

INSTITUT FÜR INFORMATIK

**On the Analysis and Decomposition of
Intrinsically One-Dimensional Signals
and their Superpositions**

Oliver Fleischmann

Bericht Nr. 1216

Dezember 2012

ISSN 2192-6247



CHRISTIAN-ALBRECHTS-UNIVERSITÄT
ZU KIEL

Institut für Informatik der
Christian-Albrechts-Universität zu Kiel
Olshausenstr. 40
D – 24098 Kiel

**On the Analysis and Decomposition of
Intrinsically One-Dimensional Signals and their
Superpositions**

Oliver Fleischmann

Bericht Nr. 1216
Dezember 2012
ISSN 2192-6247

e-mail: ofl@informatik.uni-kiel.de

Dieser Bericht gibt den Inhalt der Dissertation wieder, die der Verfasser im
September 2012 bei der Technischen Fakultät der
Christian-Albrechts-Universität zu Kiel eingereicht hat.
Datum der Disputation: 23.11.2012

1. Gutachter: Prof. Dr.-Ing. Gerald Sommer

2. Gutachter: Prof. Dr.-Ing. Michael Felsberg

3. Gutachter: Prof. Dr.-Ing. Reinhard Koch

Datum der mündlichen Prüfung: 23.11.2012

Abstract

Computer and machine vision tasks can roughly be divided into a hierarchy of processing steps applied to input signals captured by a measuring device. In the case of image signals, the first stage in this hierarchy is also referred to as low-level vision or low-level image processing. The field of low-level image processing includes the mathematical description of signals in terms of certain local signal models. The choice of the signal model is often task dependent. A common task is the extraction of features from the signal. Since signals are subject to transformations, for example camera movements in the case of image signals, the features are supposed to fulfill the properties of invariance or equivariance with respect to these transformations. The chosen signal model should reflect these properties in terms of its parameters.

This thesis contributes to the field of low-level vision. Local signal structures are represented by (sinusoidal) intrinsically one-dimensional signals and their superpositions. Each intrinsically one-dimensional signal consists of certain parameters such as orientation, amplitude, frequency and phase. If the affine group acts on these signals, the transformations induce a corresponding action in the parameter space of the signal model. Hence, it is reasonable, to estimate the model parameters in order to describe the invariant and equivariant features.

The first and main contribution studies superpositions of intrinsically one-dimensional signals in the plane. The parameters of the signal are supposed to be extracted from the responses of linear shift invariant operators: the generalized Hilbert transform (Riesz transform) and its higher-order versions and the partial derivative operators. While well known signal representations, such as the monogenic signal, allow to obtain the local features amplitude, phase and orientation for a *single* intrinsically one-dimensional signal, there exists no general method to decompose *superpositions* of such signals into their corresponding features. A novel method for the decomposition of an arbitrary number of sinusoidal intrinsically one-dimensional signals in the plane is proposed. The responses of the higher-order generalized Hilbert transforms in the plane are interpreted as symmetric tensors, which allow to restate the decomposition problem as a symmetric tensor decomposition. Algorithms, examples and applications for the novel decomposition are provided.

The second contribution studies curved intrinsically one-dimensional signals in the plane. This signal model introduces a new parameter, the curvature, and allows the representation of curved signal structures. Using the inverse stereographic projection to the sphere, these curved signals are locally identified with intrinsically one-dimensional signals in the three-dimensional Euclidean space and analyzed in terms of the generalized Hilbert transform and partial derivatives therein.

The third contribution studies the generalized Hilbert transform in a non-Euclidean space, the two-sphere. The mathematical framework of Clifford analysis proposes a further generalization of the generalized Hilbert transform to the two-sphere in terms of the corresponding Cauchy kernel. Nonetheless, this transform lacks an intuitive interpretation in the frequency domain. A decomposition of the Cauchy kernel in terms of its spherical harmonics is provided. Its coefficients not only provide insights to the generalized Hilbert transform on the sphere, but also allow for fast implementations in terms of analogues of the convolution theorem on the sphere.

Contents

1. Introduction	1
1.1. Motivation	1
1.2. Contributions of this thesis	4
2. Preliminaries	7
3. Signal models	17
3.1. Intrinsically one-dimensional signals	17
3.2. Superimposed intrinsically one-dimensional signals	19
3.3. Transformation behavior of superimposed signals	20
4. Classical operators	23
4.1. The Hilbert transform on the real line	23
4.2. The Hilbert transform in \mathbb{R}^n	25
4.3. Differential operators in \mathbb{R}^n	28
4.4. Multiscale operators	30
5. Decomposition of superimposed intrinsically one-dimensional signals in \mathbb{R}^2	35
5.1. From symmetric tensors to homogeneous polynomials	36
5.2. Symmetric tensor decomposition	44
5.3. Orientation estimation	47
5.4. Phase estimation	49
5.5. Examples and practical considerations	55
5.5.1. Resolution of the kernels	55
5.5.2. Choosing the scale	57
5.5.3. Automatic selection of the order	59
5.5.4. Singularities	62
5.6. Averaging the operator responses	63
5.6.1. Automatic selection of the order	68
5.6.2. Choosing the neighborhood	69
5.7. Application: Characteristic scale of superimposed signals	70
5.7.1. Scale selection for superimposed signals	73
5.7.2. Estimation of the single scales	75
5.7.3. Experiments	77
5.8. Summary	82

6. Analysis of curved intrinsically one-dimensional signals in the plane	85
6.1. Curved signal model	85
6.2. Conformal embedding	87
6.2.1. Properties of the stereographic projection	88
6.2.2. Circles as plane-sphere intersections	89
6.2.3. Geometric interpretation of the normal vector	90
6.2.4. Isophote properties	92
6.2.5. Isophote classification	92
6.2.6. Approximating intrinsically one-dimensional signals in \mathbb{R}^3 with curved intrinsically one-dimensional signals in the plane	93
6.3. Experiments	99
6.3.1. Curvature estimation of planar curves	100
6.3.2. Ridge curvature estimation for digital images	102
6.4. Outlook: The three-dimensional conformal monogenic signal	109
6.5. Summary	112
7. Analysis of intrinsically one-dimensional signals on the two-sphere	113
7.1. Clifford algebra	114
7.2. Basic Fourier analysis on the sphere	115
7.2.1. Spherical harmonics	115
7.2.2. Wigner-D functions	115
7.3. The Cauchy and the Hilbert transform for subsets of \mathbb{R}^n	116
7.4. The Cauchy and the Hilbert transform on the sphere	118
7.4.1. The Cauchy transform on \mathbb{S}^2 as a directional correlation	119
7.4.2. Fourier series expansion of the Cauchy kernel	120
7.5. Applications	129
7.5.1. Poisson scale space	129
7.5.2. Plane wave analysis	130
7.6. Summary	133
8. Conclusion and outlook	135
A. Algorithms	137
Bibliography	137
Index	150

1. Introduction

1.1. Motivation

One of the greatest unsolved problems investigated by man is the understanding of the human visual system and visual perception in general. The nature of this problem seems simple to address, since the apparatus of the visual system and the concept of visual perception are present and used all the time by almost every human being. Nonetheless, to this day, although being a basic building block for the human perception in general, human visual perception to a large amount remains an unsolved problem.

Although being even unaware of the biological part of visual perception, the image formation process, i.e. the formation of a retinal image in the human visual system, the basic question of how visual perception works has already been addressed by ancient Greek philosophers such as Democritus, Epicurus, Plato or Aristotle. Despite the unawareness of the image formation process, their theories already stated, that some sort of copy of the world, as seen by a human being, must be present in the human visual system. These theories can be regarded as the first theories involving some kind of template matching, an idea which is still present in today's theories of visual perception.

One of the most influential early theories of vision dates back to Alhazen and his most famous monography, *The Book of Optics* (1011 - 1021). Although focusing on the image formation process rather than the psychological part of visual perception, Alhazen was the first to give a complete description of geometrical optics and the first to give a correct description of the *camera obscura*, the *pinhole camera*, the basis of every camera today. This understanding of the pinhole camera was the fundament of understanding the image formation process in the human visual system. Nonetheless, it should take almost another 600 years until Johannes Kepler for the first time correctly described the formation of the retinal image in the eye in *Ad Vitellionem paralipomena* (1604).

From that point on, several theories have evolved to explain how the internal representation of the world in the human mind is generated from the physical input, i.e. the intensity pattern of the retina.

Among the most important theories, at least in the context of this thesis, is the *computational approach* to visual perception based on the pioneering work by Gibson (1979) and Marr (1982). The computational approach assumes, that a description of the world can be obtained by mathematical computations acting just on the retinal input image. With the rise of more sophisticated computational capabilities of modern com-

puter hardware, this approach gained a lot of interest. Although today's psychologists dealing with visual perception are aware that our perception of the world cannot be completely described just from these intensity patterns (Mausfeld and Knuesel (2002), Mausfeld (2005)), the computational approach is the theory which, to a large amount, covers the fields today known as *computer vision* and *machine vision*.

According to Marr, the processing of an image consists of three main representations:

1. The *primal sketch* dealing with the description of intensity changes and local image geometry.
2. The 2.5-dimensional sketch dealing with orientation, contour and depth of visible surfaces
3. The 3-dimensional model, dealing with three-dimensional objects

This thesis will contribute to the first of Marr's three representations, the primal sketch. This representation is also referred to as *low-level image processing*. It is one of the first stages in the processing of visual information. Marr himself roughly described this stage in Marr (1976) as a rich description of the intensity changes in the retinal image while asking, which intensity changes might provide such a rich description. This led to the question what kind of components to choose as the most low-level structures, the visual system might process the information as.

Mathematicians, although not motivated by the understanding of the visual system, dealt with similar questions: How can a function be decomposed into a series of simpler *basis functions* with respect to certain properties? While there is no single answer to this question without specifying properties such as the domain of the function or the function space, the basis functions are supposed to span, one of the most prominent examples is the theory of Joseph Fourier who decomposed certain functions on the real line into (infinite) series of trigonometric orthonormal basis functions, the *Fourier series*. Using these complex valued basis functions, Fourier was able to solve certain partial differential equations arising from natural phenomena, such as the heat equation. It turned out that the theory of Joseph Fourier was not limited to the real line but applicable in a straightforward manner to n -dimensional Euclidean spaces and hence the two-dimensional plane, the domain of definition for retinal images.

Biological and psychological investigations of the first steps of information processing in the human visual system such as the investigations by Campbell and Robson (1968) or Langley et al. (1996) resulted in the assumption, that the human visual system processes information by *filtering* the input image in terms of functions, which are very similar to the Fourier basis: Localized frequency- and orientation-selective *bandpass filters*. These filters are realized in terms of certain neurons in the primary visual cortex which only respond if certain frequencies and orientations are present in the input signal, in this context the retinal image. Thus, the human visual system

at a first level performs a decomposition of the input signal into a discrete set of those basis functions and analyzes the presence of each component in the image signal. Using such a decomposition, the information content of the signal is transformed such that a description in terms of the *parameter space* of the basis functions is possible. These parameters, as stated above, are the *orientation*, the *frequency* and the *phase* of each basis function. The presence or strength of each basis function in the signal is measured in terms of the *amplitude*.

The decomposition behavior of the visual front-end hence corresponds to a projection of the signal to a set of (possibly overcomplete) basis functions. Unfortunately, the parameter space of the Fourier basis is continuous and an analysis in terms of projections to basis functions with respect to all possible parameters would result in an infinite number of projections. A strategy to avoid an infinite number of projections is on the one hand a discretization of the parameter space, and on the other hand a choice of upper and lower bounds for unbounded parameters like the frequency. Experiments have shown that the human visual system operates in a similar way, resulting in a limited orientation selectivity and a limited resolution of the human visual system. Nonetheless, even if the parameter space of the basis functions is discretized and bounded, the decomposition in terms of the basis function can be quite costly.

If the human visual system is supposed to be imitated by a computer, it is therefore reasonable, to use some kind of linear filtering operations as the first processing stage. But what is the computer supposed to do with the linear filter responses? While operators might, in the best case, establish an isomorphism between different image representations, they do not establish a *meaning* for the image content. If a human is asked, what the important components of an image are, the answer might include *lines*, *edges*, *contours*, etc. These are concepts our visual perception system automatically attaches to certain intensity patterns. Based on the low-level filter responses, a higher level mechanism performs a segmentation of an image into concepts and categories as mentioned above, also known as *features* (Koenderink (1993)). If the computer is supposed to segment an image analogously, it has to know which intensity patterns or operator responses of the first processing step constitute a feature.

Further, features should be designed in such a way, that they can be identified as such under certain transformations occurring in the world. Suppose a feature has been designed to represent an edge. Further suppose an edge is detected in an image and the image is rotated. Then the same edge should also be detected in the rotated version of the image. This property is also known as *invariance* of a feature. In addition features are often supposed to represent the transformation they have undergone with respect to some identity element. Suppose for example an edge is detected in an image and the orientation angle of that edge is used as a feature. Further suppose the image rotates. Then the orientation angle at the transformed point in the second image rotates accordingly. This property is also known as *equivariance* of a feature.

Within this thesis, features in terms of parameters of certain *signal models* are studied. A priori it is assumed that a signal, such as an image, locally looks like the signal model.

The parameters of the model are supposed to be estimated from the responses of linear shift-invariant operators, similar to the preprocessing in terms of linear filters of the visual front-end. The parameters of the models allow to study the transformation behavior of the model with respect to certain transformation groups and allow to identify the invariance and equivariance properties of the parameters.

1.2. Contributions of this thesis

This thesis contributes to the field of low-level image processing as follows: In chapter 5 the signal model under consideration is constituted by *intrinsically one-dimensional* signals in the plane and their superpositions. A novel method for the decomposition of the superpositions of an arbitrary number of intrinsically one-dimensional signals in the plane into its orientation and phase components is proposed. The decomposition of the orientations is based on the generalized Hilbert transform or partial derivative operators and higher-order compositions of these operators. The phase estimation is only valid for sinusoidal intrinsically one-dimensional signals and requires the generalized Hilbert transform in the plane and its higher order compositions. Conditions for the minimum order of the operators will be provided. Apart from the algorithm using the theoretical optimal operator responses, a practical implementation using discrete convolution kernels with respect to different scales is provided and its problems and shortcomings will be discussed. To eliminate especially the problems related to singularities of the operator responses, an averaging procedure of the operator responses in a local neighborhood is proposed, leading to a *generalized structure tensor*. Based on this construction, an affine equivariant feature is proposed which has possible applications in the field of image matching. The feature detector is compared to state of the art methods.

In chapter 6, intrinsically one-dimensional signals with respect to a non-Cartesian coordinate system are introduced. The chapter proposes a novel way to analyze curved image structures in terms of the new curved intrinsically one-dimensional signal model. Using the inverse stereographic projection to the sphere in \mathbb{R}^3 , it will be shown that the projection of this signal model yields an approximate intrinsically one-dimensional signal in \mathbb{R}^3 restricted to the sphere. This signal is then analyzed in terms of the classical generalized Hilbert transform and partial derivative operators in \mathbb{R}^3 . The orientation vector of this signal encodes, apart from the classical orientation in the plane, the curvature of the signal isophote passing through the current point of interest. Practical experiments will emphasize the validity of the proposed method as a novel way to estimate the isophote curvature yielding the same results as the classical differential geometric method. The experiments include the curvature estimation of digital curves, as well as the isophote curvature estimation in digital images.

Chapter 7 is of more theoretical nature. It aims to provide insights into the Hilbert transform on non Euclidean spaces, in this case the sphere. While it is possible to

generalize the Hilbert transform from the real line to \mathbb{R}^n , further generalizations are possible. In the mathematical context of *Clifford analysis* such generalizations have been proposed for arbitrary subsets of \mathbb{R}^n with smooth boundary. Similar to the Fourier domain interpretation of the Hilbert transform in \mathbb{R}^n , a spectral characterization of the Hilbert transform on the unit sphere in \mathbb{R}^3 will be derived in terms of its spherical harmonic decomposition. The derived Fourier multipliers allow for easy implementations in the spherical harmonic domain and provide an intuitive interpretation of the Hilbert transform on the sphere.

Parts of this work have been published in Fleischmann and Sommer (2012), Fleischmann et al. (2011), Fleischmann et al. (2010a), Fleischmann et al. (2010b), Fleischmann et al. (2009), Wietzke, Fleischmann, Sedlazeck, and Sommer (2010), Wietzke, Fleischmann, and Sommer (2009), Wietzke, Sommer, and Fleischmann (2009), Wietzke, Fleischmann, and Sommer (2008b) and Wietzke, Fleischmann, and Sommer (2008a).

2. Preliminaries

Throughout this thesis, the main subject is the analysis of so called *signals*. These for example might include audio signals in the one-dimensional case, images captured by a camera in the two dimensional case, or geophysical data captured around the earth in the case of the sphere. Although images captured by a measuring device, such as a camera, are discrete entitites, we will consider images as continuous functions:

Definition 2.1 (Signal). Let $V \subseteq \mathbb{R}^n$. Let f be a an element of the function space $L^2(V; \mathbb{K}) \cap L^1(V; \mathbb{K})$ with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then we call f a *signal* on V . Signal

Two-dimensional signals, i.e. $V = \mathbb{R}^2$, will be referred to as *images*. These image signals will appear most frequently within this thesis. The treatment of the signals in the continuous framework is valid in the sense, that as long as certain sampling conditions, such as the Nyquist conditions in \mathbb{R}^n , are satisfied, the continuous function within the chosen function space can always be recovered from its discrete samples in terms of interpolation.

The most common signals arising in the field of computer vision are images captured by cameras. Movements of the camera or movements of objects in a scene, captured by the camera, induce transformations of the captured image. In order to extract information from the images which is unaffected (invariant) with respect to these transformations or to obtain the parameters of a transformation between two images, it is helpful to describe these image transformations in a formal mathematical framework known *group theory*. In the following very basic group theoretic terms are defined. For further details we refer the reader to Olver (2000) and Serre (1977).

The central algebraic object describing certain transformations is the concept of a *group*:

Definition 2.2 (Group). Group

Given a set G and a binary operation \circ , G is called a *group* if the following holds:

- i) The operation \circ is associative: $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$ for all $g_1, g_2, g_3 \in G$.
- ii) There exists an identity element $e \in G$ such that $g \circ e = e \circ g = g$ for all $g \in G$.
- iii) For each $g \in G$ there exists an inverse element $g^{-1} \in G$ such that $g \circ g^{-1} = g^{-1} \circ g = e$.

A group by itself is an abstract algebraic object. In order to describe certain transformations in terms of group elements, the notion of a *group action* has to be introduced. It establishes a way, to apply the transformation described by the group elements to elements of a set.

Group homomorphism

Definition 2.3 (Group homomorphism).

Given two groups G, H with identity elements e_G, e_H and operations \circ, \bullet a map $\rho : G \rightarrow H$ is called a *group homomorphism* if the following holds:

- i) $\rho(e_G) = e_H$
- ii) $\rho(g \circ h) = \rho(g) \bullet \rho(h) \quad \forall g, h \in G$
- iii) $\rho(g^{-1}) = \rho(g)^{-1}$.

Group action

Definition 2.4 (Group action).

Given a set X a *group action* is a group homomorphism ρ from G to the invertible one-to-one maps on X .

The notion of a group action relates the group elements to invertible maps on the set X . Consequently, it is possible to transform the elements of X using the group action. Groups acting on a set X are also referred to as *transformation groups*. A set X together with a group action of G is also called a *G-set*. Common examples of group actions are the general linear group $GL(2; \mathbb{R})$ acting on the real plane \mathbb{R}^2 , e.g:

Example 2.5 ($SO(2)$ action on \mathbb{R}^2). Let $G = SO(2)$ and $X = \mathbb{R}^2$. Then G acts on \mathbb{R}^2 as

$$\rho(g) \circ \mathbf{x} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (2.1)$$

with $\theta \in [0, 2\pi)$.

The G -sets considered in this thesis will have an additional structure, the structure of a vector space. Thus, group homomorphisms defining an action on these linear spaces are of special interest:

Representation

Definition 2.6 (Representation). Let G denote a group and V a vector space. Then a group homomorphism ρ from G to the group of invertible linear transformations $GL(V)$ in V is called a *representation* of G in V .

The term of the representation refines the group action. Whereas the homomorphism in the case of a group action acts on arbitrary sets and maps to the set of invertible one-to-one maps, a representation maps to the set of *linear* invertible maps on a *vector space*. Given a vector space V over a field \mathbb{K} with $\dim V = n$, the group of all linear invertible matrices together with the operation of matrix multiplication is known as the *general linear group* and will be denoted by $GL(n; \mathbb{K})$. Often it is obvious which field we will refer to such that we abbreviate the notation as $GL(n)$. The fields we will deal with in this thesis are the real and the complex numbers, $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Groups

whose elements are invertible matrices are also referred to as *matrix groups* which we will most often encounter in these thesis.

The most prominent subgroups of the general linear group arising in the context of this thesis are the *classical groups*. They allow to describe common transformations such as rotations, dilations or translations.

Definition 2.7 (Special linear group). The group of all invertible $n \times n$ matrices with determinant 1 Special linear group

$$SL(n; \mathbb{K}) = \{\mathbf{M} \in GL(n; \mathbb{K}) \mid \det(\mathbf{M}) = 1\} \quad (2.2)$$

is referred to as the *special linear group*.

Definition 2.8 (Orthogonal group). Orthogonal group
The group

$$O(n; \mathbb{K}) = \{\mathbf{M} \in GL(n; \mathbb{K}) \mid \mathbf{M}\mathbf{M}^T = \mathbf{M}^T\mathbf{M} = \mathbf{I}\} \quad (2.3)$$

is referred to as the *orthogonal group*.

Definition 2.9 (Special orthogonal group). Special orthogonal group
The group

$$SO(n; \mathbb{K}) = SL(n; \mathbb{K}) \cap O(n; \mathbb{K}) \quad (2.4)$$

is referred to as the *special orthogonal group* or the *rotation group*.

Definition 2.10 (Affine group). Affine group
The group

$$\text{Aff}(n; \mathbb{K}) = V \rtimes GL(n; \mathbb{K}) \quad (2.5)$$

is referred to as the *affine group*, where \rtimes denotes the *semidirect product*. In addition to the general linear group, it includes translations in V described by the standard addition operation in V .

According to definition (2.1), images are defined as elements of a certain vector space whose elements are functions. In order to describe transformations acting on these functions, we define a representation on the space of signals by

Action on functions

Definition 2.11 (Action on functions).

Let $f : V \rightarrow \mathbb{K}$ be a function from a (possibly infinite dimensional) Hilbert space of functions H and suppose V is a G -set. Then we define an action on H as

$$(g \circ f)(\mathbf{x}) = f(g^{-1} \bullet \mathbf{x}). \quad (2.6)$$

Here the group action \circ acts on the function space H , while the group action \bullet acts on V .

In addition to basic terms of group theory, basic functional analysis terms have to be introduced.

(Linear) functional

Definition 2.12 ((Linear) Functional).

Let H denote a Hilbert space over a field $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Then a map $f : H \mapsto \mathbb{K}$ is called a *functional*. A functional f is *linear* if the following holds:

$$f(a\mathbf{x} + b\mathbf{y}) = a f(\mathbf{x}) + b f(\mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in H, a, b \in \mathbb{K}. \quad (2.7)$$

Evaluation functional

Definition 2.13 (Evaluation functional).

Let H denote a Hilbert space of functions. Then the functional

$$\text{ev}_{\mathbf{x}} f = f(\mathbf{x}) \quad (2.8)$$

is called the *evaluation functional*.

Throughout this thesis, linear maps acting on Hilbert spaces are of special interest. They arise in many different ways. We have already seen such linear maps in terms of group representations acting on the standard Euclidean spaces as well as function spaces, containing for example signals. We call a linear map acting on a Hilbert space a *linear operator*:

Linear operator

Definition 2.14 (Linear operator, Akhiezer et al. (1993)). Let H be a Hilbert space over a field \mathbb{K} . Then a *linear operator* is a map $\mathcal{A} : H \mapsto H$ fulfilling

$$\mathcal{A}(af + bg) = a\mathcal{A}f + b\mathcal{A}g \quad \text{for all } f, g \in H, a, b \in \mathbb{K}. \quad (2.9)$$

Operators map signals to signals. We will focus on *integral operators*, arising from certain kernels:

Definition 2.15 (Integral operator). Let H be a Hilbert space of functions with $f : V \rightarrow \mathbb{K}$ for $f \in H$. Further let $K : V \times V \rightarrow \mathbb{K}$ denote a *kernel*. Then the *integral operator* defined by

$$\mathcal{A}f(\mathbf{x}) = \text{ev}_{\mathbf{x}} \mathcal{A}f = \int_{\mathcal{V}} f(\mathbf{y})K(\mathbf{x}, \mathbf{y})d\mathbf{y} \quad (2.10)$$

is a linear operator.

An important class of operators is the subset of linear operators which are additionally *shift invariant*:

Definition 2.16 (Linear shift-invariant operator). A linear operator \mathcal{A} is *shift-invariant* LSI operator if it commutes with the translations operator \mathcal{T} , i.e.

$$\mathcal{A}\mathcal{T}f = \mathcal{T}\mathcal{A}f. \quad (2.11)$$

An integral operator is shift invariant if there exists a convolution kernel $k : \mathcal{V} \rightarrow \mathbb{K}$ Convolution kernel with $K(\mathbf{x}, \mathbf{y}) = k(\mathbf{x} - \mathbf{y})$. It follows that integral operators with shift invariant kernels can be expressed as a *convolution*

$$\mathcal{A}f(\mathbf{x}) = \int_{\mathcal{V}} f(\mathbf{y})k(\mathbf{x} - \mathbf{y})d\mathbf{y} = (f * k)(\mathbf{x}). \quad (2.12)$$

The convolution evaluated at a point \mathbf{x} corresponds to a projection of f to a reflection of the kernel k at the origin translated by \mathbf{x} . This thesis will focus on integral operators with shift invariant kernels.

Linear feature

Following the notion of Otsu (1986), we will also call the convolution $(f * k)$ evaluated at a point \mathbf{x} a *linear feature* at \mathbf{x} . It results from the application of the evaluation functional to the convolution of f and k , such that a linear feature itself is a linear functional. Given $m \in \mathbb{N}$ linear shift invariant integral operators $\mathcal{A}_1, \dots, \mathcal{A}_m$ induced by the convolution kernels k_1, \dots, k_m , the vector

$$(\mathcal{A}_1 f(\mathbf{x}), \dots, \mathcal{A}_m f(\mathbf{x}))^T = ((f * k_1)(\mathbf{x}), \dots, (f * k_m)(\mathbf{x}))^T \quad (2.13)$$

is called a *linear feature vector* of f at \mathbf{x} . The space \mathbb{K}^m as a vector space will be called a *feature space*. Note that the convolutions with the kernels k_1, \dots, k_m evaluated at a point \mathbf{x} correspond to coefficients of the projections to the reflected and translated kernels. If the kernels are orthonormal, these projections yield the best approximation of f in $\text{span}\{k_1, \dots, k_m\}$ in the least-squares sense, minimizing the error

Linear feature vector

$$E(f) = \|f - \sum_i \langle f, k_i \rangle k_i\|_2^2. \quad (2.14)$$

The features extracted from signals, which correspond to linear operator responses, are supposed to represent certain structures in the signal in a way, which simplifies their detection and analysis. In the case of images, prominent examples of such structures are corners, edges or junctions which depend on certain parameters. If these local structures transform due to an image transformation induced by a group, their parameters possibly change. Nonetheless, the class of the image structure should be preserved. For example a line in an image changes its orientation due to a rotation but is still an element of the set of all lines, even after the rotation. The group theoretical framework allows us to formalize these requirements:

Orbit **Definition 2.17** (Orbit).

Let G denote a group and X a G -set. For $x \in X$ the set

$$Gx = \{g \circ x : g \in G\} \quad (2.15)$$

is called the *orbit* of x .

Example 2.18 (Unit circle). For $x = (\cos(t), \sin(t))^T, t \in \mathbb{R}$, the unit circle is the orbit of x with respect to the rotation group $SO(2)$.

From the set of all orbits a new set can be defined:

Quotient **Definition 2.19.** Let G be a group and X a G -set. Then the set

$$X/G = \{Gx : x \in X\} \quad (2.16)$$

is called the *quotient* of the action on X .

Signals are considered as elements of a function space, carrying a Hilbert space structure. Signals can therefore be considered as points in this Hilbert space. Using the action on functions (2.11) the set

$$\{g \circ f \mid g \in G\} \quad (2.17)$$

is the corresponding orbit of f in the function space H . The quotient partitions X into equivalence classes where two elements $x, y \in X$ are considered as equivalent if there exists an orbit $O \in X/G$ such that $x \in O$ and $y \in O$.

Example 2.20. A simple example is again obtained from the special orthogonal group $SO(2)$ acting on the real plane \mathbb{R}^2 . The elements of the quotient $\mathbb{R}^2/SO(2)$ are the circles with radius $r \in \mathbb{R}_+$ together with the origin, which is a fixed point under the action of $SO(2)$.

Given a certain signal which has been transformed by an action of a transformation group, the representation in the canonical basis might not reveal the transformation, the signal has undergone, in an accessible way. Nonetheless, it is often possible to construct a map in terms of certain operators, which maps the signal in a feature space respecting the group action, such that the feature vector lies on an orbit of the group in the feature space. Such maps are known as *equivariant maps*:

Definition 2.21 (Equivariant map).

Given a group G acting on vector spaces V, W with actions $\circ : G \times V \rightarrow V$ and $\bullet : G \times W \rightarrow W$ a map $f : V \rightarrow W$ is called *equivariant* if for all $\mathbf{x} \in V$

Equivariant
map

$$f(g \circ \mathbf{x}) = g \bullet f(\mathbf{x}) \quad (2.18)$$

holds.

A special case of an equivariant map is an *invariant map*:

Definition 2.22 (Invariant map).

Given a group G acting on vector spaces V, W with actions $\circ : G \times V \rightarrow V$ and $\bullet : G \times W \rightarrow W$ a map $f : V \rightarrow W$ is called *invariant* if for all $\mathbf{x} \in V$

Invariant map

$$f(g \circ \mathbf{x}) = f(\mathbf{x}) \quad (2.19)$$

holds, i.e. it is equivariant with respect to the trivial representation in W .

Obviously, invariant maps are constant on the orbits in the target space. If we consider the space V in the definition above as the space of signals and the space W as the feature space, invariants allow to distinguish between certain types of signals in the feature space. If we know all possible invariants, we would know all the orbits in the feature space. The knowledge of the orbits would then allow to distinguish between different signals corresponding to the different orbits. Further, if additionally the group action in the feature space is known, an element of an orbit can always be transformed back to a certain identity element, such that an inference about the underlying transformation of an input signal is possible. In addition to the shift invariance of the corresponding operators, it seems therefore natural to study the transformation behavior of the resulting features obtained from an input signal, which has been transformed by the action of a transformation group. Since the features are supposed to describe the

image structure, they should also capture its transformation behavior. In special cases, the features transform *equivariantly* in the feature space. To be precise, if $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is an input signal and G is a transformation group acting on \mathbb{R}^n , then the responses of the linear shift invariant operators $\mathcal{A}_1, \dots, \mathcal{A}_m$ transform according to definition 2.21 as

$$(\mathcal{A}_1, \dots, \mathcal{A}_m)(g \circ f)(\mathbf{x})^\top = g \bullet (\mathcal{A}_1 f(\mathbf{x}), \dots, \mathcal{A}_m f(\mathbf{x}))^\top \quad (2.20)$$

where \circ is the action of G in H and \bullet the action in the feature space. The condition states, that there exists a representation of G in the feature space and the features obtained from a transformed signal transform accordingly in the feature space. A basic problem is the construction of operators, which fulfill the required equivariance condition. A major requirement is, that the kernels k_i of the operators \mathcal{A}_i span an *equivariant function system*:

Equivariant
function system

Definition 2.23 (Equivariant function system, Teo and Hel-Or (1998)). Let V, W denote two vector spaces over a field \mathbb{K} with $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and dimensions $|V| = n$, $|W| = m$. Further let $f : V \rightarrow W$

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^\top \quad (2.21)$$

be a map which is equivariant with respect to a group G where $f_i(\mathbf{x}) : V \rightarrow \mathbb{K}$. Then it holds that for all $g \in G$

$$f(g \circ \mathbf{x}) = (f_1(g \circ \mathbf{x}), \dots, f_m(g \circ \mathbf{x}))^\top \quad (2.22)$$

$$= g \bullet (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^\top \quad (2.23)$$

$$= g \bullet \mathbf{f}(\mathbf{x}). \quad (2.24)$$

We call the set $\{f_1, \dots, f_m\}$ an *equivariant function system*.

Suppose a signal $f : \mathbb{R}^m \rightarrow \mathbb{R}$ and set of linear shift-invariant kernels k_1, \dots, k_m are given, constituting an equivariant function system. Then the corresponding integral operators \mathcal{A}_i applied to a transformed version of the input signal f yield

$$\mathcal{A}_i(g \circ f)(\mathbf{x}) = \int_{\mathcal{V}} f(g^{-1} \bullet \mathbf{y}) k_i(\mathbf{x} - \mathbf{y}) d\mathbf{y} \quad (2.25)$$

$$= \frac{1}{|\det(\rho(g))|} \int_{\mathcal{V}} f(g^{-1} \bullet \mathbf{x} - \mathbf{z}) k_i(g \bullet \mathbf{z}) d\mathbf{z} \quad (2.26)$$

$$= \frac{1}{|\det(\rho(g))|} \sum_{j=1}^m a_{i,j} \int_{\mathcal{V}} f(g^{-1} \bullet \mathbf{x} - \mathbf{z}) k_j(\mathbf{z}) d\mathbf{z} \quad (2.27)$$

$$= \frac{1}{|\det(\rho(g))|} \sum_{j=1}^m a_{i,j} (f * k_j)(g^{-1} \bullet \mathbf{x}) \quad (2.28)$$

where $\rho(g)$ is the representation of g in \mathbb{R}^m . From this relation the transformation of the whole feature vector $(\mathcal{A}_1 f(\mathbf{x}), \dots, \mathcal{A}_m f(\mathbf{x}))^T$ is obtained as

$$\begin{pmatrix} \mathcal{A}_1(g \circ f)(\mathbf{x}) \\ \vdots \\ \mathcal{A}_m(g \circ f)(\mathbf{x}) \end{pmatrix} = \frac{1}{|\det(\rho(g))|} \begin{bmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,m} \end{bmatrix} \begin{pmatrix} (f * k_1)(\mathbf{x}) \\ \vdots \\ (f * k_m)(\mathbf{x}) \end{pmatrix}. \quad (2.29)$$

If the group G is the special linear group, the determinant of the transformation is 1. If Gf denotes the orbit of f with respect to G in H , the equivariance of the kernels ensures, that the linear feature vectors lie on an orbit in the feature space.

Operators \mathcal{A}_i induced by kernels, which constitute an equivariant function system are also known as *steerable filters*. Classically, as for example proposed in Freeman and Adelson (1991), steerable filters have been used from a *synthesis viewpoint* to obtain transformed filters by a linear combination of certain basis filters which are actually the kernels constituting an equivariant function system. In that case, the group transformation is known a priori. A signal might supposed to be analyzed in terms of the operator responses along different orientations as it is for example the case in Yu et al. (2000). The subject of this thesis is instead an *analytical* one. From linear features we want to obtain the group transformation that has caused a certain transformation with respect to an identity signal. This is equivalent to determining the orbit of the features in the feature space and searching for the transformation which maps the features to a certain identity element of the orbit.

Steerable filter

The construction of steerable filters is well understood in the case of compact Lie groups, where the steerable filters corresponds to the basis functions of the Fourier series expansion on the Lie group itself. In the case of non-compact groups there exists no straightforward construction method which prohibits the construction of affine or projective steerable filters. For these groups, the determinant in (2.28) might

be arbitrary small which might lead to unbounded operators (compare Perona (1995)). Further, even if a set of steerable filters can be constructed, it is still unknown what the orbits in the feature space are. In Sommer et al. (1998); Perona (1995) and Hel-Or and Teo (1998) it has been proposed to approximate the steerable filters as the best linear approximation of a set of transformed patterns in terms of their singular value decomposition. While this leads to a set of steerable filters for the specified pattern it has been optimized for, it reveals no knowledge about the orbits, different patterns may lie on, if they are projected to the steerable filters. Consequently, from an analytic point of view, the approach of approximating a set of steerable filters from a set of transformed patterns seems inappropriate.

In some situations we will study signals and operators in their *frequency domain* representations obtained by the *Fourier transform*:

Fourier
transform

Definition 2.24 (Fourier transform, Bracewell (1978)).

Given a square integrable function $f \in L^2(\mathbb{R}^n)$ the integral transform

$$\mathcal{F}f(\mathbf{u}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{x} \quad (2.30)$$

is called the *Fourier transform* of f . The *inverse Fourier transform* is given by

$$f(\mathbf{x}) = \mathcal{F}^{-1}(\mathcal{F}f)(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \mathcal{F}f(\mathbf{u}) e^{i\langle \mathbf{u}, \mathbf{x} \rangle} d\mathbf{u}. \quad (2.31)$$

A special property of the frequency domain representation in conjunction with LSI operators is given by the convolution theorem which states that convolution in the spatial domain corresponds to a multiplication in the Fourier domain:

Proposition 2.25 (Convolution theorem, Bracewell (1978)).

$$(f * h)(\mathbf{x}) = \mathcal{F}^{-1}(\mathcal{F}f \mathcal{F}h)(\mathbf{x}). \quad (2.32)$$

A further property we will frequently use is the commutation relation of convolutions and differentiation:

Proposition 2.26 (Differentiation of convolutions, Bracewell (2003)). *Let \mathcal{D} be a partial differential operator. Then the following holds:*

$$\mathcal{D}(f * h)(\mathbf{x}) = (\mathcal{D}f * h)(\mathbf{x}) = (f * \mathcal{D}h)(\mathbf{x}). \quad (2.33)$$

3. Signal models

In the following, structures in signals, such as low-level image structures, will *locally* be represented by a certain signal model. To be precise, we assume that in a local (in theory infinitesimal) neighborhood of a certain point of interest, e.g. a pixel in a captured image, the signal coincides with the model or is at least approximatively equal to the model up to a negligible error. If a signal f coincides with the signal model at a point $\mathbf{x}_0 \in \mathbb{R}^n$, then we will frequently assume, that this point, the *point of interest*, is chosen as the origin of our coordinate system.

A model assumption always requires a justification of its validity with respect to a certain task. This justification might include mathematical reasoning, experimental validation or biologically inspired reasoning. Before the model is introduced and a justification for the model is provided, we will recapitulate the tasks which are supposed to be solved.

