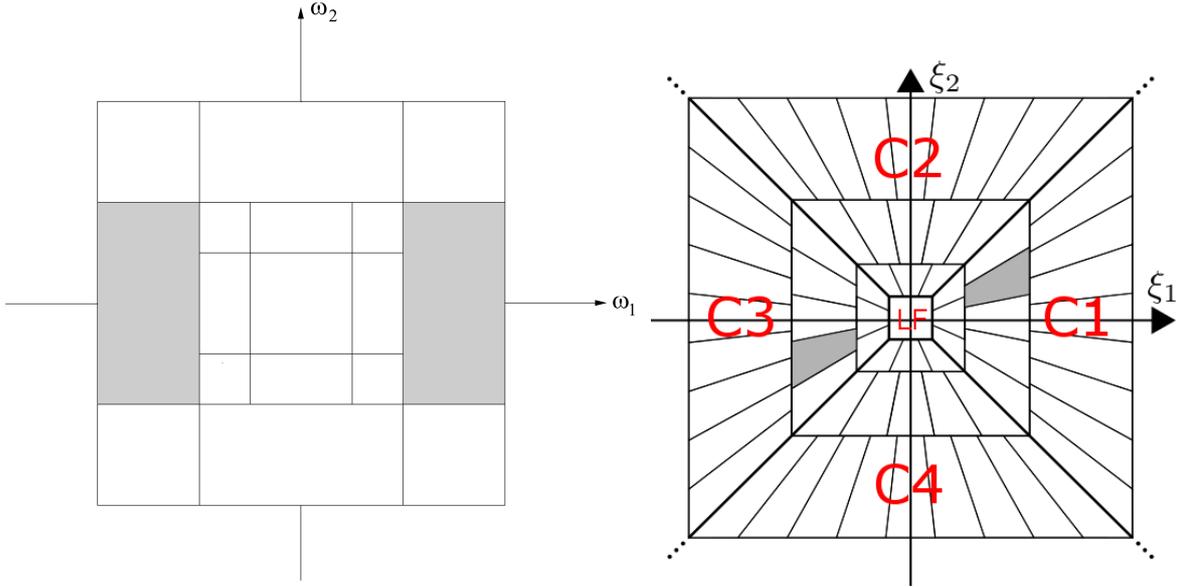


Figure 2.9: Tilings of the frequency domain induced by wavelet-based (left) and shearlet-based (right) systems. Gray areas indicate the essential frequency support of a single atom. Furthermore, the partition of the frequency plane into four cones and a low-frequency square used in definition 2.1.16 is depicted on the right side (the images are modified but originally taken from [13] and <http://en.wikipedia.org/wiki/Shearlet>).

The cone-adapted discrete shearlet system $SH(\phi, \psi, \tilde{\psi})$ is then given by

$$SH(\phi, \psi, \tilde{\psi}) = \Phi(\phi) \cup \Psi(\psi) \cup \tilde{\Psi}(\tilde{\psi}),$$

where

$$\begin{aligned} \Phi(\phi) &= \{\phi_m = \phi(\cdot - m) : m \in \mathbb{Z}^2\}, \\ \Psi(\psi) &= \left\{ \psi_{j,k,m} = 2^{j\frac{3}{4}} \psi(S_k A_j \cdot -m) : j \in \mathbb{N}_0, |k| < \left\lceil 2^{\frac{j}{2}} \right\rceil, m \in \mathbb{Z}^2 \right\}, \\ \tilde{\Psi}(\tilde{\psi}) &= \left\{ \tilde{\psi}_{j,k,m} = 2^{j\frac{3}{4}} \tilde{\psi}(S_k^T \tilde{A}_j \cdot -m) : j \in \mathbb{N}_0, |k| < \left\lceil 2^{\frac{j}{2}} \right\rceil, m \in \mathbb{Z}^2 \right\}. \end{aligned}$$

Again, the question comes up whether there exist functions $\psi, \tilde{\psi}$ and ϕ such that the system $SH(\phi, \psi, \tilde{\psi})$ forms a frame for $L^2(\mathbb{R}^2)$. As we would expect by now, the answer is yes. In fact, with an aptly chosen scaling function ϕ and by projecting the sets $\Psi(\psi)$ and $\tilde{\Psi}(\tilde{\psi})$ onto their respective frequency cones, a frame with framebounds $A = B = 1$ (compare (2.2)) can be constructed from any classical shearlet ψ .

Theorem 2.1.17 (Shearlet-based frame for $L^2(\mathbb{R}^2)$). *Let $\psi \in L^2(\mathbb{R}^2)$ be a classical shearlet and let a scaling function $\phi \in L^2(\mathbb{R}^2)$ be chosen such that*

$$\left| \hat{\phi}(\xi) \right|^2 + \sum_{j \geq 0} \sum_{|k| \leq \lceil 2^{j/2} \rceil} \left| \hat{\psi}(S_{-k}^T A_{-j} \xi) \right|^2 \mathbf{1}_C + \sum_{j \geq 0} \sum_{|k| \leq \lceil 2^{j/2} \rceil} \left| \hat{\tilde{\psi}}(S_{-k}^T \tilde{A}_{-j} \xi) \right|^2 \mathbf{1}_{\tilde{C}} = 1,$$

for almost every $\xi \in \mathbb{R}^2$, where the set C denotes the horizontal frequency cones $C = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \left| \frac{\xi_2}{\xi_1} \right| \leq 1\}$ and \tilde{C} denotes the vertical frequency cones $\tilde{C} = \mathbb{R}^2 \setminus C$. Let furthermore the sets $\Phi(\phi), \Psi(\psi), \tilde{\Psi}(\tilde{\psi}) \subset L^2(\mathbb{R}^2)$ be given as in definition 2.1.16 and $P_C\Psi(\psi)$ be the projection of $\Psi(\psi)$ onto C and $P_{\tilde{C}}\tilde{\Psi}(\tilde{\psi})$ be the projection of $\tilde{\Psi}(\tilde{\psi})$ onto \tilde{C} . Then, the union

$$\Phi(\phi) \cup P_C\Psi(\psi) \cup P_{\tilde{C}}\tilde{\Psi}(\tilde{\psi})$$

is a frame for $L^2(\mathbb{R}^2)$ with frame bounds $A = B = 1$ (compare (2.2)).

Proof. See [24]. □

We now hope that the error of N -term approximations of cartoon-like image functions provided by such shearlet frames will decay faster than it does with wavelet based constructions (compare example 2.1.14) and maybe even optimally fast in the sense of definition 2.1.4. Indeed, when repeating the computation carried out with wavelets in figure 2.7 with shearlets, the coefficients decay visibly faster in a neighborhood of the discontinuity when increasing the scaling parameter j (see figure 2.10). It was in fact shown by Guo and Labate in 2007 in the case of bandlimited⁷ shearlet generators [25] and by Kutyniok and Lim in 2011 for shearlet generators with compact support in the time domain [26] that shearlet frames of the form 2.1.16 can provide optimally sparse approximations of cartoon-like image functions up to a log factor.

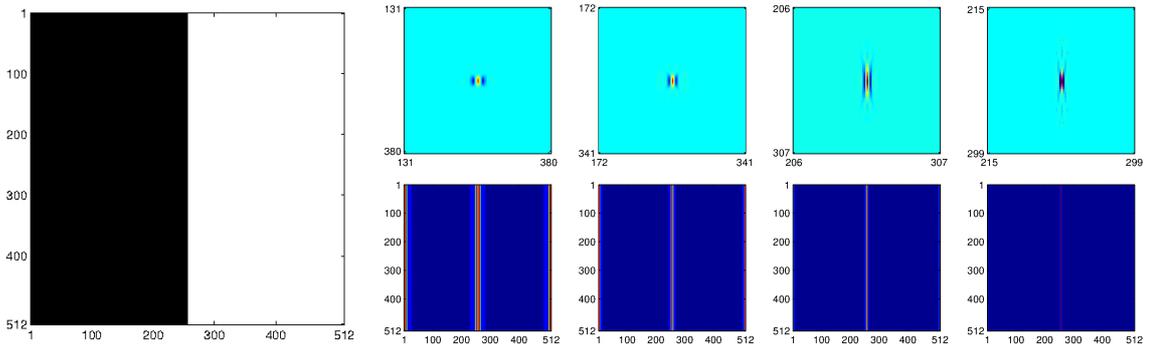


Figure 2.10: The lower right row shows the magnitudes of the shearlet coefficients corresponding to the digital two-dimensional shearlets displayed above. The coefficients were obtained from a digital shearlet transform of the two-dimensional step function shown left, computed with ShearLab 3D [23]. In contrast to the wavelet case (see figure 2.7), the shearlet atoms are scaled anisotropically and the thereby increased directional sensitivity when going up the ladder of scales induces a visibly faster decay of the shearlet coefficients in the neighborhood of the discontinuity.

Theorem 2.1.18 (Optimally sparse shearlet-based approximations of cartoon-like images). *Let $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be aptly chosen shearlet generators and $\phi \in L^2(\mathbb{R}^2)$ be an aptly chosen scaling function (e.g. like in theorem 2.1.17, for more details see [2, p.175])*

⁷A function is called bandlimited if its Fourier transform is compactly supported, i.e., the used frequency bands are limited.

and let the cone-adapted discrete shearlet system $SH(\phi, \psi, \tilde{\psi}) \subset L^2(\mathbb{R}^2)$ (see definition 2.1.16) be a frame for $L^2(\mathbb{R}^2)$. Then, apart from a log-factor, the frame $SH(\phi, \psi, \tilde{\psi})$ provides optimally sparse approximations of cartoon-like image functions (see definition 2.1.2) in the sense of definition 2.1.4, that is

$$\|f - f_N\| \lesssim N^{-1} \log(N)^{3/2}, \quad (2.36)$$

for all cartoon-like image functions $f \in L^2(\mathbb{R}^2)$, where f_N denotes the N -term approximation associated with the N largest shearlet coefficients in magnitude provided by $SH(\phi, \psi, \tilde{\psi})$.

Proof. See [2, p. 175ff.]. □

So up to the negligible logarithmic factor in (2.36), shearlet frames do in fact provide in some sense optimal representations of cartoon-like image functions and hence can - in accordance with the argument laid out at the very beginning of this chapter - be considered an interesting model of the functional behavior of the primary visual cortex, where certain neurons in fact encode a time-frequency decomposition.

2.2 Wavelet and Shearlet Transforms

Having defined wavelet orthonormal bases in section 2.1.2 and shearlet frames in section 2.1.3, we now formally fix the notion of shearlet and wavelet transforms. In principal, performing a shearlet or wavelet transform is nothing but the application of the synthesis operator (see equation 2.4) associated with the respective basis or frame to a function in $L^2(\mathbb{R}^d)$.

Definition 2.2.1 (One-dimensional discrete wavelet transform). *Let $\psi \in L^2(\mathbb{R})$ be a wavelet associated with a multiresolution analysis (see definition 2.1.9 and theorem 2.1.10), then the one-dimensional discrete wavelet transform associated with ψ is given by*

$$f \mapsto (\mathcal{W}_\psi f)(j, m) = \langle f, \psi_{j,m} \rangle, \quad (2.37)$$

where $f \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$, $m \in \mathbb{Z}$ and the wavelets $\psi_{j,m}$ are given as in theorem 2.1.10.

Definition 2.2.2 (Two-dimensional discrete wavelet transform). *Let $\phi \in L^2(\mathbb{R})$ be the scaling function of a multiresolution analysis (see definition 2.1.9) and $\psi \in L^2(\mathbb{R})$ be the associated wavelet (see theorem 2.1.10), then the discrete two-dimensional wavelet transform associated with ϕ and ψ is given by*

$$f \mapsto (\mathcal{W}_{\phi,\psi} f)(k, j, m) = \langle f, \psi_{j,m}^{(k)} \rangle, \quad (2.38)$$

where $f \in L^2(\mathbb{R}^2)$, $k \in \{1, 2, 3\}$, $j \in \mathbb{Z}$, $m \in \mathbb{Z}^2$ and the wavelets $\psi_{j,m}^{(k)}$ are given as in theorem 2.1.13.

Definition 2.2.3 (Discrete cone-adapted shearlet transform). *Let $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be shearlet generators and $\phi \in L^2(\mathbb{R}^2)$ be a scaling function such that $SH(\phi, \psi, \tilde{\psi})$ (see definition 2.1.16) is a frame for $L^2(\mathbb{R}^2)$, then the discrete cone-adapted shearlet transform associated with ϕ, ψ and $\tilde{\psi}$ is given by*

$$f \mapsto (\mathcal{SH}_{\phi, \psi, \tilde{\psi}} f)(m', (j, k, m), (\tilde{j}, \tilde{k}, \tilde{m})) = \left(\langle f, \phi_{m'} \rangle, \langle f, \psi_{j,k,m} \rangle, \langle f, \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}} \rangle \right), \quad (2.39)$$

where $j, \tilde{j} \in \mathbb{N}_0$, $k, \tilde{k} \in \left\{ -\left\lceil 2^{\frac{j}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{j}{2}} \right\rceil \right\}$, $m', m, \tilde{m} \in \mathbb{Z}^2$ and the functions $\phi_{m'}, \psi_{j,k,m}$ and $\tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}$ are given as in definition 2.1.16.

2.3 Complex Wavelet and Shearlet Transforms

In section 2.1, we have argued rather extensively that a discrete cone-adapted shearlet transform can in some sense be considered optimal to decompose natural images and hence might just be the time-frequency decomposition we are looking for. But as we will see in chapter 3, the primary visual cortex is not only populated by cells essentially encoding linear filters, so-called simple cells, but also cells whose behavior seems to be a nonlinear integration of simple cell responses, so-called complex cells. While these complex cells have been found to be selective to changes in rotation and frequency (modeled by the shear matrices S_k and the scaling matrices A_j in the context of shearlets), their responses are to a certain extent invariant to translations, that is, to spatial phase modulations [27, 28].

Naturally, this raises the question of how the responses of such shift invariant complex cells can be modeled, assuming that a discrete cone-adapted shearlet transform in fact captures the essential functional behavior of the more basal simple cells. As we have already discussed at the very beginning of this chapter and formally shown in remark 2.1.7, the complex valued Fourier transform (see theorem 2.1.6) is perfectly shift invariant in the absolute value, that is, when considering a sine and a cosine wave having the same frequency ξ and nonlinearly combining them by $\sqrt{\sin(\xi \cdot)^2 + \cos(\xi \cdot)^2}$, we always end up with the radius of the unit circle, that is 1. Furthermore, both the wavelet admissibility condition (2.16) and the shearlet admissibility condition (2.33) are solely based on the absolute value of the Fourier transform of a candidate function, whereas in definition 2.1.15, the classical shearlet is purely described by the structure of the magnitude response of the Fourier transforms of functions ψ_1 and ψ_2 . Also, the action of scaling or shearing a shearlet generator can be characterized by the respective changes in the frequency plane (see figures 2.8 and 2.9). Hence, a worthwhile approach to the construction of shift invariant responses from wavelet and shearlet atoms seems to be to again exploit the special relationship of sine and cosine waves. That is, we will generalize the structure of the Fourier modes $e^{i\xi \cdot} = \cos(\xi \cdot) + i \sin(\xi \cdot)$ to wavelets and shearlets in the sense that for each shearlet (or wavelet) ψ , we will consider a second shearlet ψ^* for which the roles of sine and cosine are reversed. Meaning that both ψ and ψ^* will have the same Fourier transform in the absolute value but the amount of

cosine waves present in ψ^* will be equal to the amount of sine waves in ψ and vice versa.

When considering the one-dimensional case (i.e. the wavelet case), a bounded linear operator on $L^2(\mathbb{R})$ performing just that is given by the Hilbert transform.

Definition 2.3.1 (Hilbert transform). *Let $f \in L^2(\mathbb{R})$, then the Hilbert transform of f is given by*

$$(\mathcal{H}f)(x) = \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(t)}{t-x} dt \quad (2.40)$$

or equivalently in the Fourier domain

$$\mathcal{F}(\mathcal{H}f)(\xi) = -i \operatorname{sgn}(\xi) \hat{f}(\xi), \quad (2.41)$$

where i denotes the imaginary unit and sgn the sign function.

Note that in equation (2.40), the denominator in the integrand is zero for $x = t$. Hence, the integral should be interpreted as a Cauchy principal value. Furthermore, equation (2.40) can also be written as a convolution, namely $-\frac{1}{\pi} * f$. The equivalence of (2.40) and (2.41) therefore follows from $\mathcal{F}\left(\frac{1}{\cdot}\right) = -\pi i \operatorname{sgn}(\cdot)$ and the Fourier convolution theorem.

The Hilbert transform is a bounded and linear operator on $L^2(\mathbb{R})$ and commutes with translations and positive dilations (see [29, p. 55]), that is

$$(\mathcal{H}f(\cdot - m)) = (\mathcal{H}f)(\cdot - m) \quad (2.42)$$

and

$$(\mathcal{H}f(a\cdot)) = (\mathcal{H}f)(a\cdot) \quad (2.43)$$

for all $f \in L^2(\mathbb{R})$, $m \in \mathbb{R}$ and $a \in \mathbb{R}^+$. In fact, any bounded linear operator on $L^2(\mathbb{R})$ which commutes with translations, positive dilations and anticommutes with the reflection $f(\cdot) \mapsto f(-\cdot)$ is a constant multiple of the Hilbert transform [29, p. 55]. Of course, these commutation properties will become important when defining complex wavelet or shearlet systems, as they indicate that it will suffice to apply the Hilbert transform solely to generator functions. But firstly, let us elaborate on the earlier stated claim that the Hilbert transform interchanges sine and cosine waves.

Remark 2.3.2 (The Hilbert transform interchanges sine and cosine). *Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ with f being real-valued, then*

$$(\mathcal{F}^{-1}\hat{f})(x) = \frac{1}{\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) \cos(\xi x) - \operatorname{Im}(\hat{f}(\xi)) \sin(\xi x) d\xi \quad (2.44)$$

and

$$(\mathcal{F}^{-1}\widehat{\mathcal{H}f})(x) = \frac{1}{\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) \sin(\xi x) + \operatorname{Im}(\hat{f}(\xi)) \cos(\xi x) d\xi. \quad (2.45)$$

Proof. We know that for each real-valued function, \hat{f} is Hermitian, i.e. $\hat{f}(\xi) = \overline{\hat{f}(-\xi)}$ and hence compute

$$\begin{aligned}
(\mathcal{F}^{-1}\hat{f})(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\xi) e^{ix\xi} d\xi \\
&= \frac{1}{2\pi} \int_0^\infty \hat{f}(\xi) (\cos(x\xi) + i \sin(x\xi)) + \frac{1}{2\pi} \int_0^\infty \overline{\hat{f}(\xi)} (\cos(x\xi) - i \sin(x\xi)) d\xi \\
&= \frac{1}{2\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) (\cos(x\xi) + i \sin(x\xi)) + \operatorname{Im}(\hat{f}(\xi)) (i \cos(x\xi) - \sin(x\xi)) d\xi \\
&\quad + \frac{1}{2\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) (\cos(x\xi) - i \sin(x\xi)) - \operatorname{Im}(\hat{f}(\xi)) (i \cos(x\xi) + \sin(x\xi)) d\xi \\
&= \frac{1}{\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) \cos(\xi x) - \operatorname{Im}(\hat{f}(\xi)) \sin(\xi x) d\xi
\end{aligned}$$

and with the definition of the Hilbert transform in the frequency domain (2.41),

$$\begin{aligned}
(\mathcal{F}^{-1}\widehat{\mathcal{H}f})(x) &= \frac{1}{2\pi} \int_{\mathbb{R}} -i \operatorname{sgn}(\xi) \hat{f}(\xi) e^{ix\xi} d\xi \\
&= \frac{1}{2\pi} \int_{-\infty}^0 i \hat{f}(\xi) e^{ix\xi} d\xi - \frac{1}{2\pi} \int_0^\infty i \hat{f}(\xi) e^{ix\xi} d\xi \\
&= \frac{1}{2\pi} \int_0^\infty \overline{i \hat{f}(\xi)} (\cos(x\xi) - i \sin(x\xi)) d\xi - \frac{1}{2\pi} \int_0^\infty i \hat{f}(\xi) (\cos(x\xi) + i \sin(x\xi)) d\xi \\
&= \frac{1}{2\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) (i \cos(x\xi) + \sin(x\xi)) + \operatorname{Im}(\hat{f}(\xi)) (\cos(x\xi) - i \sin(x\xi)) d\xi \\
&\quad + \frac{1}{2\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) (-i \cos(x\xi) + \sin(x\xi)) + \operatorname{Im}(\hat{f}(\xi)) (\cos(x\xi) + i \sin(x\xi)) d\xi \\
&= \frac{1}{\pi} \int_0^\infty \operatorname{Re}(\hat{f}(\xi)) \sin(\xi x) + \operatorname{Im}(\hat{f}(\xi)) \cos(\xi x) d\xi
\end{aligned}$$

□

Motivated by the above computation and given the fact that the Hilbert transform acts nicely on functions in $L^2(\mathbb{R})$, that is it commutes with scalings and translations, we will - analogous to the Fourier modes - now consider so-called complex wavelets of the form

$$\psi^c = \psi + i\mathcal{H}\psi. \quad (2.46)$$

This construction, also denoted the 'complex signal' of ψ , will allow us to decompose the underlying wavelet ψ into an amplitude $|\psi^c|$ and a phase $\operatorname{Re}\left(\frac{\psi^c}{|\psi^c|}\right)$, that is,

$$\psi(x) = |\psi^c(x)| \operatorname{Re}\left(\frac{\psi^c(x)}{|\psi^c(x)|}\right) \quad (2.47)$$

for $x \in \mathbb{R}$, where the amplitude serves as a slowly varying envelope and the oscillatory behavior of ψ is encoded in the phase (for an illustration, see figure 2.11). Another interesting observation is that the complex signal ψ^c only consists of positive frequencies, that is $\psi^c(\xi) = 0$ for almost every $\xi < 0$.