Signals, such as images captured by a camera, are supposed to be described in terms of certain linear features resulting from LSI operators. From these linear features, the parameters of a certain signal model are supposed to be estimated. The basic questions are: What are the invariants of the model with respect to a certain transformation group? If parameters are invariant with respect to a certain transformation, hence being constant on the orbits in the parameter space, they can be used to identify certain points in the signals as *interest-points* or *key-points*. These points can for example be used as a first step for matching different signals which differ by certain transformations. The second question is: Which parameters are equivariant with respect to a certain transformation group? This knowledge allows to infer about how a signal has transformed with respect to some reference signal. This knowledge also allows to assign equivariant regional features to interest-points, which can be used to describe the underlying image content in an invariant way.

3.1. Intrinsically one-dimensional signals

The signal model we will consider in the following chapters consists of sinusoidal intrinsically one-dimensional signals and their superpositions:

Definition 3.1 (Sinusoidal intrinsically one-dimensional signal). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a signal. The signal f is called *sinusoidal intrinsically one-dimensional* (i1D) if there exist $A \in \mathbb{R}^+$, $k \in \mathbb{R}$, $\mathbf{n} \in \mathbb{S}^{n-1}$, $\varphi \in [0, 2\pi)$ such that

Sinusoidal i1D
signal

$$f(\mathbf{x}) = A \cos(k \langle \mathbf{x}, \mathbf{n} \rangle + \varphi). \quad (3.1)$$

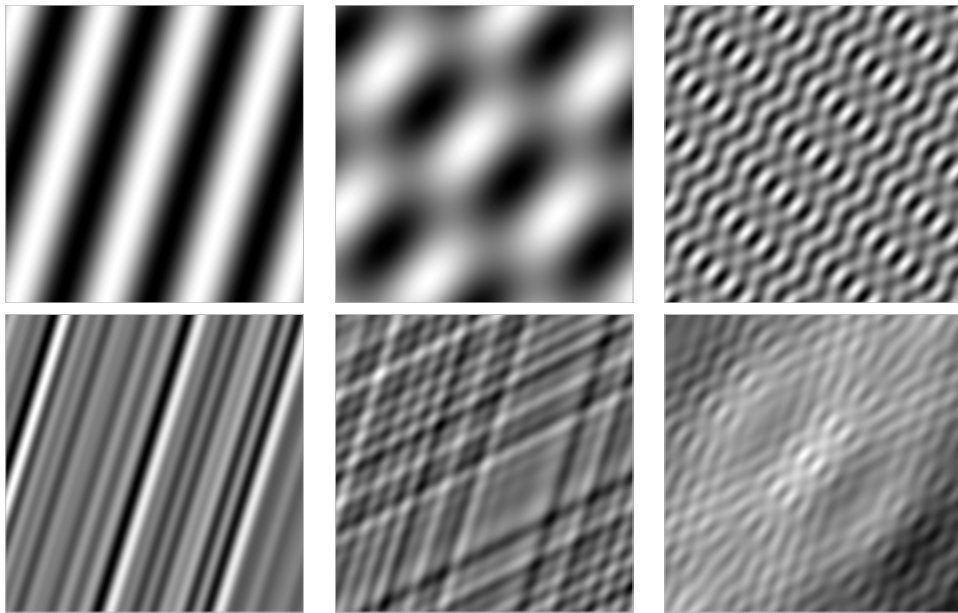


Figure 3.1.: Top row: Superpositions of sinusoidal intrinsically one-dimensional signals. The orders are $d = 1$ (left), $d = 2$ (middle), $d = 3$ (right). Bottom row: Superpositions of intrinsically one-dimensional signals.

The parameters $A, k, \mathbf{n}, \varphi$ are referred to as *amplitude, frequency, orientation* and *linear phase shift* of the signal.

From a mathematical point of view, this model seems appropriate, since every band-limited function can be decomposed into a sum of these signals (compare e.g. Rudin (1976)). Apart from this mathematical argument, the spatial receptive fields of simple cells in the human mammalian striate cortex are, according to Olshausen and Field (1997), localized, oriented and bandpass, which basically states, that the human visual system analyses images in terms of such plane waves. Why is this the case? One argument is: Natural images are *sparse* with respect to these functions. Most of the time only few of these functions are needed to encode the content of natural images. Based on these biologically and mathematical arguments we argue, that the choice of the sinusoidal plane wave as a low-level signal model is (at least in the case of images) appropriate. Further, this model allows an analytical treatment of the transformations with respect to the affine group which facilitates to study the transformation properties in the parameter space of the model.

The definition of the signal model contained another term, the *intrinsic dimension*. If a signal is represented in terms of its Fourier transform, the intrinsic dimension according to Krieger and Zetsche (1996) describes the lowest dimensional subspace, the data lies on. Obviously in the case of images, there are only three possible intrinsic dimensions: i0D, i1D, i2D:

Definition 3.2 (Intrinsically zero-dimensional signal, Krieger and Zetsche (1996)). i0D signal

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a signal. The signal f is called *intrinsically zero-dimensional* (i0D) if it is constant, i.e. there exists $c \in \mathbb{R}$ such that

$$f(\mathbf{x}) = c \quad \forall \mathbf{x} \in \mathbb{R}^n \quad (3.2)$$

Definition 3.3 (Intrinsically one-dimensional signal, Krieger and Zetsche (1996)). i1D signal

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a signal. The signal f is called *intrinsically one-dimensional* (i1D) if there exists a function $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ and $\mathbf{n} \in \mathbb{R}^n, \|\mathbf{n}\| = 1$ such that

$$f(\mathbf{x}) = \tilde{f}(\langle \mathbf{x}, \mathbf{n} \rangle) \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (3.3)$$

Note that a sinusoidal intrinsically one-dimensional signal is a special case of a intrinsically one-dimensional signal. An intrinsically one-dimensional signal may consist of a (possibly infinite) sum of sinusoidal intrinsically one-dimensional signals. Several properties for intrinsically one-dimensional signals in this thesis will hold in general for all intrinsically one dimensional signals while others, such as phase estimation, are only valid in the case of sinusoidal intrinsically one-dimensional functions. Sinusoidal intrinsically one-dimensional function will also be written in the shorter notation

$$f(\mathbf{x}) = \tilde{f}(k \langle \mathbf{x}, \mathbf{n} \rangle + \varphi) \quad \forall \mathbf{x} \in \mathbb{R}^n. \quad (3.4)$$

From the context and the used parameters it will most of the time be obvious, if a certain property is restricted to sinusoidal intrinsically one-dimensional signals. Otherwise it will be mentioned explicitly.

3.2. Superimposed intrinsically one-dimensional signals

Definition 3.4 (Superimposed signal). A signal f is called *superimposed signal* of order d , if it is the superposition of d intrinsically one-dimensional signals, i.e. Superimposed signal

$$f(\mathbf{x}) = \sum_{i=1}^d f_i(\mathbf{x}) = \sum_{i=1}^d \tilde{f}_i(\langle \mathbf{x}, \mathbf{n}_i \rangle). \quad (3.5)$$

In analogue a signal f is called sinusoidal superimposed signal of order d , if it is the superposition of d sinusoidal intrinsically one-dimensional signals, i.e.

$$f(\mathbf{x}) = \sum_{i=1}^d f_i(\mathbf{x}) = \sum_{i=1}^d A_i \cos(k_i \langle \mathbf{x}, \mathbf{n}_i \rangle + \varphi_i). \quad (3.6)$$

i2D signal **Definition 3.5** (Intrinsically two-dimensional (i2D) signal).

A signal f is called *intrinsically two-dimensional* if it is neither constant nor intrinsically one-dimensional.

Obviously superimposed signals are intrinsically two-dimensional if at least two orientations of the components signals differ.

Figure 3.1 illustrates the difference between sinusoidal intrinsically one-dimensional signals and intrinsically one-dimensional signals. The top row shows superpositions of sinusoidal intrinsically one-dimensional signals of orders $d = 1$, $d = 2$ and $d = 3$. The bottom row shows examples for non-sinusoidal intrinsically one-dimensional signals and their superpositions. While the sinusoidal intrinsically one-dimensional signals consist of a *single* real plane wave, the intrinsically one-dimensional signals in the bottom row consist of 10 superimposed sinusoidal intrinsically one-dimensional signals with the same orientation but random frequency, amplitude and phase.

3.3. Transformation behavior of superimposed signals

Sinusoidal intrinsically one-dimensional signals essentially depend on the parameters amplitude, frequency, orientation and phase constituting the *parameter space*

Parameter space

$$\Theta = \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{S}^{n-1} \times \mathbb{S} \equiv \mathbb{R} \times \mathbb{R}^n \times \mathbb{S} \quad (3.7)$$

where \mathbb{S}^{n-1} denotes the unit (hyper-)sphere in \mathbb{R}^n and \mathbb{R}^+ denotes the positive real numbers. Since the amplitude of these signals is unaffected by geometric transformations of the signal, we will not include it in the following investigations, yielding the reduced parameter space

$$\Theta = \mathbb{R}^+ \times \mathbb{S}^{n-1} \times \mathbb{S} \equiv \mathbb{R}^n \times \mathbb{S}. \quad (3.8)$$

In the case of a superimposed signal of order d , the parameter space is the product of d copies of Θ

$$\Theta = \underbrace{\Theta \times \Theta \times \dots \times \Theta}_{d\text{-times}}. \quad (3.9)$$

Suppose an intrinsically one-dimensional signal f with parameter vector $(k\mathbf{n}, \varphi) \in \Theta$ is given. The general linear group $GL(n; \mathbb{K})$ acts on f as:

Proposition 3.6 (Transformation of intrinsically one-dimensional signals).

$$(g \circ f)(\mathbf{x}) = f(g^{-1}\mathbf{x}) = \tilde{f}(k \langle \mathbf{n}, \mathbf{A}^{-1}\mathbf{x} \rangle + \varphi) = \tilde{f}(\langle (\mathbf{A}^{-1})^* k\mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (3.10)$$

$$= \tilde{f}(k' \langle \mathbf{n}', \mathbf{x} \rangle + \varphi) \quad (3.11)$$

for $\mathbf{A} \in GL(n; \mathbb{K})$ and $k'\mathbf{n}' = \mathbf{A}k\mathbf{n}$. \mathbf{A}^{-1*} denotes the conjugate transpose (adjoint operator) of \mathbf{A}^{-1} which corresponds to the transpose if \mathbf{A} is real.

The action of the general linear group on the signal f induces an action on the parameter space Θ . The parameter vector θ transforms as $g \bullet \theta = (\mathbf{A}^{-1*}k\mathbf{n}, \varphi)$. Obviously an intrinsically one-dimensional signal is mapped to another intrinsically one-dimensional signal, which yields

Corollary 3.7. *Intrinsically one dimensional signals are closed under actions of the general linear group.*

Further it is worth noticing, that under an action of the general linear group, only the frequency and the orientation are affected and transformed, i.e.:

Corollary 3.8. *The phase shift φ of an intrinsically one-dimensional signal is invariant under actions of the general linear group.*

Note that if the action is unitary as it is the case for pure rotations, i.e. actions of the rotation group $SO(n)$, it holds that $\mathbf{A}^{-1*} = \mathbf{A}$ such that $\mathbf{A}(k\mathbf{n}) = k\mathbf{n}'$, the frequency is therefore unaffected by pure rotations.

Proposition 3.9 (Translations of intrinsically one-dimensional signals). *Under the action of the translation group, an intrinsically one-dimensional signal transforms as*

$$(g \circ f)(\mathbf{x}) = f(g^{-1}\mathbf{x}) = \tilde{f}(k \langle \mathbf{n}, \mathbf{x} - \mathbf{t} \rangle + \varphi) = \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle - \langle \mathbf{n}, \mathbf{t} \rangle + \varphi) \quad (3.12)$$

$$= \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle + \varphi') \quad (3.13)$$

Corollary 3.10. *The order of a superimposed signal f is preserved under the action of the general linear group if $0 \leq \|\mathbf{A}\| \leq \infty$ for $\mathbf{A} \in GL(n; \mathbb{K})$.*

This basically states that *real world* transformations, i.e. excluding infinite scalings, preserve the order of the superimposed signal. Thus, the order of a superimposed signal is an invariant with respect to such real world transformations.

Why is knowledge about the transformation behavior of the model parameters of interest in this context? In the last section we saw that invariants allow to distinguish

between different orbits in the feature space. If we regard the parameter space of the superpositional model as the feature space with respect to a fixed order d of superpositions, then the phases of the single component functions describe certain orbits with respect to the general linear group, due to their invariance. If it would be possible to estimate the phases of the signals, they would allow a classification of the signals. For example assuming a superpositional model of $d = 2$ intrinsically one-dimensional signals, all points in an image where the two phases of the local signal model are zero could be used as key-points, which are invariant to the actions of the general linear group in the plane. On the other hand, the orientation and frequency parameters allow an inference about the transformation of the local signal. Given two signals which differ by a transformation of the general linear group in the plane and an estimate of the orientations of the signal model, one could ask if it is possible to determine the transformation between the two images.

Motivated by these possible application scenarios, a major goal is the estimation of the orientation and phase parameters. The novel estimation method which will be proposed in chapter 5 is based on certain linear shift-invariant operators. These operators will be introduced in the next section.

4. Classical operators

While the previous section introduced the general signal models we will work within this thesis, this section will introduce the operators yielding linear features of the signals. This thesis will focus on two kinds of operators: differential operators and generalized Hilbert transform operators and their multiscale variants. In the following an overview of the operators and their behavior with respect to the intrinsically one-dimensional signal model is given.

4.1. The Hilbert transform on the real line

When Dennis Gabor in Gabor (1946) proposed the *analytical signal* as a natural complex representation of a real sinusoidal signal, he solely focused on one-dimensional signals in the context of transmitting information over communication channels. His complex representation of a signal, which he obtained from the real signal using the one-dimensional *Hilbert transform*, allowed him to treat signals in the mathematical formalism of quantum mechanics. Although he was aware of the simplified access to the *phase* and mean *frequency* of the signals using his complex representation, his investigations did not use these parameters for the description of structures in the signals. Nonetheless, when the fields of image processing and computer vision began to grow with more and more powerful computing capabilities, scientists working in these fields recognized the capabilities of the analytic signal to describe certain structures in signals (see e.g. Granlund and Knutsson (1995); Fleet and Jepson (1993); Fleet et al. (1991)). The basic question was: Can certain structures in images be approximately described by single sinusoidal intrinsically one-dimensional signals, such that an identification and classification based on its parameters, for example the phase, is possible? This question not only included the actual application of the classical analytic signal to the field of image processing, but also its generalization from one-dimensional signals to two-dimensional signals. The following sections will define the corresponding Hilbert transform operators. Further, several properties of these operators acting on intrinsically one-dimensional signals and their superpositions will be studied.

The analytic signal is a complex representation of a real valued one-dimensional signal. It uses a bounded linear shift-invariant operator, the *Hilbert transform*, to shift the input signal by 90° resulting in a corresponding sine wave with the same amplitude and frequency such that the two waves are said to be *in quadrature*. Due to its linearity and its shift-invariance, the Hilbert transform can be written as the

principal value of the convolution with the following singular kernel in the sense of distributions:

Hilbert transform

Definition 4.1 (One-dimensional Hilbert transform, Stein (1971)).

$$\mathcal{H}f(x) = \text{P.V.}(f * \frac{1}{\pi x})(x). \quad (4.1)$$

From the Hilbert transform, the analytic signal is then derived as the complex valued signal

Analytic signal

Definition 4.2 (Analytic signal, Gabor (1946)).

$$f_A(x) = \mathcal{J}f(x) + \mathbf{i}\mathcal{H}f(x) = f(x) + \mathbf{i}\mathcal{H}f(x) \quad (4.2)$$

where \mathcal{J} denotes the identity operator.

Example 4.3 (Analytic signal). Given a cosine wave $f(x) = A \cos(kx + \varphi)$ its analytic signal representation reads

$$f_A(x) = A \cos(kx + \varphi) + \mathbf{i}A \sin(kx + \varphi) = Ae^{\mathbf{i}kx + \varphi}. \quad (4.3)$$

Proposition 4.4. *The analytic signal representation of a single cosine wave in \mathbb{R} is equivariant with respect to translations modulo 2π .*

Proof. The proposition follows directly from the relation

$$g \circ f_A(x) = f_A(g^{-1}x) = f_A(x - t) = A(\cos(kx + \varphi - t) + \mathbf{i}\sin(kx + \varphi - t)) \quad (4.4)$$

$$= e^{-\mathbf{i}t} f_A(x) = g \bullet f_A(x) \quad (4.5)$$

where \circ denotes the action on the function f_A and \bullet the action in the complex numbers. \square

Local phase

Due to the shift equivariance of the analytic signal representation, the *local phase* of f at every x can be obtained as (see Cohen (1995)):

$$\varphi(x) = \arctan\left(\frac{\mathcal{H}f(x)}{f(x)}\right) \quad \text{with} \quad \varphi(x) \in [0, 2\pi). \quad (4.6)$$

Albeit its theoretical properties, the Hilbert transform applied to one-dimensional real world signals bears certain problems. In general, real world signals do not consist of single cosine waves but rather consist of a superposition of cosine waves with

different frequency, phase and amplitude. Since the concept of phase defined in the sense above is only valid for single cosine waves, the different components have to be extracted from a multicomponent signal by certain filtering operations.

Further, although called *local* phase, the Hilbert transform itself is not a local but a global operator, as pointed out by Boashash (1992). A localization has to be incorporated in the transform by choosing a suitable windowing function of (at least in the limit) compact support, regularizing the Hilbert transform operator.

4.2. The Hilbert transform in \mathbb{R}^n

Inspired by the analytic signal, Felsberg and Sommer (2001) constructed an analogue representation of the analytic signal for intrinsically one-dimensional cosine waves in the plane, the *monogenic signal*. The monogenic signal is based on the *Riesz transform* or *generalized Hilbert transform*, the higher-dimensional generalization of the Hilbert transform in \mathbb{R} to \mathbb{R}^n :

Definition 4.5 (Generalized Hilbert transform (Riesz transform), Stein (1971)).

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a signal in \mathbb{R}^n . Then the j -th component of the generalized Hilbert transform is given by

Generalized
Hilbert
transform

$$\mathcal{H}_j f(\mathbf{x}) = (f * h_j)(\mathbf{x}) \quad \text{for} \quad h_j(\mathbf{x}) = \frac{1}{\pi \omega_{n-1}} \frac{x_j}{\|\mathbf{x}\|}, \quad j \in \{1, \dots, n\} \quad (4.7)$$

and ω_{n-1} denotes the volume of the $(n - 1)$ -ball in \mathbb{R}^n . The generalized Hilbert transform in \mathbb{R}^n is then the vector valued operator

$$\mathcal{H}f(\mathbf{x}) = (\mathcal{H}_1 f(\mathbf{x}), \dots, \mathcal{H}_n f(\mathbf{x}))^\top. \quad (4.8)$$

The frequency domain representation of the generalized Hilbert transform has a particular simple form which will be used later on for practical purposes:

Definition 4.6 (Fourier multiplier of the generalized Hilbert-transform, Stein (1971)).

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a signal in \mathbb{R}^n . Then the Fourier transform of the j -th component of the generalized Hilbert transform in the sense of distributions is given by

Fourier
multiplier

$$\mathcal{F}\mathcal{H}_j f(\mathbf{u}) = \mathbf{i} \frac{u_j}{\|\mathbf{u}\|} \mathcal{F}f(\mathbf{u}). \quad (4.9)$$

Using the generalized Hilbert transform, the monogenic signal is constructed as the vector valued signal

Monogenic
signal

Definition 4.7 (Monogenic signal, Felsberg and Sommer (2001)).

$$\mathbf{f}_M(\mathbf{x}) = (f(\mathbf{x}), \mathcal{H}_1 f(\mathbf{x}), \dots, \mathcal{H}_n f(\mathbf{x}))^\top. \quad (4.10)$$

While Felsberg used an embedding of the monogenic signal representation using a geometric algebra in Felsberg (2002), this will not be necessary in this context and does not change the properties of the signal representation. The vector valued notation with values in \mathbb{R}^n follows the variant introduced in Stein (1971).

Proposition 4.8. *The generalized Hilbert transform in \mathbb{R}^n and consequently the monogenic signal representation is equivariant with respect to rotations, i.e actions of $SO(n)$.*

Proof. This property follows from the definition of the convolution kernels. The kernels according to definition 4.5 decompose into a product of a radial function $\frac{1}{\|\mathbf{u}\|}$ and the canonical Cartesian coordinate functions u_j in the frequency domain. Consequently, kernels are equivariant function system with respect to the group actions of $SO(n)$. \square

Of particular interest are the linear features obtained from the generalized Hilbert transform in the case of an underlying intrinsically one-dimensional signal:

Proposition 4.9 (Generalized Hilbert transform acting on intrinsically one-dimensional signals). *Let $f(\mathbf{x}) = \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle)$ denote a real-valued intrinsically one-dimensional signal in \mathbb{R}^n . Then its generalized Hilbert transform reads*

$$\mathcal{H}f(\mathbf{x}) = (\mathcal{H}_1 \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle), \mathcal{H}_2 \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle), \dots, \mathcal{H}_n \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle)) \quad (4.11)$$

$$= \mathbf{n} \operatorname{ev}_{\langle \mathbf{n}, \mathbf{x} \rangle} \mathcal{H} \tilde{f} \quad (4.12)$$

$$= \mathbf{n} \underline{\mathcal{H}} \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle). \quad (4.13)$$

In the equations above, the notation $\underline{\mathcal{H}} \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle)$ is supposed to be understood as the one-dimensional Hilbert transform of \tilde{f} evaluated at $\langle \mathbf{n}, \mathbf{x} \rangle$.

In the case of a sinusoidal intrinsically one-dimensional signal, the Hilbert transform is explicitly known:

Proposition 4.10 (Generalized Hilbert transform acting on sinusoidal intrinsically one-dimensional signals). *Let $f(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)$ denote a sinusoidal intrinsically one-dimensional signal in \mathbb{R}^n . Then its generalized Hilbert transform reads*

$$\mathcal{H}f(\mathbf{x}) = \mathbf{n} A \sin(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (4.14)$$

Apart from its rotational equivariance, the monogenic signal resembles the behavior of its one-dimensional counterpart, the analytic signal, in the sense that it is equivariant with respect to translations along its orientation. Nevertheless, this property

only holds for sinusoidal intrinsically one-dimensional signals, where the phase shift modulo 2π makes sense:

Proposition 4.11. *Given a sinusoidal intrinsically one-dimensional signal in \mathbb{R}^n , the monogenic signal $(f(\mathbf{x}), \mathcal{H}f(\mathbf{x}))^\top$ is equivariant with respect to translations along \mathbf{n} modulo 2π .*

Proof. Suppose the group of translations along \mathbf{n} mod 2π acts on a sinusoidal intrinsically one-dimensional $f(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle)$ signal in \mathbb{R}^n as:

$$(g \circ f)(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} - t\mathbf{n} \rangle) \quad (4.15)$$

$$= A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle - k \langle \mathbf{n}, t\mathbf{n} \rangle) \quad (4.16)$$

$$= A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle - kt) \quad (4.17)$$

$$= A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle - \varphi). \quad (4.18)$$

Then its generalized Hilbert transform reads

$$\mathcal{H}(g \circ f)(\mathbf{x}) = \mathbf{n} A \sin(k \langle \mathbf{n}, \mathbf{x} \rangle - \varphi). \quad (4.19)$$

Projecting the monogenic signal representation $f_M(\mathbf{x}) = (f(\mathbf{x}), \mathcal{H}f(\mathbf{x}))$ to the plane spanned by $e_1 = (1, 0, 0, \dots, 0)$ and $e_2 = (0, n_1, n_2, \dots, n_n)$ yields

$$\langle f_M(\mathbf{x}), e_1 \rangle = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle - \varphi) \quad (4.20)$$

and

$$\langle f_M(\mathbf{x}), e_2 \rangle = A \sin(k \langle \mathbf{n}, \mathbf{x} \rangle - \varphi). \quad (4.21)$$

We notice that the situation is essentially the same as in the case of the analytic signal. If we consider the complex representation

$$\langle f_M(\mathbf{x}), e_1 \rangle + \mathbf{i} \langle f_M(\mathbf{x}), e_2 \rangle = e^{\mathbf{i}k \langle \mathbf{n}, \mathbf{x} \rangle} e^{-\mathbf{i}\varphi} \quad (4.22)$$

the projections rotate accordingly to the translation modulo 2π in the e_1, e_2 plane. Obviously, for the projection the orientation \mathbf{n} has to be known. Nonetheless, the projection can be obtained up to sign as

$$\|\mathcal{H}A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)\| = |A \sin(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)|. \quad (4.23)$$

The translation parameter φ modulo 2π can then be obtained by the modified inverse tangent function as

$$\varphi = \arctan 2 (|\langle f_M(\mathbf{x}), \mathbf{e}_2 \rangle|, \langle f_M(\mathbf{x}), \mathbf{e}_1 \rangle) \quad (4.24)$$

$$= \arctan 2 (\|\mathcal{H}f(\mathbf{x})\|, f(\mathbf{x})). \quad (4.25)$$

□

Local phase Since the parameter φ describes the phase shift of the sinusoidal intrinsically one-dimensional signal, it is also known as the *local phase* of f at \mathbf{x} , which already appeared in the context of the introduced signal models.

Definition 4.12 (Composition of generalized Hilbert transforms, Stein (1971)). Let $\mathcal{H}_i, \mathcal{H}_j$ denote the Hilbert transform with respect to the x_i and x_j axis in \mathbb{R}^n . Then their composition is given by

$$\mathcal{H}_i \circ \mathcal{H}_j = \mathcal{H}_i \mathcal{H}_j. \quad (4.26)$$

The composition of the transforms induces a corresponding behavior in the frequency domain:

Definition 4.13 (Fourier domain composition of generalized Hilbert transforms, Stein (1971)). Let $\mathcal{H}_i, \mathcal{H}_j$ denote the Hilbert transform with respect to the x_i and x_j axis in \mathbb{R}^n . Then their composition in the frequency domain is given by

$$\mathcal{F}(\mathcal{H}_i \mathcal{H}_j) = -\frac{\mathbf{u}_i \mathbf{u}_j}{\|\mathbf{u}\|^2}. \quad (4.27)$$

Due to the commutativity of multiplication in the frequency domain following commutation relation is derived:

Proposition 4.14 (Commutation relation of generalized Hilbert transform, Stein (1971)). Let $\mathcal{H}_i, \mathcal{H}_j$ denote the Hilbert transform with respect to the x_i and x_j axis in \mathbb{R}^n . Then the two transforms commute such that

$$\mathcal{H}_i \mathcal{H}_j = \mathcal{H}_j \mathcal{H}_i. \quad (4.28)$$

4.3. Differential operators in \mathbb{R}^n

In addition to the generalized Hilbert transforms, signals are supposed to be analyzed in terms of linear features, resulting from partial differential operators. To achieve a common notation, these will often be denoted as follows:

Definition 4.15. The operators

Partial
Derivatives

$$\mathcal{D}_i = \frac{\partial}{\partial x_i} \quad (4.29)$$

for $i \in \{1, \dots, n\}$ are called the first order partial derivatives in \mathbb{R}^n .

Definition 4.16. The vector valued operator

Gradient
operator

$$\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n)^\top = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right)^\top \quad (4.30)$$

is called the *gradient* operator in \mathbb{R}^n .

Several properties which have already been introduced for the generalized Hilbert transforms also hold for the partial derivative operators.

Proposition 4.17 (Partial derivatives transform acting on intrinsically one-dimensional signals). *Let $f(\mathbf{x}) = \tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle)$ denote a real-valued intrinsically one-dimensional signal in \mathbb{R}^n . Then its gradient reads*

$$\mathcal{D}f(\mathbf{x}) = \mathbf{n} \underline{\mathcal{D}}\tilde{f}(\langle \mathbf{n}, \mathbf{x} \rangle) \quad (4.31)$$

While the behavior with respect to the orientation \mathbf{n} is essentially the same as in the case of the Hilbert transforms, the difference between the operators is apparent for sinusoidal intrinsically one-dimensional signals:

Proposition 4.18 (Partial derivatives acting on sinusoidal intrinsically one-dimensional signals). *Let $f(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)$ denote a sinusoidal intrinsically one-dimensional signal in \mathbb{R}^n . Then its gradient reads*

$$\mathcal{D}f(\mathbf{x}) = \mathbf{n} k A \sin(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (4.32)$$

The application of the gradient operator yields a multiplication by the frequency of the signal, showing the *high-pass* characteristic of the derivatives. In contrast to the generalized Hilbert transform, they amplify high frequencies.

Analogously to the Hilbert transform, the partial derivatives along the Cartesian coordinate axes commute:

Proposition 4.19 (Commutation relation of partial derivatives).

$$\mathcal{D}_i \mathcal{D}_j = \mathcal{D}_j \mathcal{D}_i. \quad (4.33)$$

4.4. Multiscale operators

The analysis of signals in terms of partial derivatives is ill-posed according to Lindeberg (1994) in the sense, that even small disturbances in the input signal, e.g. due to noise, may lead to arbitrary large disturbances in the operator response of the partial derivatives. Further, the partial derivatives and the generalized Hilbert transforms have infinite support in the Fourier domain which implies that the corresponding impulse responses are not band-limited. In practice, we are only able to implement filters with band-limited kernels. It is therefore mandatory to enforce a compact support (at least in the limit) in the Fourier domain. Both of these problems can be addressed by a *regularization* of the operators. The regularization we will consider in this context is achieved by a convolution of the original signal with a suitable positive regularization kernel, such as the Gaussian. Since the convolution with such kernels acts like a low-pass filter in the Fourier domain, i.e. it suppresses high frequencies, the convolved signal is a blurred or smoothed version of the original signal. The kernels depend on a new parameter, the *scale*, which controls the amount of blurring. Using blurred versions of the original signal for different scales, the image can be analyzed in terms of different resolutions. Depending on the chosen kernel, a so called *scale-space* can be established.

Linear
scale-space

Definition 4.20 (Linear scale-space representation, Lindeberg (1997)). Let f be a signal. Let $K_s(\mathbf{x}, \mathbf{y})$ be a kernel. Then K induces a *linear scale-space representation*

$$\mathcal{K}_s f(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y}) K_s(\mathbf{x}, \mathbf{y}) d\mathbf{y} \quad (4.34)$$

iff

- i) K is a linear kernel.
- ii) K is shift invariant.
- iii) K is scale invariant.
- iv) K is positive.
- v) K fulfills the semigroup property.

Due to the linearity and the shift invariance, all kernels fulfilling these axioms give rise a to convolution operator. Let K_s be a kernel fulfilling the axioms above. Then there exists a k_s such that $K_s(\mathbf{x}, \mathbf{y}) = k_s(\mathbf{x} - \mathbf{y})$ and

$$\mathcal{K}_s f(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y}) K_s(\mathbf{x}, \mathbf{y}) d\mathbf{y} = \int_{\mathbb{R}^n} f(\mathbf{y}) k_s(\mathbf{x} - \mathbf{y}) d\mathbf{y} = (f * k_s)(\mathbf{x}). \quad (4.35)$$

The kernels considered throughout this thesis are the Gaussian and the Poisson kernel. The Gaussian scale-space representation has first been proposed by Iijima in Japan (see Weickert et al. (1999) for a survey article on Iijimas work). Nonetheless, the scale-space representation has been rediscovered several times for example by Witkin (1983) or Koenderink (1984). The Poisson kernel arises naturally if the Hilbert transform is considered as the non-tangential boundary value of the Cauchy transform as it will be considered in chapter 7. In his thesis in 2002, Felsberg showed that the Poisson kernel fulfills the properties of a linear scale-space inducing kernel.

Despite its regularizing character, it has turned out that the scale-space representation is similar to the processing of spatial patterns in the mammalian visual systems as it has for example been determined in Field (1993). This justifies the scale-space approach as a biologically inspired processing method for signals at different resolutions.

The Gaussian kernel is defined as

Definition 4.21 (Gaussian kernel in \mathbb{R}^n , Lindeberg (1997)).

$$g_s(\mathbf{x}) = \frac{1}{(2\pi s)^{n/2}} e^{-\sum_{i=1}^n x_i^2/(2s)} \quad (4.36)$$

yielding the scale-space representation

Definition 4.22 (Gaussian scale-space representation in \mathbb{R}^2 , Lindeberg (1997)).

Gaussian
scale-space

$$\mathcal{G}_s f(\mathbf{x}) = (f * g_s)(\mathbf{x}) \quad (4.37)$$

Analogously the Poisson kernel is given by

Definition 4.23 (Poisson kernel in \mathbb{R}^n , Felsberg (2002)).

$$p_s(\mathbf{x}) = c_n \frac{s}{(\sum_{i=1}^n x_i^2 + s^2)^{(n+1)/2}} \quad (4.38)$$

with $c_n = \frac{\Gamma[(n+1)/2]}{\pi^{(n+1)/2}}$.

The Poisson kernel yields the scale-space representation

Definition 4.24 (Poisson scale-space representation in \mathbb{R}^2 (*Poisson transform*), Felsberg (2002)).

Poisson
scale-space

$$\mathcal{P}_s f(\mathbf{x}) = (f * p_s)(\mathbf{x}). \quad (4.39)$$

The operators considered so far, i.e. the generalized Hilbert transform and the partial derivative operators, can be applied to a scale-space representation of a signal f , yielding a set of *multiscale* operators. The operators can be derived by applying the operators to the kernels and convolving the signal by the kernels subsequently:

Multiscale
partial
derivatives

Proposition 4.25 (Multiscale differential operators).

$$\mathcal{D}\mathcal{K}_s f(\mathbf{x}) = \mathcal{D}(f * k_s)(\mathbf{x}) \quad (4.40)$$

$$= (\mathcal{D}f * k_s)(\mathbf{x}) \quad (4.41)$$

$$\stackrel{(2.33)}{=} (f * \mathcal{D}k_s)(\mathbf{x}) \quad (4.42)$$

The same properties also hold for the generalized Hilbert transforms applied to a linear multiscale representation of f :

Multiscale
Hilbert
transforms

Proposition 4.26 (Multiscale Hilbert transform operators).

$$\mathcal{H}\mathcal{K}_s f(\mathbf{x}) = \mathcal{H}(f * k_s)(\mathbf{x}) \quad (4.43)$$

$$= (\mathcal{H}f * k_s)(\mathbf{x}) \quad (4.44)$$

$$\stackrel{(2.33)}{=} (f * \mathcal{H}k_s)(\mathbf{x}) \quad (4.45)$$

Definition 4.27 (Conjugate Poisson kernel and conjugate Poisson transform, Felsberg (2002)). In the special case of the Hilbert transforms applied to the Poisson kernel, we call the kernels

$$(q_{1,s}(\mathbf{x}), q_{2,s}(\mathbf{x}), \dots, q_{n,s}(\mathbf{x}))^\top = (\mathcal{H}_1 p_s(\mathbf{x}), \mathcal{H}_2 p_s(\mathbf{x}), \dots, \mathcal{H}_n p_s(\mathbf{x}))^\top \quad (4.46)$$

Conjugate
Poisson kernel

the *conjugate Poisson kernels* and the corresponding operators

$$\mathcal{Q}_{i,s} f(\mathbf{x}) = \mathcal{H}_i \mathcal{P}_s f(\mathbf{x}) = (f * q_{i,s})(\mathbf{x}) \quad (4.47)$$

Conjugate
Poisson
transform

the *conjugate Poisson transform* along the x_i axis.

This definition yields the vector valued conjugate Poisson transform along all axes:

$$\mathcal{H}\mathcal{P}_s f(\mathbf{x}) = \mathcal{Q}_s f(\mathbf{x}) = (\mathcal{Q}_{1,s} f(\mathbf{x}), \mathcal{Q}_{2,s} f(\mathbf{x}), \dots, \mathcal{Q}_{n,s} f(\mathbf{x}))^\top. \quad (4.48)$$

If real world signals are analyzed in terms of the proposed signal model, it is rarely the case that a signal is locally a superposition of the assumed number of intrinsically one-dimensional signals. Rather a frequency band of interest has to be extracted from the signal using a *bandpass filter*. Bandpass filters suppress frequency contents in the Fourier domain which are outside of a certain pass band. The kernels which were used to obtain the linear scale-space representation of a signal turned out to be low-pass filters, suppressing higher frequencies. Nonetheless, applying the Laplace

operator to these kernels yields a band-pass filter kernel whose pass-band depends on the scale of the kernel:

$$w_s(\mathbf{x}) = \Delta k_s(\mathbf{x}) = \frac{\partial^2}{\partial x_1^2} k_s(\mathbf{x}) + \cdots + \frac{\partial^2}{\partial x_n^2} k_s(\mathbf{x}). \quad (4.49)$$

Band-pass filter
kernel

A band-pass filtered representation of the original signal is then obtained as

$$\mathcal{W}_s f(\mathbf{x}) = (f * w_s)(\mathbf{x}). \quad (4.50)$$

Note that according to Mallat (1999), this representation is equal to a continuous wavelet transform representation, since the corresponding band-pass filter kernel corresponds to an isotropic shift invariant wavelet. Fast approximations of the Laplacian of the kernels for certain bandwidths can be obtained as the difference of two Gaussian / Poisson kernels leading to fast pyramid schemes for their implementations and have for example been applied and derived in Lowe (1999) and Felsberg (2002). The band-pass filtered representation yields a further advantage: The generalized Hilbert transforms (Riesz transforms) exhibit a singular behavior at the origin in the Fourier domain and prohibit a practical implementation. The band-pass filter kernels introduced above are isotropic wavelets with two vanishing moments. The vanishing moments imply, that their Fourier transform at the origin are zero, i.e.

$$\mathcal{F} w_s(\mathbf{0}) = 0 \quad (4.51)$$

such that the singularity of the generalized Hilbert transform are tempered at the origin (compare Chenouard and Unser (2011))

$$\mathcal{F} \mathcal{H}_i \mathcal{W}_s f(\mathbf{0}) = 0. \quad (4.52)$$

If the (higher-order) generalized Hilbert transforms are supposed to be applied in the spatial domain by convolutions, the convolution kernels have to be derived analytically. While the kernels up to order three have been derived in Felsberg (2002); Wietzke (2011), the computations are complicated for any higher order. A simpler strategy is the usage of the Fourier domain representations of the operators. Starting from an isotropic band-pass filter such as the Laplacian-of-Gaussian or Laplacian-of-Poisson, the convolution kernels are generated by Fourier transforming the kernel, applying the operator in the Fourier domain and inverting the Fourier transformed kernel.

5. Decomposition of superimposed intrinsically one-dimensional signals in \mathbb{R}^2

The assumption, that local signal structures can be represented in terms of superimposed intrinsically one-dimensional signals has not only been useful to study the transformation behavior and invariants of the model but additionally allows the analysis of these signals in terms of the partial differential and generalized Hilbert transform operators in an analytic fashion. The analysis of the operator justifies as a useful approach, since it allows a geometric interpretation of the operator responses using algebraic tools.