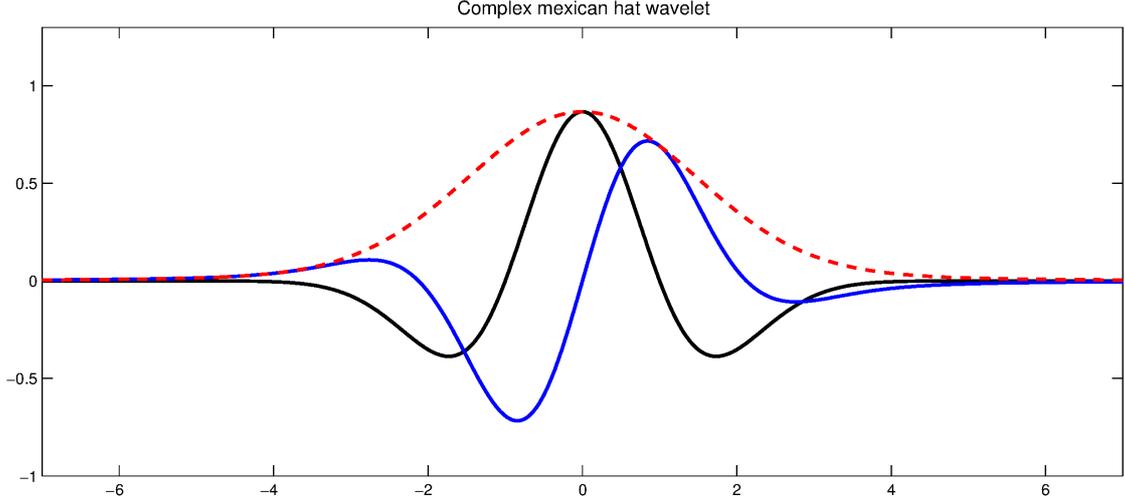


Figure 2.11: The Mexican hat wavelet (black), its Hilbert transform (blue), and the amplitude envelope induced by the corresponding complex wavelet, defined in equation (2.46) (red). The Mexican hat wavelet is symmetric, hence its Fourier transform will be real-valued, i.e. only cosine waves are present. By remark 2.3.2, the Hilbert transform of the Mexican hat wavelet therefore only consists of sine waves and is hence odd-symmetric (computed using the MATLAB Wavelet Toolbox).

Remark 2.3.3 (The complex signal has no negative frequencies). *Let $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then*

$$\widehat{\psi^c}(\xi) = \mathcal{F}(\psi + i\mathcal{H}\psi)(\xi) = \begin{cases} 2\widehat{\psi}(\xi) & \text{if } \xi > 0 \\ 0 & \text{else} \end{cases} \quad (2.48)$$

for almost every $\xi \in \mathbb{R}$.

Proof. Using the Fourier domain definition of the Hilbert transform (see (2.41)), we simply compute

$$\begin{aligned} \mathcal{F}(\psi + i\mathcal{H}\psi)(\xi) &= \widehat{\psi}(\xi) + i\widehat{\mathcal{H}\psi}(\xi) \\ &= \widehat{\psi}(\xi)(1 + \text{sgn}(\xi)) \\ &= \begin{cases} 2\widehat{\psi}(\xi) & \text{if } \xi > 0 \\ 0 & \text{else} \end{cases}. \end{aligned}$$

□

We are now ready to define one-dimensional complex wavelet systems and the corresponding transforms.

Definition 2.3.4 (One-dimensional discrete complex wavelet transform). *Let $\psi \in L^2(\mathbb{R})$ be a real-valued wavelet associated with a multiresolution analysis (see definition*

2.1.9 and theorem 2.1.10), then the discrete complex wavelet system associated with ψ is given by the set

$$\{\psi_{j,m}^c = \psi_{j,m} + i(\mathcal{H}\psi)_{j,m} : (j, m) \in \mathbb{Z} \times \mathbb{Z}\}, \quad (2.49)$$

where \mathcal{H} denotes the Hilbert transform (see definition 2.3.1) and the functions $\psi_{j,m}$ and $(\mathcal{H}\psi)_{j,m}$ are defined as in theorem 2.1.10.

Furthermore, the corresponding one-dimensional discrete complex wavelet transform is given by

$$f \mapsto (\mathcal{W}_\psi^c f)(j, m) = \langle f, \psi_{j,m}^c \rangle, \quad (2.50)$$

where $f \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$ and $m \in \mathbb{Z}^2$.

Note that the analysis of square integrable functions via complex-valued transforms similar to (2.50) has already been considered by Gabor [14], Grossmann and Morlet [30], Kingsbury [31] and many others. A great introduction and comprehensive analysis of the theory and applications of amplitude and sign decompositions induced by complex wavelet transforms was given by Martin Storath in his PhD thesis [3].

Having defined complex wavelets, the next step is to generalize the Hilbert transform to two dimensions and construct complex shearlets analogously. The most prominent generalization of the Hilbert transform to higher dimensions is based on the Fourier-based definition (2.41) and termed partial Hilbert transform.

Definition 2.3.5 (Partial Hilbert transform). *Let $\xi_0 \in \mathbb{R}^2 \setminus \{0\}$ and $f \in L^2(\mathbb{R}^2)$, then the partial Hilbert transform of f in direction ξ_0 is given by*

$$\mathcal{F}(\mathcal{H}_{\xi_0} f)(\xi) = -i \operatorname{sgn}(\langle \xi, \xi_0 \rangle) \hat{f}(\xi), \quad (2.51)$$

where sgn the sign function.

For a real-valued shearlet generator $\psi \in L^2(\mathbb{R}^2)$, we will therefore consider the complex shearlet generator

$$\psi^c = \psi + i\mathcal{H}_{\xi_0}\psi \quad (2.52)$$

with $\xi_0 \in \mathbb{R}^2 \setminus \{0\}$ (later, we will choose ξ_0 to be $(1, 0)^T$ or $(0, 1)^T$), which allows for the same phase/amplitude decomposition we already observed in the one-dimensional case (see equation (2.47)).

While the Fourier transform of complex wavelets induced by the one-dimensional Hilbert transform vanishes at all negative frequencies, the same is true for the partial Hilbert transform with respect to the half space defined by $\langle \xi, \xi_0 \rangle < 0$.

Remark 2.3.6 (The two-dimensional complex signal has all frequencies contained in a half space). *Let $\psi \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ and $\xi_0 \in \mathbb{R}^2 \setminus \{0\}$ then*

$$\widehat{\psi^c}(\xi) = \mathcal{F}(\psi + i\mathcal{H}_{\xi_0}\psi)(\xi) = \begin{cases} 2\widehat{\psi}(\xi) & \text{if } \langle \xi, \xi_0 \rangle > 0 \\ 0 & \text{else} \end{cases} \quad (2.53)$$

for almost every $\xi \in \mathbb{R}^2$.

Proof. Analogous to remark 2.3.3. □

We have already discussed that in the one-dimensional case, the Hilbert transform commutes with scalings and translations. For the partial Hilbert transform, we will consider a more general result given by Storath in [3, p. 19].

Lemma 2.3.7 (The partial Hilbert transform commutes with linear invertible mappings). *Let $\xi_0 \in \mathbb{R}^2 \setminus \{0\}$ and $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a linear and invertible mapping. Then for any $f \in L^2(\mathbb{R}^2)$, A commutes with the partial Hilbert transform, that is*

$$A(\mathcal{H}_{\xi_0} f) = \mathcal{H}_{A\xi_0}(Af). \quad (2.54)$$

Proof. See Storath [3, p. 19]. □

In particular, due to

$$\operatorname{sgn}(\langle \xi, \lambda \xi_0 \rangle) = \operatorname{sgn}(\langle \xi, \xi_0 \rangle)$$

for all coefficients $\lambda > 0$, it follows that

$$A(\mathcal{H}_{\xi_0} f) = \mathcal{H}_{\xi_0}(Af)$$

if ξ_0 is an eigenvector of A associated with a positive eigenvalue. Hence,

$$S_k \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where S_k denotes the shear matrix defined in equation (2.34), implies that the partial Hilbert transform $\mathcal{H}_{(1,0)^T}$ commutes with the shearing operator S_k for all $k \in \mathbb{Z}$. So the partial Hilbert transform commutes with all three actions performed on shearlet generators, that is translations, scalings and shearings and we are ready to define discrete cone-adapted complex shearlet systems and the associated transforms.

Definition 2.3.8 (Discrete cone-adapted complex shearlet transform). *Let $\psi, \tilde{\psi} \in L^2(\mathbb{R}^2)$ be real-valued shearlet generators and $\phi \in L^2(\mathbb{R}^2)$ be a scaling function such that $SH(\phi, \psi, \tilde{\psi})$ (see definition 2.1.16) is a frame for $L^2(\mathbb{R}^2)$, then the discrete cone-adapted complex shearlet system associated with ϕ, ψ and $\tilde{\psi}$ is given by the set*

$$SH^c(\phi, \psi, \tilde{\psi}) = \Phi(\phi) \cup \Psi^c(\psi) \cup \tilde{\Psi}^c(\tilde{\psi}),$$

where

$$\begin{aligned} \Phi(\phi) &= \{ \phi_m = \phi(\cdot - m) : m \in \mathbb{Z}^2 \}, \\ \Psi^c(\psi) &= \left\{ \psi_{j,k,m}^c = \psi_{j,k,m} + i(\mathcal{H}_{(1,0)^T} \psi)_{j,k,m} : j \in \mathbb{N}_0, |k| < \left\lceil 2^{\frac{j}{2}} \right\rceil, m \in \mathbb{Z}^2 \right\}, \\ \tilde{\Psi}^c(\tilde{\psi}) &= \left\{ \tilde{\psi}_{j,k,m}^c = \tilde{\psi}_{j,k,m} + i(\mathcal{H}_{(0,1)^T} \tilde{\psi})_{j,k,m} : j \in \mathbb{N}_0, |k| < \left\lceil 2^{\frac{j}{2}} \right\rceil, m \in \mathbb{Z}^2 \right\}. \end{aligned}$$

The functions $\psi_{j,k,m}, (\mathcal{H}_{(1,0)^T} \psi)_{j,k,m}, \tilde{\psi}_{j,k,m}$ and $(\mathcal{H}_{(0,1)^T} \tilde{\psi})_{j,k,m}$ are given as in definition 2.1.16.

Furthermore, for $f \in L^2(\mathbb{R}^2)$ the corresponding discrete complex cone-adapted shearlet transform is given by

$$f \mapsto \left(\mathcal{SH}_{\phi, \psi, \tilde{\psi}}^c f \right) (m', (j, k, m), (\tilde{j}, \tilde{k}, \tilde{m})) = \left(\langle f, \phi_{m'} \rangle, \langle f, \psi_{j,k,m}^c \rangle, \langle f, \tilde{\psi}_{\tilde{j}, \tilde{k}, \tilde{m}}^c \rangle \right), \quad (2.55)$$

with $j, \tilde{j} \in \mathbb{N}_0$, $k, \tilde{k} \in \left\{ -\left\lceil 2^{\frac{j}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{j}{2}} \right\rceil \right\}$ and $m', \tilde{m} \in \mathbb{Z}^2$.

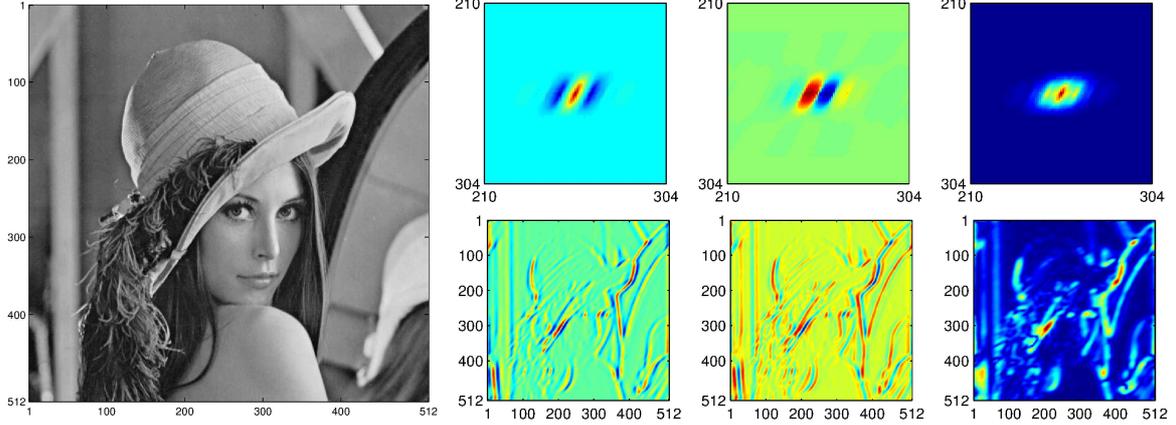


Figure 2.12: The upper right row shows a symmetric digital shearlet filter, its odd-symmetric Hilbert transform and the absolute value of the corresponding complex shearlet. Below, the respective coefficients - obtained from a digital complex shearlet transform of the image shown on the left - are plotted (computed using ShearLab 3D [23]).

As the absolute value of the Fourier transform is invariant under the Hilbert transform, $\psi \in L^2(\mathbb{R}^2)$ satisfying the shearlet admissibility condition 2.33 implies that $\mathcal{H}_{\xi_0} \psi$ is an admissible shearlet as well. Furthermore, it holds for any classical shearlet (see definition 2.1.15) that its Hilbert transform is again a classical shearlet and if a cone-adapted shearlet system $\text{SH}(\phi, \psi, \tilde{\psi})$ is a frame for $L^2(\mathbb{R}^2)$ according to theorem 2.1.17, then so is $\text{SH}(\phi, \mathcal{H}_{(1,0)^\top} \psi, \mathcal{H}_{(0,1)^\top} \tilde{\psi})$.

Finally, as it was already mentioned in figure 2.11, by virtue of interchanging cosines and sines, any symmetric shearlet generator $\psi \in L^2(\mathbb{R}^2)$ will have an odd-symmetric Hilbert transform $\mathcal{H}_{\xi_0} \psi$ and we will make good use of this symmetry relationship when considering complex shearlet transforms for the task of edge detection in section 4.1.

For an illustration of coefficients computed using a complex discrete cone-adapted shearlet transform, please refer to figure 2.12.

Chapter 3

Complex Shearlet Transforms and the Visual Cortex

At the beginning of the previous chapter, we have claimed that at the early stages of human visual processing, some kind of time-frequency decomposition is performed. In this chapter, we will briefly examine the neurophysiological evidence supporting this claim. That is, we will give a short introduction into what is known about the shape and structure of the receptive fields¹ of neurons in the primary visual cortex (V1).

The first breakthrough concerning the description of the receptive fields of V1 neurons was made in 1960 when Nobel prize winners Hubel and Wiesel quite accidentally discovered that neurons in V1 respond to moving elongated bars of light rather than blobs and that their response highly depends on the orientation of such bars [32]. In their experiments, electrodes measured the firing rate of neurons in the primary visual cortex of a cat who was exposed to visual stimuli presented via a slide projector. While at first, they failed to elicit any significant responses, they eventually discovered that when changing the slides, the thereby induced shadow line caused measurable excitations in the neurons connected to the electrode. After successfully continuing their experiments with accordingly adapted stimuli, they made another interesting discovery. While for some neurons, elongated light bars at the right orientation had an excitatory effect throughout their entire receptive field, for others, the light bars seemed to act excitatory only in the center of their receptive field but antagonistic when moving through the outer regions of their receptive field (i.e. the neurons started firing after the light bars had passed, see figure 3.1). That is, some neurons reacted to elongated light bars within their receptive field invariant to the precise location of the bars while the response of others was indeed sensitive to translations. Hubel and Wiesel conjectured that translation invariant neurons combined the responses of translation dependent neurons and hence termed the former complex cells and the latter simple cells.

¹The receptive field of a neuron is the area for which the neuron will respond to stimuli. For cells in the primary visual cortex, this is typically a portion of the visual field whereas for sensory neurons, the receptive field is a region of the human body.

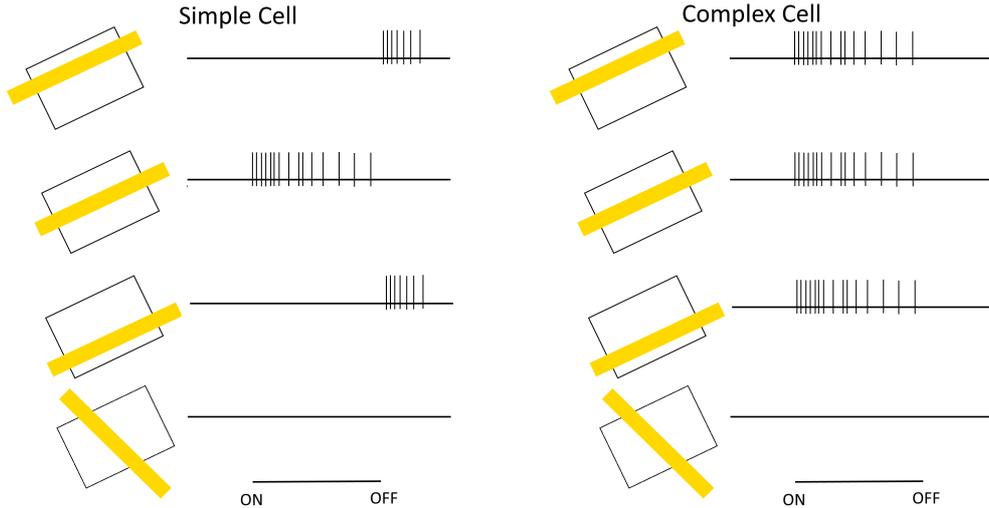


Figure 3.1: While the response of the complex cell is invariant to the exact location of the light bar, a light bar placed on the outer regions of the receptive field of the simple cell acts inhibitory and a light bar placed in the center of the receptive field of the simple cell acts excitatory.

If one wants to model the functional behavior of the primary visual cortex, the fact that V1 neurons are orientationally selective already suggests the use of transforms based on possibly anisotropic atoms, such as the shearlet transforms defined in the previous chapter. Furthermore, it was shown that neurons in the primary visual cortex are not only orientationally sensitive but also selective with respect to spatial frequencies [33], suggesting a multiscale approach. In fact, already in 1985, Daugman successfully fitted two-dimensional Gabor wavelets of the form

$$\psi(x, y) = e^{-\pi(x-x_0)^2\alpha^2+(y-y_0)^2\beta^2} e^{-2\pi i(u_0(x-x_0)+v_0(y-y_0))}, \quad (3.1)$$

to experimentally measured receptive fields of V1 neurons [34], where the parameters (x_0, y_0) determine the location, the parameters (α, β) the level of anisotropy and the parameters (u_0, v_0) the spatial frequency and the orientation of ψ (see figure 3.2).