As it has been shown in the previous chapter, the monogenic signal is able to decompose a signal into its equivariant and invariant parameters with respect to the general linear group, i.e. the orientation and the phase of the signal. Nonetheless, the descriptive power of a single sinusoidal 1D signal is limited to the concept of a line or an edge in the case of image signals in the plane. In order to model more complex image structures it is essential to extend the model in terms of superimposed signals. This chapter deals with the decomposition of such superimposed signals into its orientation and phase components using a symmetric tensor representation consisting of the partial derivatives and generalized Hilbert transform operators. The signals are restricted to images, therefore the domain of definition of the signals in this chapter will always be \mathbb{R}^2 if not otherwise mentioned.

Using a symmetric tensor representation consisting of quadrature filter outputs to encode local signal orientation is not a new concept. Knutsson already proposed a tensor representation for local image structures using quadrature filters in Knutsson (1989). His representation is able to represent the orientation in the case of intrinsically one-dimensional signals and distinguish between intrinsically one- and two-dimensional signals using the quadrature filter responses. A deep analysis of the invariance and equivariance properties of this tensor representation has been provided in Nordberg et al. (1993). Recently Knutsson extended the tensor concept in Knutsson et al. (2011) to tensors of higher order. These tensors are supposed to represent multiple orientations in the case of superimposed intrinsically one-dimensional signals. Nonetheless, no method to obtain the orientation or the phases of the underlying intrinsically one-dimensional signals from the higher-order tensor representation has been provided.

Felsberg studied superpositions of intrinsically one-dimensional signals in terms of the generalized Hilbert transforms in his thesis in year 2002. The result was the *structure multivector*, a decomposition algorithm based on the generalized Hilbert-

Structure
multivector

transforms up to order three, to decompose a superposition of two intrinsically one-dimensional signals into its single component functions. However, the approach was limited to superpositions of intrinsically one-dimensional signals, whose orientations are orthogonal. An extending approach was proposed in Zang (2007). Here, the orthogonality constraint was removed by estimating the orientations of the underlying signals using the generalized structure tensor as introduced in Aach et al. (2006). Nonetheless, the structure tensor uses a local averaging of the operator responses in a local neighborhood introducing a novel parameter, the neighborhood size. Based on the orientations, the single phases of the component functions were estimated.

Signal
multivector

Wietzke and Sommer (2010) succeeded in estimating the single orientations of two superimposed signals without the local averaging of the operator responses, resulting in the *signal multivector*. While this approach does not depend on a local averaging of the operator responses, it is limited to the decomposition of *two* superimposed 1D signals.

In this chapter, a novel method to decompose a superposition of an *arbitrary* number of intrinsically one-dimensional signals in terms of the generalized Hilbert transforms will be proposed. The decompositions consists of two steps: The orientation estimation and the phase estimation of the single 1D signals. While the orientation estimation procedure is not limited to the generalized Hilbert transforms but can also be performed in terms of the partial derivative operators, the phase estimation demands for the generalized Hilbert transforms.

In contrast to the previously mentioned related work, no embedding in a geometric algebra is used explicitly. Instead, the classical tensor notation will be used. We will first establish a correspondence between tensors and homogeneous polynomials.

5.1. From symmetric tensors to homogeneous polynomials

The tensorial notation is greatly simplified by the multi-index notation:

Multi-index

Definition 5.1 (Multi-index). A multiindex α of order n is a n -tuple

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}_0^n \quad (5.1)$$

with $|\alpha| = \sum_{i=1}^n \alpha_i$.

Using the multi-index notation, tensors are defined as follows:

Tensor

Definition 5.2 (Cartesian tensor, Comon et al. (2008)).

Let $\mathbf{u}_1, \dots, \mathbf{u}_m \in \mathbb{R}^n$ denote m vectors in \mathbb{R}^n . Then their *outer product* is defined as

$$\mathbf{U} = [\mathbf{u}_{\alpha_1, \alpha_2, \dots, \alpha_m}]_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n} \quad (5.2)$$

$$= \mathbf{u}_1 \otimes \mathbf{u}_2 \otimes \dots \otimes \mathbf{u}_m \quad (5.3)$$

$$= [\mathbf{u}_{1, \alpha_1} \mathbf{u}_{2, \alpha_2} \dots \mathbf{u}_{m, \alpha_m}]_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n}. \quad (5.4)$$

The quantity $\mathbf{U} \in \mathbb{R}^{n \times n \times \dots \times n}$ is called a n -dimensional *tensor* of order m or just *tensor*.

Strictly speaking, this quantity is a m -way array or m -dimensional hypermatrix, which corresponds to a tensor in the tensor product space $\mathbb{R}^n \otimes \mathbb{R}^n \otimes \dots \otimes \mathbb{R}^n$. With respect to its canonical basis

$$\{\mathbf{e}_{\alpha_1}^{(1)} \otimes \mathbf{e}_{\alpha_2}^{(2)} \otimes \dots \otimes \mathbf{e}_{\alpha_m}^{(m)} \mid 1 \leq \alpha_1 \leq n, 1 \leq \alpha_2 \leq n, \dots, 1 \leq \alpha_m \leq n\} \quad (5.5)$$

there exists an isomorphism Λ between the two spaces, such that

$$\Lambda \left(\sum_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n} \mathbf{u}_{1, \alpha_1} \mathbf{u}_{2, \alpha_2} \dots \mathbf{u}_{m, \alpha_m} \mathbf{e}_{\alpha_1}^{(1)} \otimes \mathbf{e}_{\alpha_2}^{(2)} \otimes \dots \otimes \mathbf{e}_{\alpha_m}^{(m)} \right) = \mathbf{V} \quad (5.6)$$

where $\{\mathbf{e}_{\alpha_1}^{(1)}, \mathbf{e}_{\alpha_2}^{(1)}, \dots, \mathbf{e}_{\alpha_m}^{(1)}\}$ denote the canonical basis vectors in \mathbb{R}^n . This correspondence allows us to speak of arrays like \mathbf{U} as *tensors* (see e.g. Comon et al. (2008)).

Definition 5.3 (Symmetric tensor).

A tensor

Symmetric
tensor

$$\mathbf{U} = [\mathbf{U}_\alpha] = [\mathbf{u}_{1, \alpha_1} \mathbf{u}_{2, \alpha_2} \dots \mathbf{u}_{m, \alpha_m}]_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n} \quad (5.7)$$

is called *symmetric* if

$$\mathbf{U}_{\alpha_{\sigma(1)} \dots \alpha_{\sigma(m)}} = \mathbf{U}_{\alpha_1, \dots, \alpha_m} \quad \text{for } \alpha_1, \dots, \alpha_m \in \{1, \dots, n\} \quad (5.8)$$

holds for all permutations σ from the symmetric group of $\{1, \dots, m\}$.

Proposition 5.4 (Comon and Mourrain (1996)). *The set of homogeneous polynomials of degree m in \mathbb{R}^n is isomorphic to the set of symmetric n -tensors of order m via the isomorphism*

$$P(\mathbf{y}) = \Psi(\mathbf{U}) = \sum_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n} \mathbf{U}_\alpha \mathbf{y}_\alpha. \quad (5.9)$$

Two examples are supposed to illustrate this simple, yet important relationship between symmetric tensors and homogeneous polynomials:

Example 5.5. Consider the symmetric two-dimensional tensor of order two

$$\mathbf{u} = \begin{bmatrix} u_{1,1} & u_{1,2} \\ u_{2,1} & u_{2,2} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \quad (5.10)$$

Then Ψ maps \mathbf{U} to a homogeneous polynomial in two variables of degree two as

$$P(\mathbf{y}) = \Psi(\mathbf{U}) = 2y_1y_1 + y_1y_2 + y_2y_1 + 3y_2y_2 \quad (5.11)$$

$$= 2y_1^2 + 2y_1y_2 + 3y_2^2. \quad (5.12)$$

Example 5.6. Consider the symmetric two-dimensional tensor of order three

$$\mathbf{u} = \left[\begin{array}{cc|cc} u_{1,1,1} & u_{1,2,1} & u_{1,1,2} & u_{1,2,2} \\ u_{2,1,1} & u_{2,2,1} & u_{2,1,2} & u_{2,2,2} \end{array} \right] = \left[\begin{array}{cc|cc} 1 & 2 & 2 & 3 \\ 2 & 3 & 3 & 4 \end{array} \right]. \quad (5.13)$$

Then Ψ maps \mathbf{U} to a homogeneous polynomial in two variables of degree three as

$$P(\mathbf{y}) = \Psi(\mathbf{U}) = y_1y_1y_1 + 2y_1y_2y_1 + 2y_1y_1y_2 + 3y_1y_2y_2 \quad (5.14)$$

$$+ 2y_2y_1y_1 + 3y_2y_2y_1 + 3y_2y_1y_2 + 4y_2y_2y_2 \quad (5.15)$$

$$= y_1^3 + 6y_1^2y_2 + 9y_1y_2^2 + 4y_2^3. \quad (5.16)$$

In the last section, the gradient and the generalized Hilbert transform operators in \mathbb{R}^n have been introduced, to obtain certain signal representations in terms of linear features of the operator responses. Using the notion of tensors, we are now able to introduce compositions of these operators, also called *higher-order operators*, to obtain a tensor valued signal representation for a signal f at every $\mathbf{x} \in \mathbb{R}^n$.

Higher order
derivatives

Given a multi-index α with $1 \leq \alpha_i \leq n$ and α of order m , we define the partial derivative operator with respect to α as

$$\mathcal{D}_\alpha = \mathcal{D}_{\alpha_1} \mathcal{D}_{\alpha_2} \cdots \mathcal{D}_{\alpha_m} = \frac{\partial^m}{\partial x_{\alpha_1} \partial x_{\alpha_2} \cdots \partial x_{\alpha_m}} \quad (5.17)$$

and analogously the generalized Hilbert transform operator with respect to α as

Higher order
generalized
Hilbert
transform

$$\mathcal{H}_\alpha = \mathcal{H}_{\alpha_1} \mathcal{H}_{\alpha_2} \dots \mathcal{H}_{\alpha_m}. \quad (5.18)$$

Figure 5.1 shows the multiscale generalized Hilbert transform kernels up to order 5 for a certain scale using a Laplacian-of-Gaussian kernel regularization kernel.

Example 5.7. Consider the case of two variables, i.e. $n = 2$ and order 2. Let $\alpha = (1, 1)$. Then the corresponding operator reads.

$$\mathcal{D}_\alpha = \frac{\partial^2}{\partial x_1 \partial x_1} = \frac{\partial^2}{\partial x_1^2} \quad (5.19)$$

Example 5.8. Consider the partial differential operator of order $m = 4$ in three variables, i.e. $n = 3$:

$$\mathcal{D}_\alpha = \frac{\partial^4}{\partial x_1 \partial x_2 \partial x_1 \partial x_3}. \quad (5.20)$$

Then the corresponding multi-index α is $\alpha = (1, 2, 1, 3)$.

Consider now the vector valued operators

$$\mathcal{D} = (\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_n)^\top \quad \text{and} \quad \mathcal{H} = (\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_n)^\top \quad (5.21)$$

which correspond to the gradient and the generalized Hilbert transform in \mathbb{R}^n respectively. Then we define the m -th order *outer product* with respect to composition as the multiplication operation as

$$\mathcal{D}^m = \underbrace{\mathcal{D} \otimes \mathcal{D} \otimes \dots \otimes \mathcal{D}}_{m\text{-times}} = [\mathcal{D}_\alpha] = [\mathcal{D}_{(\alpha_1, \alpha_2, \dots, \alpha_m)}]_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n} \quad (5.22)$$

and

$$\mathcal{H}^m = \underbrace{\mathcal{H} \otimes \mathcal{H} \otimes \dots \otimes \mathcal{H}}_{m\text{-times}} = [\mathcal{H}_\alpha] = [\mathcal{H}_{(\alpha_1, \alpha_2, \dots, \alpha_m)}]_{\alpha_1, \alpha_2, \dots, \alpha_m=1}^{n, n, \dots, n}. \quad (5.23)$$

Since the partial differential operators commute, any permutation of the indices of a multi-index α describes the same operator \mathcal{D}_α . The same holds for the operator \mathcal{H}^m . Hence, \mathcal{D}^m and \mathcal{H}^m are operator valued symmetric n -dimensional tensors of order m . The application of \mathcal{D}^m or \mathcal{H}^m to a signal f results in a *tensor field*, attaching a tensor to each $x \in \mathbb{R}^n$ as

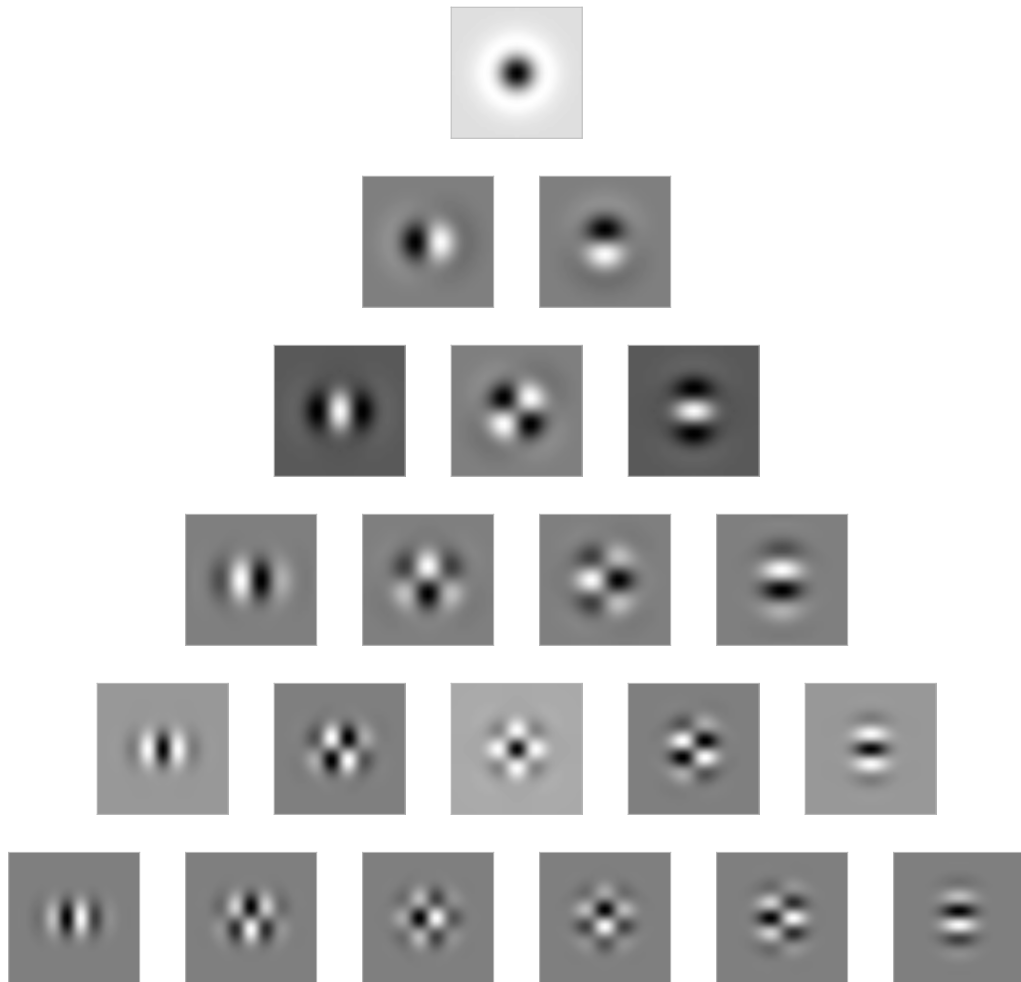


Figure 5.1.: Generalized Hilbert transform kernels up to order 5 of an isotropic wavelet with two vanishing moments (here: Laplacian-of-Gaussian, top row).

$$\mathcal{D}^m f(\mathbf{x}) = [\mathcal{D}_\alpha f(\mathbf{x})] \quad \text{or} \quad \mathcal{H}^m f(\mathbf{x}) = [\mathcal{H}_\alpha f(\mathbf{x})]. \quad (5.24)$$

Example 5.9. Let f be a signal in \mathbb{R}^2 . Then

$$\mathcal{D}^2 f(\mathbf{x}) = ((\mathcal{D}_1, \mathcal{D}_2)^\top \otimes (\mathcal{D}_1, \mathcal{D}_2)^\top) f(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(\mathbf{x}) & \frac{\partial^2}{\partial x_1 \partial x_2} f(\mathbf{x}) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(\mathbf{x}) & \frac{\partial^2}{\partial x_2 \partial x_2} f(\mathbf{x}) \end{bmatrix} \quad (5.25)$$

assigns a symmetric 2-dimensional tensor of order 2 to every $\mathbf{x} \in \mathbb{R}^2$ which is just the Hessian of f at \mathbf{x} .

The tensors resulting from the application of the operators $\mathcal{D}^m, \mathcal{H}^m$ exhibit a particular interesting structure, if the underlying signal is an intrinsically one-dimensional signal:

Lemma 5.10. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a (sinusoidal) intrinsically one-dimensional signal with $f(\mathbf{x}) = \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)$. Then the application of the operators \mathcal{D}^m and \mathcal{H}^m yield*

$$\underbrace{(\mathcal{D} \otimes \mathcal{D} \otimes \cdots \otimes \mathcal{D})}_{m\text{-times}} f(\mathbf{x}) = \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \cdots \otimes \mathbf{n})}_{m\text{-times}} k^m \underline{\mathcal{D}}^m \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.26)$$

$$\underbrace{(\mathcal{H} \otimes \mathcal{H} \otimes \cdots \otimes \mathcal{H})}_{m\text{-times}} f(\mathbf{x}) = \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \cdots \otimes \mathbf{n})}_{m\text{-times}} \underline{\mathcal{H}}^m \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.27)$$

where $\underline{\mathcal{D}}^m$ and $\underline{\mathcal{H}}^m$ denote the m -th order one-dimensional derivative and Hilbert transform applied to \tilde{f} such that

$$\underline{\mathcal{D}}^m \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) = e^{v_{k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi}} \underbrace{\underline{\mathcal{D}} \circ \underline{\mathcal{D}} \circ \cdots \circ \underline{\mathcal{D}}}_{m\text{-times}} \tilde{f} \quad (5.28)$$

and

$$\underline{\mathcal{H}}^m \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) = e^{v_{k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi}} \underbrace{\underline{\mathcal{H}} \circ \underline{\mathcal{H}} \circ \cdots \circ \underline{\mathcal{H}}}_{m\text{-times}} \tilde{f}. \quad (5.29)$$

Proof. It is sufficient to prove the result for one of the two operators since the situation

is analogous. The proof is obtained by induction over m . Suppose $m = 1$. Then

$$\mathcal{D}f(\mathbf{x}) \stackrel{4.18}{=} \mathbf{n} k \underline{\mathcal{D}}\tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.30)$$

Now suppose the assumption holds for m . It is left to show that

$$\underbrace{(\mathcal{D} \otimes \mathcal{D} \otimes \dots \otimes \mathcal{D})}_{(m+1)\text{-times}} f(\mathbf{x}) = \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \dots \otimes \mathbf{n})}_{(m+1)\text{-times}} k^{(m+1)} \underline{\mathcal{D}}^{(m+1)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.31)$$

Let $\mathcal{D}^m = \underbrace{(\mathcal{D} \otimes \mathcal{D} \otimes \dots \otimes \mathcal{D})}_{m\text{-times}}$. Then for $i \in \{1, \dots, n\}$

$$\mathcal{D}_i \mathcal{D}^m f(\mathbf{x}) = \mathcal{D}_i \left(\underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \dots \otimes \mathbf{n})}_{m\text{-times}} k^m \underline{\mathcal{D}}^{(m)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \right) \quad (5.32)$$

$$= n_i \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \dots \otimes \mathbf{n})}_{m\text{-times}} k^{(m+1)} \underline{\mathcal{D}}^{(m+1)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.33)$$

Consequently,

$$\mathcal{D} \otimes \mathcal{D}^m f(\mathbf{x}) = \mathbf{n} \otimes \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \dots \otimes \mathbf{n})}_{m\text{-times}} k^{(m+1)} \underline{\mathcal{D}}^{(m+1)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.34)$$

$$= \underbrace{(\mathbf{n} \otimes \mathbf{n} \otimes \dots \otimes \mathbf{n})}_{(m+1)\text{-times}} k^{(m+1)} \underline{\mathcal{D}}^{(m+1)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.35)$$

□

Thus, the symmetric tensor of order m at \mathbf{x} resulting from the operators \mathcal{D}^m or \mathcal{H}^m is just the m -th order outer product of the normal vector \mathbf{n} times a scalar.

Proposition 5.4 states, that the n -dimensional tensors of order m are isomorphic to the homogeneous polynomials of degree m in n variables. For symmetric tensors, which arise as the m -th order outer product of a single vector in \mathbb{R}^n , the corresponding homogeneous polynomials reveal a special structure:

Lemma 5.11. *Let $\mathbf{v} \in \mathbb{R}^n$ and denote by Ψ the isomorphism from proposition 5.4. Then the m -th order outer product of \mathbf{v} with itself results in the polynomial*

$$\Psi \left(\underbrace{\mathbf{v} \otimes \mathbf{v} \otimes \dots \otimes \mathbf{v}}_{m\text{-times}} \right) = (v_1 y_1 + v_2 y_2 + \dots + v_n y_n)^m. \quad (5.36)$$

Proof. By induction over m . Suppose $m = 1$. Then $\Psi(\mathbf{v}) = v_1 y_1 + v_2 y_2$. Now suppose the assumption holds for m . Then for $i \in \{1, \dots, n\}$

$$\Psi(\underbrace{\mathbf{v} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{v}}_{m\text{-times}}) = v_i y_i (v_1 y_1 + v_2 y_2 + \cdots + v_n y_n)^m \quad (5.37)$$

such that

$$\Psi(\underbrace{\mathbf{v} \otimes \mathbf{v} \otimes \cdots \otimes \mathbf{v}}_{(m+1)\text{-times}}) = \sum_{i=1}^n v_i y_i (v_1 y_1 + v_2 y_2 + \cdots + v_n y_n)^m \quad (5.38)$$

$$= (v_1 y_1 + v_2 y_2 + \cdots + v_n y_n)(v_1 y_1 + v_2 y_2 + \cdots + v_n y_n)^m \quad (5.39)$$

$$= (v_1 y_1 + v_2 y_2 + \cdots + v_n y_n)^{(m+1)}. \quad (5.40)$$

□

The resulting homogeneous polynomial corresponding to the m -th order outer product factors into m linear polynomials, which are homogeneous polynomials of degree 1. With the preceding two lemmas, the tensor resulting from the m -th order operator \mathcal{D}^m or \mathcal{H}^m applied to an intrinsically one-dimensional signal f at a point $\mathbf{x} \in \mathbb{R}^n$ is related to a corresponding homogeneous polynomial, which factors into m linear polynomials, as follows:

Theorem 5.12.

$$\Psi(\mathcal{D}^m f(\mathbf{x})) = (n_1 y_1 + n_2 y_2 + \cdots + n_n y_n)^m k^m \underline{\mathcal{D}}^{(m)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.41)$$

$$\Psi(\mathcal{H}^m f(\mathbf{x})) = (n_1 y_1 + n_2 y_2 + \cdots + n_n y_n)^m \underline{\mathcal{H}}^{(m)} \tilde{f}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.42)$$

Up to now, the m -th order operators applied to a *single* intrinsically one-dimensional signal in \mathbb{R}^n did not provide more information about its orientation parameter \mathbf{n} than the ordinary first order operators \mathcal{D}, \mathcal{H} . Nonetheless, the character of the problem changes, if *superpositions* of intrinsically one-dimensional signals are considered:

Theorem 5.13. Let f_1, \dots, f_d be d (sinusoidal) intrinsically one-dimensional signals with $f_i(\mathbf{x}) = \tilde{f}_i(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)$.

$$\Psi \left(\mathcal{D}^m \sum_{i=1}^d f_i(\mathbf{x}) \right) \quad (5.43)$$

$$= \sum_{i=1}^d (\mathbf{n}_{i,1}y_1 + \mathbf{n}_{i,2}y_2 + \cdots + \mathbf{n}_{i,n}y_n)^m k_i^m \underline{\mathcal{D}}^{(m)} \tilde{f}_i(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i) \quad (5.44)$$

$$\Psi \left(\mathcal{H}^m \sum_{i=1}^d f_i(\mathbf{x}) \right) \quad (5.45)$$

$$= \sum_{i=1}^d (\mathbf{n}_{i,1}y_1 + \mathbf{n}_{i,2}y_2 + \cdots + \mathbf{n}_{i,n}y_n)^m \underline{\mathcal{H}}^{(m)} \tilde{f}_i(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i). \quad (5.46)$$

Proof. The statement immediately follows from theorem 5.12 and the linearity of the operators. \square

5.2. Symmetric tensor decomposition

As already mentioned in the introduction of this chapter, the goal is the decomposition of the a superimposed signal of arbitrary order into the orientation and phases of the underlying single intrinsically one-dimensional signals. As a first step, the orientations \mathbf{n}_i will be obtained. From the previous theorem 5.13 it follows, that this problem can now be stated as the following symmetric tensor decomposition problem according to Brachat et al. (2010) and Kolda and Bader (2009):

Given a superposition of d superimposed intrinsically one-dimensional signals in \mathbb{R}^n , find $m \in \mathbb{N}$, $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_d \in \mathbb{S}^{(n-1)}$ and $\lambda_1, \lambda_2, \dots, \lambda_d \in \mathbb{R}$ such that the symmetric tensor $\mathcal{A}^m f(\mathbf{x})$ with $\mathcal{A}^m \in \{\mathcal{D}^m, \mathcal{H}^m\}$ decomposes into the sum of d symmetric tensors as

$$\mathcal{A}^m f(\mathbf{x}) = \sum_{i=1}^d \lambda_i \overbrace{(\mathbf{n}_i \otimes \mathbf{n}_i \otimes \cdots \otimes \mathbf{n}_i)}^{m\text{-times}}. \quad (5.47)$$

For arbitrary dimensions, i.e. an arbitrary number of variables, this problem is in general not directly solvable and NP-hard (Hillar and Lim (2009)). Nonetheless, as we are dealing with image processing problems and therefore signals in \mathbb{R}^2 we will concentrate in the following on the two-variable case. For the two variable case, an algorithm exists for the decomposition which has already been found by Sylvester in 1886:

Theorem 5.14 (Sylvester (1886)). Let $P(\mathbf{y}) = \sum_{i=0}^m \binom{m}{i} c_i y_1^{m-i} y_2^i$ denote a binary form. Then $P(\mathbf{y})$ can be written as a sum of m -th powers of d distinct linear binary forms as

$$P(\mathbf{y}) = \sum_{i=1}^d \lambda_i (n_{i,1} y_1 + n_{i,2} y_2)^m \quad (5.48)$$

with $\|\mathbf{n}_i\| = 1$ and $\lambda_i \in \mathbb{R}$ if and only if

$$\begin{bmatrix} c_0 & c_1 & \cdots & c_d \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-d} & \cdots & c_{m-1} & c_m \end{bmatrix} \begin{pmatrix} p_0 \\ \vdots \\ p_d \end{pmatrix} = \mathbf{C} \mathbf{p} = \mathbf{0} \quad (5.49)$$

and the polynomial

$$Q(\mathbf{y}) = \sum_{i=0}^d p_i y_1^{d-i} y_2^i \quad (5.50)$$

admits d distinct real roots such that

$$Q(\mathbf{y}) = \prod_{i=1}^d (-n_{i,2} y_1 + n_{i,1} y_2). \quad (5.51)$$

Proof. For the original proof of the theorem in French language we refer the reader to Sylvester (1886). An alternative and more recent version in English can be found in Comon and Mourrain (1996). \square

In which cases does the system $\mathbf{C} \mathbf{p} = \mathbf{0}$ admit a unique nontrivial solution up to a scalar?

Proposition 5.15. For a fixed $d \in \mathbb{N}$, the above system $\mathbf{C} \mathbf{p} = \mathbf{0}$ admits a unique nontrivial solution up to a scalar, if \mathbf{C} is of full rank and $m = 2d - 1$.

Proof. $m - d + 1$ linear independent vectors in \mathbb{R}^{d+1} define a d -dimensional hyperplane if $m - d + 1 = d$. Hence, \mathbf{C} has to be a $d \times (d + 1)$ matrix and $m = 2d - 1$. Since m corresponds to the order of the operators \mathcal{D}^m or \mathcal{H}^m , it follows that a decomposition of the corresponding homogeneous polynomial into a sum of d homogeneous polynomials of degree m always requires operators of order $2d - 1$, they are always of *odd* order. \square

Suppose now that given a binary form $P(\mathbf{y}) = \sum_{i=0}^m \binom{m}{i} c_i y_1^{m-i} y_2^i$, the corresponding system $\mathbf{Cp} = 0$ has been solved for $\mathbf{p} \in \mathbb{R}^{d+1}$. Obviously \mathbf{p} is the normal vector of a d dimensional subspace in \mathbb{R}^{d+1} . It is left to show how the orientation vectors \mathbf{n}_i can be extracted from \mathbf{p} . We will first interpret the corresponding homogeneous polynomial

$$Q(\mathbf{y}) = \sum_{i=0}^d p_i y_1^{d-i} y_2^i = \prod_{i=1}^d (-n_{i,2} y_1 + n_{i,1} y_2) \quad (5.52)$$

geometrically. Suppose $P(\mathbf{y})$ decomposes into a single linear form i.e. $d = 1$ and $P(\mathbf{y}) = \lambda_1 (n_{1,1} y_1 + n_{1,2} y_2)$. Then the corresponding polynomial $Q(\mathbf{y}) = (-n_{1,2} y_1 + n_{1,1} y_2)$ admits a single homogeneous root given by

$$Y_0 = \{\mathbf{y} : (-n_{1,2} y_1 + n_{1,1} y_2) = 0\} \quad (5.53)$$

which geometrically corresponds to a straight line through the origin perpendicular to \mathbf{n}_1 . Now suppose $P(\mathbf{y})$ decomposes into the sum of two homogeneous polynomials which factor into 3 linear forms. Then the homogeneous roots of the corresponding $Q(\mathbf{y})$ yield

$$Y_0 = \{\mathbf{y} : (-n_{1,2} y_1 + n_{1,1} y_2)(-n_{2,2} y_1 + n_{2,1} y_2) = 0\} \quad (5.54)$$

which corresponds to two straight lines through the origin perpendicular to $\mathbf{n}_1, \mathbf{n}_2$. Thus, for arbitrary d , the distinct real roots of Q are d distinct lines through the origin. Every line in \mathbb{R}^2 is uniquely characterized by its intersection with the horizontal line $L_H = \{\mathbf{y} : \langle (0, 1)^T, \mathbf{y} \rangle = 0\}$ parallel to the y_1 axis. Hence, given \mathbf{p} we may also solve for the roots of the corresponding inhomogeneous polynomial

$$\tilde{Q}(y_1) = Q(y_1, 1) = \sum_{i=0}^d p_i y_1^{d-i} \quad (5.55)$$

which is the restriction of $Q(\mathbf{y})$ to L_H . Then the inhomogeneous roots \tilde{Y}_0 of \tilde{Q} are the y_1 coordinates of the points where the lines perpendicular to \mathbf{n}_i intersect L_H . The inhomogeneous polynomial \tilde{Q} of degree d admits at most d distinct real roots. Figure 5.2 illustrates the homogeneous roots of Q and the inhomogeneous roots of \tilde{Q} geometrically for three superimposed signals of order $d = 1, d = 2$ and $d = 3$. For each of the roots $r_i \in \tilde{Y}_0$ the corresponding angles θ_i of the orientations vectors \mathbf{n}_i is recovered as

$$\theta_i = \begin{cases} \frac{\pi}{2} & \text{if } r_i = 0 \\ \arctan(1/r_i) & \text{else} \end{cases} \quad (5.56)$$

such that the orientation vectors are obtained as

$$\mathbf{n}_i = (\cos(\theta_i), \sin(\theta_i))^T. \quad (5.57)$$

If r_1, \dots, r_k are the real roots of \tilde{Q} with $k < d$ and additionally $\tilde{Q}(0) = 0$ holds, then \tilde{Q} has an additional root at infinity, i.e. $\theta_{k+1} = 0$ and $\mathbf{n}_{k+1} = (1, 0)^T$.

Once the orientations \mathbf{n}_i have been recovered from \mathbf{p} it is possible to solve for the coefficients λ_i of the decomposition (5.47) by solving the system

$$\begin{bmatrix} \mathbf{n}_{1,1}^m & \dots & \mathbf{n}_{d,1}^m \\ \mathbf{n}_{1,1}^{m-1} \mathbf{n}_{1,2} & \dots & \mathbf{n}_{d,1}^{m-1} \mathbf{n}_{d,2} \\ \mathbf{n}_{1,1}^{m-2} \mathbf{n}_{1,2}^2 & \dots & \mathbf{n}_{d,1}^{m-2} \mathbf{n}_{d,2}^2 \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{1,2}^m & \dots & \mathbf{n}_{d,2}^m \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_d \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}. \quad (5.58)$$

Note that unless $d = m - 1$ this system is in general overdetermined, so it has to be solved in the least squares sense as the solution of

$$\arg \min_{(\lambda_1, \dots, \lambda_d)^T} \left\| \begin{bmatrix} \mathbf{n}_{1,1}^m & \dots & \mathbf{n}_{d,1}^m \\ \mathbf{n}_{1,1}^{m-1} \mathbf{n}_{1,2} & \dots & \mathbf{n}_{d,1}^{m-1} \mathbf{n}_{d,2} \\ \mathbf{n}_{1,1}^{m-2} \mathbf{n}_{1,2}^2 & \dots & \mathbf{n}_{d,1}^{m-2} \mathbf{n}_{d,2}^2 \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{1,2}^m & \dots & \mathbf{n}_{d,2}^m \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_d \end{pmatrix} - \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix} \right\|^2. \quad (5.59)$$

5.3. Orientation estimation

In the following we will show, how the algorithm can be applied to decompose a tensor resulting from the application of \mathcal{D}^m or \mathcal{H}^m to a superposition of intrinsically one-dimensional signals in the plane. Suppose $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a superposition of d intrinsically one-dimensional signals in the plane. From proposition 5.15 it follows, that the minimum order m of the required operators for the decomposition of the component functions is $2d - 1$. According to proposition 5.4, the symmetric tensor resulting from the application of $\mathcal{A}^m \in \{\mathcal{D}^m, \mathcal{H}^m\}$ is isomorphic to a homogeneous polynomial of degree m in two variables (a *binary form*)

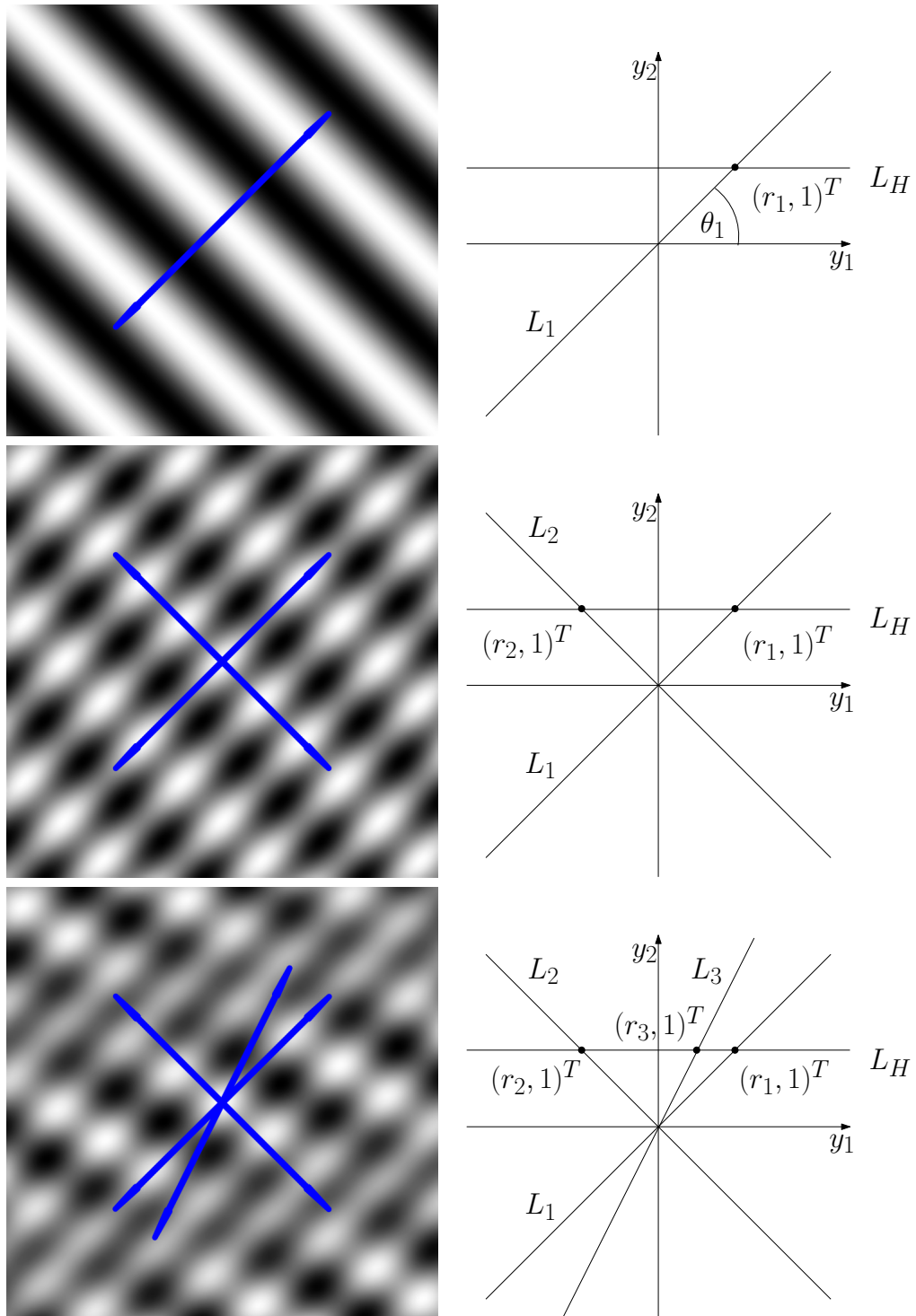


Figure 5.2.: Left column: Superpositions of 1,2 and 3 intrinsically one dimensional signals and the estimated orientations depicted in blue color. Right column: Geometric illustration of the homogenized roots of the polynomials Q and the roots of the inhomogeneous polynomials \tilde{Q} for each signal.

$$\Psi(\mathcal{A}^m f(\mathbf{u})) = \sum_{i=0}^m \binom{m}{i} c_i x^{m-i} y^i \quad (5.60)$$

with $c_i = \mathcal{A}_1^{m-i} \mathcal{A}_2^i f(\mathbf{u})$. From the $m + 1$ coefficients c_0, \dots, c_m of the m -th order operators, the matrix \mathbf{C} according to theorem 5.14 can be constructed as

$$\mathbf{C} = \begin{bmatrix} \frac{\partial^m}{\partial x_1^m} f(\mathbf{x}) & \frac{\partial^m}{\partial x_1^{m-1} \partial x_2} f(\mathbf{x}) & \dots & \frac{\partial^m}{\partial x_1^{m-d} \partial x_2^d} f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^m}{\partial x_1^{m-d+1} \partial x_2^{d-1}} f(\mathbf{x}) & \frac{\partial^m}{\partial x_1^{m-d} \partial x_2^d} f(\mathbf{x}) & \dots & \frac{\partial^m}{\partial x_2^m} f(\mathbf{x}) \end{bmatrix} \quad (5.61)$$

or

$$\mathbf{C} = \begin{bmatrix} \mathcal{H}_1^m f(\mathbf{x}) & \mathcal{H}_1^{m-1} \mathcal{H}_2^1 f(\mathbf{x}) & \dots & \mathcal{H}_1^{m-d} \mathcal{H}_2^d f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{H}_1^{m-d+1} \mathcal{H}_2^{d-1} f(\mathbf{x}) & \mathcal{H}_1^{m-d} \mathcal{H}_2^d f(\mathbf{x}) & \dots & \mathcal{H}_2^m f(\mathbf{x}) \end{bmatrix}. \quad (5.62)$$

The vector \mathbf{p} is then obtained by minimizing the system

$$\arg \min_{\mathbf{p}} \|\mathbf{C} \mathbf{p}\|^2 \quad (5.63)$$

$$\text{s.t. } \|\mathbf{p}\| = 1 \quad (5.64)$$

whose solution is the right singular vector of \mathbf{C} corresponding to the smallest singular value. Using \mathbf{p} , the polynomial \tilde{Q} in one variable is constructed according to (5.55) and solved for its roots. If the roots are real, the angles and consequently the orientations \mathbf{n}_i are obtained by (5.56) and (5.57).