Furthermore, since Adelson and Bergen in 1985 proposed using the magnitude response of complex-valued filters constructed from Hilbert transform pairs of even- and odd-symmetrical linear filters to measure motion energy [35], approximating the response of complex cells in the primary visual cortex in such a way has become a common practice. Finally, it should be noted that Polat and Norcia in 1998 collected experimental evidence that orientation information is pooled in V1 over considerable distances along the axis of the preferred orientation of neurons in the primary visual cortex [36], suggesting a high sensitivity with respect to strongly anisotropic features.

While this short compendium of neurophysiological findings does not imply that a complex discrete shearlet transform as given in definition 2.3.8 is carried out in the human visual cortex, the fact that V1 neurons are known to be orientationally selective, sensitive to spatial frequencies and that the primary visual cortex is sensitive to anisotropic

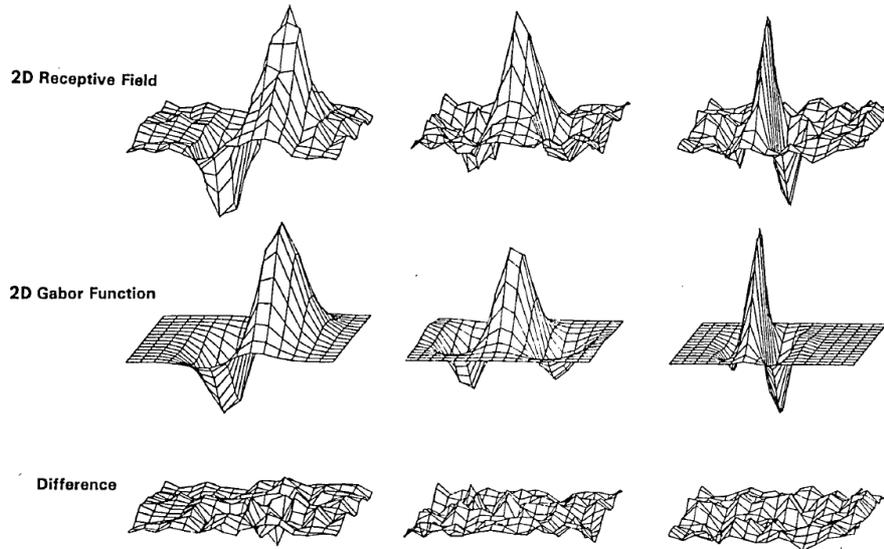


Figure 3.2: Gabor wavelets of the form (3.1) fitted to experimentally measured receptive fields of V1 simple cells (taken from [34]).

features at the very least suggests that assuming a connection on a functional level is a valid approach.

Chapter 4

Applications

In this chapter, we will examine two applications of the complex discrete shearlet transforms defined in section 2.3, namely edge detection and image quality assessment. As it was already stated at the beginning of chapter 2, our main intuition why applying complex shearlet transforms in such tasks might be an idea worth pursuing is the assumption that - at least on a functional level - they provide a good model for the behavior of certain neurons in the primary visual cortex.

4.1 Edge Detection

The history of both the mathematical analysis of edges and their computational detection is long and has brought forward many models and algorithms. One of the most prominent methods developed by Canny in 1986 [37] is essentially based on the idea that edges are characterized by large gradients when considering images to be differentiable functions. Multiscale approaches based on wavelets have been proposed by Mallat and Zhong [38] while the use of magnitude responses of complex wavelet transforms was already considered by Tu, Hwang and Ho in 2005 [39]. A contrast invariant approach to edge detection based on the concept of phase congruency, which will be discussed in detail in section 4.1.1, was introduced by Kovessy in 1999 [4, 40] and further improved by Storath in his PhD thesis [3], while a similar concept using so-called monogenic wavelets (i.e. complex wavelets based on the Riesz transform) was introduced by Felsberg and Sommer [41]. Furthermore, a multiscale approach based on band-limited shearlets was proposed by Yi, Labate, Easley and Krim in 2009 [42].

In most edge detection algorithms, it is assumed that points lying on an edge can be found by looking for local maxima of the absolute value of the gradient¹. Naturally, such an approach requires the image to be differentiable, which is typically enforced by beforehand convolving it with a Gaussian kernel. While on the plus side, this will

¹For a more elaborate edge model based on the connection between the boundary of a geometric object and the wavefront set of its associated function, see Storath [3, p. 56ff.].

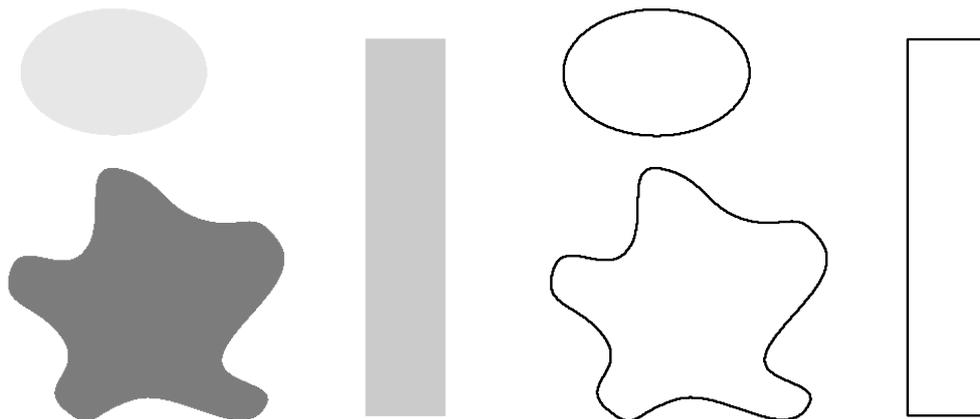


Figure 4.1: For edge detection, we will assume images to be of a cartoon-like nature, as is the image depicted on the left. Our goal is to assert a value ranging from 0 to 1 to any pixel, such that 1 indicates the presence of an edge and 0 a smooth neighborhood. In particular, we want this value to be independent of the intensity of contrast between two regions separated by an edge. The image to the right shows a handmade desirable outcome of such an algorithm where black equals 1 and white equals 0.

inhibit an algorithm's sensitivity to noise, to a certain extent, it will also decrease the sharpness of edges present in the image.

In chapter 2, we have exhaustively made the case that discrete shearlet transforms are in some sense optimal for analyzing cartoon-like image functions (see definition 2.1.2 and theorem 2.1.18), a class of functions we assume to provide a good model for natural images. Following this line of thought, we will again consider images to be of a cartoon-like nature and hence define edges simply as the (possibly piecewise) C^2 -smooth boundary curves, separating two smooth regions, as given in the definition of cartoon-like image functions. Hence, to us, edges will be characterized by the geometrical structure induced by a boundary curve and not by the amount of contrast between the two regions the curve separates, as it is the case in gradient-based methods. That is, the algorithm we have in mind should be invariant to changes in contrast, as long as they leave the shape of the separating boundary curves intact. Furthermore, it should be noted that the definition of cartoon-like image functions does not account for point singularities, or in general, types of discontinuities which are not separating two regions but only ripping through an otherwise smooth function, like a single line drawn through a picture. When going to the digital realm, this means that an edge detection algorithm based on the cartoon-like image model might not be very sensitive to structures only extending over very few pixels, like lines or points.

Also, we would like our algorithm to give us the possibility of discriminating between edges of a very sharp and clear nature on the one hand and edges with a less structured geometry on the other hand. To sum this up, for a given image, we will look for a

computational rule asserting a value ranging from 0 to 1 to each point on the real plane (or the pixel domain in the digital case), where a value close to 1 indicates the presence of an edge, whereas a value close to 0 indicates a smooth neighborhood. This value should hence depend on features describing the structure and geometry of an edge, namely the smoothness of the separating boundary curve and the smoothness of the regions it separates but it should be invariant to contrast, that is, to the step size between two neighboring regions.

Eventually, we will derive such a rule by considering the different behavior of coefficients associated with symmetric and odd-symmetric shearlets aligned with an edge when going up the ladder of scales. But before, we will examine Kovese's application of phase congruency to edge detection, which significantly inspired the complex shearlet-based algorithm developed later and already provides us with a contrast invariant way of detecting features in images.

4.1.1 Phase Congruency

Let us consider a real-valued function $f \in L^2(-\pi, \pi)$, then by theorem 2.1.5, we know that f can be written as

$$f(x) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} \langle f, e^{in\cdot} \rangle e^{inx}$$

for almost every $x \in [-\pi, \pi]$. Let us denote $c_n = \langle f, e^{in\cdot} \rangle$ and consider the polar form, i.e. $c_n = |c_n| e^{i\varphi_n}$, where φ_n denotes the respective phase angle. We get

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} |c_n| e^{i\varphi_n} e^{inx} \\ &= \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} |c_n| e^{i(nx + \varphi_n)} \\ &= \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} |c_n| (\cos(nx + \varphi_n) + i \sin(nx + \varphi_n)) \\ &= \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} |c_n| \cos(nx + \varphi_n) \\ &= \frac{|c_0|}{2\pi} + \frac{1}{\pi} \sum_{n \in \mathbb{N}^+} |c_n| \cos(nx + \varphi_n) \end{aligned}$$

as f is real-valued. Hence, at each point x , each term of the Fourier reconstruction is associated with an angle $nx + \varphi_n$, which depends on the reconstructed function f via the phase angle of the complex-valued inner product $\langle f, e^{in\cdot} \rangle$. It was observed by Morrone et al. [43, 44] that at the location of features like jump-discontinuities, the values $nx + \varphi_n$ are roughly the same for all $n \in \mathbb{N}^+$ and proposed the following measure

termed phase congruency

$$\text{PC}(x) = \max_{\mu \in [0, 2\pi]} \frac{\sum_{n=1}^{\infty} |c_n| \cos(nx + \varphi_n - \mu)}{\sum_{n=1}^{\infty} |c_n|}, \quad (4.1)$$

where c_n and φ_n are defined as above.

To give an example, let us again consider the step function $f = \mathbf{1}_{[-\pi, 0]}$, defined on the interval $[-\pi, \pi]$. As we have already computed in example 2.1.8, the Fourier coefficients are given by $c_n = \langle \mathbf{1}_{[-\pi, 0]}, e^{in\cdot} \rangle = \frac{2i}{n}$ if n is odd and $c_n = 0$ if n is even. Hence, for n odd and positive, we have $c_n = |c_n| e^{i\frac{\pi}{2}}$. That is, at $x = 0$, the values

$$nx + \varphi_n = \varphi_n = \frac{\pi}{2}$$

are the same for all odd integers, implying phase congruency at this location. This is illustrated in figure 4.2.

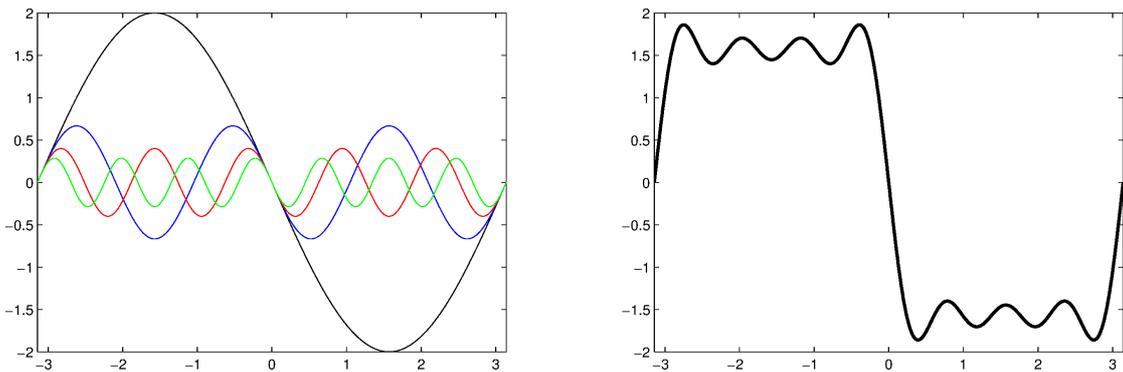


Figure 4.2: The left plot shows the functions $|c_n| \cos(\cdot n + \varphi_n)$ for $n \in \{1, 3, 5, 7\}$, where $c_n = \langle \mathbf{1}_{[-\pi, 0]}, e^{in\cdot} \rangle$ and φ_n denotes the phase angle of c_n . The right plot displays the sum of the functions shown left. Please note that the phase congruency at 0 on the left coincides with the formation of a jump discontinuity at 0 on the right.

While phase congruency is indeed a contrast invariant measure, finding the angle $\mu \in [0, 2\pi]$ for which (4.1) attains its maximum is a tedious task. Note that by computing (4.1), one tries in fact to observe the local effects of the ratio between sine and cosine waves of the same frequency, as given by the phase angle φ_n . Due to the already discussed nature of the Fourier modes however, this phase angle is an inherently unlocalized property. Hence, it should be no surprise to us that an easily computable estimate of the phase congruency measure can be defined by considering complex wavelets, which are not only localized but whose complex-valued coefficients also encode the local relationship between sine and cosine.

These considerations lead us directly to the complex wavelet-based phase congruency estimate introduced by Kovessi in [40].

Definition 4.1.1 (Complex wavelet-based phase congruency measure). *Let $f \in L^2(\mathbb{R})$, $J \in \mathbb{N}$, and $\psi_{a_j,x}^c$ be complex wavelets scaled by the parameter $a_j \in \mathbb{R}^+$ and centered around $x \in \mathbb{R}$ given by*

$$\psi_{a_j,x}^c = \psi_{a_j,x} + i\mathcal{H}\psi_{a_j,x},$$

where

$$\psi_{a_j,x} = a_j^{-\frac{1}{2}}\psi\left(\frac{\cdot - x}{a_j}\right)$$

with $\psi \in L^2(\mathbb{R})$ being a real-valued symmetric wavelet and $\mathcal{H}\psi$ being its odd-symmetric Hilbert transform (see section 2.3). Then, the one-dimensional complex wavelet-based phase congruency estimate of f at the point $x \in \mathbb{R}$ is defined as

$$\text{PC}_{\psi^c}(f, x) = \frac{\left| \sum_{j=1}^J \langle f, \psi_{a_j,x}^c \rangle \right|}{\sum_{j=1}^J \left| \langle f, \psi_{a_j,x}^c \rangle \right| + \epsilon}, \quad (4.2)$$

where $\epsilon > 0$ prevents division by zero.

Note that the symmetry condition imposed on ψ implies that $\hat{\psi}$ will be purely real-valued, that is, ψ will be constructed solely from cosine waves, whereas $\mathcal{H}\psi$ only consists of sine waves. Hence, assuming perfect phase congruency at a point x , we have $\langle f, \psi_{a_j,x}^c \rangle = \left| \langle f, \psi_{a_j,x}^c \rangle \right| e^{i\varphi}$ for some fixed angle $\varphi \in [0, 2\pi]$ and all $j \in \mathbb{N}$ and get

$$\begin{aligned} \text{PC}_{\psi^c}(f, x) &= \frac{\left| \sum_{j=1}^J \langle f, \psi_{a_j,x}^c \rangle \right|}{\sum_{j=1}^J \left| \langle f, \psi_{a_j,x}^c \rangle \right|} \\ &= \frac{|e^{i\varphi}| \sum_{j=1}^J \left| \langle f, \psi_{a_j,x}^c \rangle \right|}{\sum_{j=1}^J \left| \langle f, \psi_{a_j,x}^c \rangle \right|} \\ &= 1. \end{aligned}$$

Naturally, the more the phase angles differ on various scales j , the more the complex-valued coefficients $\langle f, \psi_{a_j,x}^c \rangle$ will cancel each other out, moving $\text{PC}_{\psi^c}(f, x)$ closer to zero.

Hence, $\text{PC}_{\psi^c}(f, x)$ is a contrast invariant measure ranging from 0 to 1 and thus already satisfies two of the desiderata stated at the beginning of section 4.1. Furthermore, by considering two-dimensional complex wavelets not only on different scales but - similarly to the shearlet case - also with different orientations, definition 4.1.1 can readily be generalized to the two-dimensional case [40].

Definition 4.1.2 (Two-dimensional complex wavelet-based phase congruency measure). Let $f \in L^2(\mathbb{R}^2)$, $J \in \mathbb{N}$, $K \in \mathbb{N}$ and $\psi_{a_j, \varphi_k, x}^c$ be complex wavelets scaled by the parameter $a_j \in \mathbb{R}^+$, rotated by $\varphi_k \in [0, 2\pi]$ and centered around $x \in \mathbb{R}^2$ given by

$$\psi_{a_j, \varphi_k, x}^c = \psi_{a_j, \varphi_k, x} + i\mathcal{H}\psi_{a_j, \varphi_k, x},$$

where

$$\psi_{a_j, \varphi_k, x} = a_j^{-1} \psi \left(R_{\varphi_k} \frac{\cdot - x}{a_j} \right)$$

with $\psi \in L^2(\mathbb{R}^2)$ being a real-valued symmetric wavelet, $\mathcal{H}\psi$ being its odd-symmetric Hilbert transform (see section 2.3) and a rotation matrix

$$R_{\varphi_k} = \begin{pmatrix} \cos(\varphi_k) & -\sin(\varphi_k) \\ \sin(\varphi_k) & \cos(\varphi_k) \end{pmatrix}.$$

Then, the two-dimensional complex wavelet-based phase congruency estimate of f at the point $x \in \mathbb{R}^2$ is defined as

$$\text{PC}_{\psi^c}^{2D}(f, x) = \frac{\sum_{k=1}^K \left| \sum_{j=1}^J \langle f, \psi_{a_j, \varphi_k, x}^c \rangle \right|}{\sum_{k=1}^K \sum_{j=1}^J \left| \langle f, \psi_{a_j, \varphi_k, x}^c \rangle \right| + \epsilon}, \quad (4.3)$$

where $\epsilon > 0$ prevents division by zero.

Please note that for implementing the phase congruency measure just defined, Kovessi proposes to use two-dimensional polar-separable Gaussian wavelets considered on four scales (i.e. $J = 4$) and with six possible orientations (i.e. $K = 6$) separated by 30° [40]. Furthermore, a MATLAB function computing the two-dimensional complex wavelet-based phase congruency measure can be downloaded from [45].

It would seem that by replacing complex wavelets with complex shearlets and rotations with shearings, the phase congruency measure given in definition 4.1.2 could easily be generalized to shearlets. One might even think that due to the superiority of shearlets over wavelets regarding the representation of two-dimensional geometric features, a shearlet-based phase congruency measure would provide better and sharper localizations of features in images. Sadly, while in principle the described generalization is of course possible, figure 4.3 illustrates that using complex shearlets for phase congruency measurements does not bring any improvements. In fact, the contrary is true.

The main reason why - against our first intuition - shearlets are inferior to wavelets with respect to phase congruency measurements is that phase congruency only indicates the presence of geometrical features but is in itself not a geometrical property. The objective of the complex wavelet-based formula (4.3) is to estimate the local behavior of the Fourier atoms when changing the frequency and it succeeds in doing so by iteratively considering only small portions of the frequency plane, given by the essential frequency support of the complex wavelets $\psi_{a_j, \varphi_k, x}^c$. When applying shearlets, on the other hand,

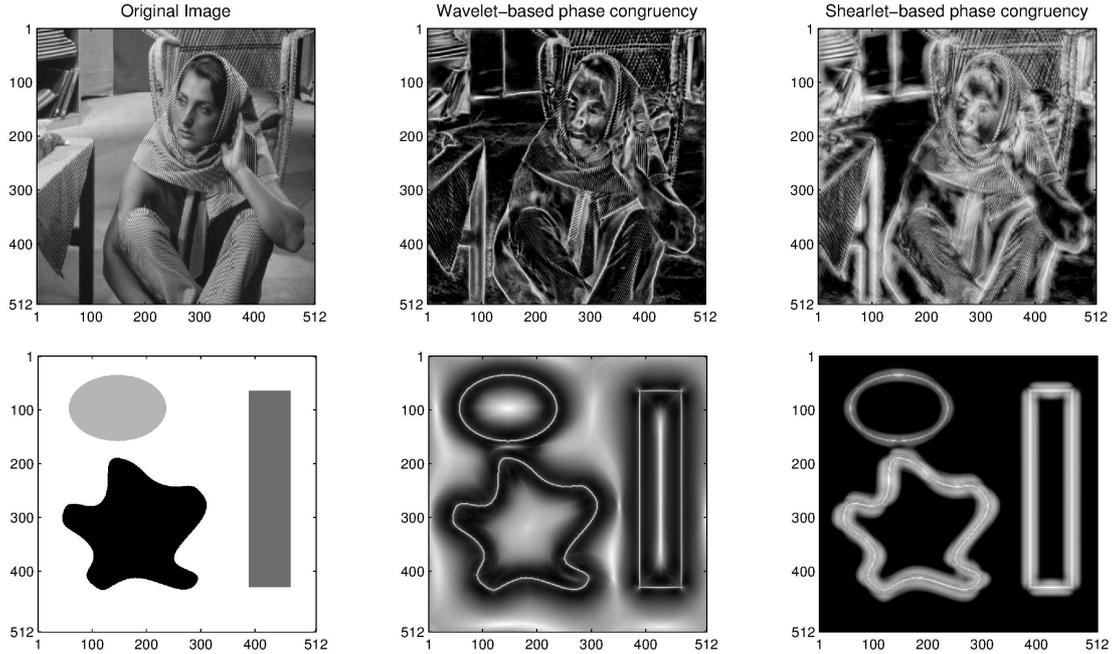


Figure 4.3: We compare the complex wavelet-based phase congruency measure (see definition 4.1.2) to its complex shearlet-based counterpart. With both the natural and the artificial image, the wavelet-based phase congruency measure provides a much more precise display of the significant features present in the images. Note that in both methods, soft thresholding was applied to the values $|\sum_{j=1}^J \langle f, \psi_{a_j, \varphi_k, x}^c \rangle|$ in the numerator of equation (4.3) for the purpose of noise reduction. For more details on phase congruency and noise, see [40, p. 8]. (computed using Peter Kovesis’s *phasecong2* method [45] and ShearLab 3D [23])

anisotropic scaling will cause the frequency support of the complex shearlet atoms to be increasingly elongated (see figure 2.9), resulting in much coarser phase congruency estimates.

In the following two sections, we will develop a more geometry-driven edge detection algorithm, which - in the spirit of the phase congruency measure - will also be contrast invariant and based on the interplay of real-valued symmetric atoms and their odd-symmetric Hilbert transforms. Again, we first consider the one-dimensional case, that is, the detection of jump singularities in piecewise smooth functions by applying complex wavelets.

4.1.2 Detection of Singularities With Pairs of Symmetric and Odd-Symmetric Wavelets

Our goal is to derive a computational rule which, given a piecewise smooth function $f \in L^2(\mathbb{R})$ (see definition 2.1.1), asserts a value ranging from 0 to 1 to any point in

the interval $[0, 1]$ in accordance with the desiderata laid out at the beginning of this chapter. That is, a value close to 1 should indicate the presence of a discontinuity, a value close to 0 should indicate a smooth neighborhood and the assignment of these values should be independent of the amount of contrast, i.e. the jump size.

To develop an intuition, let us again consider a simple step function $f = c\mathbf{1}_{[0, \frac{1}{2})} + d\mathbf{1}_{[\frac{1}{2}, 1]}$ defined on the unit interval with coefficients $c, d \in \mathbb{R}$ (for an illustration, see figure 4.4). Let furthermore $\psi^{\text{even}} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a real-valued and even-symmetric wavelet such that its odd-symmetric Hilbert transform $\mathcal{H}\psi^{\text{even}}$ also belongs to $L^1(\mathbb{R})$. While $g \in L^p(\mathbb{R})$ implies $\mathcal{H}g \in L^p(\mathbb{R})$ for $1 < p < \infty$, it should be noted at this point that an integrable function does not necessarily have an integrable Hilbert transform. It is, however, a necessary condition for $\mathcal{H}\psi^{\text{even}} \in L^1(\mathbb{R})$ that $\widehat{\psi^{\text{even}}}$ has a vanishing mean value [46, p. 211f.], which is implied by the continuity of $\widehat{\psi^{\text{even}}}$ and the wavelet admissibility condition (see equation (2.18)). Furthermore, we assume that ψ^{even} is normalized with respect to the L^1 -norm and set $\psi^{\text{odd}} = \frac{\mathcal{H}\psi^{\text{even}}}{\|\mathcal{H}\psi^{\text{even}}\|_{L^1}}$, that is,

$$\|\psi^{\text{even}}\|_{L^1} = \int_{\mathbb{R}} |\psi^{\text{even}}(x)| dx = 1 \text{ and } \|\psi^{\text{odd}}\|_{L^1} = \int_{\mathbb{R}} |\psi^{\text{odd}}(x)| dx = 1.$$

We now denote

$$\psi_{a,x}^{\text{even}}(\cdot) = a^{-1}\psi^{\text{even}}\left(\frac{\cdot - x}{a}\right) \text{ and } \psi_{a,x}^{\text{odd}}(\cdot) = a^{-1}\psi^{\text{odd}}\left(\frac{\cdot - x}{a}\right) \quad (4.4)$$

for a scaling parameter $a \in \mathbb{R}^+$ and a translation parameter $x \in \mathbb{R}$. Note that contrary to the systems of wavelets considered so far, the functions $\psi_{a,x}^{\text{even}}$ and $\psi_{a,x}^{\text{odd}}$ are normalized in L^1 and not in L^2 .

Let us now examine the behavior of coefficients associated with even- and odd-symmetric wavelets aligned with the jump singularity, that is, we will consider $\langle f, \psi_{a, \frac{1}{2}}^{\text{even}} \rangle$ and $\langle f, \psi_{a, \frac{1}{2}}^{\text{odd}} \rangle$ for different choices of $a \in \mathbb{R}^+$. First, we take a look at $\psi_{a, \frac{1}{2}}^{\text{even}}$.

$$\begin{aligned} \langle f, \psi_{a, \frac{1}{2}}^{\text{even}} \rangle &= \int_{\mathbb{R}} \left(c\mathbf{1}_{[0, \frac{1}{2})}(x) + d\mathbf{1}_{[\frac{1}{2}, 1]}(x) \right) \psi_{a, \frac{1}{2}}^{\text{even}}(x) dx \\ &= \frac{c}{a} \int_0^{\frac{1}{2}} \psi^{\text{even}}\left(\frac{x - \frac{1}{2}}{a}\right) dx + \frac{d}{a} \int_{\frac{1}{2}}^1 \psi^{\text{even}}\left(\frac{x - \frac{1}{2}}{a}\right) dx \\ &= c \int_{-\frac{1}{2a}}^0 \psi^{\text{even}}(x) dx + d \int_0^{\frac{1}{2a}} \psi^{\text{even}}(x) dx \\ &= (c + d) \int_{-\frac{1}{2a}}^0 \psi^{\text{even}}(x) dx. \end{aligned}$$

Furthermore, we know that ψ^{even} has a vanishing mean value, that is

$$\int_{\mathbb{R}} \psi^{\text{even}}(x) dx = 0,$$

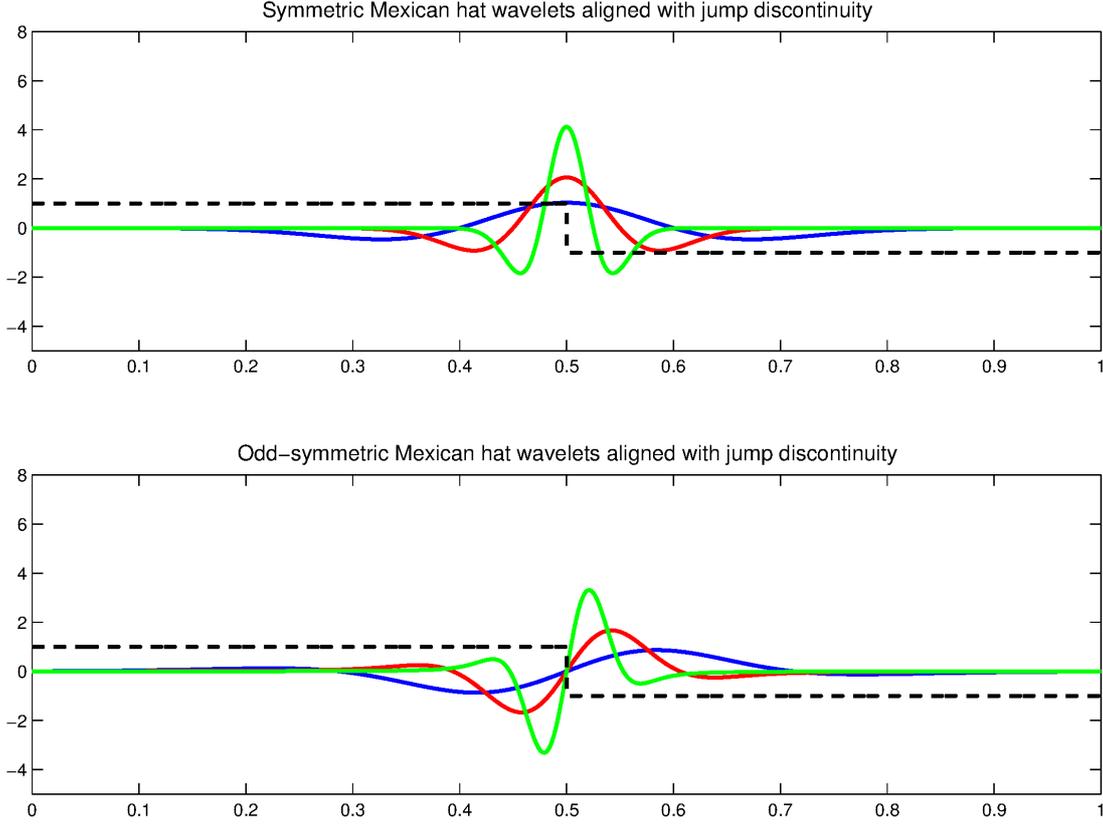


Figure 4.4: Differently scaled symmetric and odd-symmetric wavelets located at a jump discontinuity. All wavelets have the same L^1 -norm, namely 0.25.

which, due to symmetry, implies that

$$\int_{-\infty}^0 \psi^{\text{even}}(x) dx = 0.$$

Thus, we get

$$\begin{aligned} \langle f, \psi_{a, \frac{1}{2}}^{\text{even}} \rangle &= (c + d) \left(\int_{-\infty}^0 \psi^{\text{even}}(x) dx - \int_{-\infty}^{-\frac{1}{2a}} \psi^{\text{even}}(x) dx \right) \\ &= -(c + d) \int_{-\infty}^{-\frac{1}{2a}} \psi^{\text{even}}(x) dx. \end{aligned}$$

Obviously, the last term tends to 0 as a gets smaller, but from a practical perspective, even more is true. Given a wavelet with sufficient decay properties in the time domain or a wavelet having compact support, for all scales typically considered in a digital implementation, the wavelet coefficients will almost be 0 at the location of a jump discontinuity separating two constant functions. Of course, this will also be true for wavelet coefficients located inside a neighborhood, in which the analyzed function is smooth. To distinguish between these cases, we will now consider the behavior of the coefficients associated with $\psi_{a, \frac{1}{2}}^{\text{odd}}$.

Repeating the above computation yields

$$\begin{aligned}\langle f, \psi_{a, \frac{1}{2}}^{\text{odd}} \rangle &= c \int_{-\frac{1}{2a}}^0 \psi^{\text{odd}}(x) dx + d \int_0^{\frac{1}{2a}} \psi^{\text{odd}}(x) dx \\ &= (c - d) \int_{-\frac{1}{2a}}^0 \psi^{\text{odd}}(x) dx.\end{aligned}\tag{4.5}$$

Let us now assume that for ψ^{odd} , the integral over one half of the real line does not vanish, i.e. we denote

$$C_{\psi^{\text{odd}}} = \int_{-\infty}^0 \psi^{\text{odd}}(x) dx$$

and assume $|C_{\psi^{\text{odd}}}| > 0$. Then, equation (4.5) can be written as

$$\langle f, \psi_{a, \frac{1}{2}}^{\text{odd}} \rangle = (c - d)C_{\psi^{\text{odd}}} - (c - d) \int_{-\infty}^{-\frac{1}{2a}} \psi^{\text{odd}}(x) dx.$$

Again, the last term will go towards 0 for small a and similar to the symmetric case, assuming ψ^{odd} to have a nice decay behavior or even compact support implies that the coefficients $\langle f, \psi_{a, \frac{1}{2}}^{\text{odd}} \rangle$ will be almost exactly the jump size times $C_{\psi^{\text{odd}}}$ at the location of a jump discontinuity separating two constant functions for all practically relevant values a .

In conclusion, we can assume that at the location of a jump discontinuity, the coefficients associated with the even-symmetric wavelets remain close to zero on all scales of practical interest, while the coefficients associated with the odd-symmetric wavelets stay constant but non-zero. This observation motivates the singularity measure based on pairs of even-symmetric and odd-symmetric wavelets given below.

Definition 4.1.3 (A singularity measure based on pairs of even-symmetric and odd-symmetric wavelets). *Let $\psi^{\text{even}} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a real-valued even-symmetric wavelet and $\psi^{\text{odd}} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ be a real-valued odd-symmetric wavelet such that*

$$C_{\psi^{\text{odd}}} = \int_{-\infty}^0 \psi^{\text{odd}}(x) dx \neq 0,$$

and $\|\psi^{\text{even}}\|_{L^1} = \|\psi^{\text{odd}}\|_{L^1} = 1$. Let furthermore $J \in \mathbb{N}$ and $(a_j)_{j \in \{1, \dots, J\}} \subset \mathbb{R}^+$ be a sequence of scaling parameters. Then, for $f \in L^2(\mathbb{R})$ and a point $x \in \mathbb{R}$, a singularity measure is given by

$$\tilde{S}_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x) = \frac{\left| \sum_{j=1}^J \langle f, \psi_{a_j, x}^{\text{odd}} \rangle \right| - \sum_{j=1}^J \left| \langle f, \psi_{a_j, x}^{\text{even}} \rangle \right|}{J \max_{j \in \{1, \dots, J\}} \left| \langle f, \psi_{a_j, x}^{\text{odd}} \rangle \right| + \epsilon},\tag{4.6}$$

where $\epsilon > 0$ prevents division by zero and

$$\begin{aligned}\psi_{a_j,x}^{even}(\cdot) &= a_j^{-1} \psi^{even}\left(\frac{\cdot - x}{a_j}\right), \\ \psi_{a_j,x}^{odd}(\cdot) &= a_j^{-1} \psi^{odd}\left(\frac{\cdot - x}{a_j}\right),\end{aligned}$$

for a scaling parameter $a_j \in \mathbb{R}^+$ and a translation parameter $x \in \mathbb{R}$.

To ensure a mapping ranging from 0 to 1, we finally set

$$S_{\psi^{even},\psi^{odd}}(f, x) = \max \left\{ \tilde{S}_{\psi^{even},\psi^{odd}}(f, x), 0 \right\}. \quad (4.7)$$

A closer look at the numerator in equation (4.6) reveals that large coefficients associated with odd-symmetric atoms on different scales are considered as evidence for the presence of a singularity while large coefficients associated with even-symmetric wavelets count as counter-evidence. Keeping in mind the idea that wavelet- or shearlet-based atoms can serve as functional models of certain cells in the primary visual cortex, this means that in the context of singularity detection, the odd-symmetric wavelets can be seen as excitatory cells whereas the even-symmetric wavelets assume the role of inhibitory cells. Finally, the denominator in equation (4.6) enforces our intuition that the coefficients associated with odd-symmetric atoms not only should be large but also need to stay constant when changing the scale and thereby provides the desired contrast invariance.

While our considerations started with wavelets forming a Hilbert transform pair, we do not require this quite special relationship in definition 4.1.3. The Hilbert transform is a useful tool for constructing pairs of even-symmetric and odd-symmetric wavelets having the same frequency support but beyond that, there is no reason why Hilbert transform pairs should in any sense be optimal for this specific measure. So far, the best results have been achieved with a pair of one even-symmetric and one odd-symmetric real-valued wavelet generator, which is associated with a tight frame² for $L^2(\mathbb{R})$ and was proposed by Selesnick and Abdelnour in [47].

Two examples of singularity detections using the measure (4.6) are given in figure 4.5. The first example, analyzing a piecewise polynomial function, demonstrates that the measure is indeed invariant different jump sizes and provides a precise localization of singularities. In the second example, in which we analyze a piecewise smooth function, our singularity measure is compared to the discrete signature proposed by Storath in [3, p. 79ff.], which is conceptually closely related to the phase congruency measure (4.2). While our method provides a much sharper localization of jump singularities, it is indeed completely non-sensitive to cusps, which might be considered a disadvantage. However, one can argue that when moving to the two-dimensional setting and considering only images of a cartoon-like nature, cusps or point singularities are not the things one is looking for when trying to detect edges.

²That is, the frame bounds (see equation (2.2)) A and B are equal.

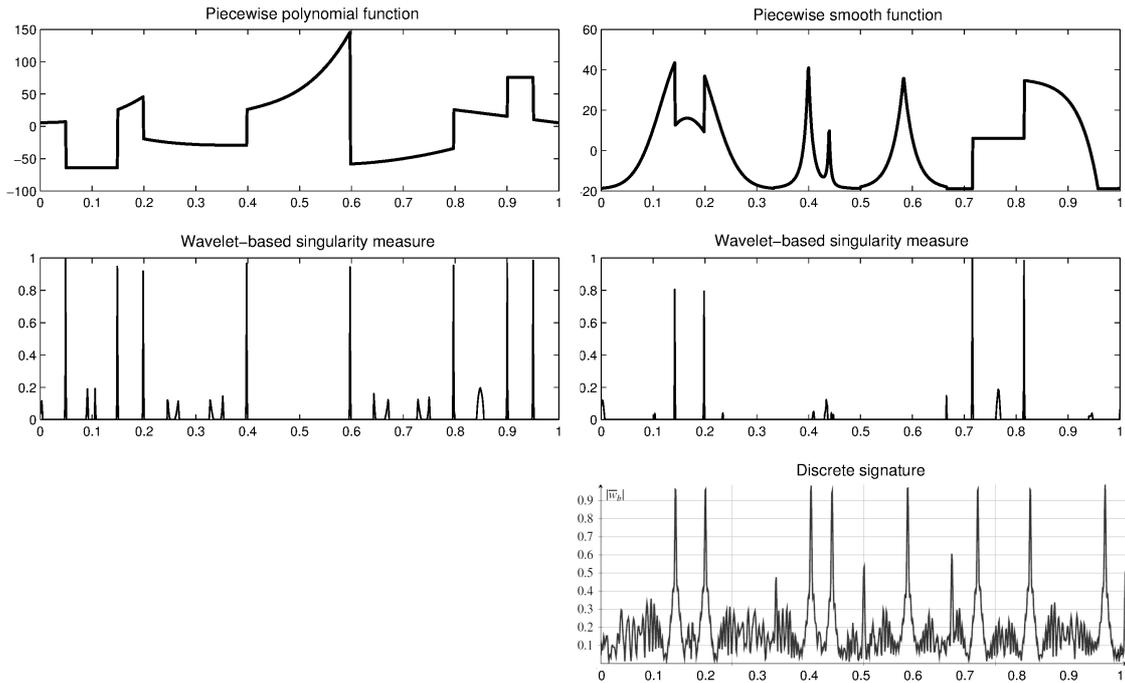


Figure 4.5: The singularity measure (4.6) is applied to a piecewise polynomial function and a piecewise smooth function with $J = 5$ and ψ^{even} and ψ^{odd} forming a pair of generator wavelets associated with a tight frame for $L^2(\mathbb{R})$ as proposed in [47]. For the piecewise smooth function, we also compare our measure to the discrete signature introduced by Storath in [3], where the discrete signature values are taken directly from [3, p. 85]. Both analyzed functions were computed with [9].