It is worth noting, that in the case of the orientation estimation the differential operators and the generalized Hilbert transform operators yield the same results:

Proposition 5.16. *For superimposed intrinsically one-dimensional signals, both, the differential and generalized Hilbert transform operators yield the same orientations of the signals, using the described estimation procedure.*

5.4. Phase estimation

While both, the partial derivatives and the generalized Hilbert transform, yield the same orientation information \mathbf{n}_i , the nature of the decomposition coefficients λ_i is

of different quality. This difference is especially evident, if the superimposed signal consists of sinusoidal intrinsically one-dimensional signals, i.e.

$$f(\mathbf{x}) = \sum_{i=1}^d A_i \cos(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i). \quad (5.65)$$

In the case of a signal consisting of a single sinusoidal intrinsically one-dimensional signal, the monogenic signal (compare definition 4.7) did not only reveal the orientation information of the plane wave, but also the phase φ of the plane wave. As the phase contains structural information about the underlying signal and is invariant with respect to actions of the general linear group in the plane, it is desirable to extend the phase estimation concept to an arbitrary number of superimposed signals. From lemma 5.10 we know, that the coefficients λ_i of the decomposition (5.47) resulting from the generalized Hilbert transform of order $m = 2d - 1$ yields

$$\lambda_i = \begin{cases} (-1)^{\lceil \frac{m}{2} \rceil} A_i \cos(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i) & \text{if } m \text{ even} \\ (-1)^{\lceil \frac{m}{2} \rceil} A_i \sin(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i) & \text{if } m \text{ odd} \end{cases}. \quad (5.66)$$

Depending on the order m , the coefficients therefore either yield a sine or a cosine wave evaluated at the current point of interest, which can be identified with the origin. A phase estimation would be possible, if the sinusoidal intrinsically one-dimensional signals in quadrature phase relationship with the λ_i could be obtained. Since the orientations \mathbf{n}_i are already known, this is indeed possible using the generalized Hilbert transforms of order $m - 1 = 2d - 2$. Comparing the two decompositions using the operators of order m and order $m - 1$ given by

$$\Psi(\mathcal{H}^m f(\mathbf{x})) = (-1)^{\lceil \frac{m}{2} \rceil} \sum_{i=1}^d \underbrace{A_i \sin(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)}_{\lambda_i} \underbrace{(n_{i,1} y_1 + n_{i,2} y_2)^m}_{\Psi(\otimes_{j=1}^m \mathbf{n}_i)} \quad (5.67)$$

$$\Psi(\mathcal{H}^{m-1} f(\mathbf{x})) = (-1)^{\lceil \frac{m}{2} \rceil} \sum_{i=1}^d \underbrace{A_i \cos(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)}_{\bar{\lambda}_i} \underbrace{(n_{i,1} y_1 + n_{i,2} y_2)^{m-1}}_{\Psi(\otimes_{j=1}^{m-1} \mathbf{n}_i)} \quad (5.68)$$

we notice that the coefficients λ_i and $\bar{\lambda}_i$ correspond to two sinusoidal intrinsically one-dimensional signals in quadrature phase relationship evaluated at \mathbf{x} . Remember that the coefficients λ_i were obtained by solving the $m \times d = 2d - 1 \times d$ system

$$\begin{bmatrix} \mathbf{n}_{1,1}^m & \dots & \mathbf{n}_{d,1}^m \\ \mathbf{n}_{1,1}^{m-1}\mathbf{n}_{1,2} & \dots & \mathbf{n}_{d,1}^{m-1}\mathbf{n}_{d,2} \\ \mathbf{n}_{1,1}^{m-2}\mathbf{n}_{1,2}^2 & \dots & \mathbf{n}_{d,1}^{m-2}\mathbf{n}_{d,2}^2 \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{1,2}^m & \dots & \mathbf{n}_{d,2}^m \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_d \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}. \quad (5.69)$$

with $c_i = \mathcal{H}_1^{m-i}\mathcal{H}_2^i f(\mathbf{x})$ in the least squares sense. Suppose the orientations \mathbf{n}_i of the superimposed signal have been obtained by the procedure described in the previous sections. Using the orientations not only the coefficients λ_i can be obtained, but also the coefficients $\bar{\lambda}_i$ using the operators of degree $m - 1$ by solving the corresponding $m - 1 \times d = 2d - 2 \times d$ system for the coefficients $\bar{\lambda}_i$:

$$\begin{bmatrix} \mathbf{n}_{1,1}^{m-1} & \dots & \mathbf{n}_{d,1}^{m-1} \\ \mathbf{n}_{1,1}^{m-2}\mathbf{n}_{1,2} & \dots & \mathbf{n}_{d,1}^{m-2}\mathbf{n}_{d,2} \\ \mathbf{n}_{1,1}^{m-3}\mathbf{n}_{1,2}^2 & \dots & \mathbf{n}_{d,1}^{m-3}\mathbf{n}_{d,2}^2 \\ \vdots & \ddots & \vdots \\ \mathbf{n}_{1,2}^{m-1} & \dots & \mathbf{n}_{d,2}^{m-1} \end{bmatrix} \begin{pmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \\ \bar{\lambda}_3 \\ \vdots \\ \bar{\lambda}_d \end{pmatrix} = \begin{pmatrix} \bar{c}_0 \\ \bar{c}_1 \\ \bar{c}_2 \\ \vdots \\ \bar{c}_{m-1} \end{pmatrix}. \quad (5.70)$$

with $\bar{c}_i = \mathcal{H}_1^{m-1-i}\mathcal{H}_2^i f(\mathbf{x})$ for $i \in 0, \dots, m - 1$. The operators differ by one degree and m is always of odd order. It follows from (5.66), that the coefficients λ_i and $\bar{\lambda}_i$ are in quadrature phase relationship, evaluated at the origin. This yields the phases as

$$\varphi_i = \arctan 2(\lambda_i, \bar{\lambda}_i). \quad (5.71)$$

The proposed procedure for orientation and phase estimation is supposed to be illustrated in detail by an example.

Example 5.17 (Orientation and phase decomposition of two superimposed 1D signals). Let

$$f(\mathbf{x}) = A_1 \cos(k_1 \langle \mathbf{n}_1, \mathbf{x} \rangle + \varphi_1) + A_2 \cos(k_2 \langle \mathbf{n}_2, \mathbf{x} \rangle + \varphi_2) \quad (5.72)$$

denote a superposition of $d = 2$ real-valued intrinsically one-dimensional signals in \mathbb{R}^2 . Using the generalized Hilbert transform operators and the previously introduced algorithm, the single orientation and phase parameters of the superimposed signal f are supposed to be obtained. According to proposition 5.15, the minimum order m of the operator \mathcal{H}^m allowing for an orientation estimate is $m = 2d - 1 = 3$. The corresponding tensor-valued operators of order $m = 3$ are given by

$$\mathcal{H}^3 = \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} = \left[\begin{array}{cc|cc} \mathcal{H}_1\mathcal{H}_1\mathcal{H}_1 & \mathcal{H}_1\mathcal{H}_2\mathcal{H}_1 & \mathcal{H}_1\mathcal{H}_1\mathcal{H}_2 & \mathcal{H}_1\mathcal{H}_2\mathcal{H}_2 \\ \mathcal{H}_2\mathcal{H}_1\mathcal{H}_1 & \mathcal{H}_2\mathcal{H}_2\mathcal{H}_1 & \mathcal{H}_2\mathcal{H}_1\mathcal{H}_2 & \mathcal{H}_2\mathcal{H}_2\mathcal{H}_2 \end{array} \right]. \quad (5.73)$$

The application of the tensor-valued operators results in the tensor field

$$\begin{aligned} \mathcal{H}^3 f(\mathbf{x}) = & -A_1 \sin(k_1 \langle \mathbf{n}_1, \mathbf{x} \rangle + \varphi_1) \left[\begin{array}{cc|cc} n_{1,1}^3 & n_{1,1}^2 n_{1,2} & n_{1,1}^2 n_{1,2} & n_{1,1} n_{1,2}^2 \\ n_{1,1}^2 n_{1,2} & n_{1,1} n_{1,2}^2 & n_{1,1} n_{1,2}^2 & n_{1,2}^3 \end{array} \right] \\ & -A_2 \sin(k_2 \langle \mathbf{n}_2, \mathbf{x} \rangle + \varphi_2) \left[\begin{array}{cc|cc} n_{2,1}^3 & n_{2,1}^2 n_{2,2} & n_{2,1}^2 n_{2,2} & n_{2,1} n_{2,2}^2 \\ n_{2,1}^2 n_{2,2} & n_{2,1} n_{2,2}^2 & n_{2,1} n_{2,2}^2 & n_{2,2}^3 \end{array} \right]. \end{aligned} \quad (5.74)$$

We seek for decomposition of the corresponding homogeneous polynomial

$$\Psi(\mathcal{H}^3 f(\mathbf{x})) = \sum_{i=1}^2 \underbrace{-A_i \sin(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)}_{\lambda_i} \underbrace{(n_{i,1} y_1 + n_{i,2} y_2)^3}_{\Psi(\mathbf{n}_i \otimes \mathbf{n}_i \otimes \mathbf{n}_i)}. \quad (5.75)$$

Due to the commutation relations of the generalized Hilbert transform operators it follows that

$$\mathcal{H}_1\mathcal{H}_2\mathcal{H}_1 = \mathcal{H}_1\mathcal{H}_1\mathcal{H}_2 = \mathcal{H}_2\mathcal{H}_1\mathcal{H}_1 \quad (5.76)$$

$$\mathcal{H}_1\mathcal{H}_2\mathcal{H}_2 = \mathcal{H}_2\mathcal{H}_2\mathcal{H}_1 = \mathcal{H}_2\mathcal{H}_1\mathcal{H}_2. \quad (5.77)$$

It is therefore sufficient to compute the four operator responses

$$(c_0, c_1, c_2, c_3) = (\mathcal{H}_1\mathcal{H}_1\mathcal{H}_1 f(\mathbf{x}), \mathcal{H}_1\mathcal{H}_1\mathcal{H}_2 f(\mathbf{x}), \mathcal{H}_2\mathcal{H}_2\mathcal{H}_1 f(\mathbf{x}), \mathcal{H}_2\mathcal{H}_2\mathcal{H}_2 f(\mathbf{x})) \quad (5.78)$$

$$= (\mathcal{H}_1^3 f(\mathbf{x}), \mathcal{H}_1^2 \mathcal{H}_2 f(\mathbf{x}), \mathcal{H}_2^2 \mathcal{H}_1 f(\mathbf{x}), \mathcal{H}_2^3 f(\mathbf{x})) \quad (5.79)$$

to construct the Hankel matrix \mathbf{C} as

$$\mathbf{C} = \begin{bmatrix} c_0 & c_1 & c_2 \\ c_1 & c_2 & c_3 \end{bmatrix} = \begin{bmatrix} \mathcal{H}_1^3 f(\mathbf{x}) & \mathcal{H}_1^2 \mathcal{H}_2 f(\mathbf{x}) & \mathcal{H}_1 \mathcal{H}_2^2 f(\mathbf{x}) \\ \mathcal{H}_1^2 \mathcal{H}_2 f(\mathbf{x}) & \mathcal{H}_1 \mathcal{H}_2^2 f(\mathbf{x}) & \mathcal{H}_2^3 f(\mathbf{x}) \end{bmatrix}. \quad (5.80)$$

We find \mathbf{p} as the solution of

$$\arg \min_{\mathbf{p}} \left\| \begin{bmatrix} c_0 & c_1 & c_2 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \end{pmatrix} \right\|^2 \quad (5.81)$$

$$\text{s.t. } \|\mathbf{p}\| = 1 \quad (5.82)$$

which is the right singular vector corresponding to the smallest singular value of the singular value decomposition of \mathbf{C} . The vector \mathbf{p} is used to construct the homogeneous polynomial

$$Q(\mathbf{y}) = p_0 y_1^2 + p_1 y_1 y_2 + p_2 y_2^2 \quad (5.83)$$

whose inhomogeneous version ($y_2 = 1$) reads

$$\tilde{Q}(y_1) = p_0 y_1^2 + p_1 y_1 + p_2. \quad (5.84)$$

The polynomial has at most two distinct real roots where the roots are only equal if the parameters of the two intrinsically one-dimensional component functions coincide. Let r_1, \dots, r_k with $k \leq d$ denote the real roots of \tilde{Q} such that

$$\tilde{Q}(r_i) = 0 \quad \text{for } i \in \{1, \dots, k\}. \quad (5.85)$$

Then the angles θ_i for $i \in \{1, \dots, k\}$ are obtained as

$$\theta_i = \begin{cases} \frac{\pi}{2} & \text{if } r_i = 0 \\ \arctan\left(\frac{1}{r_i}\right) & \text{else.} \end{cases} \quad (5.86)$$

The orientation vectors of the two intrinsically one-dimensional signals are then obtained as

$$\mathbf{n}_i = (\cos(\theta_i), \sin(\theta_i))^T. \quad (5.87)$$

If $k < d$ and additionally $\tilde{Q}(0) = 0$ holds, then \tilde{Q} has an additional root at infinity such that $\theta_{k+1} = 0$ and $\mathbf{n}_{k+1} = (1, 0)^T$.

In addition to the orientations \mathbf{n}_i , the phases φ_i are supposed to be determined. Using the responses of the operators of order $m = 3$, the system

$$\begin{bmatrix} n_{1,1}^3 & n_{2,1}^3 \\ n_{1,1}^2 n_{1,2} & n_{2,1}^2 n_{2,2} \\ n_{1,1} n_{1,2}^2 & n_{2,1} n_{2,2}^2 \\ n_{1,2}^3 & n_{2,2}^3 \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} \quad (5.88)$$

is solved by minimizing

$$\arg \min_{(\lambda_1, \lambda_2)^T} \left\| \begin{bmatrix} n_{1,1}^3 & n_{2,1}^3 \\ n_{1,1}^2 n_{1,2} & n_{2,1}^2 n_{2,2} \\ n_{1,1} n_{1,2}^2 & n_{2,1} n_{2,2}^2 \\ n_{1,2}^3 & n_{2,2}^3 \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} - \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} \right\|^2. \quad (5.89)$$

In addition, the phase estimation requires the operators of order $m - 1 = 2$. We therefore calculate the operator responses $(\bar{c}_0, \bar{c}_1, \bar{c}_2) = (\mathcal{H}_1^2 f(\mathbf{x}), \mathcal{H}_1 \mathcal{H}_2 f(\mathbf{x}), \mathcal{H}_2^2 f(\mathbf{x}))$ which allow a decomposition as

$$\Psi(\mathcal{H}^2 f(\mathbf{x})) = \sum_{i=1}^2 \underbrace{-A_i \cos(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)}_{\bar{\lambda}_i} \underbrace{(\mathbf{n}_{i,1} y_1 + \mathbf{n}_{i,2} y_2)^2}_{\Psi(\mathbf{n}_i \otimes \mathbf{n}_i)}. \quad (5.90)$$

Since the orientations have already been obtained, it is possible to solve for the coefficients $\bar{\lambda}_i$ of the second order decomposition by solving the system

$$\begin{bmatrix} n_{1,1}^2 & n_{2,1}^2 \\ n_{1,1} n_{1,2} & n_{2,1} n_{2,2} \\ n_{1,2}^2 & n_{2,2}^2 \end{bmatrix} \begin{pmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \end{pmatrix} = \begin{pmatrix} \bar{c}_0 \\ \bar{c}_1 \\ \bar{c}_2 \end{pmatrix} \quad (5.91)$$

in the least squares sense as

$$\arg \min_{(\bar{\lambda}_1, \bar{\lambda}_2)^T} \left\| \begin{bmatrix} n_{1,1}^2 & n_{2,1}^2 \\ n_{1,1} n_{1,2} & n_{2,1} n_{2,2} \\ n_{1,2}^2 & n_{2,2}^2 \end{bmatrix} \begin{pmatrix} \bar{\lambda}_1 \\ \bar{\lambda}_2 \end{pmatrix} - \begin{pmatrix} \bar{c}_0 \\ \bar{c}_1 \\ \bar{c}_2 \end{pmatrix} \right\|^2. \quad (5.92)$$

From the coefficients $\lambda_i, \bar{\lambda}_i$ the phases φ_i are obtained as

$$\varphi_i = \arctan 2 \left(\frac{A_i \sin(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)}{A_i \cos(k_i \langle \mathbf{n}_i, \mathbf{x} \rangle + \varphi_i)} \right) = \arctan 2(\lambda_i, \bar{\lambda}_i). \quad (5.93)$$

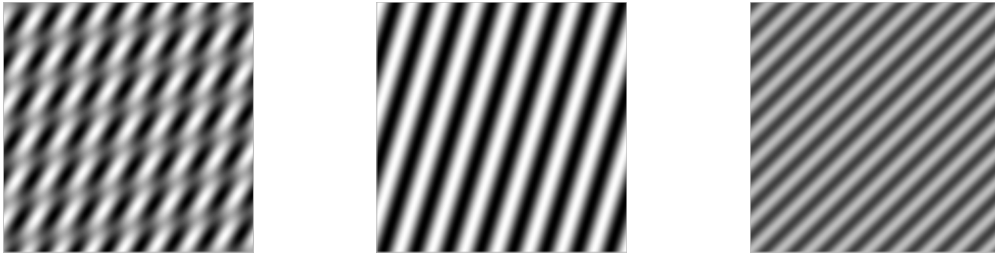


Figure 5.3.: Left: Superposition of the two sinusoidal intrinsically one-dimensional signals in the middle and on the right.

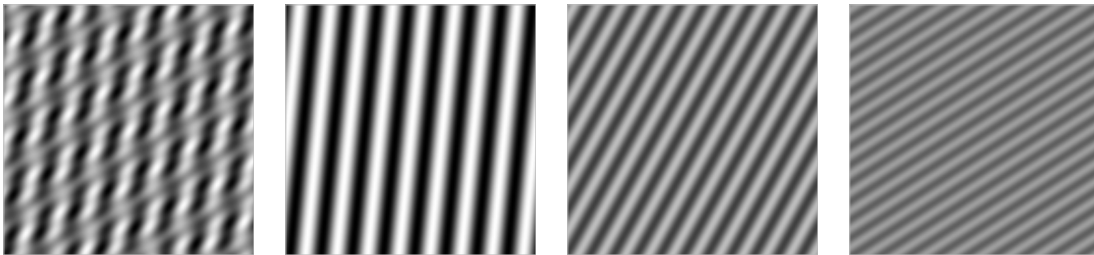


Figure 5.4.: Left: Superposition of the three sinusoidal intrinsically one-dimensional signals in the remaining columns.

5.5. Examples and practical considerations

The complete estimation procedure for the orientations and phases is summarized by the algorithms Algorithm 1, Algorithm 2, Algorithm 3 and Algorithm 4 in Appendix A. Figure 5.3 shows a superposition of two sinusoidal intrinsically one-dimensional signals with different amplitude, orientation and frequency. Figure 5.5 shows the orientation and phase estimation for the underlying component signals. The algorithm has been applied using the ideal operator responses from section 4.2 and the discrete convolution kernels obtained from the generalized Hilbert transforms of order $m = 3$ and $m - 1 = 2$ and a Laplacian of Gaussian at scale $s = 3$. In the case of the phase estimates the border pixels have been set to zero. Figure 5.4 shows a superposition of three sinusoidal intrinsically one-dimensional signals with different amplitude, orientation and frequency. In figure 5.6 the estimated parameters using the perfect operator responses of order $m = 5$ and $m = 4$ are depicted. Figure 5.7 shows the estimated parameters using the discrete convolution kernels at scale $s = 3$.

5.5.1. Resolution of the kernels

The order of the used operators for the parameter estimation depends on the number of superpositions in the signal. If the number of superpositions increases, the order of the operators has to be increased as well. Nonetheless, an increasing order requires

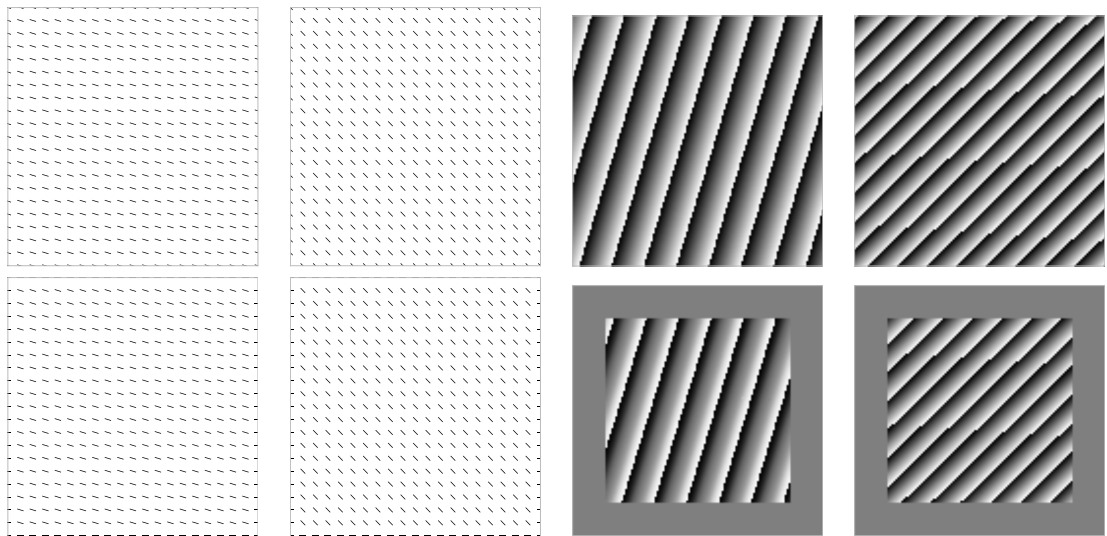


Figure 5.5.: Estimated orientations (first two columns) and phases (last two columns) with respect to the input signal in figure 5.3. Top row: Results using the perfect operator responses. Bottom row: Results using the proposed kernels at scale $s = 3$.

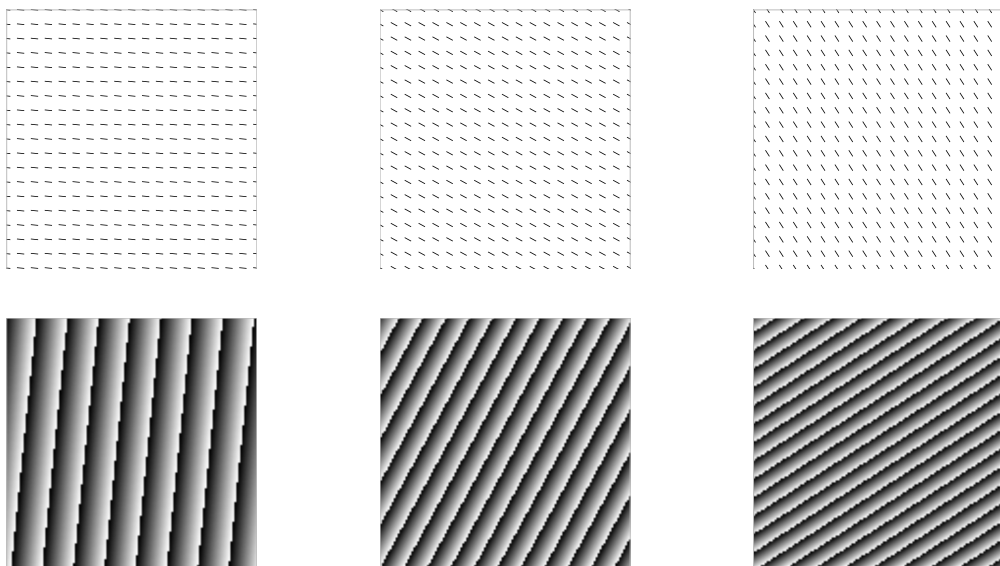


Figure 5.6.: Estimated orientations (top) and phases (bottom) with respect to the input signal in figure 5.4. The algorithm has been applied using the perfect generalized Hilbert transforms.

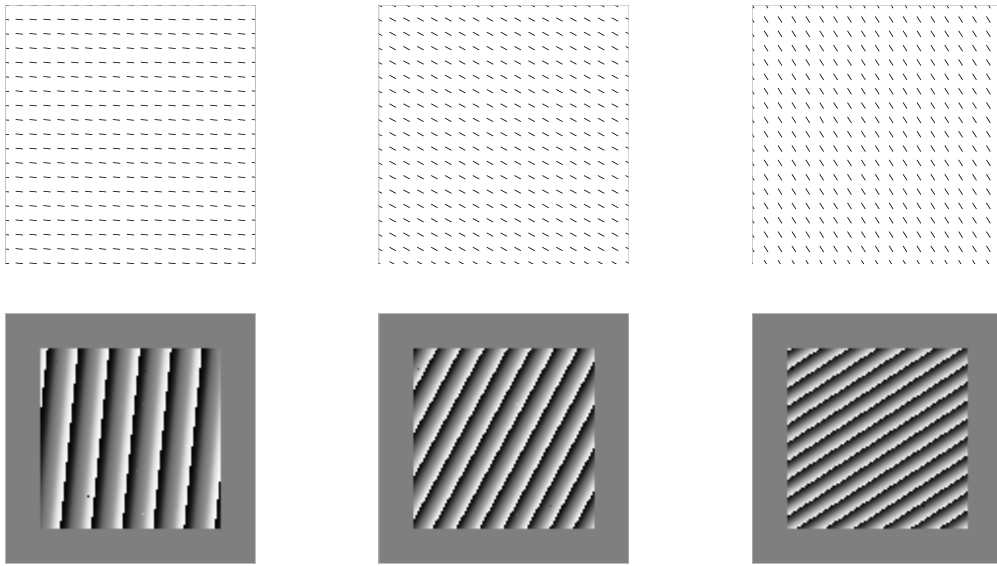


Figure 5.7.: Estimated orientations (top) and phases (bottom) with respect to the input signal in figure 5.4. The algorithm has been applied using the proposed convolution kernels at scale $s = 3$.

an increasing angular resolution of the corresponding convolution kernels to avoid aliasing. If the size of the convolution masks is not large enough, errors are introduced in the estimated orientations and correspondingly the phase estimates will be defective, since they depend on the order. Figure 5.8 illustrates the effect of different convolution mask sizes for a fixed scale with respect to the input signal in figure 5.3.

5.5.2. Choosing the scale

It has been mentioned in section 4.4, that the application of the generalized Hilbert transform in terms of a convolution operator requires an additional filtering with a suitable wavelet with at least one vanishing moment due to the singularity of the Hilbert transform kernel at the origin in the Fourier domain. As a suitable kernel for the task, the isotropic Laplacian-of-Gaussian or Laplacian-of-Poisson kernels have been proposed. These kernels have vanishing moments of order two. While the filtering with these kernels eliminates the singularity, it introduces an additional frequency dependence in the amplitude of the filter responses. Mallat (1999) has pointed out, that every wavelet with scale parameter s and vanishing moments of order d in \mathbb{R}^n acts as a (regularized) differential operator of order d at scale s . According to Unser et al. (2008), the generalized Hilbert transform of order d applied to the Laplacian of Gaussian or Laplacian of Poisson yields wavelets with vanishing moments of order d . Consequently, the convolution with these wavelets and therefore the generalized Hilbert transforms act as differential operators of order d at scale s .

5. DECOMPOSITION OF SUPERIMPOSED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN \mathbb{R}^2

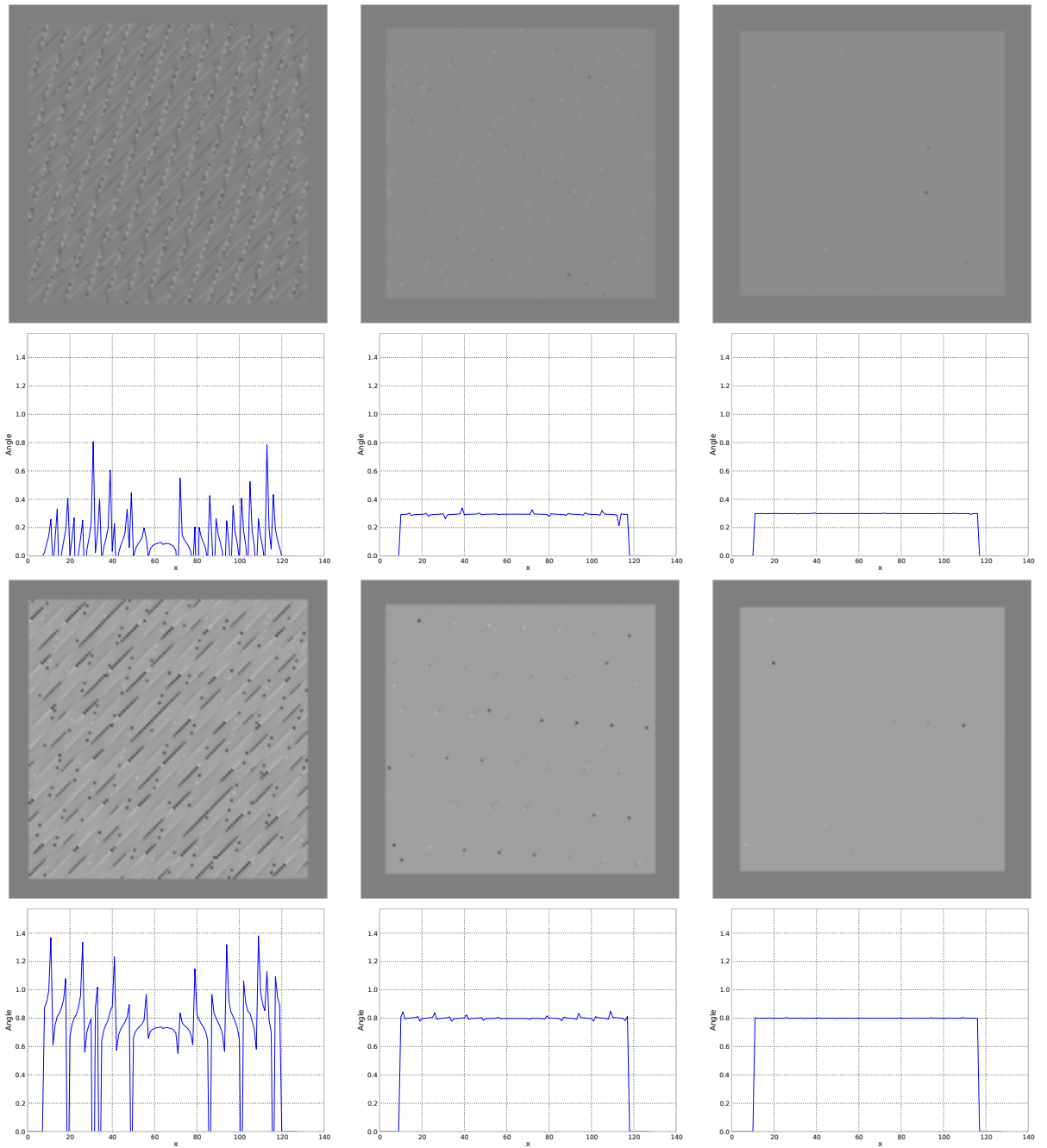


Figure 5.8.: The estimated orientations (the angles are plotted as intensity values) with respect to the input signal depicted in figure 5.3 at scale $s = 2$ for convolution mask sizes 17, 21, 25. Further the slice along $(x_1, 0)^T$ has been plotted explicitly. First two rows: first angle, last two rows: second angle.

In section 4.3 it has been shown, that a partial derivative operator of order d acts on a sinusoidal intrinsically one-dimensional signal as

$$\frac{\partial^d}{\partial x^{d-i} \partial y^i} \mathcal{W}_s A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) = k^d n_1^{d-i} n_2^i \mathcal{W}_s \cos^{(d)}(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.94)$$

for $i \in \{0, \dots, d\}$. The additional frequency factor is k^d . While this factor is not important for the estimation of the orientations, the algorithm for the phase decompositions assumes that no additional frequency factors are introduced, which is the case for the theoretical perfect generalized Hilbert transform. If the original signal is a superimposed signal, each component signal introduces an additional factor for its frequency. Since these frequencies are not known in advance, it is not possible to compensate for these introduced factors. A possible strategy is to assume that interesting structures consist of intrinsically one-dimensional signals, which have roughly the same frequency and consequently the same scale. This scale can for example be found by searching for maxima of the modulus of the Laplacian-of-Gaussian responses. If the underlying component signals do not have the same frequency (scale), errors are introduced in the phase estimates for the single components. Figure 5.9 illustrates the effect of different scales s on the phase estimates of the component signals for a superposition of two intrinsically one-dimensional signals.

Another problem arising, illustrated in Figure 5.10, is the different order of the odd and even filters. Since they both differ in their order, the additional introduced frequency factors differ. It is possible to compensate for this effect by adjusting the amplitudes of the filter kernels, such that the maximal and minimal amplitudes of both orders coincide.

5.5.3. Automatic selection of the order

The algorithm introduced so far relied on à priori knowledge about the number d of the superpositions. If the number d of superpositions is known, the algorithm is able to obtain the orientations and the phases using the generalized Hilbert transform operators of order m and $m - 1$. If the parameters are supposed to describe local low-level image structures in real world images, the number of superpositions is in general unknown. Nonetheless, it is possible to decide, if the number of superpositions is lower than or equal to a chosen maximum number d_{\max} .

Proposition 5.18. *Let $d_{\max} \in \mathbb{N}$ and let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ denote a superimposed signal in the plane consisting of $d \leq d_{\max}$ intrinsically one-dimensional signals. Then the Hankel matrix \mathbf{C} obtained from the operator responses $\mathcal{H}^m f(\mathbf{x})$ or $\mathcal{D}^m f(\mathbf{x})$ with $m = 2d_{\max} - 1$ has rank d .*

Since the singular value decomposition of \mathbf{C} is computed anyway for the decomposition, the additional amount to compute the number of superpositions is minimal as

5. DECOMPOSITION OF SUPERIMPOSED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN \mathbb{R}^2

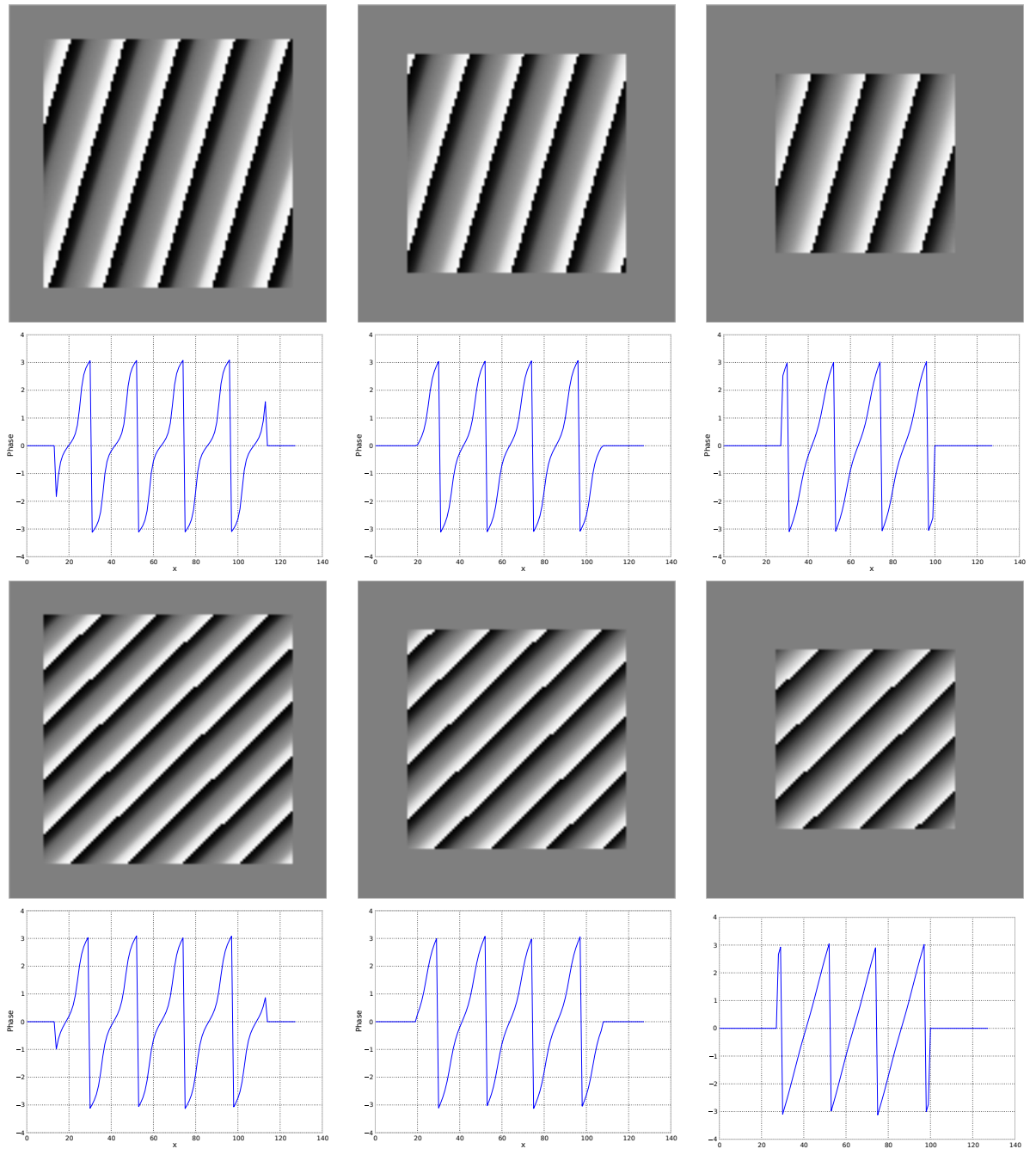


Figure 5.9.: The estimated phases with respect to the input signal 5.3 at scales $s = 2$, $s = 2\sqrt{2}$, $s = 4$. Further the center slice along $(x_1, 0)^T$ has been plotted explicitly. The phase estimates slightly depend on the chosen scale.

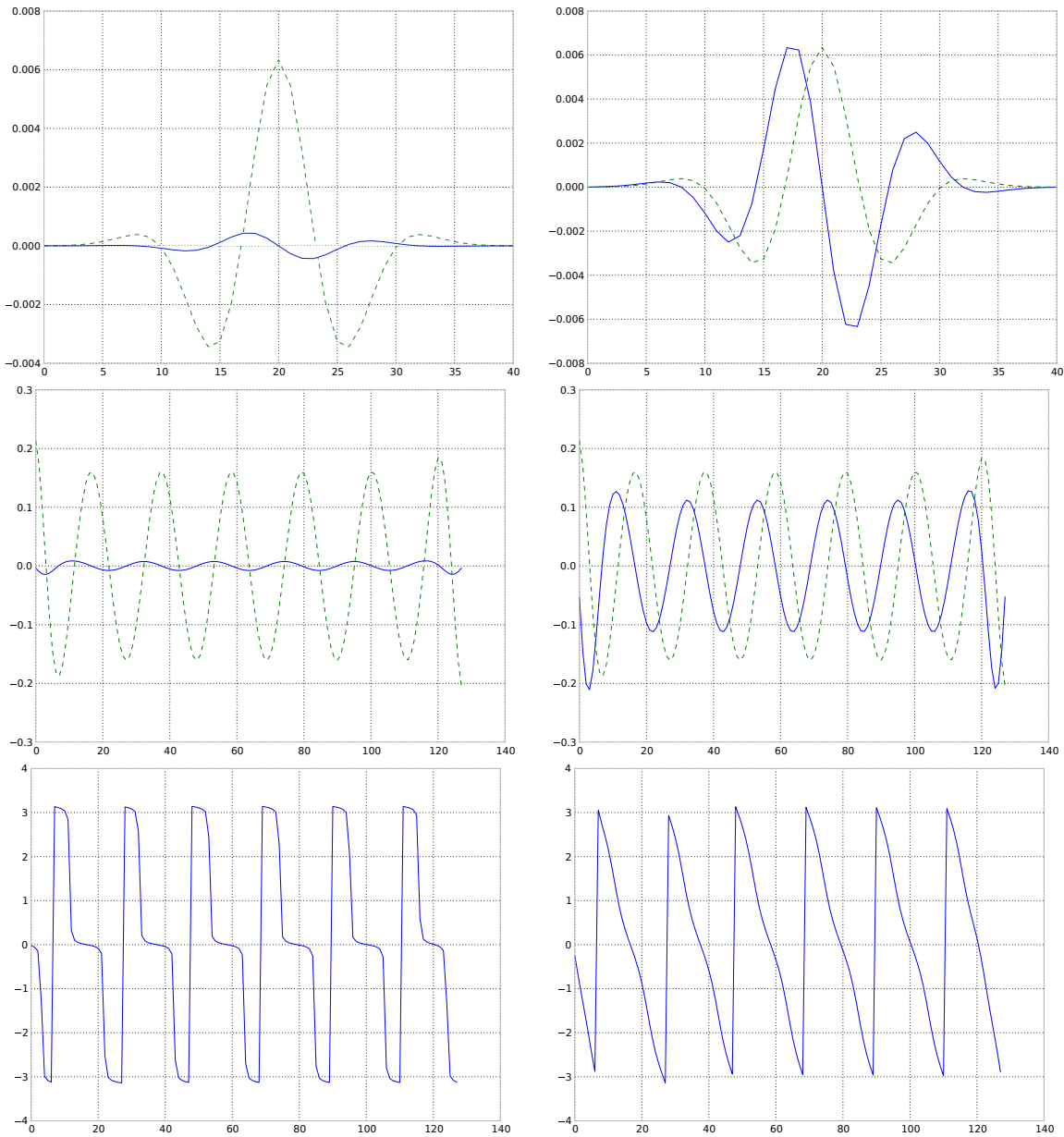


Figure 5.10.: Top row: First kernels of the generalized Hilbert transform of order three (solid) and two (dashed). Left: Without normalization. Right: With normalization. Middle row: Amplitude of one of the decomposed components of a superposition of two i1D signals. Bottom row: Phase estimated for this component.

the rank can be determined from its eigenvalues. In a strict sense, the rank will be equal to $d_{\max} - l$ where l is the number of zero eigenvalues. Working with real world images which are possibly disturbed by noise, matrix \mathbf{C} will usually be of full rank in a strict sense. It is therefore necessary to measure the rank with respect to a tolerance threshold for the eigenvalues which assumes that the eigenvalue is zero if it is lower than the threshold.

5.5.4. Singularities

While the proposed algorithm for the decomposition of the signal is purely local, this property bears a major problem caused by singularities of the operator responses.

Singularity **Definition 5.19.** Let $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ denote a intrinsically one-dimensional signal in the plane. Then a point $\mathbf{x}_0 \in \mathbb{R}^2$ is called a *singularity* of f with respect to the operator $\mathcal{A}^m \in \{\mathcal{D}^m, \mathcal{H}^m\}$ if

$$\mathcal{A}^m f(\mathbf{x}_0) = 0. \quad (5.95)$$

Suppose now that the intrinsically one-dimensional component signal f_j of the superpositional signal f has a singularity at \mathbf{x}_0 with respect to \mathcal{A}^m . From the linearity of the operators \mathcal{A}^m we know that the operators act on the component functions as

$$\mathcal{A}^m f(\mathbf{x}) = \sum_{i=1}^d \mathcal{A}^m f_i(\mathbf{x}). \quad (5.96)$$

Since $\mathcal{A}^m f_j$ is zero at \mathbf{x}_0 , the superpositional signal f locally looks like a superposition of $d - 1$ intrinsically one-dimensional signals, i.e.

$$\mathcal{A}^m f(\mathbf{x}) = \sum_{i=1}^{d-1} \mathcal{A}^m f_{j_i}(\mathbf{x}) \quad \text{with } j_i \in \{1, \dots, d\} \setminus \{j\}. \quad (5.97)$$

At the singularity \mathbf{x}_0 it will not be possible to recover the orientation of the component signal f_j . Consequently, a phase estimate will be impossible.

A point \mathbf{x}_0 can obviously be a singularity of more than one intrinsically one-dimensional signal, such that it is even possible that the operator response for a superpositional signal f is zero, i.e

$$\mathcal{A}^m f(\mathbf{x}_0) = (0, 0, \dots, 0). \quad (5.98)$$

In such a case it would neither be possible to estimate any of the orientations \mathbf{n}_i , nor the phases φ_i at \mathbf{x}_0 . To eliminate the problem of singularities, a method using averages of the operator responses in a neighbourhood of a point of interest will be introduced in the following section.

5.6. Averaging the operator responses

The considerations from the previous section have been purely local in the sense, that the orientation and the phases of the superimposed signal model have been estimated from the operator responses (linear features) evaluated at a single point of interest. This led to two problems: As it has been shown in proposition 5.15, the required order of the operators for d superimposed signals is $2d - 1$. Consequently, the more signals are supposed to be decomposed, the higher the angular resolution of the filter kernels has to be. This also leads to a larger neighborhood which has to be taken into account, if the overall resolution of the input signal is fixed. Further, the complete decomposition is only possible, if none of the underlying intrinsically one-dimensional signals vanishes at the current point of interest, i.e. the current point of interest is not a singularity of any of the underlying intrinsically one-dimensional signals.

The estimation of the single phases requires a prior estimation of the underlying orientations. If the current point of interest is a singularity of one of the underlying intrinsically one-dimensional component functions, the orientation estimation will discard this component, leading to local orientation estimation for a number of signals, which does not reflect the number of component signals in the local neighborhood.

To circumvent this problem we will consider features evaluated in a local neighborhood of a point of interest \mathbf{x} . While this method requires to introduce a further parameter, it also introduces new constraints for the orientation estimation, reducing the order of the required operators, and further also allows the estimation at points which are singularities of the underlying intrinsically one-dimensional signals.

The introduced method is based on an alternative characterization of intrinsically one-dimensional signals. We will focus the investigation on the two-dimensional case such that the considered signals are all real-valued signals (images) in the plane.

In contrast to its original definition (3.3), an intrinsically one-dimensional signal allows for an alternative characterization in terms of first order differential operators. Let

$$f(\mathbf{x}) = \tilde{f}(\langle \mathbf{x}, \mathbf{n} \rangle) \quad \mathbf{n} \in \mathcal{S} \quad (5.99)$$

denote an intrinsically one-dimensional signal. Then the signal f is constant along the lines in \mathbb{R}^2 perpendicular to \mathbf{n} . This implies that the gradient of f vanishes along these lines.

5. DECOMPOSITION OF SUPERIMPOSED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN \mathbb{R}^2

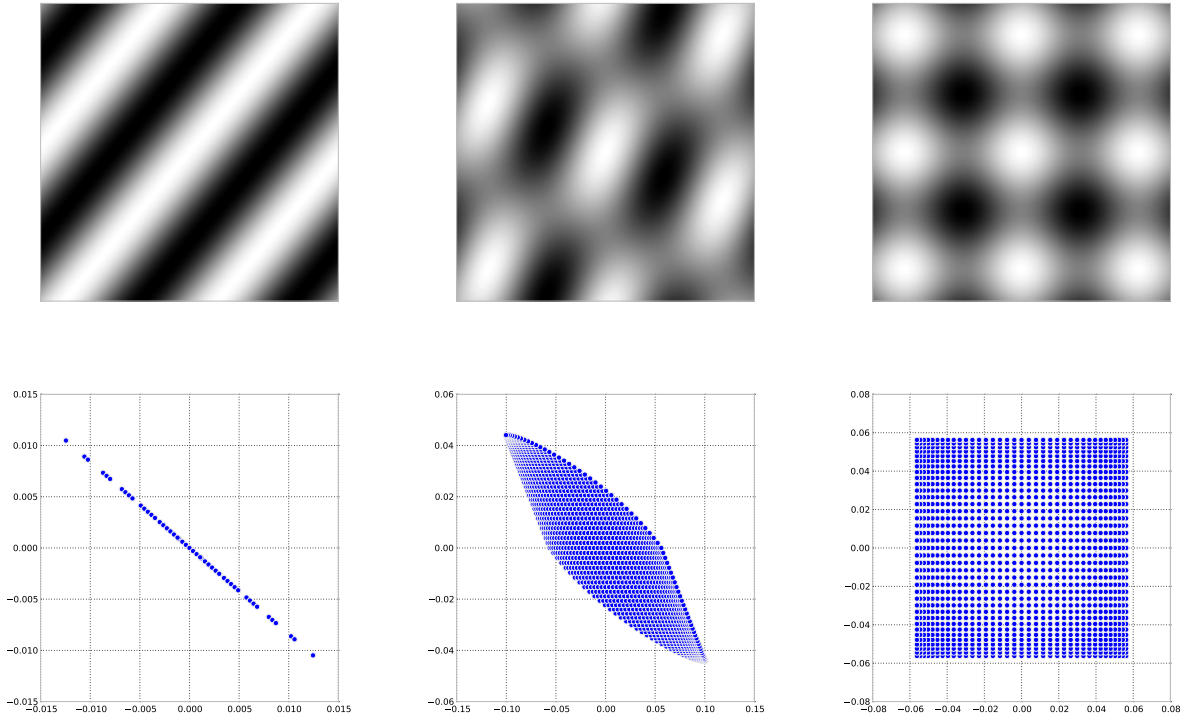


Figure 5.11.: Left column: Intrinsically one-dimensional signal and two superpositions of intrinsically one-dimensional signals with varying orientations. Right column: First order derivatives interpreted as points in \mathbb{R}^2 .

Proposition 5.20 (Differential characterization of i1D signals, Aach et al. (2006)).

$$\mathcal{L}f(\mathbf{x}) = \langle \mathcal{D}f(\mathbf{x}), \bar{\mathbf{n}} \rangle = \langle \nabla f(\mathbf{x}), \bar{\mathbf{n}} \rangle = 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^2 \text{ with } \langle \mathbf{n}, \bar{\mathbf{n}} \rangle = 0. \quad (5.100)$$

Now suppose f is a superposition of d intrinsically one-dimensional functions:

$$f(\mathbf{x}) = \sum_{i=1}^d f_i(\mathbf{x}) = \sum_{i=1}^d \tilde{f}_i(\langle \mathbf{x}, \mathbf{n}_i \rangle). \quad (5.101)$$

Theorem 5.21 (Differential characterization of superimposed signals). *The superimposed signal f of order d vanishes under a composition of the differential operators \mathcal{L}_i with $\mathcal{L}_i f = \langle \nabla f, \bar{\mathbf{n}}_i \rangle$ with $i \in \{1, \dots, d\}$:*

$$\mathcal{L}_1 \mathcal{L}_2 \dots \mathcal{L}_d f(\mathbf{x}) = 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^2. \quad (5.102)$$

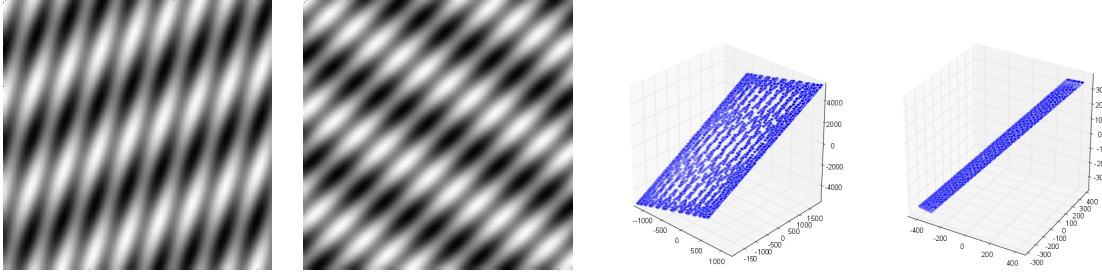


Figure 5.12.: Left: Superpositions of intrinsically one-dimensional signals with varying orientations. Right: Corresponding second order derivatives interpreted as points in \mathbb{R}^3 . Notice that all the points lie in a plane which changes its normal if the orientations of the signals change.

Proof. The differential operators \mathcal{L}_i are first order differential operators. It is possible to identify these operators with vector fields in the plane as

$$V_i(x, y) = \bar{n}_{i,1} \frac{\partial}{\partial x} + \bar{n}_{i,2} \frac{\partial}{\partial y} = -n_{i,2} \frac{\partial}{\partial x} + n_{i,1} \frac{\partial}{\partial y} \quad (5.103)$$

where the partial derivatives $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ represent the basis vectors of the tangent space of \mathbb{R}^2 . Without loss of generality, the vector fields V_i are rotated versions of the canonical Cartesian coordinate vector field $V_x = \frac{\partial}{\partial x}$. This vector field is the infinitesimal generator of group of translation along the x axis. Thus, the vector fields V_i are the infinitesimal generators for the group of translations along \bar{n}_i . The group of translations in the plane is commutative, following from the commutativity of addition within the real numbers. Hence, the application of any two vector fields V_i, V_j and therefore for any two operators $\mathcal{L}_i, \mathcal{L}_j$ with $i, j \in \{1 \dots, d\}$ commutes:

$$\mathcal{L}_i \mathcal{L}_j f = \mathcal{L}_j \mathcal{L}_i f. \quad (5.104)$$

Since the differential operators are linear operators and f is a superposition of intrinsically one-dimensional signals, the composition of the differential operators acts on each intrinsically one-dimensional signal f_i separately as $\mathcal{L}_1 \mathcal{L}_2 \dots \mathcal{L}_d f_i(x)$. Due to the commutation relation of the operators, the order of the operators can be changed arbitrarily, such that

$$\mathcal{L}_1 \mathcal{L}_2 \dots \mathcal{L}_d f_i(x) = \mathcal{L}_{j_1} \dots \mathcal{L}_{j_{n-2}} \mathcal{L}_i f_i(x) \quad j_k \in \{1, \dots, d\} \setminus \{i\}. \quad (5.105)$$

But the operator \mathcal{L}_i applied to f_i is zero everywhere, due to the differential char-

acterization of intrinsically one-dimensional signals in (5.20) which concludes the proof. \square

The composition of the differential operators \mathcal{L}_i yields the single operator

$$\mathcal{L} = \mathcal{L}_1 \mathcal{L}_2 \dots \mathcal{L}_d = \prod_{i=1}^d (\bar{n}_{i,1} \frac{\partial}{\partial x} + \bar{n}_{i,2} \frac{\partial}{\partial y}). \quad (5.106)$$

such that the application of \mathcal{L} on f can be rewritten as a differential operator of order d consisting of the $d + 1$ partial derivatives of order d

$$\mathcal{L}f(\mathbf{x}) = \mathcal{L}_1 \mathcal{L}_d \dots \mathcal{L}_d f(\mathbf{x}) = \sum_{i=0}^d \alpha_i \frac{\partial^d}{\partial x^{d-i} \partial y^i} f(\mathbf{x}) = 0 \quad (5.107)$$

for all $\mathbf{x} \in \mathbb{R}^2$. We want to derive a geometric interpretation for this relationship. Suppose we identify the partial derivatives applied to f with points in \mathbb{R}^{d+1} as

$$\mathbf{v}(\mathbf{x}) = \left(\frac{\partial^d}{\partial x^d} f(\mathbf{x}), \frac{\partial^d}{\partial x^{d-1} \partial y} f(\mathbf{x}), \dots, \frac{\partial^d}{\partial y^d} f(\mathbf{x}) \right)^T. \quad (5.108)$$

Since the operator \mathcal{L} applied to f vanishes everywhere its integration over the plane yields

$$\int_{\mathbb{R}^2} \langle \mathbf{v}(\mathbf{x}), \mathbf{a} \rangle d\mathbf{x} = 0. \quad (5.109)$$

Thus, all the points $\mathbf{v}(\mathbf{x})$ in \mathbb{R}^{d+1} are perpendicular to \mathbf{a} . Consequently, the relationship (5.107) geometrically states that the partial derivatives evaluated at the points $\mathbf{x} \in \mathbb{R}^2$ lie in a d -dimensional linear subspace in \mathbb{R}^{d+1} perpendicular to the vector \mathbf{a} .

Figure 5.12 illustrates this relationship for a superposition of two intrinsically one-dimensional signals in the plane. The operator responses (linear features) of order $d = 2$ interpreted as points in the feature space \mathbb{R}^3 lie in a two-dimensional subspace, a plane through the origin. If the signals change their orientations, the normal vector of the plane changes accordingly.

The coefficients $\mathbf{a} = (\alpha_0, \dots, \alpha_d)^T$ can be identified with the coefficients of the polynomial

$$Q(\mathbf{y}) = \prod_{i=1}^d Q_i(\mathbf{y}) = \prod_{i=1}^d (\bar{n}_{i,1}y_1 + \bar{n}_{i,2}y_2) = \sum_{i=0}^d \alpha_i y_1^{d-i} y_2^i. \quad (5.110)$$

Following the basic idea from the previous section, we will show, how the orientation vectors \mathbf{n}_i can be recovered from the coefficient vector \mathbf{a} . $Q(\mathbf{y})$ is a homogeneous polynomial of degree d , which factors into d linear binary forms Q_i . Of special interest are the roots of the binary form Q . Analogously to (5.52), we will again first consider the homogenized real roots of the linear binary forms Q_i . The homogenized real roots of the forms Q_i are the lines

$$L_i = \{\mathbf{y} : Q_i(\mathbf{y}) = (\bar{n}_{i,1}y_1 + \bar{n}_{i,2}y_2) = 0\} \quad (5.111)$$

perpendicular to $\bar{\mathbf{n}}_i$ passing through the origin (compare Fig. 5.2). Since every line L_i is uniquely determined by its intersection with the horizontal line $L_H = \{(y_1, 1)^T, y_1 \in \mathbb{R}\}$, the homogenized root L_i of Q_i intersects L_H in a real root of the inhomogeneous polynomial $\tilde{Q}(y_1) = Q(y_1, 1)$. Let r_1, \dots, r_k with $k \leq d$ denote the roots of \tilde{Q} such that

$$\tilde{Q}(r_i) = 0 \quad \text{for } i \in \{1, \dots, k\}. \quad (5.112)$$

Then the angles θ_i for $i \in \{1, \dots, k\}$ are obtained as

$$\theta_i = \begin{cases} \frac{\pi}{2} & \text{if } r_i = 0 \\ \arctan\left(\frac{1}{r_i}\right) & \text{else} \end{cases} \quad (5.113)$$

yielding

$$\mathbf{n}_i = (\cos(\theta_i), \sin(\theta_i))^T. \quad (5.114)$$

If $k < d$ and additionally $\tilde{Q}(0) = 0$ holds, then \tilde{Q} has an additional root at infinity such that $\theta_{k+1} = 0$ and $\mathbf{n}_{k+1} = (1, 0)^T$.

Now that we know, how the orientations can be recovered from the coefficient vector \mathbf{a} , it is left to show how \mathbf{a} is estimated from an image I . Suppose that in a local neighborhood $\Omega(\mathbf{x}_0)$ around a point $\mathbf{x}_0 \in \mathbb{R}^2$ the image I can be modeled as a superposition f of d intrinsically one-dimensional signals. As already mentioned above, the coefficient vector \mathbf{a} is perpendicular to the hyperplane in \mathbb{R}^{d+1} spanned by the partial derivatives of order d , interpreted as points in \mathbb{R}^{d+1} . The coefficient vector \mathbf{a} is

then found by a classical least squares fit of a d -dimensional hyperplane to the partial derivatives, leading to the optimization problem

$$\min_{\mathbf{a}} \int_{\Omega(\mathbf{x}_0)} \langle \mathbf{v}(\mathbf{x}), \mathbf{a} \rangle^2 d\mathbf{x} \quad (5.115)$$

$$\text{s.t. } \|\mathbf{a}\| = 1 \quad (5.116)$$

where $\mathbf{v}(\mathbf{x})$ are the d -th order partial derivatives of f at \mathbf{x} . The solution to the least squares problem above is obtained as the eigenvector corresponding to the smallest eigenvalue of the matrix (see e.g. Kanatani (1996))

Generalized
structure tensor

$$\mathbf{S}(\mathbf{x}_0) = \begin{bmatrix} S_{0,0}(\mathbf{x}_0) & \cdots & S_{0,d-1}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ S_{d-1,0}(\mathbf{x}_0) & \cdots & S_{d-1,d-1}(\mathbf{x}_0) \end{bmatrix} \quad (5.117)$$

with

$$S_{ij}(\mathbf{x}_0) = \int_{\Omega(\mathbf{x}_0)} \frac{\partial^d}{\partial x^{d-i} \partial y^i} f(\mathbf{x}) \frac{\partial^d}{\partial x^{d-j} \partial y^j} f(\mathbf{x}) d\mathbf{x}. \quad (5.118)$$

In the case of $d = 1$ the matrix \mathbf{S} corresponds to the classical *structure tensor* introduced in Harris and Stephens (1988) and Förstner and Gülch (1987). \mathbf{S} generalizes the notion of the structure tensor to arbitrary superpositions. In the following we will call \mathbf{S} the *generalized structure tensor* of order d .

5.6.1. Automatic selection of the order

In the case of d superimposed intrinsically one-dimensional signals, the linear features of the d -th order operators span a d -dimensional subspace in \mathbb{R}^{d+1} . In this case the generalized structure tensor is of full rank. In analogue to proposition 5.18 it is not possible to decide, if the signal in the local neighborhood consists of more than d superimposed intrinsically one-dimensional signals using the operators of order d . Nonetheless, it is possible to decide, if the signal consists of *less* than d intrinsically one-dimensional signals. If the signal consists of $c \leq d$, $c \in \mathbb{N}$ superimposed intrinsically one-dimensional signals, the rank of the generalized structure tensor is equal to c . Of course the same remark as in proposition 5.18 holds: Since real world signals are disturbed by noise, the generalized structure tensor will almost always be of full rank.

It is therefore mandatory to determine the rank based on some thresholding on the eigenvalues or based on the ratios of the eigenvalues.

Figure 5.11 depicts this relationship for the case of the classical structure tensor, i.e. the operators of order $d = 1$. In the case of a perfect intrinsical one-dimensional signal, the operator responses lie on a one-dimensional subspace in \mathbb{R}^2 . This is not the case anymore for a superposition of intrinsically one-dimensional signals with different orientations. Nonetheless, it is not possible to infer about the orientations of the single intrinsically one-dimensional signals in the case of a superposition. It is just possible to decide whether the underlying signal consists of one or more intrinsically one-dimensional signals. Otherwise, a structure tensor of higher order has to be used.

5.6.2. Choosing the neighborhood

The generalized structure tensor of order d averages the linear features obtained from the operators of order d in a local neighborhood Ω of each point \mathbf{x} . An open question is, how this neighborhood should be chosen.

The most common choice, introduced by Lindeberg (1998), is a weighted averaging using a Gaussian kernel, such that the neighborhood Ω corresponds to the support of the discretized Gaussian. The operators are usually evaluated at multiple scales, the *differentiation scales*. The scale of the Gaussian used for the averaging of the operator responses, the *integration scale* is then chosen depending on the differentiation scale. Suppose the partial derivatives of order d are calculated at a differentiation scale s_D as

$$\mathcal{D}_1^{d-i} \mathcal{D}_2^i \mathcal{G}_{s_D} f(\mathbf{x}) = (f * \mathcal{D}_1^{d-i} \mathcal{D}_2^i g_{s_D})(\mathbf{x}). \quad (5.119)$$

Then the components of the generalized structure tensor of order d using a Gaussian averaging are obtained at the integration scale $s_I = \alpha s_D$, $\alpha > 1$ as

$$S_{ij}(\mathbf{x}_0; s_D, s_I) = (\mathcal{D}_1^{d-i} \mathcal{D}_2^i \mathcal{G}_{s_D} f * g_{s_I})(\mathbf{x}_0) (\mathcal{D}_1^{d-j} \mathcal{D}_2^j \mathcal{G}_{s_D} f * g_{s_I})(\mathbf{x}_0) \quad (5.120)$$

$$= \left(\frac{\partial^d}{\partial x^{d-i} \partial y^i} \mathcal{G}_{s_I} f * g_{s_I} \right) (\mathbf{x}_0) \left(\frac{\partial^d}{\partial x^{d-j} \partial y^j} \mathcal{G}_{s_I} f * g_{s_I} \right) (\mathbf{x}_0). \quad (5.121)$$

While not necessarily a Gaussian has to be chosen for the weighted averaging, it is preferred in practical situations, due to its separability, which allows a fast convolution.

In contrast to the local method introduced in the previous section, the linear features resulting from the operators are not only considered at the current point of interest, but are averaged over a local neighborhood. Due to the averaging, the orientations can also be obtained at points of interest, which are singularities.

5.7. Application: Characteristic scale of superimposed signals

The detection of low-level image features is often the first step in many computer vision tasks. The major goal is a sparse representation of the image content while being invariant with respect to certain, possibly task dependent, image transformations. Image features can broadly be divided into two groups: Pure *point based features*, also known as key-points or interest points, consisting only of a spatial image coordinate and *regional features* describing a certain neighborhood around an interest point. If regional detectors obey the property of equivariance with respect to certain transformation groups, such that the estimated parameters of the detector transform according to a transformation of the underlying input, it is possible to invert the transformation and normalize the content within the region such that so called *feature descriptors* as for example introduced in Lowe (2004); Bay et al. (2008); Dalal and Triggs (2005), are able to describe the content in an invariant way.

In this section, a model based approach is used to design regional image features which are equivariant with respect to the affine group in the plane. Image structures are locally modeled as superimposed intrinsically one-dimensional signals. Based on the orientation estimation scheme introduced in the previous section, we propose an estimation scheme to obtain the *characteristic scale* of each intrinsically one-dimensional component signal. The orientations and the characteristic scales yield the final affine equivariant parameter set describing the affine equivariant regional feature.

A broad range of interest point detectors is based on the basic idea of the Harris corner detector dating back to 1988, see Harris and Stephens (1988). These detectors use linear shift invariant operators, usually discrete approximations of differential operators, to establish a cost function whose extrema indicate the presence of a certain image structure, e.g. corners, edges or blob-like structures. The detectors evaluate the distribution of the operator outputs in a local neighborhood of an image point, the most prominent example being the Harris or Foerstner operator, see Harris and Stephens (1988); Förstner and Gülch (1987). They investigate the distribution of the local gradient responses in terms of their second moment matrix \mathbf{S} , which has already been introduced as the structure tensor. Extrema of cost functions obtained from the second moment matrix such as the minima of the smallest eigenvalue or maxima of the heuristic $\det(\mathbf{S}) - \alpha \text{trace}(\mathbf{S})$, Lindeberg (1998), are then used to select spatially localized interest points. Nonetheless, in many applications such as image matching, the target images differ not only by spatial translations but also by changes in scale. Multi-scale behavior has therefore been incorporated in the structure tensor approach, replacing the derivatives with multi-scale differential operators and letting the local neighborhood depend on a certain integration scale, see Mikolajczyk and Schmid (2004). Cost functions are then designed to additionally depend on scale such that extrema of these cost functions are not only localized in space but also in scale. Localizing interest points in scale is also known as the determination of the *characteristic scale* of the image structure, see Lindeberg (1998). Equipping an image

structure with a characteristic scale directly turns a pure point based image feature into a regional image feature assigning a circular (at least in the case of isotropic kernels) region depending on the characteristic scale to the interest point. This approach has for example successfully been applied in Mikolajczyk and Schmid (2004) with the multi-scale extension of the Harris-Laplace detector. While these multi-scale methods extend the equivariance properties of the pure point based features to circular regional features which are equivariant with respect to scale changes, it is often desirable to be further equivariant with respect to the affine group in the plane. To ensure this equivariance property, Mikolajczyk et al. extended the multi-scale Harris-Laplace detector even further to the Harris-affine and Hessian-affine detectors in Mikolajczyk and Schmid (2004). Iteratively, the underlying affine transformation of the image region is estimated from the multi-scale structure tensor. Using the transformation estimate, the image content in the region described by the transformation is normalized. From the normalized image region the structure tensor is re-estimated and the steps are continued until convergence.

While our approach is related to the structure tensor based approaches, it is worth noting that one of the most successful affine equivariant region detectors is the MSER detector introduced by Matas et al. (2004). After proposing our affine equivariant region detector, we will compare our method to the Harris-affine, Hessian-affine and MSER region detectors.

In chapter 4 it has been shown, that the general linear group in the plane acts on the orientation and frequency parameters of the superimposed signal model, while leaving the phases invariant. The problem of designing an affine equivariant regional feature in terms of the superimposed signal model is reformulated as follows:

Suppose two superimposed signals $f(\mathbf{x}), f'(\mathbf{x})$ consisting of d intrinsically one-dimensional signals are given and related by a linear transformation $\mathbf{A} \in GL(2; \mathbb{R})$ as $f'(\mathbf{x}) = f(\mathbf{A}\mathbf{x})$. Can the linear transformation \mathbf{A} be recovered from the orientations \mathbf{n}_i and frequencies k_i for $i \in \{1, \dots, d\}$?

Prior to the recovery of the affine transformation (which will turn out not to be completely recoverable from the parameters without ambiguities), the frequency estimation problem has to be addressed. While the previous section introduced two methods to obtain the orientations \mathbf{n}_i of the signal model, the frequencies k_i still have to be estimated. In the case of plane waves (sinusoidal waves) the concept of frequency is closely related to the concept of scale, as the following example shows:

Example 5.22. Let $f(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)$. Suppose the group of uniform dilations in \mathbb{R}^2 acts on f as

$$(g \circ f)(\mathbf{x}) = f(\mathbf{x}/s) = A \cos((k/s) \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.122)$$

The orientation of the transformed signal stays the same, while the frequency of the signal changes to $k' = k/s$.

It seems natural to investigate, if the multiscale differential and generalized Hilbert transform operators could be used, to estimate the scale and consequently the frequency of the underlying signal components. Lindeberg has studied this topic in 1993 in terms of multiscale differential operators in the Gaussian scale space. He was interested in the selection of a *characteristic scale* for certain low-level image structures indicating the size of the structure and proposed to analyze image structures in terms of *scale normalized* derivative operators in \mathbb{R}^2

$$\mathcal{D}_{\text{norm}} = \sqrt{s} \mathcal{D} \quad (5.123)$$

yielding the higher-order scale normalized partial derivatives

$$\mathcal{D}_1^i \mathcal{D}_2^j = \sqrt{s^{i+j}} \mathcal{D}_1 \mathcal{D}_2. \quad (5.124)$$

The scale normalized Laplacian $\Delta_{\text{norm}} = s(\mathcal{D}_1^2 + \mathcal{D}_2^2)$ applied to the scale space embedding $\mathcal{G}_s f(\mathbf{x}) = (f * g_s)(\mathbf{x})$ of a single intrinsically one-dimensional signal $f(\mathbf{x}) = A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi)$ yields

$$s(\mathcal{D}_1^2 \mathcal{G}_s f(\mathbf{x}) + \mathcal{D}_2^2 \mathcal{G}_s f(\mathbf{x})) = -sk^2 e^{-k^2 s / \pi} A \cos(k \langle \mathbf{n}, \mathbf{x} \rangle + \varphi). \quad (5.125)$$

Characteristic
scale

If the function $e(s) = s(\mathcal{D}_1^2 \mathcal{G}_s f(\mathbf{x}) + \mathcal{D}_2^2 \mathcal{G}_s f(\mathbf{x}))$ is regarded for a fixed \mathbf{x} , it yields a unique extremum at (π/k^2) , called the *characteristic scale* (cmp. Lindeberg (1993)) of f at \mathbf{x} with respect to the scale normalized Laplacian which is directly related to the frequency k . If two intrinsically one-dimensional signals are related by a uniform scale change as

$$f_1(\mathbf{x}) = A \cos(k_1 \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.126)$$

$$f_2(\mathbf{x}) = (g \circ f_1)(\mathbf{x}) \quad (5.127)$$

$$= A \cos((k_1/t) \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.128)$$

$$= A \cos(k_2 \langle \mathbf{n}, \mathbf{x} \rangle + \varphi) \quad (5.129)$$

their characteristic scales at a fixed \mathbf{x} are related by

$$s_1 = (\pi/k_1^2) \quad \wedge \quad s_2 = (\pi/k_2^2) = (\pi/(k_1/t)^2) \quad \Rightarrow \quad s_2 = s_1 t^2. \quad (5.130)$$

Scale
equivariance

The characteristic scale of the intrinsically one-dimensional signal model is therefore *scale equivariant*, i.e. equivariant with respect to uniform dilations. It allows to attach a size to an intrinsically one-dimensional signal and allows an inference about how

two intrinsically one-dimensional signals might be related by a scale change.

It might be worth noting that the scale normalized Laplacian not only leads to a notion of characteristic scale for intrinsically one-dimensional signals but also for a Gaussian itself, a so called *blob-like* image structure which has been used frequently in the past, for example in Lowe (1999); Lindeberg (1993); Mikolajczyk and Schmid (2004). For a blob detector based on the scale normalized Laplacian (which also includes its approximations such as the Difference-of-Gaussian) it is therefore not sufficient to assume, that every unique extremum in scale corresponds to a blob-like structure since it might also result from an intrinsically one-dimensional signal which might be interpreted as an edge or a line. Consequently, these detectors further determine the type of the underlying structure from its differential geometric properties, such as the curvature.

Extreme points in scale space do not only yield the proposed notion of the characteristic scale, but also turned out to reflect a major amount of the signal information. In Liew and Nguyen (1995) and Mallat and Zhong (1992) it has for example been shown that signals can be reconstructed just from extrema of a multiscale representation.

5.7.1. Scale selection for superimposed signals

While the scale normalized Laplacian leads to a uniform extremum in scale for a fixed \mathbf{x} in the case of an intrinsically one-dimensional signal, this is not necessarily true anymore if a superposition

$$f(\mathbf{x}) = \sum_{i=1}^d A_i \cos(k_i \langle \mathbf{x}, \mathbf{n}_i \rangle - \varphi_i) \quad (5.131)$$

of intrinsically one-dimensional signals is considered. Further, images which are subject to transformations of the general linear group are not only subject to uniform but also non-uniform changes in scale. The scale changes which will be considered in the following are not restricted to uniform scale changes. Instead we focus on scale changes along arbitrary orthogonal axes in conjunction with rotations described by actions of

$$\mathbf{G} = \left\{ \mathbf{UDV}^T : \mathbf{UU}^T = \mathbf{I}, \mathbf{VV}^T = \mathbf{I}, \mathbf{D} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \right\} \quad (5.132)$$

where all the matrices are real-valued and $s_1, s_2 \in \mathbb{R}^+$. While the characteristic scale of the superimposed model with respect to the scale-normalized Laplacian might not be unique, every of the underlying intrinsically one-dimensional component signals of f has a unique characteristic scale. The goal is therefore to estimate the single char-

acteristic scales which are afterwards incorporated into one single affine characteristic scale which is equivariant with respect to the group actions of G .

It is convenient to illustrate our scale selection method for the superimposed signal model in the frequency domain. We consider the ideal case of d superimposed cosine waves

$$f(\mathbf{x}) = \sum_{i=1}^d A_i \cos(k_i \langle \mathbf{x}, \mathbf{n}_i \rangle - \varphi_i). \quad (5.133)$$

whose two-dimensional Fourier transform in the sense of distributions is the superposition of $2d$ Dirac delta impulses

$$\mathcal{F}f(\mathbf{u}) = \pi \sum_{i=1}^d A_i e^{-i\varphi_i(u+v)} (\delta(\mathbf{u} - k_i \mathbf{n}_i) + \delta(\mathbf{u} + k_i \mathbf{n}_i)). \quad (5.134)$$

Let $P = \{k_i \mathbf{n}_i\} \cup \{-k_i \mathbf{n}_i\}$ denote the $2d$ points which coincide with the locations of the delta impulses in the frequency domain. Let

$$\mathbf{M}_F = \frac{1}{2d} \sum_{\mathbf{p} \in P} \begin{bmatrix} p_1^2 & p_1 p_2 \\ p_1 p_2 & p_2^2 \end{bmatrix} \quad (5.135)$$

denote the second moment matrix of the points P in the frequency domain. The matrix \mathbf{M}_F is symmetric and positive-definite, describing an ellipse in the frequency domain whose axes point along the eigenvectors and whose axes lengths are given by the square-roots of the eigenvalues of \mathbf{M}_F . Figure 5.13 shows two superimposed signals for $d = 2$ and $d = 3$ and the corresponding idealized Dirac impulses in the frequency domain. Further it illustrates the ellipse described by \mathbf{M}_F obtained from the locations of the Dirac impulses.