4.1.3 Edge Detection With Pairs of Symmetric and Odd-Symmetric Shearlets

In pursuit of our original goal of developing an algorithm for edge detection, we will now generalize the singularity measure given in definition 4.1.3 to two dimensions by applying pairs of even-symmetric and odd-symmetric shearlet generators. Since in the two-dimensional realm, we will have to deal with anisotropic singularities, it seems natural to consider a value similar to (4.6) for each direction by a shearlet system based on the chosen generators. However, this raises the question of how to condense this finite sequence of values into one single coefficient. While we can easily dispose of the idea of taking some kind of mean, as this would to a certain extent demand the presence of not only one but many edges, there are at least two reasonable approaches of selecting just one value associated with a specific direction. One simple and very intuitive way to go would certainly be to maximize over all considered directions. In praxis, however, it turned out that more satisfying results can be achieved by preselecting a principle direction, which can be determined by the largest coefficient in magnitude associated with an odd-symmetric shearlet. That is, for each point in the two-dimensional plane, we will maximize over the absolute values of all coefficients associated with an odd-

symmetric shearlet and choose the direction of the corresponding shearlet to be the principle direction. It should be noted that this approach slightly undermines our goal of deriving a contrast invariant computational rule, as such a preselecting process will cause high-contrast edges to locally dominate low-contrast edges. On the plus side, however, this will also significantly improve the localization of corner points.

Another issue to be resolved is the fact that when considering discrete shearlet systems, the number of different directions (i.e. the number of differently sheared shearlets) does not remain constant over all scales. Typically, shearlet atoms associated with high-frequency regions in the frequency plane are much more elongated than their low-frequency counterparts, which causes an increased number of differently sheared atoms on higher scales (compare the relationship between the parameters k and j in the definition of a discrete cone-adapted shearlet system). We will deal with this by limiting the number of eligible directions to the number of differently sheared atoms on the lowest scale considered. In return, this means that for each eligible direction, it might be necessary to consider more than one shearlet on higher scales. In this case, we will simply choose the shearlet associated with the largest coefficient in magnitude.

For simplicity, the following definition of a shearlet-based edge measure will only use shearlets associated with the horizontal frequency cones (see definition 2.1.16). In any actual implementation, the shearlets belonging to the vertical frequency cones have to be included analogously.

Definition 4.1.4 (An edge measure based on pairs of even-symmetric and odd-symmetric shearlets). *Let $\psi^{even} \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ be a real-valued even-symmetric shearlet associated with the horizontal frequency cones and $\psi^{odd} \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ be a real-valued odd-symmetric shearlet associated with the horizontal frequency cones such that*

$$C_{\psi^{odd}} = \int_{-\infty}^0 \int_{\mathbb{R}} \psi^{odd}(x, y) dy dx \neq 0, \quad (4.8)$$

and $\|\psi^{even}\|_{L^1} = \|\psi^{odd}\|_{L^1} = 1$. Let furthermore $J_{min}, J_{max} \in \mathbb{N}_0$ with $J_{min} \leq J_{max}$ and denote

$$\begin{aligned} \psi_{j,k,x}^{even} &= 2^{\frac{3j}{2}} \psi^{even}(S_k A_j(\cdot - x)), \\ \psi_{j,k,x}^{odd} &= 2^{\frac{3j}{2}} \psi^{odd}(S_k A_j(\cdot - x)) \end{aligned}$$

for $j \in \{J_{min}, \dots, J_{max}\}$, $|k| \leq \left\lceil 2^{\frac{j}{2}} \right\rceil$ and $x \in \mathbb{R}^2$, where shear matrices S_k and scaling matrices A_j are given as in definition 2.1.16. Then, for $f \in L^2(\mathbb{R}^2)$ and a point $x \in \mathbb{R}^2$, a principle direction $k_{\psi^{even}, \psi^{odd}}^*(f, x) \in \left\{ -\left\lceil 2^{\frac{J_{min}}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{J_{min}}{2}} \right\rceil \right\}$ is given by

$$k_{\psi^{even}, \psi^{odd}}^*(f, x) = \arg \max_{\tilde{k} \in \left\{ -\left\lceil 2^{\frac{J_{min}}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{J_{min}}{2}} \right\rceil \right\}} \max_{\substack{j \in \{J_{min}, \dots, J_{max}\}, \\ k \in \left\{ -\left\lceil 2^{\frac{j}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{j}{2}} \right\rceil \right\}, \\ \left| \tilde{k} - k \left\lceil 2^{\frac{J_{min}}{2}} \right\rceil \left\lceil 2^{\frac{j}{2}} \right\rceil^{-1} \right| \leq \frac{1}{2}}} |\langle f, \psi_{j,k,x}^{odd} \rangle| \quad (4.9)$$

and an edge measure sensitive to edges not differing more than 45° from a perfectly vertical position is defined as

$$\tilde{E}_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x) = \frac{\left| \sum_{j=J_{\min}}^{J_{\max}} \langle f, \psi_{j, k_j, x}^{\text{odd}} \rangle \right| - \sum_{j=J_{\min}}^{j=J_{\max}} \left| \langle f, \psi_{j, k_j, x}^{\text{even}} \rangle \right|}{(J_{\max} - J_{\min} + 1) \max_{j \in \{J_{\min}, \dots, J_{\max}\}} \left| \langle f, \psi_{j, k_j, x}^{\text{odd}} \rangle \right| + \epsilon}, \quad (4.10)$$

where $\epsilon > 0$ prevents division by zero and

$$k_j = \arg \max_{k \in \left\{ -\left\lceil 2^{\frac{j}{2}} \right\rceil, \dots, \left\lceil 2^{\frac{j}{2}} \right\rceil \right\}} \left| \langle f, \psi_{j, k, x}^{\text{odd}} \rangle \right|. \quad (4.11)$$

$$\left| k_{\psi^{\text{even}}, \psi^{\text{odd}}}^*(f, x) - k \left\lceil 2^{\frac{J_{\min}}{2}} \right\rceil \left\lceil 2^{\frac{j}{2}} \right\rceil^{-1} \right| \leq \frac{1}{2}$$

To ensure a mapping ranging from 0 to 1, we finally set

$$E_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x) = \max \left\{ \tilde{E}_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x), 0 \right\}. \quad (4.12)$$

Please note that the arguments of the maxima in equations (4.9) and (4.11) are not necessarily unique, which can for instance be dealt with by randomly selecting one of the arguments associated with a maximum.

Furthermore, if $E_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x)$ is close to 1 for a given function $f \in L^2(\mathbb{R}^2)$ and $x \in \mathbb{R}^2$, the value $k_{\psi^{\text{even}}, \psi^{\text{odd}}}^*(f, x)$ can be used to approximate the tangential direction of the detected edge in the point x , as depicted in figure 4.6, where first results of a digital implementation of the measure (4.10) are shown. The shearlets ψ^{odd} and ψ^{even} used in this implementation were constructed from tensor products of an even-symmetric scaling function and the even- and odd-symmetric wavelets already used in the one-dimensional case and proposed in [47]. These wavelets not only have compact support and exhibit a high degree of localization, they also are associated with highpass filters of even length, which ensures that we can take full advantage of their symmetry properties in the digital realm. Finally, it should be noted that for the examples shown in figures 4.6 and 4.7, mild soft-thresholding was applied for noise removal, that is, equation (4.10) was changed to

$$\tilde{E}_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x) = \frac{\left| \sum_{j=J_{\min}}^{J_{\max}} \langle f, \psi_{j, k_j, x}^{\text{odd}} \rangle \right| - \sum_{j=J_{\min}}^{j=J_{\max}} \left| \langle f, \psi_{j, k_j, x}^{\text{even}} \rangle \right| - (J_{\max} - J_{\min} + 1)T}{(J_{\max} - J_{\min} + 1) \max_{j \in \{J_{\min}, \dots, J_{\max}\}} \left| \langle f, \psi_{j, k_j, x}^{\text{odd}} \rangle \right| + \epsilon},$$

where for grayscale images with values ranging from 0 (black) to 255 (white), T was set to $2 \left| C_{\psi^{\text{odd}}} \right| \approx 0.466$.

While in figure 4.7, the edge measure (4.12) is compared to the phase congruency estimate proposed by Peter Kovsi in [40] and the famous Canny edge detector [37],

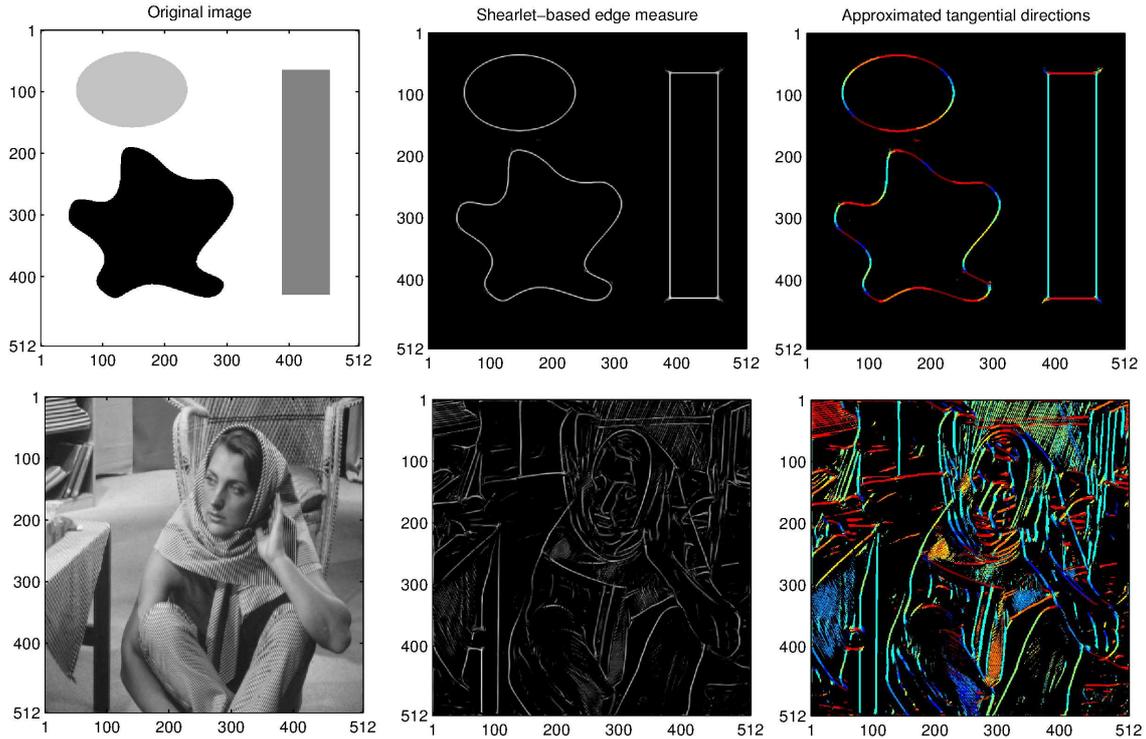


Figure 4.6: The shearlet-based edge measure defined in 4.1.4 applied to one artificial and one natural test image. The second column shows a grayscale plot of the values of (4.12) with white = 1 and black = 0 and parameters $J_{\min} = 1$ and $J_{\max} = 3$. The third column shows the approximated tangential directions as given by the values $k_{\psi^{\text{even}}, \psi^{\text{odd}}}^*(f, x)$ (see equation (4.9)), where bright red indicates a perfect horizontal and light blue a perfect vertical alignment (computed using ShearLab 3D [23]).

maybe one of its most interesting properties is illustrated in figure 4.8. In some sense, edges aren't objects themselves, but merely the structures arising at the borders of two distinct but neighboring regions. That is, an edge can be considered the manifestation of a transition from one geometrical object to another. So far in our considerations, we have always assumed these transitions to be of a very sharp nature, that is, we have modeled them as jump singularities. In practice however, for instance when analyzing parts of a photograph which are not in focus, such transitions can also be of a continuous or smooth nature. As can be seen in figure 4.8, the shearlet-based edge measure is in fact capable of discriminating between different kinds of transitions without loss of localization.

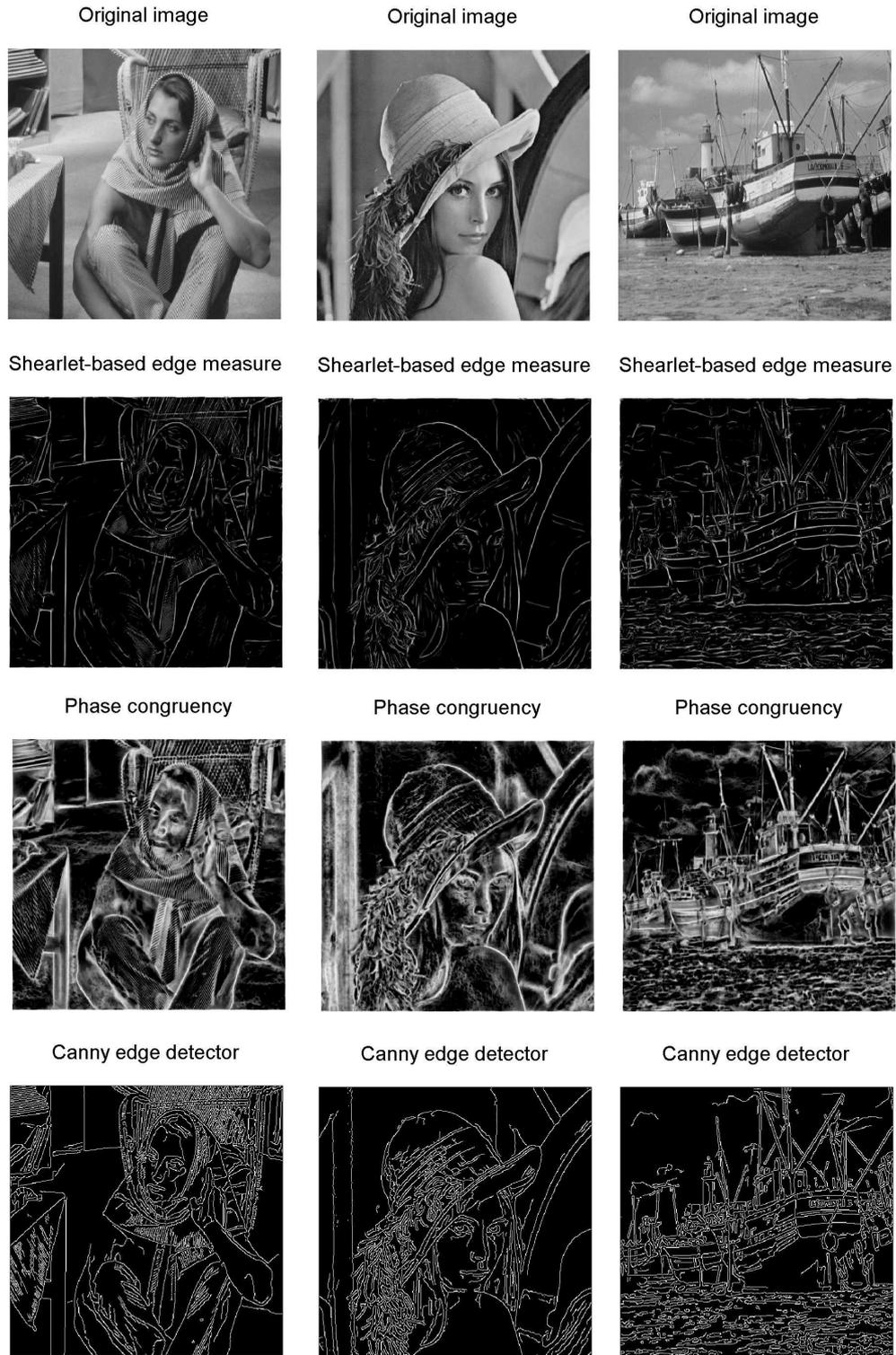


Figure 4.7: We compare the shearlet-based edge measure from definition 4.1.4 to the phase congruency estimate proposed by Kovessi [40] and the Canny edge detector [37]. The respective images were computed using ShearLab 3D [23], the *phasecong2* method from [45] and the MATLAB Image Processing Toolbox.

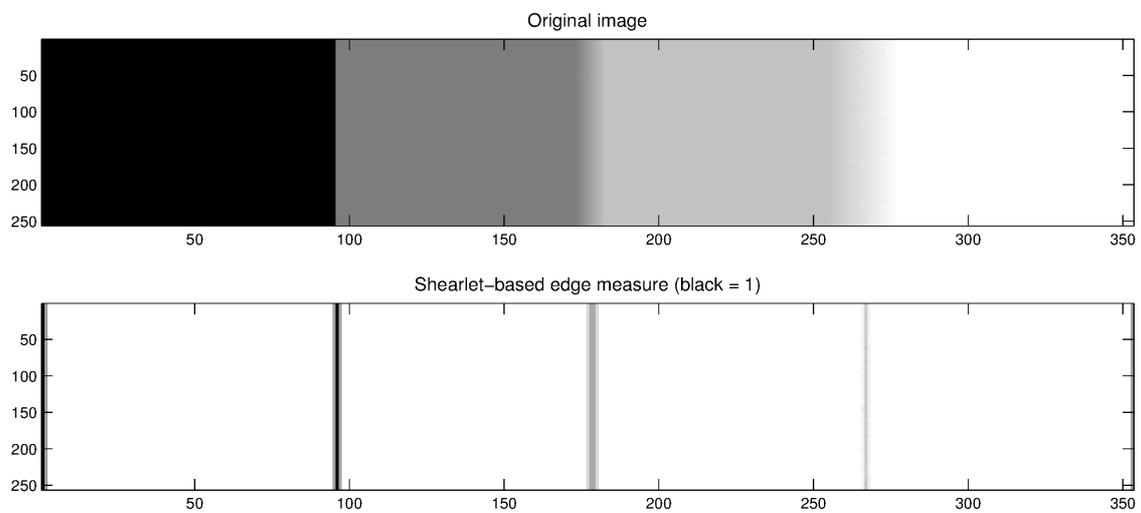


Figure 4.8: The shearlet-based edge measure (see definition 4.1.4) can discriminate between sharp and smooth transitions without loss of localization. While for a point x neighboring the jump singularity on the left, the values $E_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x)$ (see equation (4.12)) are almost exactly 1, the shearlet-based edge measure gradually decreases as the transitions get smoother. Still, all three 'edges' are detected with perfect localization (computed using ShearLab 3D [23]).

4.2 Image Quality Assessment

The second application we consider is using discrete complex shearlet transforms for computing image quality assessments. That is, when given a pair of images where one is a somehow distorted version of the other, we will try to determine the loss of image quality as subjectively perceived by a human observer via a comparison of the coefficients associated with discrete complex shearlet transforms of both images. That such an image quality assessment is not trivial at all from a computational perspective is illustrated in figure 4.9. In this example, the Lenna-image is perturbed once with Gaussian white noise and once with a black square, covering parts of the eyes and the nose. While the mean squared error (MSE) is approximately the same for both distorted images, obscuring important features of a face will certainly result in a much higher degree of subjectively perceived quality loss than uniformly adding Gaussian noise. In general, any distortion of an area containing highly structured information important to a human observer, such as clearly shaped objects or body parts will have a much stronger effect on the perceived image quality than distortions that are restricted to texture-like areas like a water surface or a gravel road.

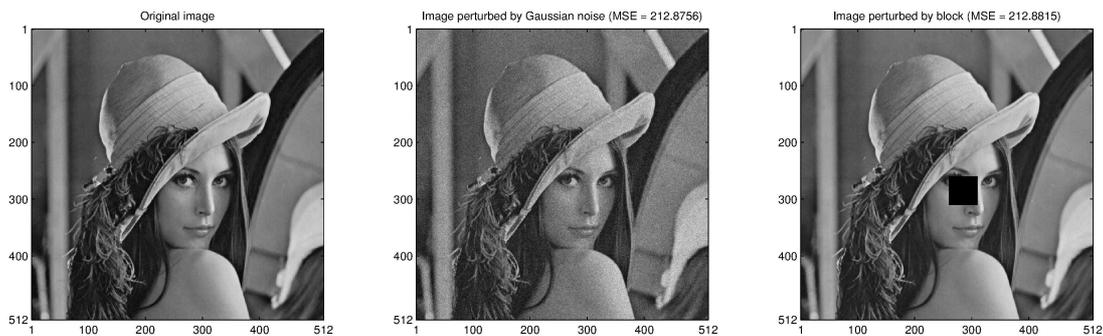


Figure 4.9: While the mean squared error is approximately the same for both distorted images, the perturbation on the right certainly has a much more severe impact on the subjectively perceived image quality.

The task laid out in the paragraph above is also denoted full-reference image quality assessment, as we always assume that both the perturbed and the undistorted reference image is available for analysis. One of the most significant potential applications of such quality assessments is the field of image compression, where the question as to which parts of an image can be subjected to a high degree of which kind of compression without severely affecting the subjectively perceived image quality is of great importance. However, the obvious problem accompanying any approach to the computational assessment of image quality is the question of how one can measure the validity and success of such a method. Given the subjective nature of a human image quality assessment, it will typically be highly dependent on factors like prior personal experiences, cultural influences or the current mood, which are extremely difficult to control for.

The standard approach, which will also be adopted in this section, to at least rudi-

mentary evaluate the accuracy of a computational rule providing full-reference image quality assessments is to compare it to so-called mean opinion scores experimentally collected for large databases of differently distorted images by combining thousands of image quality assessments provided by human test subjects. Examples of such databases include the Tampere Image Database 2008 [48] and the LIVE Image Quality Assessment Database [49], both of which will be used in the numerical analysis following later in this section.

Our motivation for applying discrete complex shearlet transforms to compute image quality assessments again is the assumption that such transforms are a means of reproducing parts of the functional behavior of the human visual system. That is, we hope that the changes of the coefficients of a discrete complex shearlet transform induced by the distortion of an image are approximating how this distortion is perceived by the primary visual cortex. But first, we will examine one of the current state of the art methods for computational image quality assessment, the so-called structural similarity index (SSIM).

4.2.1 Structural Similarity Index (SSIM)

The SSIM was introduced in 2004 by Wang et al. [5] with the ambition of providing a full-reference image quality metric better fitting the behavior of the human visual system than the then and up to now widely used mean squared error. The fundamental idea behind the SSIM is that the human visual system is highly adapted to extract structural information and hence correspondingly sensitive to the degradation of structures induced by distortions. A characteristic, which can definitely not be reproduced by computing the mean squared error, as illustrated in figure 4.9.

To locally measure the structural similarity of two discrete images at a point $x \in \mathbb{Z}^2$, the SSIM considers the discrete Euclidean inner product of vectors given for each image by a predefined window placed around x . While structural information in images is generally independent of the level of luminance and the level of contrast, the same will not be true for locally computed inner products. Hence, the SSIM does not consider the original pair of images for measuring structural similarity but first normalizes them with respect to luminance and contrast, where for luminance normalization, the local mean is subtracted and for contrast normalization, the images are divided by the local standard deviation. That is, the local mean of the images used for computing the structural similarity is always zero while their local standard deviation is one, which eventually results in the computation of the local correlation coefficient. Finally, measures for the local structural similarity, the local contrast similarity and the local luminance similarity are combined to form the structural similarity index defined below.

Definition 4.2.1 (Structural similarity index (SSIM), as in [5]). *Let images $f, g \in \ell^2(\mathbb{Z}^2)$, exponents $\alpha, \beta, \gamma > 0$ and weights $w \in \ell^2(\mathbb{Z}^2)$ be given such that $\sum_{n \in \mathbb{Z}^2} w(n) = 1$.*

Then, the structural similarity index (SSIM) at a point $x \in \mathbb{Z}^2$ is defined as

$$\text{SSIM}(f, g, x) = [l_{f,g}(x)]^\alpha [c_{f,g}(x)]^\beta [s_{f,g}(x)]^\gamma, \quad (4.13)$$

where a local luminance similarity measure l is given by

$$l_{f,g}(x) = \frac{2\mu_f(x)\mu_g(x) + C_1}{\mu_f(x)^2 + \mu_g(x)^2 + C_1}, \quad (4.14)$$

a local contrast similarity measure c is given by

$$c_{f,g}(x) = \frac{2\sigma_f(x)\sigma_g(x) + C_2}{\sigma_f(x)^2 + \sigma_g(x)^2 + C_2}, \quad (4.15)$$

and a local structural similarity measure s is given by

$$s_{f,g}(x) = \frac{\sigma_{f,g}(x) + C_3}{\sigma_f(x)\sigma_g(x) + C_3}, \quad (4.16)$$

with the statistical measures μ_f , σ_f and $\sigma_{f,g}$ defined as

$$\mu_f(x) = \sum_{n \in \mathbb{Z}^2} w(n) f(n-x), \quad (4.17)$$

$$\sigma_f(x)^2 = \sum_{n \in \mathbb{Z}^2} w(n) (f(n-x) - \mu_f(x))^2, \quad (4.18)$$

$$\sigma_{f,g}(x) = \sum_{n \in \mathbb{Z}^2} w(n) (f(n-x) - \mu_f(x)) (g(n-x) - \mu_g(x)) \quad (4.19)$$

and with constants $C_1, C_2, C_3 > 0$ prohibiting instability when the respective denominator would be close to zero and typically chosen as a function of the dynamic range of the images f and g .

Finally, assuming that both the support of g and f is finite, the global structural similarity index is given by

$$\text{SSIM}(f, g) = \frac{1}{N} \sum_x \text{SSIM}(f, g, x), \quad (4.20)$$

where N is the number of indexes taken by x and x typically runs through the indexes of a rectangle fully containing the support of both f and g .

Examples of the structural similarity maps induced by equation (4.13) and the structural similarity index (4.20) for three different types of distortions are given in figure 4.10, where the window w was chosen to be a discrete two-dimensional Gaussian filter of size 11×11 with a standard deviation of 1.5 and all three exponents occurring in equation (4.13) were set to 1, i.e. $\alpha = \beta = \gamma = 1$. Comparing these results with the corresponding mean squared errors clearly shows that the SSIM is less sensitive to changes in contrast and discriminates whether structural information is preserved in the presence of noise or not.

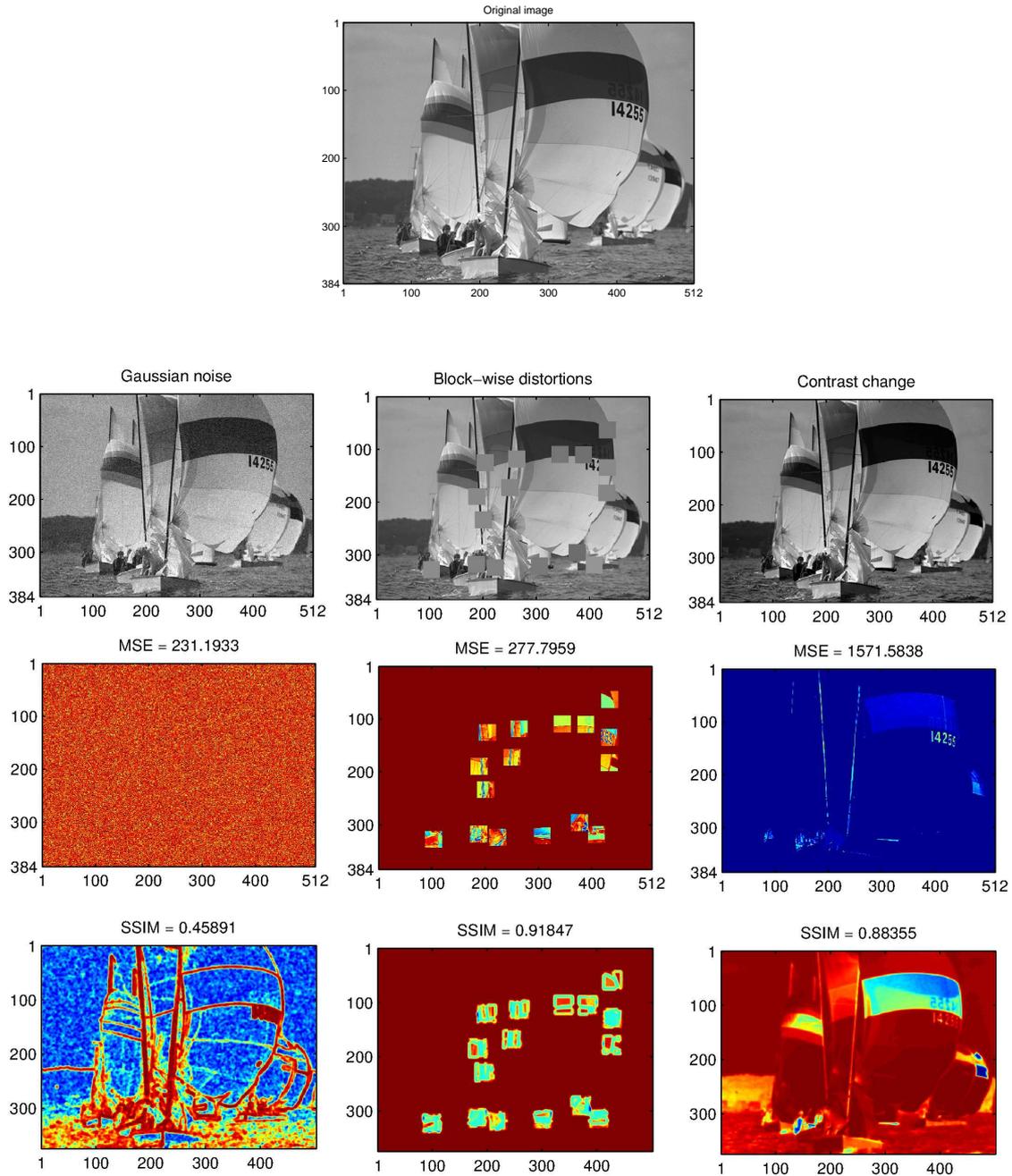


Figure 4.10: The structural similarity maps induced by equation (4.13) and the respective structural similarity indexes (see equation (4.20)) are compared to the mean squared error for three different types of distortions. While in the case of uniformly distributed Gaussian noise the MSE cannot discriminate between the different effects caused by the addition of noise on different parts of the image, the SSIM clearly recognizes that the essential structures forming the boat and the sails have mostly been preserved. Furthermore, the SSIM is considerably less sensitive to changes in contrast, as shown in the third example. The structural similarity maps and indexes were computed using the *ssim_index* method, obtained from [50] and all images were taken from the Tampere Image Database 2008 [48].

Furthermore, it should be noted that the SSIM obeys certain technical conditions one would typically expect from a meaningful image quality metric. It follows from the commutativity of the multiplication that the SSIM is symmetric and due to the construction of the luminance similarity measure (4.14), the contrast similarity measure (4.15) and the structural similarity measure (4.16), the SSIM is bounded above by 1. Additionally, the SSIM is 1 if and only if the compared images are in fact equal. These properties were first noted in [5] and are summarized in the remark below.

Remark 4.2.2 (Properties of the SSIM). *The structural similarity index given in definition 4.2.1 is symmetric, bounded and has a unique maximum, that is, it satisfies*

$$\begin{aligned} \text{SSIM}(f, g) &= \text{SSIM}(g, f), \\ \text{SSIM}(f, g) &\leq 1, \\ \text{SSIM}(f, g) &= 1 \Leftrightarrow f = g \end{aligned}$$

for all $f, g \in \ell^2(\mathbb{Z}^2)$ having finite support.

Proof. These observations are easy to verify and were first stated in [5]. □

While we have already stated that the individual characteristics of a test subject will always influence the subjectively perceived loss in quality, it is important to recognize that the same is true for the experimental setup in which subjective opinion scores are being obtained. Specifically, the subjectively perceived image quality will also depend on factors like the resolution of the images shown and the chosen viewing distance. Hence, depending on these factors, it might be useful to consider a measure similar to (4.13) for a number of several resolutions, thereby simulating different experimental setups. Furthermore, such a multiscale approach would also make it possible to discriminate between distortion-induced structural changes only visible on specific scales. A thus motivated generalization of the SSIM termed multiscale structural similarity index (MSSIM) was introduced in 2003 by Wang, Simoncelli and Bovik [6].

Definition 4.2.3 (Multiscale structural similarity index (MSSIM), as in [6]). *Let $J \in \mathbb{N}$, images $f, g \in \ell^2(\mathbb{Z}^2)$, exponents $\alpha, \beta_j, \gamma_j > 0$ with $j \in \{1, \dots, J\}$ and weights $w \in \ell^2(\mathbb{Z}^2)$ be given such that $\sum_{n \in \mathbb{Z}^2} w(n) = 1$. Then, the multiscale structural similarity index (MSSIM) is defined as*

$$\text{MSSIM}(f, g) = [l(f_J, g_J)]^\alpha \prod_{j=1}^J [c(f_j, g_j)]^{\beta_j} [s(f_j, g_j)]^{\gamma_j}, \quad (4.21)$$

where the images f_j and g_j are constructed by $(j - 1)$ times convolving f (respectively g) with a low-pass filter and applying a dyadic downsampling operator after each convolution and the luminance similarity measure $l(\cdot, \cdot)$, the contrast similarity measure

$c(\cdot, \cdot)$ and the structural similarity measure $s(\cdot, \cdot)$ are defined as

$$\begin{aligned} l(f, g) &= \frac{1}{N} \sum_x l_{f,g}(x), \\ s(f, g) &= \frac{1}{N} \sum_x s_{f,g}(x), \\ c(f, g) &= \frac{1}{N} \sum_x c_{f,g}(x) \end{aligned}$$

for $f, g \in \ell^2(\mathbb{Z}^2)$ with $l_{f,g}(\cdot)$, $c_{f,g}(\cdot)$ and $s_{f,g}(\cdot)$ defined as in 4.2.1 and N denoting the number of indexes taken by $x \in \mathbb{Z}^2$, where x typically iterates through the indexes of a rectangle fully containing the support of both f and g .

The exponents α , β_j and γ_j in above definition can be seen as weights, defining the assumed degree of influence of the different scales on the subjectively perceived image quality. For a MSSIM considering five scales (i.e. $J = 5$), Wang et al. experimentally measured the relative importance of each scale, resulting in a set of exponents given by $\beta_1 = \gamma_1 = 0.0448$, $\beta_2 = \gamma_2 = 0.2856$, $\beta_3 = \gamma_3 = 0.3001$, $\beta_4 = \gamma_4 = 0.2363$ and $\alpha = \beta_5 = \gamma_5 = 0.1333$ (see [6, p. 11]). As a discrete complex shearlet transform naturally induces a partition of a given image into several scales, we will later reuse these parameters in our numerical experiments.

4.2.2 Complex Shearlet-Based Image Similarity Measure

We will now define a measure for the similarity of two given images based on the coefficients of a discrete complex shearlet transform. As it was already stated, the fundamental reason why such a shearlet-based approach might be worth considering is the assumption that a complex shearlet transform provides a good model for the functional behavior of certain parts of the primary visual cortex. However, another intuition suggesting the application of shearlet-based systems in this context is given by the observation that the coefficients associated with a discrete shearlet transform do in fact encode information of both the local structure and the local contrast. The locality directly follows from the localization properties exhibited by shearlets, while the sensitivity to transient features induced by the underlying wavelet (see definition 2.1.15) in combination with the anisotropic nature of shearlets naturally causes shearlet coefficients to be highly responsive to structures present in an image. The sensitivity to contrast, finally, is just a consequence of the bilinearity of the L^2 -inner product.

That is, there seems to be a tight connection between the contrast similarity measure c and the structural similarity measure s used in definition 4.2.1 and the coefficients associated with a discrete shearlet transform. Additionally, similar to the multiscale structural similarity index (see definition 4.2.3), any discrete shearlet transform naturally partitions a signal into multiple scales. Hence, our definition of a shearlet-based image similarity measure will mostly uphold the structure of the computational rules given in definitions 4.2.1 and 4.2.3, with the exception that coefficients associated with

a discrete complex shearlet transform will take the places of the statistical measures μ_f , σ_f and $\sigma_{f,g}$.

However, before in such a way defining a complex shearlet-based image similarity measure, we need to decide on how to exactly use the complex-valued coefficients associated with a discrete complex shearlet transform. We have seen in chapter 3 that so-called complex cells play an important role in the primary visual cortex. Furthermore, it can be assumed that small spatial shifts will only insignificantly affect the subjectively perceived image quality, as long as the overall structure is preserved. Hence, in the following definition, we will only use the magnitude response of a discrete complex shearlet transform.

Please note that, for simplicity, the definition given below only includes shearlets associated with the horizontal frequency cones. The atoms associated with the vertical frequency cones have of course to be included in any actual implementation.

Definition 4.2.4 (Complex shearlet-based image similarity measure). *Let $J \in \mathbb{N}_0$, images $f, g \in L^2(\mathbb{R}^2)$, a real-valued shearlet generator $\psi \in L^2(\mathbb{R}^2)$ associated with the horizontal frequency cones and a scaling function $\phi \in L^2(\mathbb{R}^2)$ be given and denote*

$$\begin{aligned}\psi_{j,k,x}^c &= 2^{\frac{3j}{4}} \left(\psi(S_k A_j(\cdot - x)) + i(\mathcal{H}_{(1,0)^T} \psi)(S_k A_j(\cdot - x)) \right), \\ \phi_x &= \phi(\cdot - x)\end{aligned}$$

for $j \in \{0, \dots, J\}$, $|k| \leq \left\lceil 2^{\frac{j}{2}} \right\rceil$ and $x \in \mathbb{R}^2$. Then, a complex shearlet-based image similarity measure at a point $x \in \mathbb{R}^2$ and a scale $j \in \{0, \dots, J\}$ is given by

$$\text{Sim}_{\psi,\phi}(f, g, j, x) = \begin{cases} \prod_{|k| \leq \left\lceil 2^{\frac{j}{2}} \right\rceil} \frac{2|\langle f, \psi_{j,k,x}^c \rangle| |\langle g, \psi_{j,k,x}^c \rangle| + C}{|\langle f, \psi_{j,k,x}^c \rangle|^2 + |\langle g, \psi_{j,k,x}^c \rangle|^2 + C} & \text{if } j > 0 \\ \frac{2|\langle f, \phi_x \rangle| |\langle g, \phi_x \rangle| + C}{|\langle f, \phi_x \rangle|^2 + |\langle g, \phi_x \rangle|^2 + C} & \text{if } j = 0 \end{cases} \quad (4.22)$$

with a constant $C > 0$.

Now, let $N \in \mathbb{N}$ and $(x_i)_{i \in \{1, \dots, N\}} \subset \mathbb{R}^2$ be a finite sequence of sampling points, then the complex shearlet-based image similarity measure on a scale $j \in \{0, \dots, J\}$ is given by the geometric mean, that is

$$\text{Sim}_{\psi,\phi}(f, g, j) = \left(\prod_{i=1}^N \text{Sim}_{\psi,\tilde{\psi},\phi}(f, g, j, x_i) \right)^{\frac{1}{N}} \quad (4.23)$$

and with exponents $\alpha_j > 0$ for $j \in \{0, \dots, J\}$, the complex shearlet-based image similarity of f and g is given by

$$\text{Sim}_{\psi,\tilde{\psi},\phi}(f, g) = \prod_{j=0}^J \text{Sim}_{\psi,\phi}(f, g, j)^{\alpha_j}. \quad (4.24)$$

A closer look at (4.22) indeed reveals that the complex-valued shearlet coefficients assume the role of the statistical measures used in the definition of the SSIM (see 4.2.1). Another specificity of the above definition is that in equation (4.23), the geometric mean is used instead of the arithmetic mean, which is applied in equation (4.20). The reason behind this is the assumption that a small number of severe distortions, such as the block artefacts shown in figure 4.10, has a much stronger influence on the subjectively perceived image quality than a minor but globally occurring distortion, such as Gaussian noise.

Examples of the image similarity maps induced by equation (4.22) and the complex shearlet-based image similarity measure (4.24) for three different types of distortions are given in figure 4.11. For this example, the number of scales was set to $J = 4$, the constant C , occurring in equation (4.22), was chosen to be 1000 and the exponents from equation (4.24) were again taken from [6] and set to $\alpha_0 = 0.1333$, $\alpha_1 = 0.2363$, $\alpha_2 = 0.3001$, $\alpha_3 = 0.2856$ and $\alpha_4 = 0.0448$. The reference image as well as the distorted images were grayscale images with values ranging from 0 to 255 and taken from [48].