Affine
characteristic
scale

Definition 5.23 (Affine characteristic scale). Suppose we decompose \mathbf{M}_F by its singular value decomposition as $\mathbf{M}_F = \mathbf{U}_M \mathbf{\Sigma}^{-1} \mathbf{V}_M^T$. We define the matrix $\mathbf{M}_S = \mathbf{U}_M \mathbf{\Sigma} \mathbf{V}_M^T$ describing an ellipse with the same axes directions as \mathbf{M}_F , but reciprocal axes lengths. We call the matrix \mathbf{M}_S the *affine characteristic scale* of f .

Suppose G acts on f in the spatial domain as $(g \circ f)(\mathbf{x}) = f(g^{-1}\mathbf{x}) = f(\mathbf{A}\mathbf{x})$ with $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$. Due to the affine theorem of the Fourier transform (see e.g. Bracewell (1978)), the induced action on the Fourier transform of f is given by

$$\mathcal{F}(g \circ f)(\mathbf{u}) = \frac{1}{|\det(\tilde{\mathbf{A}})|} \mathcal{F}f(\tilde{\mathbf{A}}\mathbf{u}) \quad (5.136)$$

with $\tilde{\mathbf{A}} = \mathbf{U}\mathbf{D}^{-1}\mathbf{V}^T$. While the rotations by \mathbf{U}, \mathbf{V}^T are the same as in the spatial domain, the scale changes along the rotated axes act reciprocally. It follows that the induced action on the locations of the delta impulses and therefore the second moment matrix \mathbf{M}_F is given by $\tilde{\mathbf{M}}_F = \tilde{\mathbf{A}}\mathbf{M}_F\tilde{\mathbf{A}}^T$.

Proposition 5.24 (Affine equivariance). *The affine characteristic scale \mathbf{M}_S is equivariant with respect to the group G where G acts on \mathbf{M}_S as $\tilde{\mathbf{M}}_S = \mathbf{A}\mathbf{M}_S\mathbf{A}^T$.*

We conclude that the knowledge of the orientations \mathbf{n}_i and frequencies k_i allows the construction of a regional feature which is equivariant with respect to the group G .

5.7.2. Estimation of the single scales

Suppose now, that the orientations \mathbf{n}_i of a superimposed signal of order d have been estimated using the generalized structure tensor of order d introduced in the previous section. In order to determine the affine characteristic scale \mathbf{M}_S , the frequencies k_i still have to be determined. Rather than estimating the frequencies directly, we will follow a slightly different approach. Using the estimated orientations \mathbf{n}_i , we will estimate the characteristic scales of each intrinsically one-dimensional signal f_i separately. To obtain the characteristic scales, a family of steerable wavelets parameterized by scale $s \in \mathbb{R}^+$ is chosen. The wavelets are supposed to act as bandpass filters along the orientations \mathbf{n}_i . We will use the scale-normalized second order Gaussian derivatives as introduced in Freeman and Adelson (1991).

$$\boldsymbol{\phi}(\mathbf{x}; s) = s^2 \left(\frac{\partial^2}{\partial x^2} g_s(\mathbf{x}), \frac{\partial^2}{\partial x \partial y} g_s(\mathbf{x}), \frac{\partial^2}{\partial y^2} g_s(\mathbf{x}) \right)^T \quad (5.137)$$

to construct the bandpass filters steered along the orientation angles θ_i as

$$\psi_i(\mathbf{x}; s) = \langle \mathbf{c}(\theta_i), \boldsymbol{\phi}(\mathbf{x}; s) \rangle \quad (5.138)$$

using the steering coefficients

$$\mathbf{c}(\theta_i) = (\cos^2(\theta_i), 2 \cos(\theta_i) \sin(\theta_i), \sin^2(\theta_i))^T. \quad (5.139)$$

At each interest point \mathbf{x} , the signal f is projected to the steered filters $\psi_i(\mathbf{x}, s)$ as

5. DECOMPOSITION OF SUPERIMPOSED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN \mathbb{R}^2

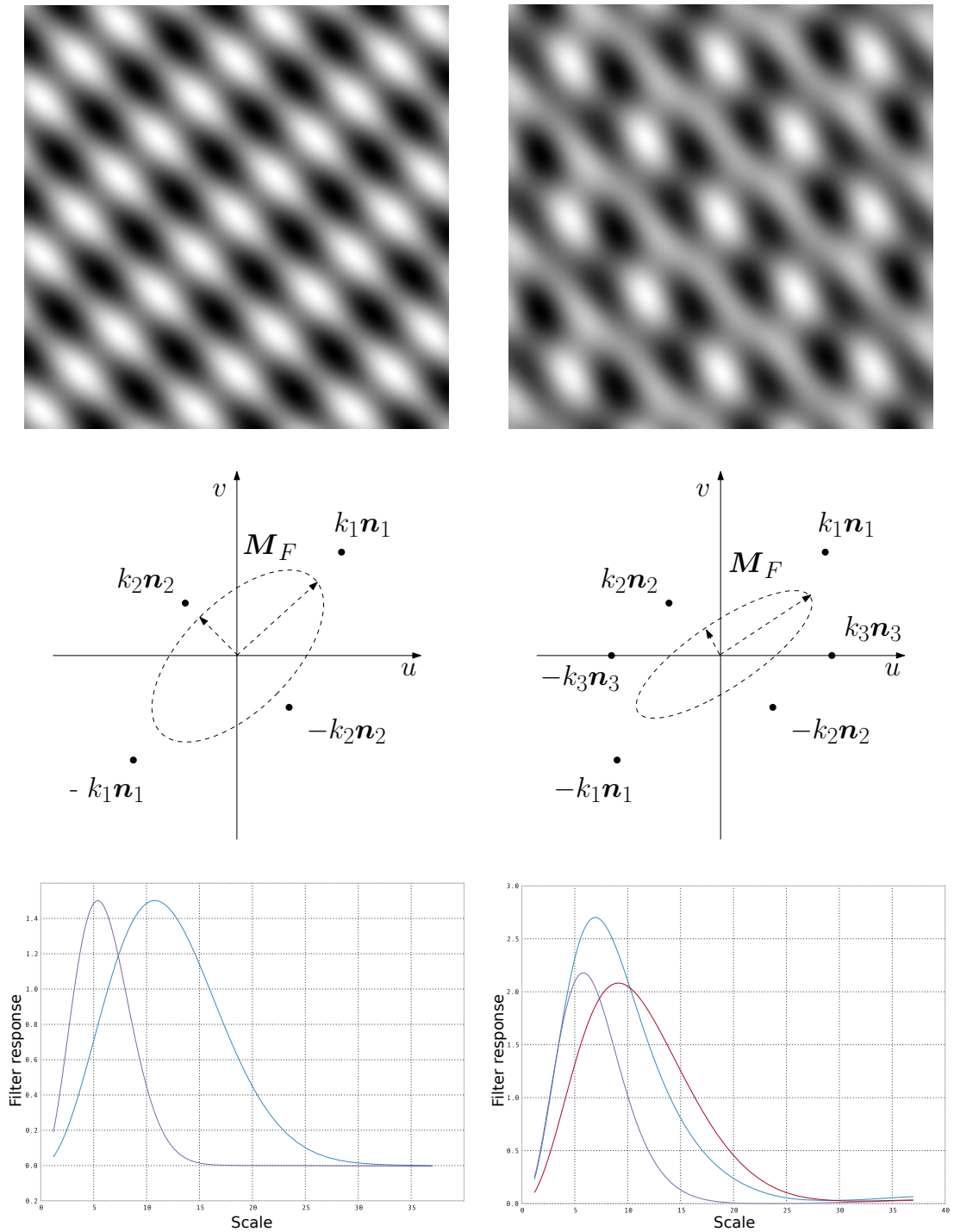


Figure 5.13.: Top: Superimposed signals consisting of $d = 2$ (left) and $d = 3$ (right) 1D signals. Middle: Schematic illustration of the Dirac delta impulses in the Fourier domain and the corresponding ellipse described by the second moment matrix M_F . Bottom: Filter responses of the steered Gaussian derivatives.

$$E_i(s) = \int_{\mathbb{R}^2} \psi_i(\mathbf{x}; s) f(\mathbf{x}) d\mathbf{x}. \quad (5.140)$$

Due to the steerability of the filters, the projections are obtained as simple linear combination of the scale normalized second order derivatives Φ .

The characteristic scales s_i for each intrinsically one-dimensional signal f_i are then chosen as the extrema of the functions $E_i(s)$ maximizing the modulus $|E_i(s)|$. If no extrema exist, the interest point is discarded. Figure 5.13 shows the functions $E_i(s)$ for two example signals, each having a unique extremum which is chosen as the characteristic scale s_i . Note that in practice, the steered filter responses have to be calculated at a finite number of scales, leading to a discretization of the scale space. Suppose that the d characteristic scales s_i of the intrinsically one-dimensional signals f_i have been estimated. Then $|E_i(s)|$ is maximal at s_i . If we consider the Fourier transform $\mathcal{F}\psi_i(\mathbf{u}; s_i)$ of the filter $\psi_i(\mathbf{x}; s_i)$, there exist two points $\mathbf{p}_{i,1}, \mathbf{p}_{i,2} = \pm(c/s_i)\mathbf{n}_i$, where c is a constant, for which the amplitude $|\mathcal{F}\psi_i(\mathbf{u}; s_i)|$ in the frequency domain attains its maximum. If the filter ψ_i is scaled along the orientation \mathbf{n}_i as $\psi_i(\mathbf{x}; r s_i), r \in \mathbb{R}^+$, the points of the maximal energy change reciprocally to $\tilde{\mathbf{p}}_{i,1}, \tilde{\mathbf{p}}_{i,2} = \pm(c/(r s_i))\mathbf{n}_i$ according to the similarity theorem of the Fourier transform. Thus from the scales s_i and the orientations \mathbf{n}_i we obtain an estimate for the locations of maximal amplitude along the orientations \mathbf{n}_i in the frequency domain, which have been idealized by Dirac delta impulses. Let $P = \{(c/s_i)\mathbf{n}_i\} \cup \{-(c/s_i)\mathbf{n}_i\}$ for $i \in \{1, \dots, d\}$ denote the estimated locations of the points of maximal amplitude in the Fourier domain. The location estimates yield the second moment matrix \mathbf{M}_F

$$\mathbf{M}_F = \frac{1}{2d} \sum_{\mathbf{p} \in P} \begin{bmatrix} p_1^2 & p_1 p_2 \\ p_1 p_2 & p_2^2 \end{bmatrix} = \mathbf{U}_M \boldsymbol{\Sigma}^{-1} \mathbf{V}_M^T \quad (5.141)$$

By inverting $\boldsymbol{\Sigma}^{-1}$ using the singular value decomposition of \mathbf{M}_F the final affine characteristic scale \mathbf{M}_S in the spatial domain reads

$$\mathbf{M}_S = \mathbf{U}_M \boldsymbol{\Sigma} \mathbf{V}_M^T. \quad (5.142)$$

5.7.3. Experiments

We have tested the proposed method with synthetic and real world images. Figure 5.14 shows two synthetic test images consisting of two and three superimposed signals. These images act as the reference frames. In red color the estimated affine characteristic scale at the origin is shown. The remaining two columns show the ref-

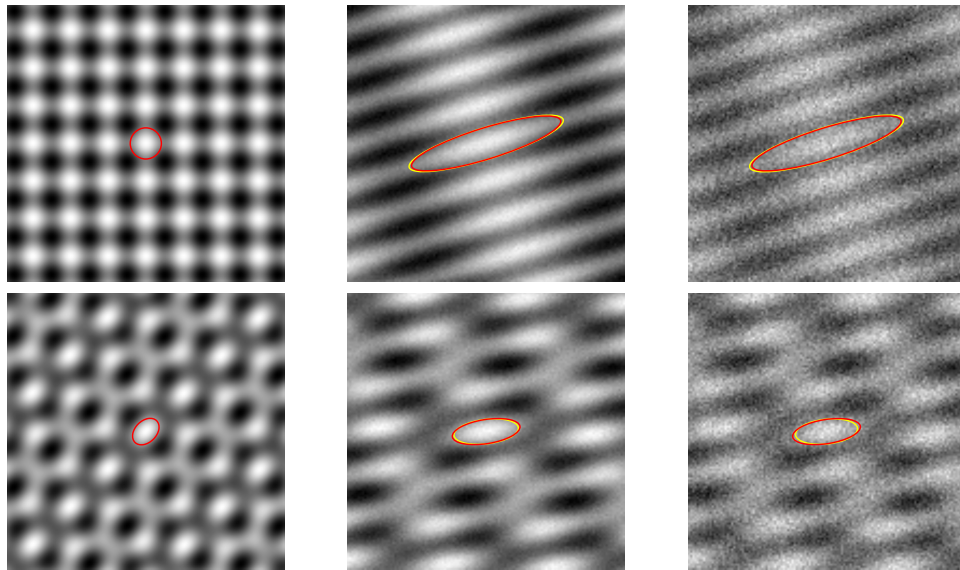


Figure 5.14.: Left column: Superimposed signals with $n = 2$ and $n = 3$ and affine characteristic scale estimated at the origin (red ellipse). Middle and right column: Images from the left column after a linear transformation with additive white noise (SNR 10 dB and 2dB), estimated affine characteristic scales (yellow) and the affine characteristic scales from the left frames mapped by the ground truth transformation (red).

reference frames after a transformation with additive white Gaussian noise and signal to noise ratios of 10 dB and 2 dB respectively. The red ellipses show the ground truth affine characteristic scale of the ellipse in the reference frames under the affine transformations. The yellow ellipses show the estimated affine characteristic scales. We notice that our proposed method is fairly exact for the chosen model, even under noisy conditions. Figure 5.15 shows the estimation of the affine characteristic scale for interest points in real world images. The images are part of the *graffiti* and *boat* sequences¹ which have also been used in e.g. Mikolajczyk and Schmid (2004). The red ellipses in the left column show the affine characteristic scales for a few manually selected interest points. They act as the reference ellipses. The remaining two columns consist of images transformed by affine transformations. Ground truth homographies have been provided by the authors of the sequence. Using the ground truth homographies, the ellipses from the reference frames are mapped to the new frames and plotted in red color. In the case of the graffiti sequence, which is subject to a projective transformation, the homography is linearized by an affine transformation. The estimated ellipses, assuming a model of $n = 3$ superimposed signals in a local neighborhood of each interest point are depicted in yellow color.

¹<http://www.robots.ox.ac.uk/~vgg/research/affine/>

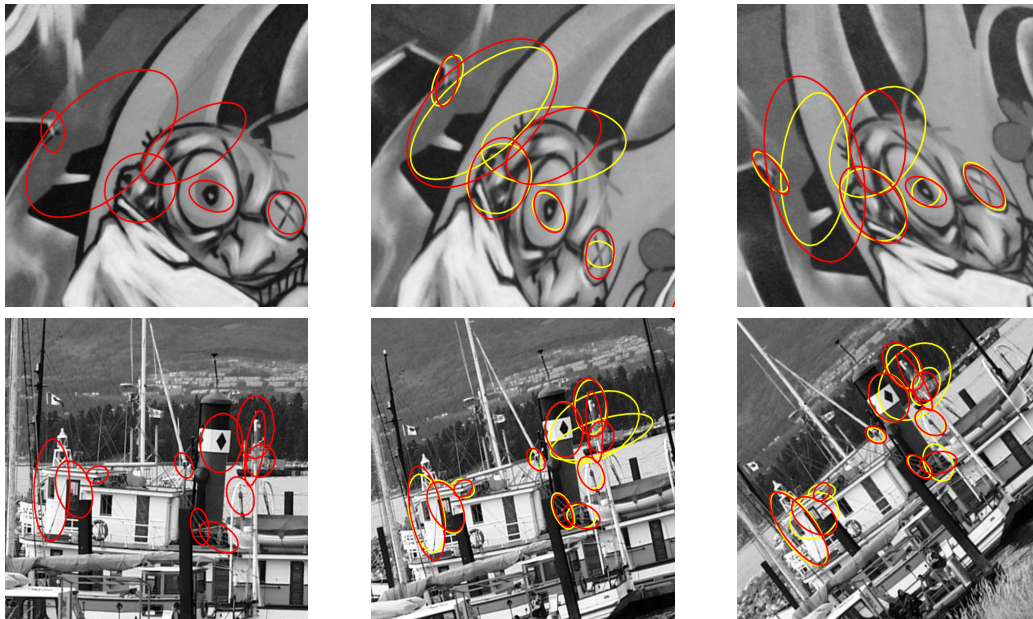


Figure 5.15.: Left column: Reference frames with estimated affine characteristic scales at manually selected points. Middle and right column: Estimated affine characteristic scales (yellow) and affine characteristics scales mapped by ground truth homographies (red).

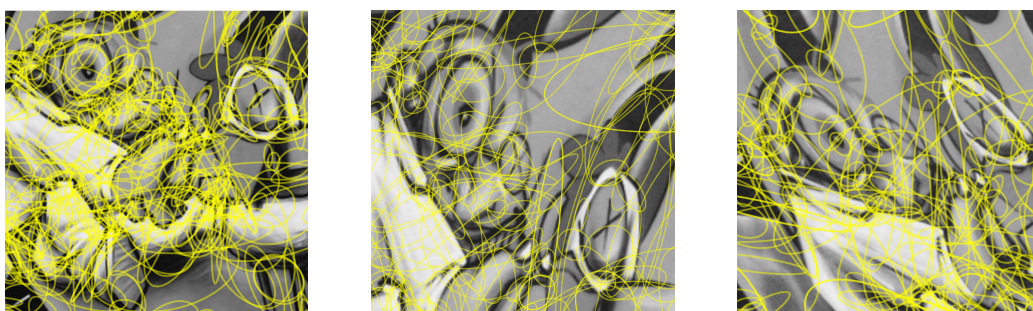


Figure 5.16.: Left: Reference frame from the graffiti sequence and the detected affine frames at DOG extrema plotted as ellipses. Middle and right: Frames three and four of the graffiti sequence and the detected affine frames.

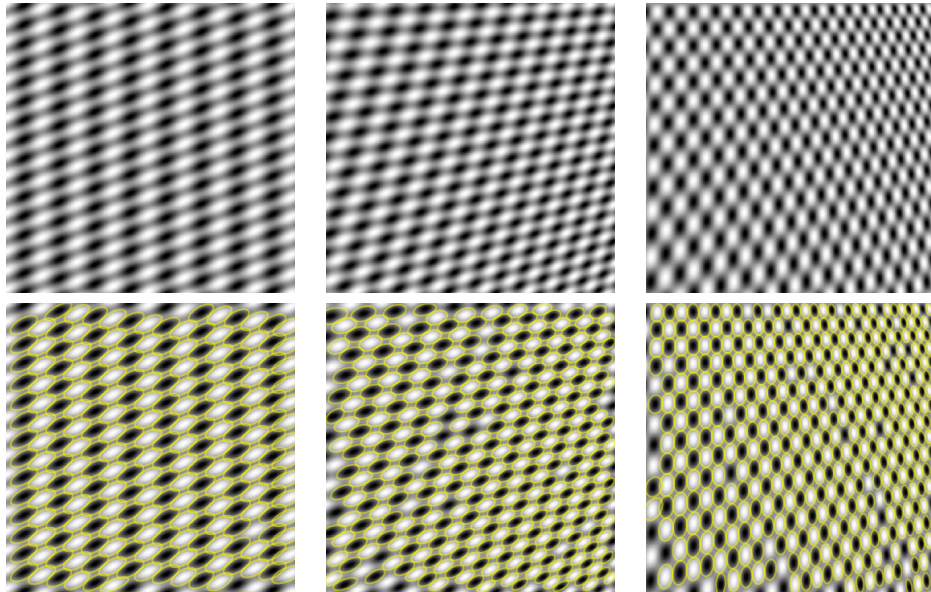


Figure 5.17.: Top row: Synthetic test sequence. Bottom row, left: Reference frame of the synthetic test sequence and the detected affine frames plotted as ellipses. Middle and right: Transformed reference frame under the same homographies as the graffiti sequence and the detected ellipses.

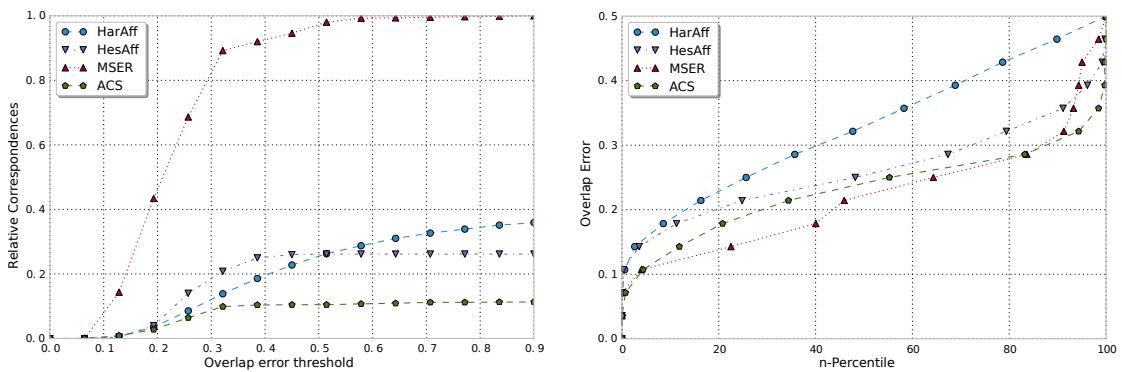


Figure 5.18.: Left: Relative correspondences of the synthetic test sequence. Right: n-percentile of the overlap error with respect to a maximum overlap error of 0.5.

Further, we have evaluated our affine equivariant region detector according to the scheme proposed in Haja et al. (2008) using the error measure introduced in Schmid et al. (2000) and Mikolajczyk and Schmid (2004). The detector is applied to three image sequences: A synthetic test sequence to confirm the theory of our detection method and the boat and graffiti sequences. These sequences consist of a set of images and a set of ground truth homographies relating the first frame to each of the remaining frames. Initial interest points are detected as maxima of the scale normalized Laplacian as it was for example proposed in Lowe (1999). At each interest point, the affine characteristic scale is calculated. The detector performance is measured in terms of the area overlap error between detected regions. For two images I_1, I_2 let E_1, E_2 denote the set of ellipses in image one and two respectively. Further suppose a homography \mathbf{H} is given which describes the transformation between the two images. Suppose two ellipses $\mathbf{a}_1 \in E_1, \mathbf{a}_2 \in E_2$ are implicitly described by $\mathbf{x}^\top \mathbf{a}_i \mathbf{x} = 1$. Further denote by \mathbf{H}' the locally linearized homography at the center of \mathbf{a}_1 . Let $A(\mathbf{a}_i) = \{\mathbf{x} : \mathbf{x}^\top \mathbf{a}_i \mathbf{x} \leq 1\}$ denote the set of points within an ellipse and $|A(\mathbf{a}_i)|$ the area of this set. Then the area overlap error between the two ellipses $\mathbf{a}_1, \mathbf{a}_2$ under the homography \mathbf{H} is defined by

$$d(\mathbf{a}_1, \mathbf{a}_2) = 1 - \frac{|A(\mathbf{H}'^\top \mathbf{a}_1 \mathbf{H}') \cap A(\mathbf{a}_2)|}{|A(\mathbf{H}'^\top \mathbf{a}_1 \mathbf{H}') \cup A(\mathbf{a}_2)|}. \quad (5.143)$$

We compared our method, *Affine Characteristic Scale* (ACS), to the Harris- and Hessian affine region detectors Mikolajczyk and Schmid (2004) as well as to the well known MSER detector Matas et al. (2004). Figure 5.18 shows the results for the synthetic test sequence. The MSER algorithm detects the largest number of correspondences with a high accuracy. Around 90% of the regions are detected with an overlap error of 0.3 and below. While determining a lower number of correspondences, the correspondences of our method are as accurate as the MSER results. About 90% are below an error of 0.3. Figure 5.19 shows the relative correspondences with respect to the overlap error as well as the n-percentile of the overlap error with respect to a maximum overlap error of 0.5 for the graffiti sequence (top) and the boat sequence (bottom). For the graffiti sequence the MSER detector is the most accurate one. Among the correspondences below 0.5, 80% are below an overlap error below of 0.3. The MSER is followed by our ACS method, performing slightly better than the Harris- and Hessian- affine detectors. In the case of the boat sequence the results of our method and MSER are the most accurate with about 50% of their overall correspondences below an error of 0.3. For the Harris affine detector around 30% and for the Hessian affine detector only 25% of their correspondences are below an error of 0.3.

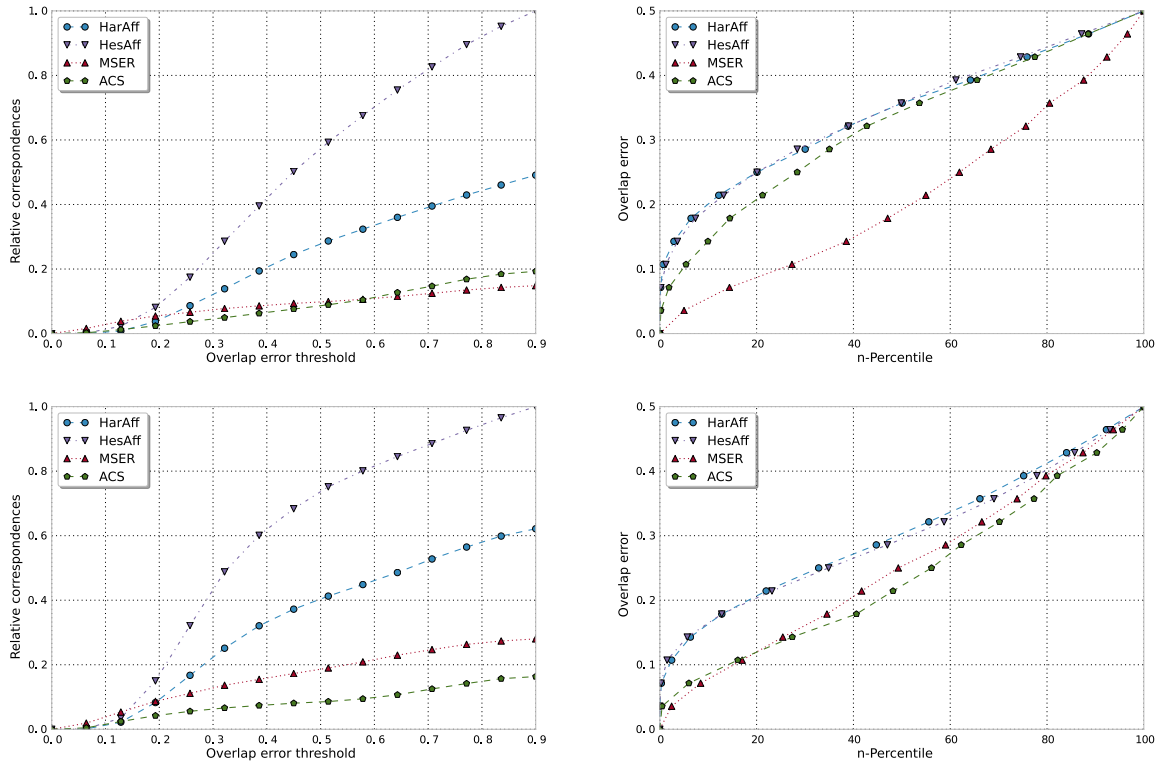


Figure 5.19.: Left column: Relative correspondences of the graffiti (top) and boat sequence (bottom). Right column: n-Percentile of the overlap errors with respect to a maximum error of 0.5 for the graffiti (top) and boat sequence (bottom)

5.8. Summary

In this chapter, a novel method for the decomposition of *arbitrary* superpositions of (sinusoidal) intrinsically one-dimensional signals into their orientation and phase components has been proposed.

The decomposition is based on (higher-order) partial derivative or generalized Hilbert transform operators whose operator responses can be interpreted as symmetric tensors. Due to an isomorphic relationship between symmetric tensors and homogeneous polynomials, the decomposition problem has been reformulated as a decomposition problem of homogenous polynomials into sums of powers of linear forms. In two dimensions, this problem can be uniquely solved. The proposed method has been developed for theoretical perfect operator responses as well as operators resulting from convolutions with discrete convolution masks, approximating the operators. It has been shown that the algorithm is exact for perfect operator responses. For the practical case, the shortcomings and problems in terms of the resolution and the scale

of the kernels, as well as the singularities have been discussed. To overcome the singularity problem, a local averaging of the operator responses has been proposed. This led to a generalized structure tensor which allows a stable orientation estimation without the problem of singularities. Further the generalized structure tensor consists of filters of lower order such that a smaller filter mask size can be chosen. Based on the generalized structure tensor, a method to estimate the characteristic scale of superimposed signals has been established, which led to an affine equivariant regional feature. The estimation procedure for these features is of non-iterative nature, depending only on convolutions with Gaussian derivative filters and one-dimensional extremum searches for the steered filter responses. It turned out that the estimation is accurate even under noisy conditions. The proposed regional feature can be integrated into existing algorithms or coupled with existing feature descriptors. In the case of feature descriptors using regional features which are only equivariant with respect to uniform scale changes, our method can easily replace existing ones to achieve equivariance for non-uniform scale changes. Further it has been shown that the model assumption is a valid image model which performs comparable to other state of the art affine regions detectors like MSER, Harris-affine or Hessian-affine.

6. Analysis of curved intrinsically one-dimensional signals in the plane

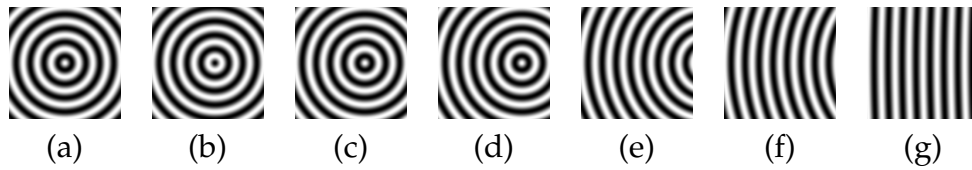


Figure 6.1.: Sinusoidal curved intrinsically one-dimensional signal with varying center \mathbf{m} . If the distance of \mathbf{m} to the origin approaches infinity, the signal corresponds to a standard intrinsically one-dimensional signal.

The previous chapter investigated in detail superimposed signals in the plane and their transformation properties with respect to the canonical Cartesian coordinate system. The transformation properties have been studied under actions of the affine group in the plane. This chapter follows a different approach. Instead of using a Cartesian coordinate system, intrinsically one-dimensional signals with respect to a nonlinear conformal embedding on the sphere in \mathbb{R}^3 are introduced. The nonlinear embedding given by the stereographic projection to the sphere in \mathbb{R}^3 and its inverse, induces a new coordinate system in the plane. Intrinsically one-dimensional signals in \mathbb{R}^3 restricted to the sphere are approximated at the origin by curved signals in plane, which are intrinsically one-dimensional signals with respect to the new coordinate system, but intrinsically two-dimensional with respect to the Cartesian coordinate system. This novel signal model allows to study the induced transformation behavior of the subgroup group $SO(3)/SO(2)$ of $SO(3)$ induced in the plane via the stereographic projection. It will be shown that the curvature of the curved signal model can be described in terms of these group actions. Further it will be possible to obtain the orientation and the phase of the projected curved signal, using the generalized Hilbert transform in \mathbb{R}^3 . In practical experiments the approach will be justified as a curvature estimator for digital curves and ridges in digital images.

6.1. Curved signal model

The main subject of this chapter are curved signals in the plane:

Definition 6.1 (Curved intrinsically one-dimensional signal). Let $\tilde{f} : \mathbb{R} \rightarrow \mathbb{R}$ be a signal and $\mathbf{m} = r_{\mathbf{m}}(\cos(\theta_{\mathbf{m}}), \sin(\theta_{\mathbf{m}}))^T \in \mathbb{R}^2$. Then the signal model Curved signal model

$$f_{\mathbf{m}}(\mathbf{x}) = \tilde{f}(\|\mathbf{x} - \mathbf{m}\|) \quad (6.1)$$

is referred to as a *curved intrinsically one-dimensional signal* or just *curved signal*.

Accordingly, sinusoidal curved intrinsically one-dimensional signals are defined:

Sinusoidal
curved signal
model

Definition 6.2 (Sinusoidal curved intrinsically one-dimensional signal). Let $r_{\mathbf{m}} \in \mathbb{R}^+$, $A \in \mathbb{R}^+$, $k \in \mathbb{R}^+$, $\phi \in [0, 2\pi)$ and $\mathbf{m} = r_{\mathbf{m}}(\cos(\theta_{\mathbf{m}}), \sin(\theta_{\mathbf{m}}))^T \in \mathbb{R}^2$. Then the signal model

$$f_{\mathbf{m}}(\mathbf{x}) = A \cos(k \|\mathbf{x} - \mathbf{m}\| + \phi) \quad (6.2)$$

is referred to as a *sinusoidal curved intrinsically one-dimensional signal*.

These signals are circular signals with amplitude A , frequency k and linear phase shift ϕ depending only on the distance from the center \mathbf{m} , which has direction $\theta_{\mathbf{m}}$ and distance $r_{\mathbf{m}}$ from the origin. The signals are constant along all circles around \mathbf{m} , which constitute the isophotes of the signal, in contrast to the i1D signals which are constant along all straight lines with a certain orientation. Figure 6.1 depicts several examples of sinusoidal curved signals in the plane with varying center \mathbf{m} .

In the following, as it can already be supposed from figure 6.1, it will turn out that this curved signal model *locally* contains the i1D signals as a subset such that the new model extends the former i1D signal model.

What are differences between the classical intrinsically one-dimensional signal and the curved signal model? The most important parameter of the signal model is its center \mathbf{m} . The angular component $\theta_{\mathbf{m}}$ of \mathbf{m} describes the *direction* of the signal. In addition to the classical orientation of an intrinsically one-dimensional signal, the direction also includes the sign of the orientation. Further, the distance $r_{\mathbf{m}}$ of the center \mathbf{m} from the origin encodes an additional property, the curvature of the isophotes. An isophote $\gamma_{\mathbf{m}}(\mathbf{x})$ of $f_{\mathbf{m}}$ at $\mathbf{x} \in \mathbb{R}^2$ is given by

Signal isophote

$$\gamma_{\mathbf{m}}(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{x} - \mathbf{m}\| = \|\mathbf{y} - \mathbf{m}\|\} \quad (6.3)$$

which is a circle around the center \mathbf{m} with radius $\|\mathbf{x} - \mathbf{m}\|$. From differential geometry it is known, that the curvature of a planar curve $\alpha(t) : \mathbb{R} \rightarrow \mathbb{R}^2$ is described by the relation $\kappa(t) = \frac{1}{r(t)}$ where $r(t)$ is the radius of osculating circle of α at t , see do Carmo (1976). Due to the definition of the signal model $f_{\mathbf{m}}$, the isophotes $\gamma_{\mathbf{m}}(\mathbf{x})$ are circles around \mathbf{m} , such that the osculating circles of the isophotes are the isophotes themselves. Therefore, the curvature $\kappa_{\mathbf{m}}(\mathbf{x})$ of an isophote $\gamma_{\mathbf{m}}(\mathbf{x})$ passing through a point of interest \mathbf{x} is obtained as

Isophote
curvature

$$\kappa_{\mathbf{m}}(\mathbf{x}) = \frac{1}{r_{\mathbf{m}}(\mathbf{x})} = \frac{1}{\|\mathbf{x} - \mathbf{m}\|}. \quad (6.4)$$

The isophote curvature of a smooth function in \mathbb{R}^2 is determined by the first and second order partial derivatives as (see e.g. Lichtenauer et al. (2005))

$$\kappa = \frac{-\frac{\partial^2}{\partial x^2} f \left(\frac{\partial}{\partial y} f\right)^2 + 2 \frac{\partial}{\partial x} f \frac{\partial}{\partial y} f \frac{\partial^2}{\partial x y} f - \frac{\partial^2}{\partial y^2} f \left(\frac{\partial}{\partial x} f\right)^2}{\left(\left(\frac{\partial}{\partial x} f\right)^2 + \left(\frac{\partial}{\partial y} f\right)^2\right)^{\frac{3}{2}}}. \quad (6.5)$$

The classical method has been used in various image processing applications, for example in van Ginkel et al. (1999) and van de Weijer et al. (2001). In the following we introduce a novel method to obtain the curvature $\kappa_{\mathbf{m}}(\mathbf{x})$ of a signal $f_{\mathbf{m}}$ at position $\mathbf{x} \in \mathbb{R}^2$ using a nonlinear embedding of the signal in a higher dimensional space. The embedding allows an intuitive estimation of the orientation and the curvature of the signal.

6.2. Conformal embedding

As a first step of our novel method we will investigate how we can obtain the radius of single isophotes of a circular input signal $f_{\mathbf{m}}$. Let us fix an isophote $\gamma_{\mathbf{m}}(\mathbf{x})$ of a circular signal $f_{\mathbf{m}}$ passing through $\mathbf{x} \in \mathbb{R}^2$. Without loss of generality, we choose a new coordinate system in \mathbb{R}^2 with \mathbf{x} as its origin. The isophote is supposed to be mapped to the sphere

$$\mathbb{S}^2 := \left\{ \mathbf{u} \in \mathbb{R}^3 : u_1^2 + u_2^2 + \left(u_3 - \frac{1}{2}\right)^2 = \frac{1}{4} \right\}. \quad (6.6)$$

In contrast to the previous chapter \mathbb{S}^2 is not the unit sphere in \mathbb{R}^3 , but the sphere with center $(0, 0, \frac{1}{2})^T$ and radius $\frac{1}{2}$. The sphere \mathbb{S}^2 touches the x_1, x_2 plane such that its south-pole coincides with the origin $(0, 0, 0)^T$. The embedding of the isophote $\gamma_{\mathbf{m}}(\mathbf{x})$ is established using the inverse stereographic projection from \mathbb{R}^2 to \mathbb{S}^2 given by

Definition 6.3 (Inverse stereographic projection, Needham (1997)).

Inverse
stereographic
projection

$$\mathcal{S}^{-1}(\mathbf{x}) = \frac{1}{1 + x_1^2 + x_2^2} \begin{pmatrix} x_1 \\ x_2 \\ x_1^2 + x_2^2 \end{pmatrix} \quad (6.7)$$

for $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$ whereas the stereographic projection from \mathbb{S}^2 to \mathbb{R}^2 is given by

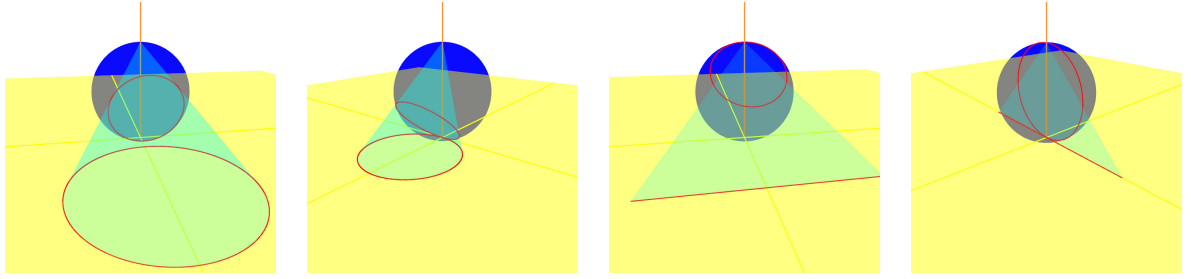


Figure 6.2.: From left to right: The first two figures illustrate circles on the plane being mapped to circles not passing through the north pole $(0, 0, 1)$ of the sphere and the next two figures illustrate lines on the plane being mapped to circles passing through the north pole of the sphere, i.e. lines are a special case of circles with infinite radius. Image source: Wietzke et al. (2008a).

Stereographic
projection

Definition 6.4 (Stereographic projection, Needham (1997)).

$$\mathcal{S}(\mathbf{u}) = \frac{1}{1 - u_3} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad (6.8)$$

for \mathbf{u} in \mathbb{S}^2 .

6.2.1. Properties of the stereographic projection

To explain the geometric idea of the curvature estimation method, we first discuss some important properties of the inverse stereographic projection. According to Needham (1997), the inverse stereographic projection is a conformal mapping from \mathbb{R}^2 to \mathbb{S}^2 . It preserves angles and maps circles in the Euclidean plane to circles on \mathbb{S}^2 . The first property is important, since the embedding preserves the direction angle θ_m of our signal f_m induced by the direction of the center \mathbf{m} of the signal. The second property is important, since it allows to treat the two signal types, intrinsically one-dimensional signals and curved signals in the plane, as the same type on the sphere: circles. This not solely includes circles of finite radius, but also circles with infinite radius which correspond to straight lines in the plane. Since the north-pole represents the point at infinity, all straight lines in \mathbb{R}^2 map to circles on \mathbb{S}^2 passing through the north-pole $(0, 0, 1)$. The south-pole of \mathbb{S}^2 coincides with the origin of \mathbb{R}^2 . Hence, straight lines through the origin map to great circles through the north- and the south-pole. Circles with finite radius through the origin in \mathbb{R}^2 map to circles through the south-pole of \mathbb{S}^2 . Figure 6.2 illustrates both scenarios.

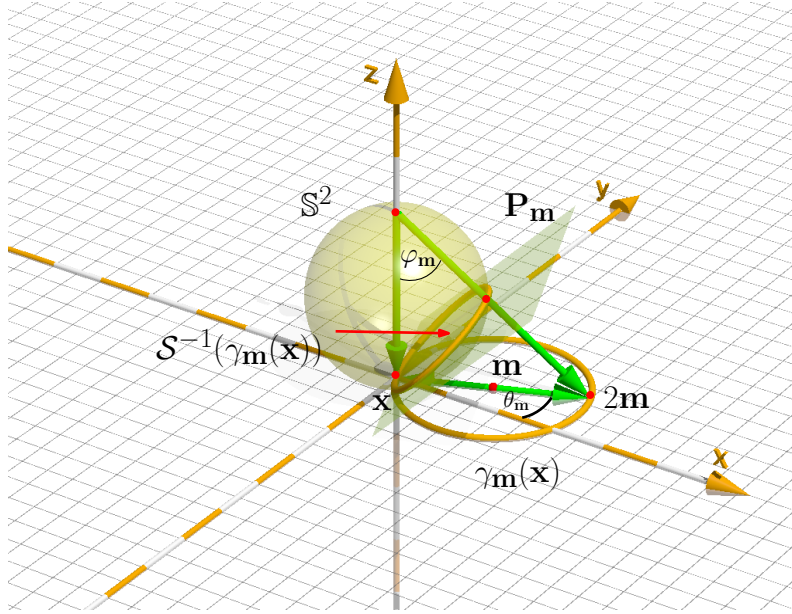


Figure 6.3.: Illustration of an isophote $\gamma_m(\mathbf{x})$ and its projection $\mathcal{S}^{-1}(\gamma_m(\mathbf{x}))$ to \mathbb{S}^2 . The radius $\|\mathbf{m} - \mathbf{x}\| = r_m$ can be calculated from the angle φ_m as $\tan \varphi_m = 2\|\mathbf{m}\|$. Therefore φ_m can be obtained from the normal vector of \mathbf{P}_m .

6.2.2. Circles as plane-sphere intersections

Straight lines and circles in the plane are described by certain circles on the sphere. Thus, it is important to observe, that both of these structures can also be described as intersections of the sphere \mathbb{S}^2 with a two-dimensional plane in \mathbb{R}^3 , since every circle on \mathbb{S}^2 results from the intersection of the sphere with a plane. If we restrict our analysis to the isophotes passing through the origin, we know that these isophotes map to circles through the south-pole of \mathbb{S}^2 . In the case of straight lines, they map to great circles additionally passing through north-pole. Now consider the isophote $\gamma_m(\mathbf{x})$ which is projected to \mathbb{S}^2 as

$$\mathcal{S}^{-1}(\gamma_m(\mathbf{x})) = \{\mathcal{S}^{-1}(\mathbf{y}) : \mathbf{y} \in \gamma_m(\mathbf{x})\}. \quad (6.9)$$

There exists a plane \mathbf{P}_m such that the projected isophote $\mathcal{S}^{-1}(\gamma_m(\mathbf{x}))$ is equal to the intersection of the sphere and \mathbf{P}_m

$$\mathcal{S}^{-1}(\gamma_m(\mathbf{x})) = \mathbf{P}_m \cap \mathbb{S}^2. \quad (6.10)$$

A plane in \mathbb{R}^3 is characterized by its unit normal vector \mathbf{n}_m and its distance d_m from the origin. The plane \mathbf{P}_m passes through the origin, since the south-pole $(0, 0, 0)^T$

is contained in $\mathbf{P}_m \cap \mathbb{S}^2$. It follows that $(0,0,0)^\top \in \mathbf{P}_m$. Hence, the distance d_m to the origin is zero. We conclude that an isophote $\gamma_m(\mathbf{x})$ is uniquely described by the normal vector \mathbf{n}_m of the corresponding plane \mathbf{P}_m . In the following we will show, how the components of the normal vector \mathbf{n}_m look like and how they relate to the isophote γ_m .

Figure 6.2 shows several possible isophotes γ_m and their projections to \mathbb{S}^2 . Notice that the first two belong to the class of curved isophotes, whereas the second two are straight lines. For all curves a Cartesian coordinate system is chosen in such a way, that the south-pole of the sphere coincides with the origin of this new coordinate system. The projection of the isophote through the origin is then described by the intersection of a plane and \mathbb{S}^2 .

6.2.3. Geometric interpretation of the normal vector

Since an isophote $\gamma_m(\mathbf{x})$ is uniquely described by the corresponding normal vector \mathbf{n}_m , all the properties of the isophote are encoded in the normal vector \mathbf{n}_m . The properties we are interested in are the direction θ_m of \mathbf{m} with respect to \mathbf{x} and the radius r_m of the isophote passing through \mathbf{x} , which is equal to the distance $\|\mathbf{x} - \mathbf{m}\|$. Figure 6.3 shows the geometric relationship between the isophote and its projection to \mathbb{S}^2 . We formulate our result as:

Theorem 6.5. *Let $\gamma_m = \{\mathbf{x} : \|\mathbf{x} - \mathbf{m}\| = r_m\}$ denote a circle through the origin with center $\mathbf{m} = r_m(\cos(\theta_m), \sin(\theta_m))^\top$ and let $\mathcal{S}^{-1}(\gamma_m)$ denote its inverse stereographic projection to \mathbb{S}^2 . Then for all $\mathbf{x} \in \gamma_m$ the projection $\mathcal{S}^{-1}(\mathbf{x})$ is perpendicular to $\mathbf{v} = (2m_1, 2m_2, -1)^\top$, i.e.*

$$\langle \mathbf{v}, \mathcal{S}^{-1}(\mathbf{x}) \rangle = 0 \quad \forall \mathbf{x} \in \gamma_m. \quad (6.11)$$

Proof. Consider a point $\mathbf{x} \in \gamma_m$ with

$$\mathbf{x} = \mathbf{m} + r_m(\cos(t), \sin(t))^\top \quad (6.12)$$

on the circle passing through the origin. Then

$$\langle \mathbf{v}, \mathcal{S}^{-1}(\mathbf{x}) \rangle = \frac{1}{1 + x_1^2 + x_2^2} (2m_1x_1 + 2m_2x_2 - x_1^2 - x_2^2) \quad (6.13)$$

with

$$2m_1x_1 = 2m_1(m_1 + r_m \cos(t)) = 2m_1^2 + 2m_1r_m \cos(t) \quad (6.14)$$

$$2m_2x_2 = 2m_2(m_2 + r_m \sin(t)) = 2m_2^2 + 2m_2r_m \sin(t) \quad (6.15)$$

$$x_1^2 = m_1^2 + 2m_1r_m \cos(t) + r_m^2 \cos^2(t) \quad (6.16)$$

$$x_2^2 = m_2^2 + 2m_2r_m \sin(t) + r_m^2 \sin^2(t). \quad (6.17)$$

such that

$$\langle \mathbf{v}, \mathcal{S}^{-1}(\mathbf{x}) \rangle = \frac{1}{1 + x_1^2 + x_2^2} (2m_1x_1 + 2m_2x_2 - x_1^2 - x_2^2) \quad (6.18)$$

$$= \frac{1}{1 + x_1^2 + x_2^2} (2m_1^2 + 2m_2^2 - m_1^2 - m_2^2 - r_m^2) \quad (6.19)$$

$$= 0. \quad (6.20)$$

□

Corollary 6.6. *There exists a plane \mathbf{P}_m in \mathbb{R}^3 perpendicular to $\mathbf{v} = (2m_1, 2m_2, -1)^\top$ such that $\mathcal{S}^{-1}(\gamma_m) \subset \mathbf{P}_m$.*

The vector \mathbf{v} is the (unnormalized) normal vector of the corresponding plane \mathbf{P}_m . Normalizing \mathbf{v} results in

$$\mathbf{n}_m = \frac{1}{\sqrt{1 + 4r_m^2}} (2m_1, 2m_2, -1)^\top \quad (6.21)$$

$$= \frac{1}{\sqrt{1 + 4r_m^2}} (2r_m \cos(\theta_m), 2r_m \sin(\theta_m), -1)^\top. \quad (6.22)$$

Corollary 6.7. *For the unit normal vector \mathbf{n}_m it holds that*

$$\mathbf{n}_m = \frac{1}{\sqrt{1 + 4r_m^2}} (2r_m \cos(\theta_m), 2r_m \sin(\theta_m), -1)^\top \quad (6.23)$$

$$= (\sin(\alpha) \cos(\beta), \sin(\alpha) \sin(\beta), \cos(\alpha))^\top. \quad (6.24)$$

with

$$\sin(\alpha) = \frac{2r_m}{\sqrt{1 + 4r_m^2}} \quad \cos(\alpha) = \frac{-1}{\sqrt{1 + 4r_m^2}} \quad (6.25)$$

$$\sin(\beta) = \sin(\theta_m) \quad \cos(\beta) = \cos(\theta_m). \quad (6.26)$$

Proof. The last component of \mathbf{n}_m implies that

$$\cos(\alpha) = \frac{-1}{\sqrt{1 + 4r_m^2}}. \quad (6.27)$$

yielding

$$\sin(\alpha) = \sqrt{1 - \cos(\alpha)^2} = \sqrt{\frac{4r_m}{1 + 4r_m^2}} = \frac{2r_m}{\sqrt{1 + 4r_m^2}}. \quad (6.28)$$

□

6.2.4. Isophote properties

The corollaries 6.6, 6.7 are the central idea of the curvature estimation method. They show, that it is possible to obtain the direction angle θ_m of the center \mathbf{m} with respect to \mathbf{x} and the radius r_m of the isophote of f_m passing through \mathbf{x} just from components of the normal vector \mathbf{n}_m of the plane \mathbf{P}_m as:

Corollary 6.8 (Isophote curvature).

$$\kappa_m = \frac{1}{r_m} = \frac{1}{|\mathbf{x} - \mathbf{m}|} = \frac{2n_{m,3}}{\sqrt{n_{m,1}^2 + n_{m,2}^2}}. \quad (6.29)$$

Corollary 6.9 (Isophote direction).

$$\theta_m = \arctan 2(n_{m,2}, n_{m,1}). \quad (6.30)$$

6.2.5. Isophote classification

The two important properties isophote curvature and isophote direction are encoded in the normal vector of the plane \mathbf{P}_m . Apart from obtaining these properties, it is also possible to decide whether an isophote is a straight line, which corresponds to a circle with an infinite radius, or a circle with a finite radius. Let γ_m denote an isophote with an infinite radius modeled by a center \mathbf{m} with infinite distance from the origin:

$$\mathbf{m} = \lim_{r_m \rightarrow \infty} r_m (\cos(\theta_m), \sin(\theta_m))^T. \quad (6.31)$$

Then the projection $\mathcal{S}^{-1}(2\mathbf{m})$ coincides with the north-pole such that the angle φ_m approaches $\pi/2$ and $\cos(\varphi_m)$ vanishes. It follows that the normal vector \mathbf{n}_m of the plane \mathbf{P}_m describing $\mathcal{S}^{-1}(\gamma_m(\mathbf{x}))$ is obtained as $\mathbf{n}_m = (\cos(\theta_m), \sin(\theta_m), 0)^T$. For

every isophote with finite radius, the angle $\cos(\varphi_m)$ does not vanish. We can therefore distinguish between straight lines and curved isophotes by examining the third component $n_{m,3}$ of the normal vector \mathbf{n}_m .

So far we developed a method which is able to obtain the curvature and the direction of the isophote by its normal vector \mathbf{n}_m , corresponding to the plane passing through the projected isophote on \mathbb{S}^2 , and consequently allows to distinguish between straight line and curved isophotes. The problem of isophote curvature estimation is therefore equivalent to the estimation of the normal vector \mathbf{n}_m .

It has been shown, how the curvature can be obtained from a single isophote if the isophote is known. If the curvature in two dimensional signals is supposed to be estimated, the proposed method can not directly be applied, since the isophotes passing through a point of interest are not known in advance. Hence, we are not able to project a single isolated isophote to \mathbb{S}^2 in the case of two dimensional signals. Instead we are faced with the problem of estimating the normal vector \mathbf{n}_m from a projected *neighborhood* with respect to a point of interest. We will introduce a method which will estimate the normal vector \mathbf{n}_m from a neighborhood projected to the sphere, using the partial derivatives or generalized Hilbert transforms in \mathbb{R}^3 such that a curvature estimation according to theorem 6.5 will be possible.

6.2.6. Approximating intrinsically one-dimensional signals in \mathbb{R}^3 with curved intrinsically one-dimensional signals in the plane

While the inverse stereographic projection embeds circles passing through the origin in the plane in linear subspaces in \mathbb{R}^3 , such that their normal vector completely characterizes the circles, it is still an open question how the curved signal model 6.1 can be analyzed in terms of the proposed method. In the following we will establish a relationship between the curved signal model and intrinsically one-dimensional signals in \mathbb{R}^3 , whose orientation vectors coincide with the normal vector of the projected isophote of a curved signal.

Suppose $f_m(\mathbf{x}) = \tilde{f}(\|\mathbf{x} - \mathbf{m}\|)$ is a curved signal in \mathbb{R}^2 . Fix $\mathbf{x} \in \mathbb{R}^2$ and suppose a coordinate system in the plane is chosen, with \mathbf{x} as the origin. Further, with respect to the fixed \mathbf{x} , define the intrinsically one-dimensional signal ψ in \mathbb{R}^3 as

$$\psi(\mathbf{u}) = \tilde{f}(\langle \mathbf{n}, \mathbf{u} \rangle + \phi') \quad (6.32)$$

with $\phi' = \|\mathbf{m}\|$ and

$$\mathbf{n} = \frac{1}{\sqrt{1 + 4r_m^2}}(2m_1, 2m_2, -1)^T \quad (6.33)$$

according to theorem 6.5. Obviously the two signals coincide at the origin, i.e

$$\psi(0,0,0) = \tilde{f}(\phi') = \tilde{f}(\|\mathbf{m}\|) = f_{\mathbf{m}}(0,0). \quad (6.34)$$

Further, the values of ψ along the whole linear subspace described by \mathbf{n} coincide with the value at the origin

$$\psi(\mathbf{u}) = \psi(0,0,0) \quad \forall \mathbf{u} : \langle \mathbf{u}, \mathbf{n} \rangle = 0. \quad (6.35)$$

Consider now the isophote $\gamma_{\mathbf{m}}(0,0)$ of $f_{\mathbf{m}}$ passing through the origin. As it has been shown in theorem 6.5, its embedding via the inverse stereographic projection is completely contained in the subspace perpendicular to \mathbf{n} . Hence,

$$\psi(\mathcal{S}^{-1}(\mathbf{x})) = f_{\mathbf{m}}(\mathbf{x}) \quad (6.36)$$

for all $\mathbf{x} \in \{r_{\mathbf{m}}(\cos(\theta), \sin(\theta))^{\top} : \theta \in [0, 2\pi)\}$.

It follows, that the orientation vector of the intrinsically one-dimensional signal ψ in \mathbb{R}^3 completely characterizes the direction and the curvature of $f_{\mathbf{m}}$. If it would be possible to recover ψ from $f_{\mathbf{m}}$, these properties could be obtained from the orientation vector \mathbf{n} vector of ψ . The question arising is: Can ψ be obtained from the values of $f_{\mathbf{m}}$ projected to \mathcal{S}^2 ?

Unfortunately the two signals do not coincide on the complete sphere. In general it holds that $\psi(\mathcal{S}^{-1}(\mathbf{x})) \neq f_{\mathbf{m}}(\mathbf{x})$. Nonetheless, the isophote curvature of a signal in \mathbb{R}^2 is a *local* property which is determined with respect to an infinitely small neighborhood. It can be shown, that in a sufficiently small neighborhood of the origin, the values of the projected curved signal $f_{\mathbf{m}}$ coincide with the values of the corresponding intrinsically one-dimensional signal ψ in \mathbb{R}^3 restricted to \mathcal{S}^2 . Suppose $\mathbf{x} = \Delta r(\cos(\theta_{\mathbf{m}}), \sin(\theta_{\mathbf{m}}))^{\top}$. Then it is sufficient to show, that the inner product of the projection of \mathbf{x} to the sphere and the orientation vector \mathbf{n} of ψ depends linearly on Δr .

Proposition 6.10. *Let $\mathbf{x} = \mathbf{m} - (r_{\mathbf{m}} + \Delta r)(\cos(\theta_{\mathbf{m}} + \Delta\theta), \sin(\theta_{\mathbf{m}} + \Delta\theta))^{\top}$. Then for $\Delta\theta, \Delta r \ll 1$ it holds that*

$$\langle \mathbf{n}, \mathcal{S}^{-1}(\mathbf{x}) \rangle = \frac{2m_1x_1 + 2m_2x_2 - x_1^2 - x_2^2}{(1 + x_1^2 + x_2^2)\sqrt{1 + 4r_{\mathbf{m}}^2}} \approx -\Delta r \frac{-2r_{\mathbf{m}}}{\sqrt{1 + 4r_{\mathbf{m}}^2}}. \quad (6.37)$$

Proof. Without loss of generality, assume that the center \mathbf{m} is located on the x_1 axis, i.e. $m_1 = r_{\mathbf{m}}, m_2 = 0$. Neglecting the quadratic terms and using the small angle approximations $\cos(\Delta\theta) \approx 1 - \Delta\theta^2/2$ and $\sin(\Delta\theta) \approx \Delta\theta$ results in

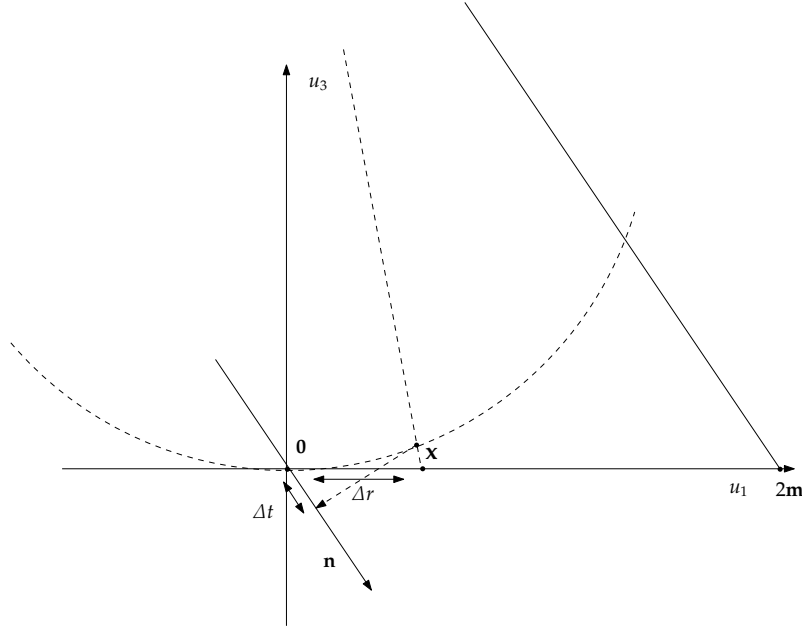


Figure 6.4.: Geometric illustration of linear behavior in the limit. For sufficiently small Δr it holds that $\Delta r \approx \Delta t$.

$$\langle \mathbf{n}, \mathcal{S}^{-1}(\mathbf{x}) \rangle \approx \frac{2m_1 x_1}{\sqrt{1 + 4r_m^2}} \quad (6.38)$$

$$\approx \frac{2r_m(r_m - r_m \cos(\Delta\theta) - \Delta r \cos(\Delta\theta))}{\sqrt{1 + 4r_m^2}} \quad (6.39)$$

$$\approx \frac{2r_m(r_m - r_m(1 - \Delta\theta^2/2) - \Delta r(1 - \Delta\theta^2/2))}{\sqrt{1 + 4r_m^2}} \quad (6.40)$$

$$\approx \frac{2r_m(r_m - r_m - \Delta r)}{\sqrt{1 + 4r_m^2}} \quad (6.41)$$

$$\approx \frac{-2r_m \Delta r}{\sqrt{1 + 4r_m^2}}. \quad (6.42)$$

□

While the two signals f_m and ψ coincide on the surface of the sphere with respect to a sufficiently small neighborhood in the plane, the signal ψ is only sampled along the surface of \mathbb{S}^2 by the projections of f_m . In the following, the function values of f_m projected to the sphere are interpreted as (irregular) samples of the intrinsically one-dimensional signal ψ .

In chapter 4 it has been shown that the orientation vector of an intrinsically one-

dimensional signal in \mathbb{R}^n can be obtained by the gradient or the generalized Hilbert transform in \mathbb{R}^n . If the gradient or the generalized Hilbert transform is supposed to be applied to the approximated ψ , the operator responses can be approximated using only the samples on the surface of \mathbb{S}^2 by

$$\mathcal{A}_i\psi(\mathbf{0}) \approx \int_{\mathbb{R}^3} \mathcal{A}_i f_m(\mathcal{S}(\mathbf{u})) d\mathbf{m}(\mathbf{u}) = \int_{\mathbb{R}^3} k_{i,s} f_m(\mathcal{S}(\mathbf{u})) d\mathbf{m}(\mathbf{u}) \quad (6.43)$$

for $\mathcal{A}_i \in \{\mathcal{H}_i \mathcal{K}_s, \mathcal{D}_i \mathcal{K}_s\}$ and corresponding convolution kernel $k_{i,s}$. \mathcal{K}_s is either the Poisson or the Gaussian scale space representation in \mathbb{R}^3 with scale parameter $s > 0$. Here $d\mathbf{m}(\mathbf{u}) = \delta(\|\mathbf{u} - \mathbf{c}\|) d\mathbf{u}$ denotes the Dirac measure where \mathbf{c} is the center of \mathbb{S}^2 and $\delta(\cdot)$ is the Delta distribution. For further details on the validity of this construction we refer the reader to Zhang et al. (2007). Since the measure is only non-zero on the surface of the sphere, the integration can also be carried out in the plane

$$\mathcal{A}_i\psi(\mathbf{0}) \approx \int_{\mathbb{R}^3} k_{i,s}(\mathbf{u}) f_m(\mathcal{S}(\mathbf{u})) d\mathbf{m}(\mathbf{u}) = \int_{\mathbb{R}^2} k_{i,s}(\mathcal{S}^{-1}(\mathbf{y})) f_m(\mathbf{y}) d\mathbf{y}. \quad (6.44)$$

The actual implementation and application of the approximations bears certain problems. Suppose the signal f_m is analyzed in a local neighborhood $\Omega(\epsilon)$ around the origin. The first question is, how the size of the neighborhood should be chosen. The correspondence between the projection of f_m to the sphere and the intrinsically one-dimensional signal ψ restricted to the sphere only holds in a sufficiently small neighborhood. Hence, the neighborhood should be chosen as small as possible. But for a small neighborhood only a small portion of the sphere near the south-pole is covered. On the other hand, the approximations of the operators are more exact, if a larger portion of the sphere is covered by samples. It turned out experimentally that with respect to the Poisson kernel, a reasonable trade-off between the two criteria is a neighborhood of 3×3 pixels which sample the Cartesian coordinate patch $[-3, 3] \times [-3, 3]$ using $s = 2$.

A further problem arises for the operator \mathcal{A}_3 , i.e. the partial derivative or generalized Hilbert transform along the u_3 axis. Due to the projection to the sphere, all the samples are contained in the upper half space of \mathbb{R}^3 which introduces a bias in the approximation of the operator. To reduce this effect, the mean of the projected u_3 coordinates is subtracted for the sampling of the corresponding kernel. Define

$$\bar{\mathbf{u}} = (0, 0, \bar{u}_3)^T \quad (6.45)$$

with the mean of the projection values along the u_3 axis given by

$$\bar{u}_3 = \frac{1}{|\Omega(\epsilon)|} \int_{\Omega(\epsilon)} \frac{x_1^2 + x_2^2}{1 + x_1^2 + x_2^2} d\mathbf{x} \quad (6.46)$$

where $|\Omega(\epsilon)|$ denotes the surface area of the local neighborhood around the origin. Then the corrected kernel is defined as

$$\widehat{k}_{3,s}(\mathbf{x}) = k(\mathcal{S}^{-1}(\mathbf{x}) - \bar{\mathbf{u}}). \quad (6.47)$$

Further the mean of the complete kernel has to be removed, yielding the final kernel

$$\widetilde{k}_{3,s}(\mathbf{x}) = \widehat{k}_{3,s}(\mathbf{x}) - \frac{1}{|\Omega(\epsilon)|} \int_{\Omega(\epsilon)} \widehat{k}_{3,s}(\mathbf{y}) d\mathbf{y}. \quad (6.48)$$

The kernels in the plane corresponding to the operators along the u_1 and the u_2 axes are defined as

$$\widetilde{k}_s(\mathbf{x}) = k_s(\mathcal{S}^{-1}(\mathbf{x})) \quad (6.49)$$

$$\widetilde{k}_{1,s}(\mathbf{x}) = k_{1,s}(\mathcal{S}^{-1}(\mathbf{x})) \quad (6.50)$$

$$\widetilde{k}_{2,s}(\mathbf{x}) = k_{2,s}(\mathcal{S}^{-1}(\mathbf{x})). \quad (6.51)$$

Since the mean of the kernels $\widetilde{k}_{1,s}, \widetilde{k}_{2,s}$ is already zero, no correction has to be applied. Convolution of the original signal f with the kernels $\widetilde{k}_{i,s}$ induces the operators $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{A}}_i$

$$\widetilde{\mathcal{A}}f_m(\mathbf{x}) = (f_m * \widetilde{k}_s)(\mathbf{x}) \quad (6.52)$$

$$\widetilde{\mathcal{A}}_i f_m(\mathbf{x}) = (f_m * \widetilde{k}_{i,s})(\mathbf{x}). \quad (6.53)$$

From the operator responses, the angles of the orientation vector \mathbf{n} of the approximated intrinsically one-dimensional signal ψ in \mathbb{R}^3 at a point \mathbf{x} are obtained as

$$\theta(\mathbf{x}) = \arctan 2(\widetilde{\mathcal{A}}_2 f_m(\mathbf{x}), \widetilde{\mathcal{A}}_1 f_m(\mathbf{x})) \quad (6.54)$$

and

$$\varphi(\mathbf{x}) = \arctan 2\left(\sqrt{(\tilde{\mathcal{A}}_2 f_{\mathbf{m}}(\mathbf{x}))^2 + (\tilde{\mathcal{A}}_1 f_{\mathbf{m}}(\mathbf{x}))^2}, \tilde{\mathcal{A}}_3 f_{\mathbf{m}}(\mathbf{x})\right). \quad (6.55)$$

While one of the parameters of the curved signal model, as pointed at the beginning of this chapter, is the direction of the center \mathbf{m} , it is not possible to recover the sign, such that the parameter $\theta(\mathbf{x})$ coincides with the orientation already known from the standard intrinsically one-dimensional signals. The sign ambiguity is similar to the two-dimensional case: $\theta(\mathbf{x})$ is obtained from the orientation vector of an intrinsically one-dimensional signal in \mathbb{R}^3 . But this orientation vector can only be obtained up to its sign, according to the two-dimensional case.

Since the orientation vector \mathbf{n} is related to the center of $f_{\mathbf{m}}$ by (6.33), corollary 6.8 yields the curvature

$$\kappa(\mathbf{u}) = \frac{2\tilde{\mathcal{A}}_3 f_{\mathbf{m}}(\mathbf{x})}{\sqrt{(\tilde{\mathcal{A}}_2 f_{\mathbf{m}}(\mathbf{x}))^2 + (\tilde{\mathcal{A}}_1 f_{\mathbf{m}}(\mathbf{x}))^2}}. \quad (6.56)$$

In section 6.2.5 we have already classified single isophotes as straight lines and curved isophotes, depending on the angle $\varphi_{\mathbf{m}}$. Since we have now established the link between single isophotes and curved signals, we are able to classify the underlying signal type. Figure 6.6 illustrates, that we are able to distinguish standard intrinsically one-dimensional and curved intrinsically one-dimensional signals, depending on the angle $\varphi_{\mathbf{m}}$. As $\varphi_{\mathbf{m}}$ approaches $\frac{\pi}{2}$, a standard intrinsically one-dimensional signal in the plane is approximated. Otherwise the underlying signal is a curved signal.

Conformal
monogenic
signal

Definition 6.11 (Conformal monogenic signal). Let k_s denote the Poisson kernel in \mathbb{R}^3 and $k_{i,s}$ the conjugate Poisson kernels, i.e. $k_s = p_s, k_{i,s} = q_{i,s}$. Further let \tilde{p}_s, \tilde{q}_s denote the corresponding (mean corrected) kernels according to (6.48), (6.49), (6.50), and (6.51). Then the signal representation

$$\mathbf{f}_{\mathbf{m}}(\mathbf{x}) = ((f_{\mathbf{m}} * \tilde{k}_s)(\mathbf{x}), (f_{\mathbf{m}} * \tilde{k}_{1,s})(\mathbf{x}), (f_{\mathbf{m}} * \tilde{k}_{2,s})(\mathbf{x}), (f_{\mathbf{m}} * \tilde{k}_{3,s})(\mathbf{x}))^T \quad (6.57)$$

$$= ((f_{\mathbf{m}} * \tilde{p}_s)(\mathbf{x}), (f_{\mathbf{m}} * \tilde{q}_{1,s})(\mathbf{x}), (f_{\mathbf{m}} * \tilde{q}_{2,s})(\mathbf{x}), (f_{\mathbf{m}} * \tilde{q}_{3,s})(\mathbf{x}))^T \quad (6.58)$$

is called the *conformal monogenic signal* representation of $f_{\mathbf{m}}$.

If the original signal $f_{\mathbf{m}}$ is a sinusoidal intrinsically one-dimensional signal, the conformal monogenic signal yields the additional properties *local amplitude* and *local phase* as

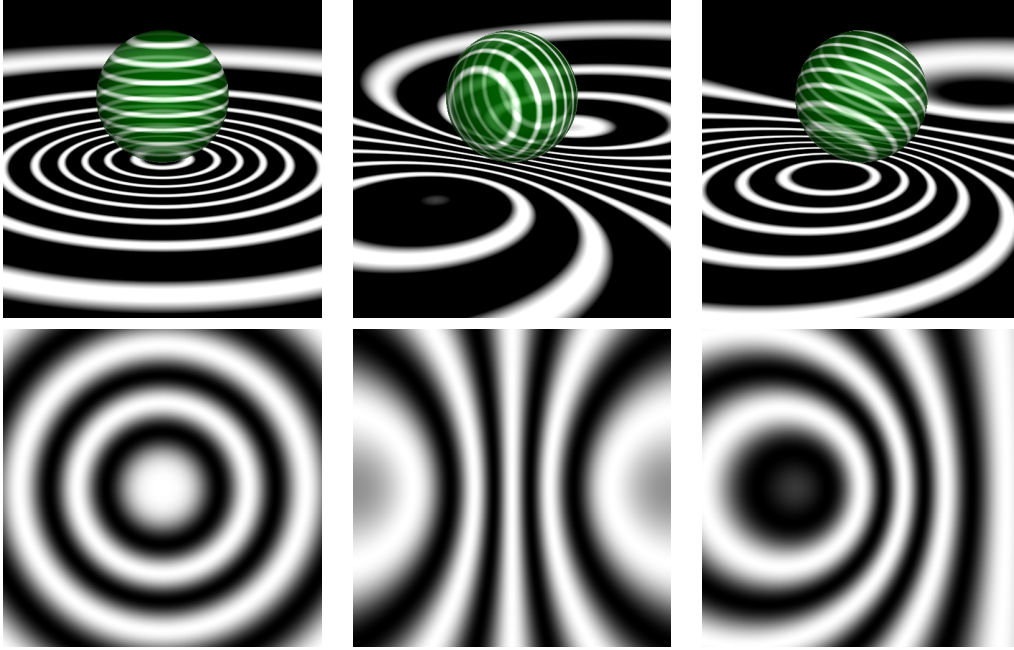


Figure 6.5.: Illustration of the conformal mapping of 2D signals to the 3D conformal space.

$$A(\mathbf{x}) = \sqrt{(f_m * \tilde{p}_s)(\mathbf{x})^2 + \sum_{i=1}^3 (f_m * \tilde{q}_{i,s})(\mathbf{x})^2}. \quad (6.59)$$

$$\phi(\mathbf{x}) = \arctan 2 \left(\sqrt{\sum_{i=1}^3 (f_m * \tilde{q}_{i,s})(\mathbf{x})^2}, (f * \tilde{p}_s)(\mathbf{x}) \right) \quad (6.60)$$

where the phase indicates a measure of parity symmetry as it is known from classical phase based signal analysis. It is worth noting that all these properties have been obtained by simple convolutions with proposed kernels, yielding linear shift-invariant features.

6.3. Experiments

We are interested in the accuracy of our curvature estimator. Since experiments regarding the orientation and the phase of the monogenic signal have already been carried out in the mentioned literature, we focus on experiments concerning the new

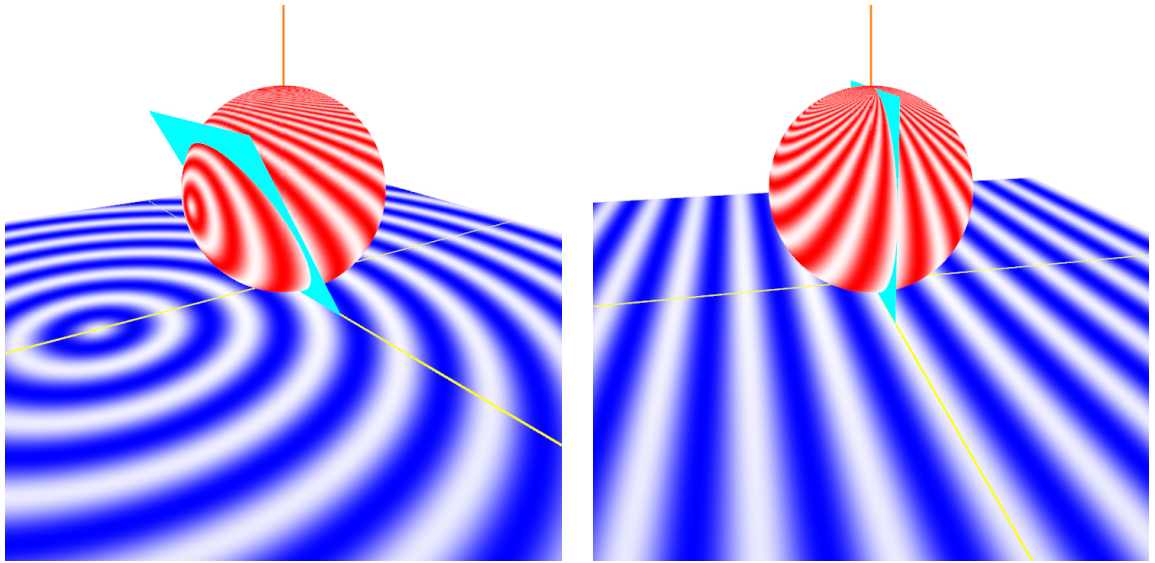


Figure 6.6.: Left figure: The plane indicates an i2D signal. Right figure: The plane of an i1D signal passes through the north pole of the sphere. Image source: Wietzke et al. (2008a)

curvature feature.

6.3.1. Curvature estimation of planar curves

We first evaluate the proposed method for planar curves. Let $\mathbf{l}(t) = (x(t), y(t))$, $t \in [a, b]$ be a part of a parameterized plane curve. Then we sample \mathbf{l} as

$$\mathbf{l}_n = (x(d_n), y(d_n))^T \text{ with } n \in \{1, \dots, N\}, d_n = n \frac{|a - b|}{N}, N \in \mathbb{N}. \quad (6.61)$$

The estimation of the curvature relies on the choice of a scale described by the neighborhood or window size $W \in \mathbb{N}$ with respect to the current point of interest. For each point \mathbf{l}_n we first shift the neighborhood $\text{NB}_W(\mathbf{l}_n) = \{\mathbf{l}_{n-W}, \dots, \mathbf{l}_n, \dots, \mathbf{l}_{n+W}\}$ to the origin and project it to the sphere \mathbb{S}^2 such that $\text{NB}'_W(\mathbf{l}_n) = \{\mathcal{S}^{-1}(\mathbf{l}_{n-W} - \mathbf{l}_n), \dots, (0, 0, 0), \dots, \mathcal{S}^{-1}(\mathbf{l}_{n+W} - \mathbf{l}_n)\}$. The curvature is then obtained as

$$\kappa_W(\mathbf{l}_i) = 2 \frac{M_W^3}{((M_W^1)^2 + (M_W^2)^2)^{1/2}} \quad (6.62)$$

with

$$M_W^i = \sum_{j=-W}^W k_{i,s}(\mathcal{S}^{-1}(\mathbf{l}_{n+j} - \mathbf{l}_n)), \quad i \in \{1, 2, 3\} \quad (6.63)$$

where $k_{i,s}$ denotes the i -th component of the conjugate Poisson kernels or the Gaussian derivative kernels in \mathbb{R}^3 with respect the scale s . We compare our method to curvature estimations obtained by circle fittings through the points of the neighborhood $\text{NB}'_W(\mathbf{l}_i)$. To fit a circle through these points we use two different distance functions which are minimized, an algebraic distance according to Gander et al. (1994) and a geometric distance according to Coope (1993).