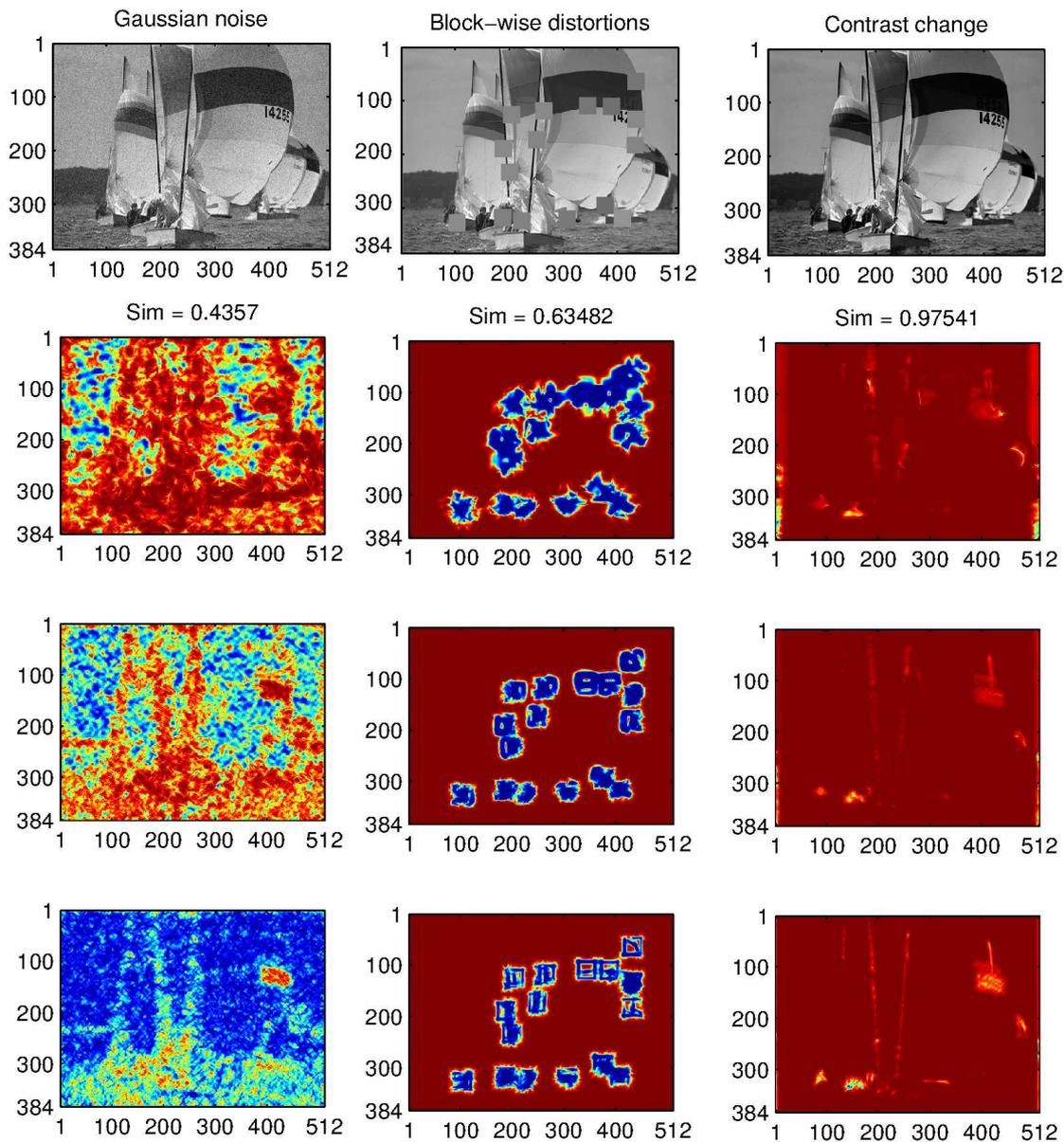


Figure 4.11: We apply the shearlet-based image similarity measure defined in 4.2.4 to assess the loss of quality induced by three different kinds of distortions, which were already considered in figure 4.10. Each row shows the similarity maps induced by equation (4.22) for a different scale, with the scaling parameter j ranging from 1 to 3. As can be seen, uniformly distributed Gaussian noise is mostly visible on the higher scales, while the measure (4.22) is basically invariant to contrast changes over all considered scales (computed using ShearLab 3D [23] on images taken from [48]).

4.2.3 Numerical Results

We now compare the complex shearlet-based image similarity measure from definition 4.2.4 to the SSIM (see definition 4.2.1), the MSSIM (see definition 4.2.3) and six other image quality metrics, namely the feature similarity index (FSIM), introduced by Zhang et al. in [51], the information fidelity criterion (IFC), proposed by Sheikh, Bovik and de Veciana in [52], the visual information fidelity criterion (VIF), also proposed by Sheikh and Bovik in [53], the wavelet-based visual signal-to-noise ratio (VSNR), introduced in [54] by Chandler and Hemami, the noise quality measure (NQM), introduced by Damera-Venkata et al. in [55] and the classical peak signal-to-noise ratio (PSNR).

In order to compare the different full-reference image quality metrics listed above, we will consider three image databases, which in total contain more than 2000 differently distorted images and the respective reference files. For each distorted image in each database, several quality assessments were experimentally collected from human test subjects in a controlled environment. These subjective quality assessments were then averaged and normalized to form a so-called mean opinion score (MOS). We now demand that a meaningful full-reference image quality metric should be capable of serving as a predictor for said mean opinion scores. To be precise, we will measure the success of an image quality metric by comparing the ranking of the distorted images within one database as given by a computational image quality metric to the ranking associated with the mean opinion scores. The degree of correlation between these rankings will then serve as a measure for the validity of the corresponding metric.

To compute the correlation of the rankings associated with an image quality metric and the mean opinion scores provided by human test subjects, we will use the rank order correlation coefficients defined below.

Definition 4.2.5 (Spearman's rank order correlation coefficient (SROCC)). *Let $N \in \mathbb{N}$ and sequences $X = (x_i)_{i \in \{1, \dots, N\}} \subset \mathbb{R}$ and $Y = (y_i)_{i \in \{1, \dots, N\}} \subset \mathbb{R}$. Let furthermore $\text{rank}: \mathbb{R}^N \rightarrow \mathbb{R}^N$ be a function computing the ranks of the elements within a sequence, that is*

$$\text{rank}(X)_i = \#\{j \in \{1, \dots, N\} : x_j \geq x_i\} - \frac{\#\{j \in \{1, \dots, N\} : x_j = x_i\} - 1}{2}. \quad (4.25)$$

Then, Spearman's rank order correlation coefficient is given by

$$\text{SROCC}(X, Y) = \frac{\sum_{i=1}^N \left(\text{rank}(X)_i - \overline{\text{rank}(X)} \right) \left(\text{rank}(Y)_i - \overline{\text{rank}(Y)} \right)}{\sqrt{\sum_{i=1}^N \left(\text{rank}(X)_i - \overline{\text{rank}(X)} \right)^2} \sqrt{\sum_{i=1}^N \left(\text{rank}(Y)_i - \overline{\text{rank}(Y)} \right)^2}}, \quad (4.26)$$

where \overline{X} denotes the arithmetic mean of the sequence X .

Definition 4.2.6 (Kendall's rank order correlation coefficient (KROCC)). *Let $N \in \mathbb{N}$ and sequences $X = (x_i)_{i \in \{1, \dots, N\}} \subset \mathbb{R}$ and $Y = (y_i)_{i \in \{1, \dots, N\}} \subset \mathbb{R}$. Let furthermore*

$$\text{Conc}(X, Y) = \{(i, j) \in \{1, \dots, N\}^2 : (x_i - x_j)(y_i - y_j) > 0\}$$

denote the set of concordant pairs and

$$\text{Disc}(X, Y) = \{(i, j) \in \{1, \dots, N\}^2 : (x_i - x_j)(y_i - y_j) < 0\}$$

denote the set of discordant pairs, then Kendall's rank order correlation coefficient is given by

$$\text{KROCC}(X, Y) = 2 \frac{\# \text{Conc}(X, Y) - \# \text{Disc}(X, Y)}{n(n-1)}. \quad (4.27)$$

Before finally stating the numerical results, we will give a short description of the considered image databases. The Tampere Image Database 2008 (TID 2008) was developed by Ponomarenko et al. [48] and contains 24 reference color images of size 512×384 , where each reference image was subjected to 17 different types of distortions with 4 different degrees of severity, resulting in a total of 1700 distorted images. The various types of distortions used for the TID 2008 are:

- Additive Gaussian noise (see figure 4.10)
- Additive Gaussian noise with different parameters for the three color components
- Spatially correlated noise
- Masked noise
- High frequency noise
- Impulse noise
- Quantization noise
- Gaussian blur
- Errors induced by an image denoising algorithm
- Errors induced by the JPEG compression algorithm
- Errors induced by the JPEG2000 compression algorithm
- JPEG transmission errors
- JPEG2000 transmission errors
- Non-eccentricity pattern noise
- Local block-wise distortions of different intensity (see figure 4.10)
- Intensity shifts (see figure 4.10)
- Contrast changes

For the numerical experiments, all images were converted to 8-bit grayscale images with values ranging from 0 to 255 via the MATLAB *rgb2gray* function.

The second database used for our experiments is the LIVE image quality assessment database, which was developed by Sheikh et al. and can be downloaded from [49]. The

LIVE database contains 29 colored reference images, typically of size 768×512 , and a total of 779 distorted images, where 5 different types of distortions were applied with varying degrees of severity. The number of distorted images is distributed between the different distortion types as follows:

- Errors induced by the JPEG2000 compression algorithm (169 images)
- Errors induced by the JPEG compression algorithm (175 images)
- Additive Gaussian noise (145 images)
- Gaussian blur (145 images)
- JPEG2000 transmission errors (145 images)

Again, for computing the considered image quality metrics, all images within the LIVE database were converted to a grayscale with integer values ranging from 0 to 255 via the MATLAB *rgb2gray* function.

Finally, the Cornell A57 database was developed by Chandler and Hemami, who also introduced the VSNR metric (see [54]), and can be downloaded from [56]. It contains just three 8-bit grayscale reference images with values ranging from 0 to 255 and considers six different types of distortions:

- Quantization noise
- Additive Gaussian noise
- Errors induced by the JPEG compression algorithm
- Errors induced by the JPEG2000 compression algorithm
- Gaussian blur

These different distortion types are distributed over a total of 54 distorted images.

Table 4.1 now compares the performance of the shearlet-based image similarity measure introduced in definition 4.2.4 to eight other image quality metrics for the Tampere Image Database 2008, the LIVE image quality assessment database and the Cornell A57 database using SROCC and KROCC values (see definitions 4.2.5 and 4.2.6). To gain a better understanding of the performance of an image quality metric with respect to a specific type of distortion, SROCC values were also computed for all respective subsets of the TID 2008 and the LIVE database. These results are compiled in table 4.2 and in both tables, the best results are colored in red, the second best results are colored in blue and the third best correlations are colored in green.

The results for the complex shearlet-based image similarity measure compiled in tables 4.1 and 4.2 were obtained with parameters $J = 4$, $\alpha_0 = 0.1333$, $\alpha_1 = 0.2363$, $\alpha_2 = 0.3001$, $\alpha_3 = 0.2856$, $\alpha_4 = 0.0448$ and a constant $C = 1000$ (compare definition 4.2.4). While the exponents α_0 to α_4 were taken from [6], the constant C was set to 1000 after comparing the image quality assessments of 17 differently distorted images from the TID 2008 for about eight different choices of C . While this is certainly not a rigorous

approach, it supports the idea that the success of an image quality metric should be determined by the underlying concepts and not by the perfect choice of parameters. The results of all other image quality metrics were copied from [51].

Table 4.1: SROCC and KROCC values of image quality metrics for three databases

		Shear	FSIM	MSSIM	SSIM	VIF	IFC	VSNR	NQM	PSNR
TID 2008	SROCC	0.8566	0.8805	0.8528	0.7749	0.7496	0.5692	0.7046	0.6243	0.5245
	KROCC	0.6633	0.6946	0.6543	0.5768	0.5863	0.4261	0.5340	0.4608	0.3696
LIVE	SROCC	0.9324	0.9634	0.9445	0.9479	0.9631	0.9234	0.9274	0.9086	0.8755
	KROCC	0.7683	0.8337	0.7922	0.7963	0.8270	0.7540	0.7616	0.7413	0.6864
A57	SROCC	0.8309	0.9181	0.8394	0.8066	0.6223	0.3185	0.9355	0.7981	0.6189
	KROCC	0.6399	0.7639	0.6478	0.6058	0.4589	0.2378	0.8031	0.5932	0.4309

Table 4.2: KROCC values of image quality metrics for specific distortion types

		Shear	FSIM	MSSIM	SSIM	VIF	IFC	VSNR	NQM	PSNR
TID 2008	gaussian noise	0.9176	0.8566	0.8094	0.8107	0.8799	0.5817	0.7728	0.7679	0.9114
	gaussian noise color	0.8943	0.8527	0.8064	0.8029	0.8785	0.5528	0.7793	0.7490	0.9068
	spatial corr-noise	0.9063	0.8483	0.8195	0.8144	0.8703	0.5984	0.7665	0.7720	0.9229
	masked noise	0.8629	0.8021	0.8155	0.7795	0.8698	0.7326	0.7295	0.7067	0.8487
	high-freq-noise	0.9084	0.9093	0.8685	0.8729	0.9075	0.7361	0.8811	0.9015	0.9323
	impulse noise	0.7754	0.7452	0.6868	0.6732	0.8331	0.5334	0.6471	0.7616	0.9177
	quantization noise	0.9017	0.8564	0.8537	0.8531	0.7956	0.5911	0.8270	0.8209	0.8699
	blur	0.9655	0.9472	0.9607	0.9544	0.9546	0.8766	0.9330	0.8846	0.8682
	denoising	0.9611	0.9603	0.9571	0.9530	0.9189	0.8002	0.9286	0.9450	0.9381
	jpg-comp	0.9608	0.9279	0.9348	0.9252	0.9170	0.8181	0.9174	0.9075	0.9011
	jpg2000-comp	0.9784	0.9773	0.9736	0.9625	0.9713	0.9445	0.9515	0.9532	0.8300
	jpg-trans-error	0.8827	0.8708	0.8736	0.8678	0.8582	0.7966	0.8056	0.7373	0.7665
	jpg2000-trans-error	0.9051	0.8544	0.8525	0.8577	0.8510	0.7303	0.7909	0.7262	0.7765
	pattern-noise	0.7228	0.7491	0.7336	0.7107	0.7608	0.8410	0.5716	0.6800	0.5931
	block-distortion	0.8478	0.8492	0.7617	0.8462	0.8320	0.6767	0.1926	0.2348	0.5852
	mean shift	0.6621	0.6720	0.7374	0.7231	0.5132	0.4375	0.3715	0.5245	0.6974
contrast	0.6787	0.6481	0.6400	0.5246	0.8190	0.2748	0.4239	0.6191	0.6126	
LIVE	jpg2000-comp	0.9593	0.9717	0.9654	0.9614	0.9683	0.9100	0.9551	0.9435	0.8954
	jpg-comp	0.9691	0.9834	0.9793	0.9764	0.9842	0.9440	0.9657	0.9647	0.8809
	gaussian noise	0.9794	0.9652	0.9731	0.9694	0.9845	0.9377	0.9785	0.9863	0.9854
	blur	0.9246	0.9708	0.9584	0.9517	0.9722	0.9649	0.9413	0.8397	0.7823
	jpg2000-trans-error	0.9316	0.9499	0.9321	0.9556	0.9652	0.9644	0.9027	0.8147	0.8907

Chapter 5

Discussion

While complex shearlets of the form (2.52) were already introduced by Storath in [3, p. 28f.], the edge measure developed in section 4.1 and the image quality metric defined in section 4.2 to my knowledge constitute the first applications of the associated discrete transforms. Throughout this thesis and especially in chapter 3, we have motivated both the construction and the use of complex shearlets with the assumption that the associated transforms provide a model for the functional behavior of certain parts of the primary visual cortex. That being said, it is important to keep in mind that this assumption should not be mistaken for the claim that there actually are neurons in the primary visual cortex, whose receptive fields are precisely shearlet-shaped or the even stronger claim, that a population of neurons put together actually performs a discrete complex shearlet transform, similar to 2.3.8. After all, within a discrete shearlet system, the atoms are not so much defined by their own structure but by the relationship induced via the anisotropic scaling matrix A_j (see definition 2.1.16) and it would require a lot of courage to claim that the receptive fields of neurons in the primary visual cortex are in fact connected the same way. Furthermore, while providing optimally sparse representations clearly constitutes a very desirable property for any time-frequency decomposition, it might not be the only quality such a decomposition should have, to optimally prepare natural images for higher level processing. For instance, when considering the task of object recognition, one of the most essential feature of any algorithm clearly is a high degree of invariance towards spatial shifts, rotations and structure-preserving deformations, as noted by Stéphane Mallat in [57], where a complex wavelet-based scattering transform is proposed as a means of computing image features for object recognition. It should be noted in this context, that the introduction of a certain degree of shift invariance to the magnitude response of a transform associated with a discrete complex shearlet system by combining Hilbert transform pairs to construct complex-valued atoms is already a first step in this direction. Still, the relatively successful application of complex shearlet-based transforms in tasks routinely carried out by the human visual system described in chapter 4 indeed supports the claim that, at least on a functional level, there is a connection between the time-frequency decomposition carried out by the primary visual cortex and the decompositions associated with a discrete complex shearlet system.

Having a closer look at the complex shearlet-based edge measure given in definition 4.1.4, it should become clear that there still are some unresolved issues and unanswered questions. As it was already discussed, the shearlet-based edge measure is almost completely insensitive to point-like singularities, which might be considered a significant drawback in many applications. On the other hand, this behavior is consistent with the cartoon-like image model and the idea that an edge is not simply defined by a random discontinuity but always associated with a structure arising at the border of two distinct geometrical regions. Furthermore, contrary to our original ambition, the complex shearlet-based edge measure is not entirely contrast independent, as the assumed direction $k_{\psi^{\text{even}}, \psi^{\text{odd}}}^*(f, x)$ is always preselected by choosing the direction associated with the largest shearlet coefficient (in magnitude) located at x (see definition 4.2.4). While doing so significantly improves to localization of corners, this approach also has the adverse effect of high-contrast edges sometimes locally dominating neighboring low-contrast edges (see for example the uncompleted line in the upper right corner of the Lenna-image in figure 4.7). Also, a rigorous mathematical analysis of the behavior of pairs of even- and odd-symmetric shearlets not only for jump discontinuities separating constant functions but also for possibly smooth transitions between smooth regions is still left to do. The same goes for a comprehensive comparison of the complex shearlet-based edge measure with other edge detection methods, going beyond a simple display of results as it was done in figure 4.7. However, even at this early stage of development, the edge measure given in definition 4.1.4 exhibits an array of desirable features. In contrast to most edge detection methods, it is almost contrast invariant and provides a surprisingly precise localization of edges, while even without thresholding, the measure returns zero with a high consistency at locations where no edge is present (see figure 4.5). Most notably, the localization of edges is as precise for smooth transitions as it is for jump discontinuities, with the smoothness of the transition being encoded in the value of $E_{\psi^{\text{even}}, \psi^{\text{odd}}}(f, x)$ (see equation (4.12) and figure 4.8). Furthermore, due to the anisotropic nature of shearlets, the value $k_{\psi^{\text{even}}, \psi^{\text{odd}}}^*(f, x)$ provides a reliable estimate of the tangential direction of an edge at the point x (see figure 4.6). Putting all of this together, the complex shearlet-based edge measure provides an extensive and reliable set of information about the locations and structures of edges present in an image, which, to my knowledge, is indeed unique in its variety. Finally, I would again like to emphasize that by considering the coefficients associated with odd-symmetric shearlets as evidence for and the coefficients associated with even-symmetric shearlets as evidence against the presence of an edge, the complex shearlet-based edge measure is mimicking the concept of inhibitory and excitatory cells.

Turning to the complex shearlet-based image quality metric given in definition 4.2.4, I think that, considering the relative simplicity of the proposed measure, the results compiled in the tables 4.1 and 4.2 are quite promising. Especially the fact that, when restricting the Tampere Image Database 2008 to a specific distortion, the shearlet-based method outperforms all other metrics in many cases suggests that this approach has some potential. However, for both other databases, the shearlet-based metric is clearly inferior to some of the other computational rules and its overall results are consistently outperformed by the feature similarity index (FSIM), introduced by Zhang

et al. in [51]. It is interesting to note here, that the success of the FSIM is in fact based on applying the phase congruency measure (see definition 4.1.2) for assessing the importance of different parts of an image as perceived by the human visual system. This suggests that the information discarded in definition 4.2.4 by only considering the magnitude response of a complex-valued shearlet transform might be crucial for correctly assessing the subjectively perceived image quality after all. While not yet leading to an improved image quality metric, this was indeed the original motivation for developing the shearlet-based edge measure, introduced in section 4.1.

Appendix A

Some Definitions and Formulas

Definition A.1.1 (L^p spaces). Let $p \geq 1$ and $d \in \{1, 2\}$, then we denote

$$\|f\|_p = \left(\int_{\mathbb{R}^d} |f(x)|^p dx \right)^{\frac{1}{p}} \quad (\text{A.1})$$

for $f: \mathbb{R}^d \rightarrow \mathbb{C}$ (even though we will almost exclusively consider real-valued functions) and set

$$L^p(\mathbb{R}^d) = \{f : \|f\|_p < \infty\}, \quad (\text{A.2})$$

where $L^p(\mathbb{R}^d)$ with the norm defined above is a Banach space of equivalence classes. Furthermore, for $p = 2$, the space $L^p(\mathbb{R}^d)$ is a separable Hilbert space with the inner product given by

$$\langle \cdot, \cdot \rangle : L^2(\mathbb{R}^d) \times L^2(\mathbb{R}^d) \rightarrow \mathbb{C} : (f_1, f_2) \mapsto \int_{\mathbb{R}^d} f_1(x) \overline{f_2(x)} dx. \quad (\text{A.3})$$

Definition A.1.2 (ℓ^p spaces). Let $p \geq 1$ and $d \in \{1, 2\}$, then we denote

$$\|f\|_p = \left(\sum_{n \in \mathbb{Z}^d} |f_n|^p \right)^{\frac{1}{p}} \quad (\text{A.4})$$

for a sequence $f \subset \mathbb{C}$ and set

$$\ell^p(\mathbb{Z}^d) = \{f : \|f\|_p < \infty\}, \quad (\text{A.5})$$

where $\ell^p(\mathbb{Z}^d)$ with the norm defined above is a Banach space.

Definition A.1.3 (C^k spaces). Let $k \in \mathbb{N}$, then $C^k(D)$ is the space of all k -times continuously differentiable functions on the set D . If D is compact, a norm is given by

$$\|f\|_{C^k} = \sup_{j \leq k} \sup_{x \in D} |f^{(j)}(x)|.$$

Lemma A.1.4 (Best N -approximations with orthonormal bases). *Let \mathcal{H} be a separable Hilbert space and $(\phi_i)_{i \in \mathbb{N}} \subset \mathcal{H}$ an orthonormal basis of \mathcal{H} , then the best N -term approximation is equivalent to the N -term approximation associated with the N largest coefficients in magnitude (see definition 2.1.3) and the error for $f \in \mathcal{H}$ is given by*

$$\|f - f_N\| = \left\| (\langle f, \phi_i \rangle)_{i \in I \setminus I_N} \right\|_{\ell^2}$$

where $I_N \subset \mathbb{N}$ is the index set associated with the N largest coefficients in magnitude and $f_N \subset \mathcal{H}$ is the corresponding N -term approximation.

Proof. We know that f can be written uniquely as

$$f = \sum_{i \in \mathbb{N}} \langle f, \phi_i \rangle \phi_i,$$

and that f_N is of the form

$$f_N = \sum_{i \in I_N}^N c_i \phi_i,$$

where $(c_i)_{i \in I_N} \subset \mathbb{C}$ is some finite sequence of coefficients. We compute by using Parseval's identity

$$\begin{aligned} \|f - f_N\|^2 &= \left\| \sum_{i \in \mathbb{N}} \langle f, \phi_i \rangle \phi_i - \sum_{i \in I_N} c_i \phi_i \right\|^2 \\ &= \left\| \sum_{i \in I_N} (\langle f, \phi_i \rangle - c_i) \phi_i + \sum_{i \notin I_N} \langle f, \phi_i \rangle \phi_i \right\|^2 \\ &= \sum_{i \in I_N} |\langle f, \phi_i \rangle - c_i|^2 + \sum_{i \notin I_N} |\langle f, \phi_i \rangle|^2 \\ &= \sum_{i \notin I_N} |\langle f, \phi_i \rangle|^2 \\ &= \left\| (\langle f, \phi_i \rangle)_{i \in I \setminus I_N} \right\|_{\ell^2}^2, \end{aligned}$$

when choosing $c_i = \langle f, \phi_i \rangle$ for all $i \in I_N$.

□

Definition A.1.5 (Meyer wavelet, originally defined in [17]). *The Meyer wavelet ψ is given by*

$$\hat{\psi}(\xi) = \begin{cases} (2\pi)^{-\frac{1}{2}} e^{i\frac{\xi}{2}} \sin\left(\frac{\pi}{2} \nu\left(\frac{3}{2\pi} |\xi| - 1\right)\right) & \text{if } \frac{2\pi}{3} \leq |\xi| \leq \frac{4\pi}{3} \\ (2\pi)^{-\frac{1}{2}} e^{i\frac{\xi}{2}} \cos\left(\frac{\pi}{2} \nu\left(\frac{3}{4\pi} |\xi| - 1\right)\right) & \text{if } \frac{4\pi}{3} \leq |\xi| \leq \frac{8\pi}{3}, \\ 0 & \text{else} \end{cases},$$

where ν is a C^k or C^∞ function satisfying

$$\nu(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ 1 & \text{if } x \geq 1 \end{cases}$$

and

$$\nu(x) + \nu(1 - x) = 1.$$

One possibility for choosing ν , given by Daubechies in [15, p.119], is

$$\nu(x) = x^4(35 - 84x + 70x^2 - 20x^3).$$

Definition A.1.6 (Lipschitz α continuity). *A function f is said to be Lipschitz α continuous over an interval $[a, b] \subset \mathbb{R}$, if there exists a constant $L \in \mathbb{R}$ and $\alpha \geq 0$, such that*

$$|f(x) - f(y)| \leq L|x - y|^\alpha$$

for all $x, y \in [a, b]$.

A.1.1 Polynomial Depth Search

Let \mathcal{H} be a separable Hilbert space and $\Phi = (\phi_i)_{i \in I} \subset \mathcal{H}$ be a frame of \mathcal{H} for some index set I , i.e. Φ satisfies the frame condition (2.2) for \mathcal{H} . This implies that for any $f \in \mathcal{H}$, a possibly infinite number of subsets $I_f \subset I$ exists such that

$$f = \sum_{i \in I_f} c_i \phi_i,$$

where $(c_i)_{i \in I_f}$ is a sequence of coefficients. When obeying the restriction of polynomial depth search, we will for each $f \in \mathcal{H}$ only consider index sets $I_f \subset I$, such that the n -th entry of I_f was selected from the $q(n)$ first entries of I , where q is some polynomial.

References

- [1] Gitta Kutyniok, Demetrio Labate, Wang-Q Lim, and Guido Weiss. Sparse multidimensional representation using shearlets. In M. Papadakis, Laine, and Unser, editors, *Wavelets XI*, volume 5914 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 254–262, aug 2005.
- [2] Gitta Kutyniok, Jakob Lemvig, and Wang-Q Lim. *Shearlets - Multiscale Analysis for Multivariate Data*, chapter Shearlets and Optimally Sparse Approximations, pages 145–199. Birkhäuser, 2012.
- [3] Martin Storath. *Amplitude and sign decompositions by complex wavelets - Theory and applications to image analysis*. PhD thesis, Technische Universität München, 2013.
- [4] Peter Kovesi. Phase congruency: A low-level image invariant. *Psychological Research*, 64:136–148, 2000.
- [5] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Proc.*, 13(4):600–612, 2004.
- [6] Zhou Wang, Eero P. Simoncelli, and Alan C. Bovik. Multi-scale structural similarity for image quality assessment. In *Proceedings of 37th IEEE Asilomar Conference on Signals, Systems and Computers*, 2003.
- [7] David L. Donoho. Sparse components of images and optimal atomic decomposition. *Constr. Approx.*, 17:353–382, 2001.
- [8] Gitta Kutyniok and Demetrio Labate. *Shearlets - Multiscale Analysis for Multivariate Data*, chapter Introduction to Shearlets, pages 1–38. Birkhäuser, 2012.
- [9] David L. Donoho, Arian Maleki, and M. Shahram. Wavelab 850. software toolkit for time-frequency analysis. 2006.
- [10] Ole Christensen. *Frames and Bases - An Introductory Course*. Birkhäuser, 2008.
- [11] T.W. Körner. *Fourier Analysis*. Cambridge University Press, 1988.
- [12] Gerald B. Folland. *Fourier Analysis and its Applications*. Brooks/Cole, 1992.
- [13] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 3rd

edition, 2009.

- [14] Dennis Gabor. Theory of communications. *J. IEE (London)*, 93 no. 3:429–457, 1946.
- [15] Ingrid Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992.
- [16] J. O. Stromberg. A modified franklin system and higher order spline systems on \mathbb{R}^n as unconditional bases for hardy spaces. In W. Beckner et al., editor, *Conference in honor of A. Zygmund*, volume 2 of *Wadsworth math. series*, pages 475–493, 1982.
- [17] Yves Meyer. Principe d’incertitude, bases hilbertiennes et algebres d’operateurs. *Seminaire Bourbaki*, 28:209–223, 1985.
- [18] Yves Meyer. Ondelettes, fonctions spline et analyses graduées. *Lectures given at the University of Torino, Italy*.
- [19] Stéphane Mallat. A theory for multiresolution signal decompositoin: the wavelet representation. *IEEE trans. PAMI*, 11:674–693, 1989.
- [20] Emmanuel Candès and David L. Donoho. Curvelets - a surprisingly effective nonadaptive representation for objects with edges. *Curves and Surfaces*, 1999.
- [21] Emmanuel Candès and David L. Donoho. New tight frames of curvelets and optimal representation of objects with c_2 singularities. *Comm. Pure. Appl. Math.*, 57:219–266, 2004.
- [22] Minh N. Do and Martin Vetterli. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans. Image Proc.*, 14:2091–2106, 2005.
- [23] Gitta Kutyniok, Wang-Q Lim, and Rafael Reisenhofer. Shearlab 3d: Faithful digital shearlet transforms based on compactly supported shearlets. *preprint*, 2014.
- [24] K. Guo, Gitta Kutyniok, and Demetrio Labate. Sparse multidimensional representations using anisotropic dilation and shear operators. *Wavelets and Splines (Athens, GA, 2005)*, Nashboro Press, Nashville, TN, pages 189–201, 2006.
- [25] K. Guo and Demetrio Labate. optimal sparse multidimensional representation using shearlets. *SIAM J. Math Anal.* s, 39:298–318, 2007.
- [26] Gitta Kutyniok and Wang-Q Lim. Compactly supported shearlets are optimally sparse. *J. Approx. Theory*, 163:1564–1589, 2011.
- [27] J. A. Movshon, I. E. Thompson, and D. J. Tollhurst. Receptive field organization of complex cells in the cat’s striate cortex. *Journal of Physiology*, 183:79–99, 1978.
- [28] R.. L. De Valois, D. G. Albrecht, and L. G. Thorell. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22:545–559, 1982.

- [29] E. M. Stein. *Singular Integrals and Differentiability Properties of Functions*. Princeton University Press, 1970.
- [30] A. Grossmann and J. Morlet. Decomposition of hardy functions in to square integrable wavelets of constant shape. *SIAM journal on mathematical analysis*, 15(4):723–736, 1984.
- [31] N. Kingsbury. Image processing with complex wavelets. *Philosophical Transactions of the Royal Society of London*, 357(1760):2543–2560, 1999.
- [32] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurons in the cat’s striate cortex. *J. Physiol.*, 148:574–591, 1959.
- [33] F. W. Campbell, G. F. Cooper, and Christina Enroth-Cugell. The spatial selectivity of the visual cells of the cat. *J. Physiol.*, 203:223–235, 1969.
- [34] John G. Daugman. Uncertainty relation for resolution in space, sptial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A*, 2(7):1160–1169, 1985.
- [35] Edward H. Adelson and James R. Bergen. Spatiotemporal energy mmodel for the perception of motion. *J. Opt. Soc. Am. A*, 2(2):284–299, 1985.
- [36] Uri Polat and Anthony M. Norcia. Elongated physiological summation pools in the human visual cortex. *Vision Research*, 38:3735–3741, 1998.
- [37] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.
- [38] Stéphane Mallat and Sifen Zhong. Charactewavelets of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:710–732, 1992.
- [39] Chun-Liang Tu, Wen-Liang Hwang, and Jinn Ho. Analysis of singularities from modulus maxima of complex wavelets. *IEEE Transactions on Information Theory*, 51(3):1049–1062, 2005.
- [40] Peter Kovesi. Image features from phase congruency. *Videre: Journal of Computer Vision Research*, 1(3):1–26, 1999.
- [41] M. Felsberg and G. Sommer. A new extension of linear signal processing for estimating local properties and detecting features. *22. DAGM Symposium Mustererkennung*, 2000.
- [42] S. Yi, Demetrio Labate, Glenn R. Easley, and H. Krim. A shearlet approach to edge analysis and detection. *IEEE Transactions on Image Processing*, 18(5):929–941, 2009.
- [43] M. C. Morrone, J. R. Ross, D. C. Burr, and R. A. Owens. Mach bands are phase dependent. *Nature*, 324(6094):250–253, 1986.

- [44] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313, 1987.
- [45] Peter D. Kovesi. Matlab and octave functions for computer vision and image processing. Centre for Exploration Targeting, School of Earth and Environment, The University of Western Australia. Available from <http://www.csse.uwa.edu.au/~pk/research/matlabfns/>.
- [46] F.W. King. *Hilbert Transforms*, volume 1. Cambridge University Press, 2009.
- [47] Ivan W. Selesnick and A. Farras Abdelnour. Symmetric wavelet tight frames with two generators. *Applied and Computational Harmonic Analysis*, 17(2):211–225, 2004.
- [48] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti. Tid2008 - a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10:30–45, 2009.
- [49] Hamid Rahim Sheikh, Zhou Wang, Lawrence Cormack, and Alan C. Bovik. Live image quality assessment database release 2. Available from <http://live.ece.utexas.edu/research/quality>.
- [50] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. The ssim index for image quality assessment. Available from <http://ece.uwaterloo.ca/~z70wang/research/ssim/>.
- [51] Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. Fsim: A feature similarity index for image quality assessment. *IEEE Trans. Image Proc.*, 20(8):2378–2386, 2011.
- [52] H. R. Sheikh, A. C. Bovik, and G. de Veciana. An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Trans. Image Proc.*, 14(12):2117–2128, 2005.
- [53] H. R. Sheikh and A. C. Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15:430–444, 2006.
- [54] D. M. Chandler and S. S. Hemami. Vsnr: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Trans. Image Proc.*, 16(9):2284–2298, 2007.
- [55] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A. Bovik. Image quality assessment based on a degradation model. *IEEE Trans. Image Proc.*, 9:636–650, 2000.
- [56] D. M. Chandler and S. S. Hemami. A57 database. Available from <http://foulard.ece.cornell.edu/dmc27/vsnr/vsnr.html>.
- [57] Stéphane Mallat. Group invariant scattering. *Communications in Pure and Applied Mathematics*, 65(10):1331–1398, 2012.

Zusammenfassung

Im Rahmen der vorliegenden Arbeit werden ein Kantenerkennungsalgorithmus sowie ein numerisches Bildqualitätsmaß mit Hilfe einer so genannten komplexen Shearlet Transformation entwickelt. Die komplexe Shearlet Transformation kann als eine komplexwertige Generalisierung der klassischen Shearlet Transformation gesehen werden, welche 2005 von Kutyniok, Labate, Lim und Weiss [1] eingeführt wurde und in welcher anisotrop skalierte, sowie gescherte Wavelet-basierte Basisfunktionen zur Zerlegung zwei- oder mehrdimensionaler Signale verwendet werden. Einerseits konnte gezeigt werden, dass Shearlet-basierte Transformationen optimal dünn besetzte Darstellungen von Elementen einer bestimmten Klasse natürlicher Bilder ermöglichen [2], andererseits besitzt die komplexe Shearlet Transformation im Absolutbetrag ähnlich der Fourier Transformation ein gewisses Maß an Translationsinvarianz. Eine Eigenschaft, welche auch von so genannten komplexen Zellen im primären visuellen Kortex gezeigt wird. Dies motiviert die Annahme, dass komplexe Shearlet Transformationen auf einer funktionalen Ebene unter Umständen ein nützliches Modell für Teile des menschlichen visuellen Systems darstellen, was wiederum ihre Anwendungen für Methoden der Bildverarbeitung nahelegt.

In Kapitel 2 werden einige grundsätzliche Ergebnisse und Konzepte aus der Fourier Analysis, der Wavelet-Theorie und der Shearlet-Theorie wiederholt, wobei die Konstruktion von Wavelets sowie Shearlets mit dem Ziel, optimal dünn besetzte Darstellungen stückweise glatter Funktionen in ein und zwei Dimensionen zu konstruieren motiviert wird. Der zentrale Part dieses Kapitels im Rahmen dieser Arbeit ist jedoch die Einführung komplexwertiger Shearlet Transformationen. Nachdem zunächst grundlegende Eigenschaften Hilbert Transformation in ein und zwei Dimensionen besprochen werden, wird die komplexe Shearlet Transformation schließlich mit Hilfe von Hilbert-Transformations-Paaren bestehend aus zwei Shearlet Generatoren definiert.

In Kapitel 3 werden kurz einige neurobiologische Erkenntnisse besprochen, welche die Grundlage für unsere Annahme, dass komplexwertige Shearlet Transformationen auf einer funktionalen Ebene ein vielversprechendes Modell bestimmter Teile des visuellen Kortex sind, bilden.

In Kapitel 4, dem zentralen Teil dieser Arbeit, werden zwei auf komplexwertigen Shearlet Transformationen basierende Methoden aus dem Bereich der Bildverarbeitung motiviert, entwickelt und untersucht. Zunächst wird eine auf komplexen Shearlets basierende Kantenerkennungsmethode definiert, welche eine hohes Maß an Kontrastinvarianz aufweist, erstaunlich präzise Lokalisierungen von Kanten ermöglicht sowie auch die Möglichkeit gibt, eine Approximation der tangentialen Richtung einer Kante in einem bestimmten Punkt zu berechnen. Da die Entwicklung dieser Methode stark von dem von Peter Kovasi eingeführten Phasenkongruenz-Maß [40] inspiriert ist, wird der Definition unseres Algorithmus eine kurze Einführung in das Konzept der Phasenkongruenz vorangestellt.

In Abschnitt 4.2 schließlich wird ein Bildqualitätsmaß definiert, welches ebenfalls auf einer komplexwertigen Shearlet Transformation basiert. Bei derartigen Bildqualitätsmaßen werden jeweils Paare von Bildern verglichen, wobei ein Bild typischerweise durch eine bestimmte Störung veränderte Version des anderen ist. Die Aufgabe eines numerischen Bildqualitätsmaßes ist es nun, den durch einen solchen Vorgang von einem menschlichen Betrachter subjektiv empfundenen Qualitätsverlust bestmöglich vorherzusehen. Eine wichtige Anwendung solcher Maße findet sich im Bereich der Bildkompressionsalgorithmen.

Die Arbeit endet mit einer kurzen Diskussion der in Kapitel 2 eingeführten Transformationen, sowie der in Kapitel 4 beschriebenen Methoden.