Algebraically a circle may be represented as the set of all $\mathbf{x} \in \mathbb{R}^2$ satisfying

$$\mathbf{a}\mathbf{x}^T\mathbf{x} + \mathbf{b}^T\mathbf{x} + c = 0 \quad (6.64)$$

with $\mathbf{a} \neq 0, \mathbf{b} \in \mathbb{R}^2, c \in \mathbb{R}$. We fit a circle through the points $\mathbf{x} \in \text{NB}(\mathbf{l}_i, W)$ by minimizing

$$\arg \min_{\mathbf{u}} \|\mathbf{A}\mathbf{u}\| \quad \text{s.t.} \|\mathbf{u}\| = 1 \quad (6.65)$$

with

$$\mathbf{A} = \begin{pmatrix} x_{11}^2 + x_{12}^2 & x_{11} & x_{12} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1}^2 + x_{m2}^2 & x_{m1} & x_{m2} & 1 \end{pmatrix} \quad (6.66)$$

and $m = 2W + 1, \mathbf{u} = (\mathbf{a}, b_1, b_2, c)^T$. The radius of curvature r is then obtained as

$$r = \sqrt{\frac{\sqrt{b_1^2 + b_2^2}}{4a^2} - \frac{c}{a}}. \quad (6.67)$$

The second method, which we refer to as the geometric method, minimizes the distance

$$\sum_{j=1}^{2W+1} (x_{j1} - \mathbf{a}_1)^2 + (x_{j2} - \mathbf{a}_2)^2 - r^2 \quad (6.68)$$

which is minimized for the unknown center $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ and the unknown radius r by solving a linear least-squares problem (see Coope (1993)). The radius serves as a curvature measure due to the already mentioned relation $\kappa = \frac{1}{r}$. Figure 6.7 shows the comparison of our method with the algebraic and geometric fitting method for three test curves with and without noise. We measured the absolute average error over all curve points for different window sizes as

$$E_W(\mathbf{l}) = \sum_{i=1}^{|\mathbf{l}|} |\kappa(\mathbf{l}_i) - \tilde{\kappa}_W(\mathbf{l}_i)| \quad (6.69)$$

where $\kappa(\mathbf{l}_i)$ denotes the ground truth curvature and $\tilde{\kappa}_W(\mathbf{l}_i)$ the estimated curvature of the curve at \mathbf{l}_i . It turns out that our novel method converges to the true radius of curvature in the case of the assumed signal model, a circle. Note that the figures show the absolute average error over all curve points depending on the window size. Since the curvature is not constant for all curve points in the case of the ellipse and the spiral there are always points on the curve for which the given window size provides a correct estimate but also points for which the given window size provides incorrect results, affecting the overall average error. Therefore there is no convergent behavior with respect to the average error for these curve types. But actually this is also a problem of the estimators we compared our method to. Notice that our estimator behaves like the standard estimators and delivers comparable results.

6.3.2. Ridge curvature estimation for digital images

The curvature estimation based on the conformal monogenic signal has the advantage that it is not limited to digital curves but can also be applied to images, where the curves are not known in advance, e.g. isophotes in gray-scale images. In these cases the curvature is often supposed to serve as a feature indicating corners or straight line segments in the case of high or low curvature. The standard method to obtain the isophote curvature in digital images uses first and second order derivatives. The conformal curvature estimator is supposed to be compared to the standard differential geometric estimator given by

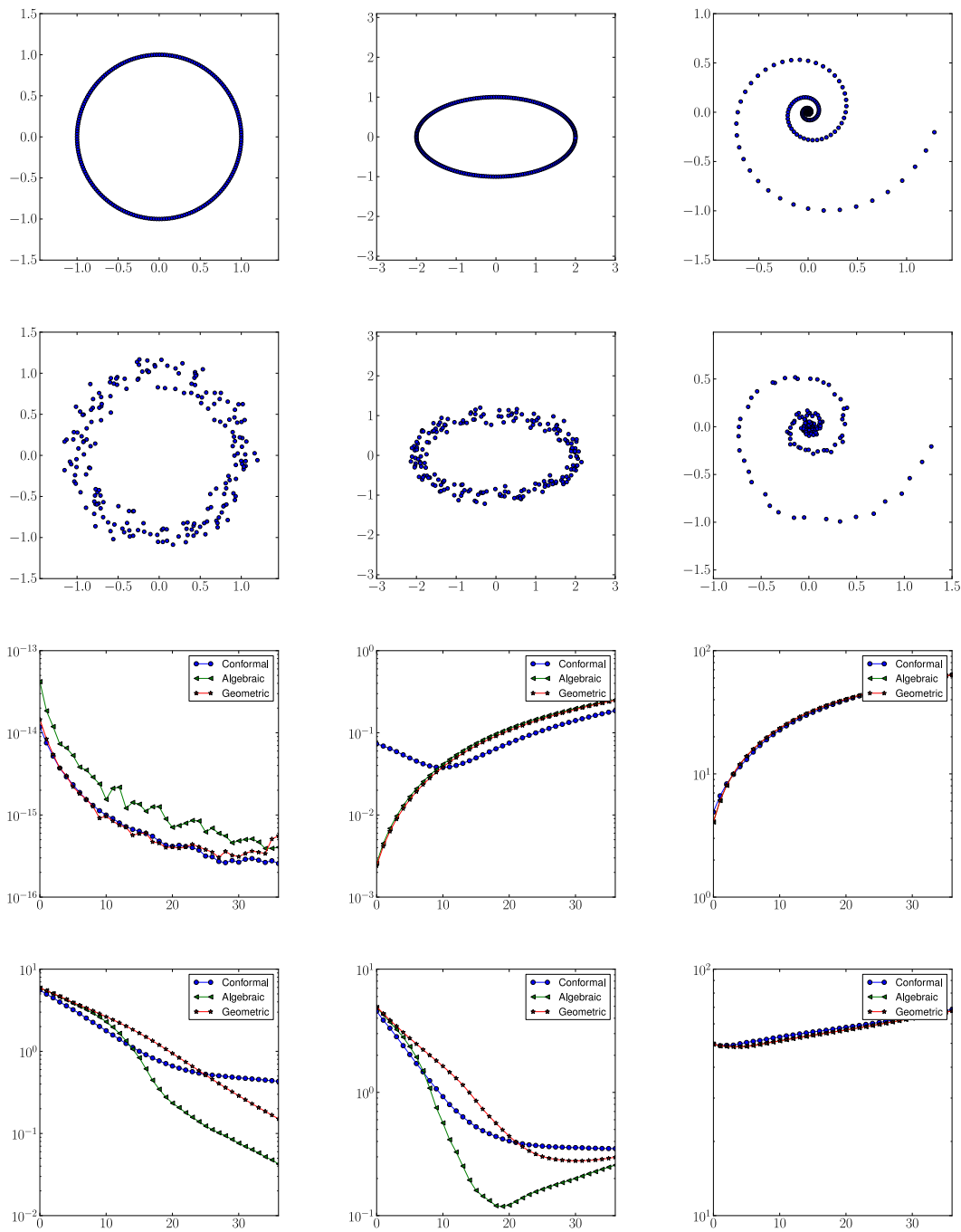


Figure 6.7.: First row: Test curves sampled at 200 points. Second row: Gaussian white noise perturbed test curves, $\sigma = 0.1$. Third and fourth row: Average absolute curvature errors E_W over all curve points depending on the window size W (abscissa) without and with noise (see also eq. (6.69)).

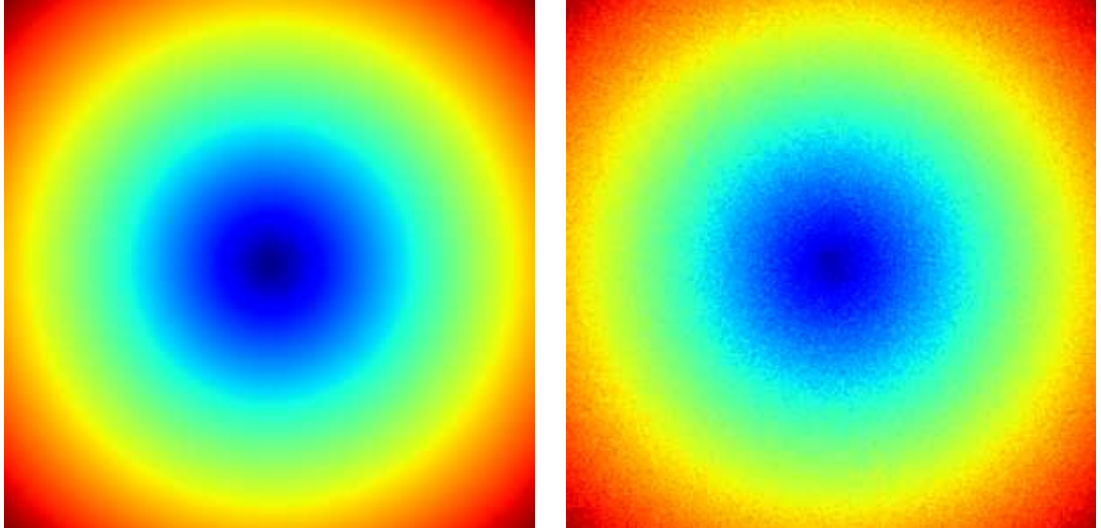


Figure 6.8.: Left : Ground truth image $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$. Right : Ground truth image perturbed with Gaussian white noise, $\sigma = 0.1$

$$\kappa = \frac{-\frac{\partial^2}{\partial x^2} f \left(\frac{\partial}{\partial y} f \right)^2 + 2 \frac{\partial}{\partial x} f \frac{\partial}{\partial y} f \frac{\partial^2}{\partial x y} f - \frac{\partial^2}{\partial y^2} f \left(\frac{\partial}{\partial x} f \right)^2}{\left(\left(\frac{\partial}{\partial x} f \right)^2 + \left(\frac{\partial}{\partial y} f \right)^2 \right)^{\frac{3}{2}}}. \quad (6.70)$$

Since natural images are often disturbed by noise, we further consider the input images in a scale-space framework, which smooths the images and suppresses the noise. The images are considered in the Gaussian scale-space with scale parameter s_G as

$$f(\mathbf{x}; s_G) = \mathcal{G}_{s_G} f(\mathbf{x}). \quad (6.71)$$

Both operators, the conformal monogenic signal and the classical derivatives in the plane, are applied to the Gaussian smoothed images at the corresponding scale. For the kernels of the conformal monogenic signal, the square $[-9, 9] \times [-9, 9]$ is sampled by 3×3 pixels with scale parameter $s = 2$. Again the reader is reminded, that the scale parameter for the conformal monogenic signal does not induce a scale space and is not to be confused with the scale-space embedding $\mathcal{G}_{s_G} f(\mathbf{x})$ of the input signals.

The signal

$$f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2} \quad (6.72)$$

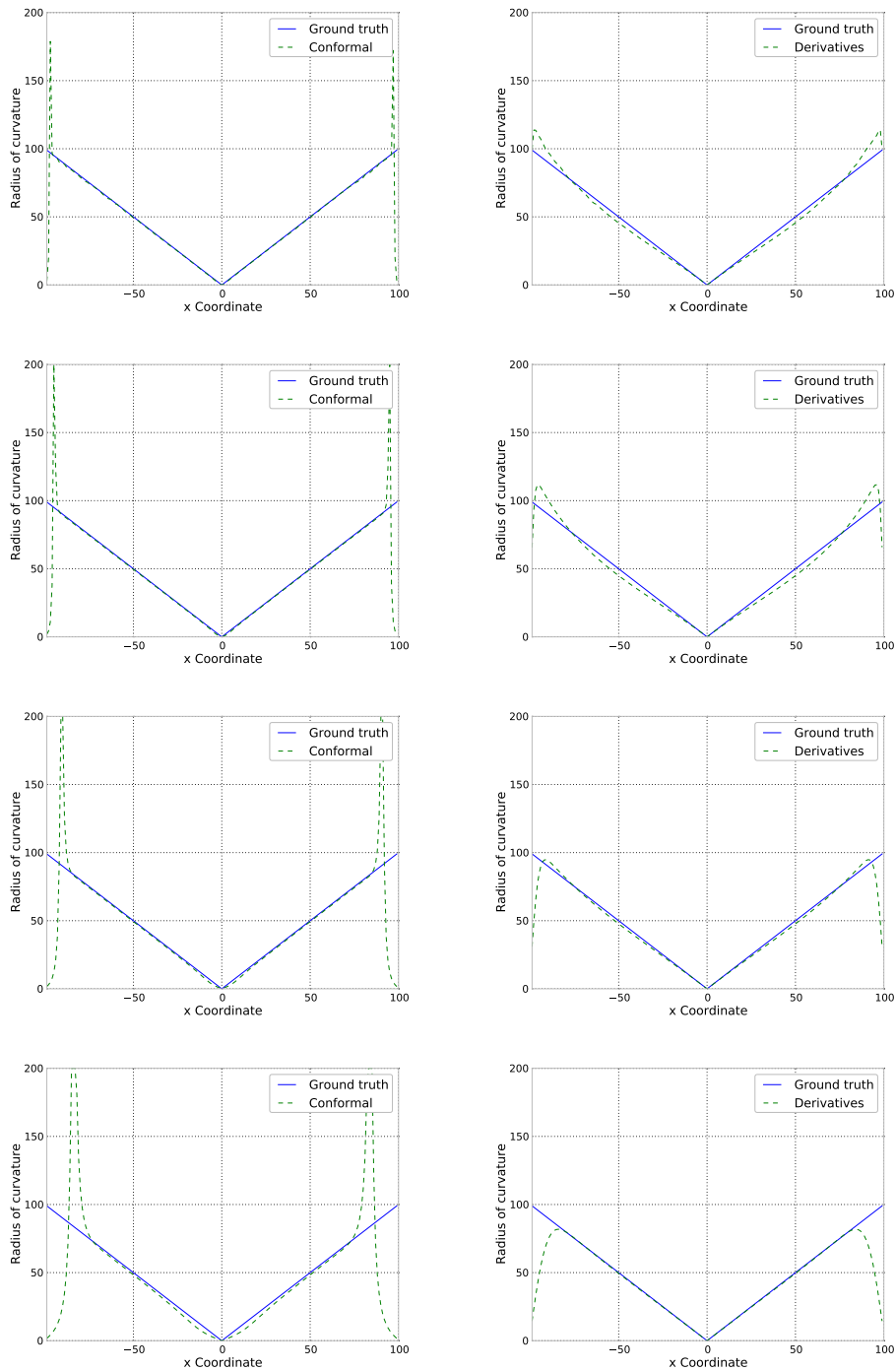


Figure 6.9.: Estimated radius of curvature for the signal $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$ with the slice $f(x_1, 0)$ with white noise added, SNR 100dB. Left column: Novel conformal estimation method vs. ground truth at scales $s_G = 1, 2, 4, 8$. Right column: Classical differential geometric method vs. ground truth at the same scales.

6. ANALYSIS OF CURVED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN THE PLANE

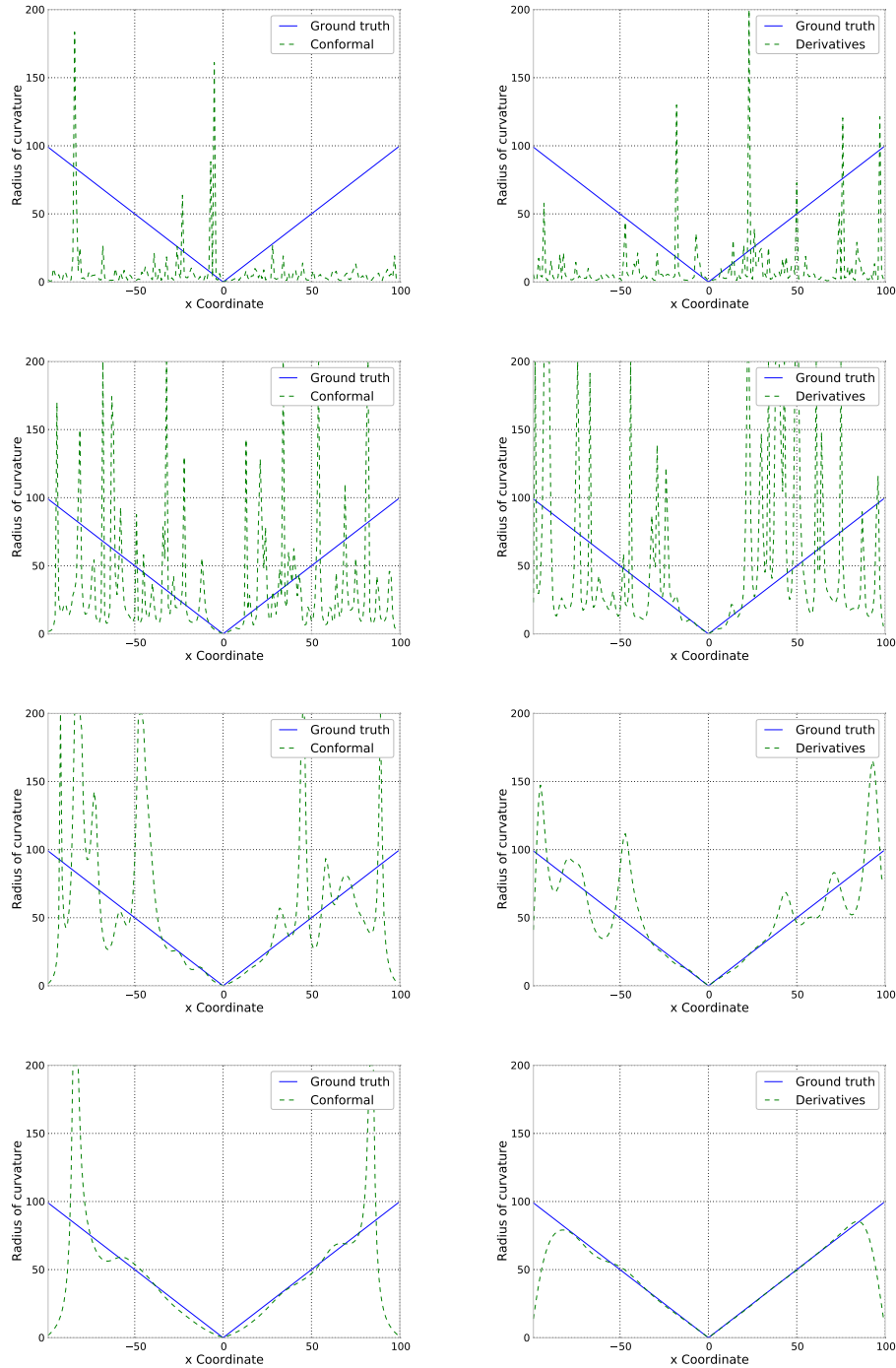


Figure 6.10.: Estimated radius of curvature for the signal $f(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}$ along the slice $f(x_1, 0)$ with white noise added, SNR 25dB. Left column: Novel conformal estimation method vs. ground truth at scales $s_G = 1, 2, 4, 8$. Right column: Classical differential geometric method vs. ground truth at the same scales.

serves as the ground truth signal for the evaluation of the estimators. For visualization purposes, we evaluate the accuracy in terms of the radius of curvature, which is just $r = 1/\kappa$. Since the ground truth signal depends linearly on the distance from the origin, its ground truth radius of curvature reads

$$r(\mathbf{x}) = \sqrt{x_1^2 + x_2^2}. \quad (6.73)$$

Figure 6.8 shows the test signal with and without noise. Figure 6.9 shows the estimated radius of curvature at scales $s_G \in \{1, 2, 4, 6\}$ for the ground truth signal f perturbed by Gaussian white noise, such that the signal to noise ratio yields 100dB. The curvature estimation is plotted along the horizontal slice $f(x_1, 0)$. The inaccuracies at the borders result from the border effects of the convolutions. Figure 6.10 shows the results for a signal to noise ratio of 25dB. As the scale increases, both estimators converge to the true radius of curvature. Figure 6.11 shows the average absolute error of the radius of curvature over the whole image for the conformal and the derivative based method for the continuum of scales $s_G \in [1, 8]$ with signal to noise ratios of 100dB, 50dB and 25dB. It justifies our proposed method as a robust and accurate curvature estimator. While the error of the conformal method is significantly lower in the case of low noise, the error of the two methods is almost identical for increasing noise at all scales. The results justify the conformal method as a valid alternative to the classical derivative based approach.

Another important aspect of the isophote curvature information pointed out in Romeny (Ed.) (1994), is the ability to obtain the ridge curves of an image. The ridge curves are the isophotes for which the gradient vanishes such that the curvature obtained by eq. (6.70) is degenerate. Due to their invariance properties concerning translation, rotation and monotonic intensity changes, ridges serve as a useful feature, especially if their evolution is considered across multiple scales. Figure 6.12 shows the ridge curves (degeneracies of the isophote curvature) of the Lenna test-image obtained by the conformal method and the curvature obtained according to the classical method using convolutions with Gaussian derivative kernels across different scales. Based on the results we conclude that our proposed estimator delivers results comparable to the standard method and therefore provides practical alternative access to the problem of isophote curvature estimation.

6. ANALYSIS OF CURVED INTRINSICALLY ONE-DIMENSIONAL SIGNALS IN THE PLANE

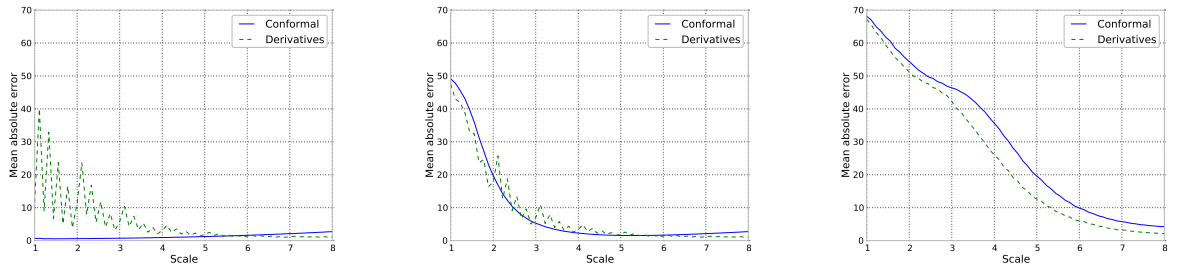


Figure 6.11.: Average absolute radius of curvature errors for the signal in Figure 6.8 perturbed by white noise with signal to noise ratios of 100dB (left), 50dB (middle), 25dB (right) for the conformal and the classical derivative based method. The errors are calculated with respect to Gaussian smoothed versions of the input image with scales depicted on the x-axis.

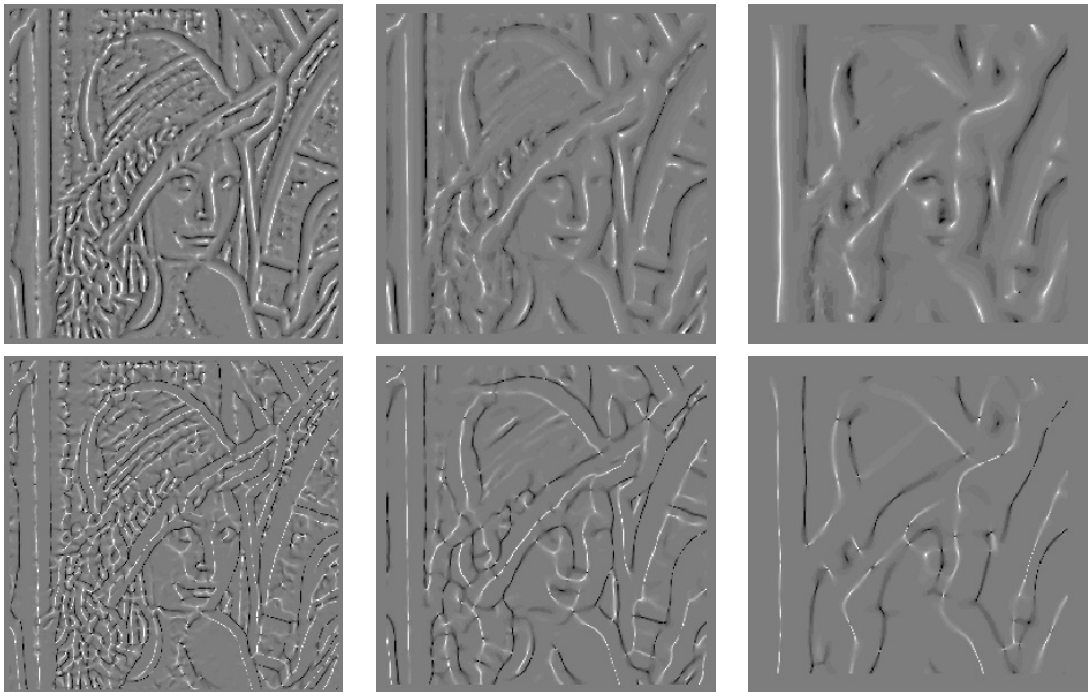


Figure 6.12.: Top row: Isophote curvatures calculated with the novel conformal method scales $s_G = 2, 4, 8$. Bottom row: Isophote curvatures calculated with the classical method in eq. (6.70) at scales $s_G = 2, 4, 8$.

6.4. Outlook: The three-dimensional conformal monogenic signal

While the previous chapter focused on the analysis of curved signal structures in the Euclidean plane using the conformal monogenic signal, this section is supposed to provide an outlook on how the proposed signal model can be generalized to three dimensions. It serves as a novel tool for volume and image-sequence processing.

The curved and sinusoidal curved intrinsically one-dimensional signals are defined analogously to the corresponding signal models in the plane as

$$f_{\mathbf{m}}(\mathbf{x}) = \tilde{f}(\|\mathbf{x} - \mathbf{m}\|) \quad (6.74)$$

and

$$f_{\mathbf{m}}(\mathbf{x}) = A \cos(k \|\mathbf{x} - \mathbf{m}\| + \phi) \quad (6.75)$$

with the difference, that the signals are now defined in \mathbb{R}^3 such that $\mathbf{x}, \mathbf{m} \in \mathbb{R}^3$ and

$$\mathbf{m} = \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = r_{\mathbf{m}} \begin{pmatrix} \sin(\theta_{m,1}) \cos(\theta_{m,2}) \\ \sin(\theta_{m,1}) \sin(\theta_{m,2}) \\ \cos(\theta_{m,1}) \end{pmatrix}. \quad (6.76)$$

These signals are spherical signals only depending on the distance from \mathbf{x} to \mathbf{m} . In contrast to the two-dimensional case, the isophotes $\gamma_{\mathbf{m}}(\mathbf{x})$ of these signals are constituted by spheres in \mathbb{R}^3 around \mathbf{m} . In order to apply concepts from the previous sections, the three-dimensional signal is projected to the hypersphere

$$\mathbb{S}^3 = \left\{ \mathbf{u} \in \mathbb{R}^4 \mid \sum_{i=1}^3 u_i^2 + (u_4 - \frac{1}{2})^2 = (\frac{1}{2})^2 \right\} \quad (6.77)$$

using the inverse stereographic projection $\mathcal{S}^{-1} : \mathbb{R}^3 \rightarrow \mathbb{S}^3$ given by

$$\mathcal{S}^{-1}(\mathbf{x}) = \frac{1}{1 + \sum_{i=1}^3 x_i^2} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \sum_{i=1}^3 x_i^2 \end{pmatrix} \quad (6.78)$$

for $\mathbf{x} \in \mathbb{R}^3$. According to the two dimensional case, the projection of the original signal $f_{\mathbf{m}}$ is identified with a plane wave in \mathbb{R}^4 given by

$$\psi(\mathbf{u}) = \tilde{f}(\langle \mathbf{n}, \mathbf{u} \rangle + \|\mathbf{m}\|) \quad (6.79)$$

where $\mathbf{n} \in \mathbb{R}^4$, $\|\mathbf{n}\| = 1$ denotes the orientation of the plane wave.

In analogue to the two-dimensional case, curved signal structures are supposed to be analyzed in terms of the orientation vector of ψ . The orientation vector is obtained from the approximations of the generalized Hilbert transform or the gradient in \mathbb{R}^4 . The type of curved signals which can be analyzed, is again determined by the geometric entities represented by the intersection of hyperplanes with the sphere. Compared to the two-dimensional case, the hyperplanes are now three-dimensional. While the intersections of planes and \mathbb{S}^2 were circles in the two-dimensional case, we now deal with spheres contained in the surface of the hypersphere \mathbb{S}^3 passing through the south-pole of \mathbb{S}^3 . Since we chose \mathbb{S}^3 such that its south-pole coincides with the origin of \mathbb{R}^4 , these hyperplanes pass through the origin in \mathbb{R}^4 . We are therefore able to characterize these hyperplanes by just using their normal vectors.

Conformal
monogenic
signal in \mathbb{R}^3

Definition 6.12 (Conformal monogenic signal in \mathbb{R}^3).

Let p_s and $q_{i,s}$ denote the Poisson and conjugate Poisson kernels in \mathbb{R}^4 with respect to the scale parameter $s > 0$ and denote by $\tilde{p}_s, \tilde{q}_{i,s}$ the corresponding (mean corrected) kernels in \mathbb{R}^3 (cmp. (6.48)). Then the signal representation

$$\mathbf{f}_m(\mathbf{x}) = \begin{pmatrix} (f_m * \tilde{p}_s)(\mathbf{x}) \\ (f_m * \tilde{q}_{1,s})(\mathbf{x}) \\ (f_m * \tilde{q}_{2,s})(\mathbf{x}) \\ (f_m * \tilde{q}_{3,s})(\mathbf{x}) \\ (f_m * \tilde{q}_{4,s})(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \tilde{\mathcal{P}}_s f_m(\mathbf{x}) \\ \tilde{\mathcal{Q}}_{1,s} f_m(\mathbf{x}) \\ \tilde{\mathcal{Q}}_{2,s} f_m(\mathbf{x}) \\ \tilde{\mathcal{Q}}_{3,s} f_m(\mathbf{x}) \\ \tilde{\mathcal{Q}}_{4,s} f_m(\mathbf{x}) \end{pmatrix} \quad (6.80)$$

is called the *conformal monogenic signal* in \mathbb{R}^3 .

We are therefore able to extract the direction angles of the original signal f_m in the three-dimensional space at $\mathbf{x} \in \mathbb{R}^3$ as

$$\theta_{m,1}(\mathbf{x}) = \arctan \left(\frac{\tilde{\mathcal{Q}}_{2,s} f_m(\mathbf{x})}{\tilde{\mathcal{Q}}_{1,s} f_m(\mathbf{x})} \right) \quad (6.81)$$

$$\theta_{m,2}(\mathbf{x}) = \arctan 2 \left(\sqrt{\tilde{\mathcal{Q}}_{1,s} f_m(\mathbf{x})^2 + \tilde{\mathcal{Q}}_{2,s} f_m(\mathbf{x})^2}, \tilde{\mathcal{Q}}_{3,s} f_m(\mathbf{x}) \right). \quad (6.82)$$

The curvature measure which corresponds to the inverse radius of the spherical signal is obtained as

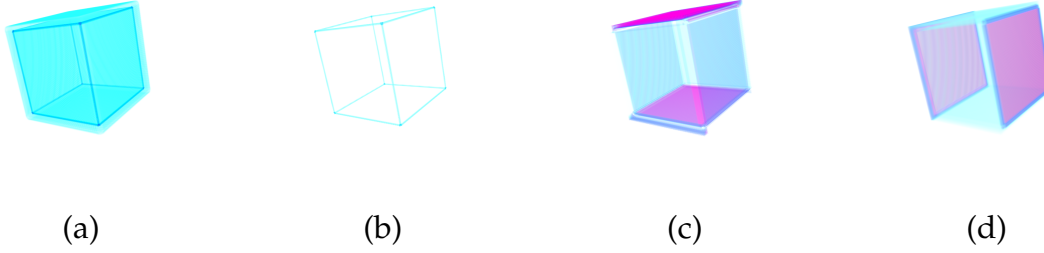


Figure 6.13.: From left to right: Volume rendering of a cube. (a) Regions with low curvature (high radius of curvature). (b) Regions with high curvature (low radius of curvature). (c) Direction with respect to $\theta_{m,2}$. (d) Direction with respect to $\theta_{m,1}$.

$$\kappa_m(\mathbf{x}) = \frac{2\tilde{Q}_{4,s}f_m(\mathbf{x})}{\sqrt{\sum_{i=1}^3 \tilde{Q}_{i,s}f_m(\mathbf{x})^2}} = \frac{2}{\tan \varphi_m(\mathbf{x})}. \quad (6.83)$$

In the case of a sinusoidal curved intrinsically one-dimensional signal the amplitude and local phase are given by

$$A(\mathbf{x}) = \sqrt{\tilde{P}_s f_m(\mathbf{x})^2 + \sum_{i=1}^4 \tilde{Q}_{i,s} f_m(\mathbf{x})^2} \quad (6.84)$$

$$\phi(\mathbf{x}) = \arctan 2 \frac{\sqrt{\sum_{i=1}^4 \tilde{Q}_{i,s} f_m(\mathbf{x})^2}}{\tilde{P}_s f_m(\mathbf{x})}. \quad (6.85)$$

The isophote curvature in the three dimensional case represents the curvature of the sphere locally approximating the signal at \mathbf{x} . Therefore, it is possible to extract the curvature information for all surfaces which may locally be approximated by a sphere. The conformal monogenic signal in \mathbb{R}^3 turns out to be useful in the analysis of volume images. Figure 6.13 shows the features of sample volume rendering of a cube. Image (a) shows the regions of the volume rendering where the curvature is low. These are planar regions, since planes correspond to spheres with infinity radius of curvature. Due to the relation $\kappa = \frac{1}{r}$ it follows that the curvature of a plane tends to 0. This can be considered as the analogue to the two-dimensional case, where straight lines represent circles with infinite radius. Image (b) shows the regions with high curvature, which are regions which can be approximated by spheres with low radius. In general this

measure of curvature allows us to distinguish between planar surfaces, edges and corners in three-dimensional images. Furthermore, as shown in Figures 6.13 (c) and (d), we are able to obtain the direction of these structures with respect to the angles $\theta_{m,1}, \theta_{m,2}$.

6.5. Summary

In this chapter, a new fundamental idea for locally analyzing curved and straight signals in one unified framework has been presented. The problem of analyzing standard intrinsically one-dimensional signals and curved signals in one framework can be solved by embedding n -dimensional signals in $(n + 1)$ -dimensional conformal spaces using the corresponding stereographic projections and their inverse in which the $(n + 1)$ -dimensional generalized Hilbert transform or gradient allows for an easy estimation of the model parameters. Local signal features such as local amplitude, phase, orientation/direction and curvature can be determined in spatial domain by simple convolution. The *conformal monogenic signal* can be computed efficiently and can be easily implemented into existing low level signal processing steps of any application. We gave formal as well as experimental proofs of our results. The *conformal monogenic signal* illustrates the natural relation of the original image domain to geometric entities such as lines, circles, hyperplanes and hyperspheres.

Future applications might use the conformal monogenic signal for the analysis of volume-images as well as image sequences. In volume image processing they may be used for preprocessing, feature detection and image registration. Possible applications in the field of image sequence analysis include the application of the conformal monogenic signal as optical flow constraints.

7. Analysis of intrinsically one-dimensional signals on the two-sphere

The previous two chapters focused on the analysis of image signals in the Euclidean plane. In this chapter, signals on the unit sphere in \mathbb{R}^3 will be analyzed. Scientific disciplines like geophysics, astrophysics and computer vision have to deal with input signals which are naturally defined on the unit sphere. These signals are for instance seismic or gravitational data captured around the earth (Crotwell (2000)), satellite data (Freeden and Hesse (2002)), cosmic microwave background (Mcewen et al. (2007)) or omnidirectional images captured by a catadioptric camera (Bogdanova et al. (2007), Makadia et al. (2004)). The signals are supposed to be analyzed in terms of the generalized Hilbert transform.

Unfortunately, the definition provided in chapter 4 only holds in the Euclidean space, not on a manifold like the two-sphere. It is therefore mandatory to introduce a more general definition of the generalized Hilbert transform. We will construct a generalization of the analytic signal on the unit sphere \mathbb{S}^2 with the help of the Hilbert and the Cauchy transforms on \mathbb{S}^2 , which have been introduced in the framework of Clifford analysis. In this mathematical framework, it can be shown that the Hilbert transform arises as the non-tangential boundary value of the *Cauchy transform*. The Cauchy transform will be introduced for arbitrary closed smooth subsets of the Euclidean space \mathbb{R}^n . Since the unit ball belongs to the class of these subsets, the Cauchy transform for the unit ball can be used to derive the generalized Hilbert transform on the unit sphere.

From the Euclidean space we know that the generalized Hilbert transform has an intuitive interpretation in the Fourier domain. We will derive an analogue interpretation with respect to the Fourier series expansion on the sphere, using the spherical harmonics as the basis functions. We derive and discuss the filter kernels for our new signal model and give an important characterization in the spherical harmonic domain in terms of their spherical harmonic coefficients. While the scalar zonal Poisson or Abel-Poisson kernel is a standard tool in multiscale and texture analysis on \mathbb{S}^2 and $SO(3)$ (Freeden et al. (1998); Schaeben (1996); Schaeben and van den Boogaart (2003); Bernstein et al. (2009)), we investigate its harmonic conjugate function arising in the Clifford algebra embedding. The key to the spherical harmonic expansion is the identification of the Cauchy transform on \mathbb{S}^2 with a directional correlation introduced by Wiaux et al. (2006). With the spectral characterization and the resulting interpretation we demonstrate the obtained scale space in Section 7.5 and analyze the results of the obtained kernels acting on the real part of a plane wave in \mathbb{R}^3 restricted to the sphere.

7.1. Clifford algebra

Clifford algebra

Throughout this chapter we work in the universal Clifford algebra $\mathbb{R}_{0,n}$ over the vector space $\mathbb{R}^{0,n}$ which is the real vector space \mathbb{R}^n equipped with a non-degenerate quadratic signature $(0, n)$. For an orthonormal basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ of $\mathbb{R}^{0,n}$ the following multiplication rules arise in $\mathbb{R}_{0,n}$:

$$\mathbf{e}_i \mathbf{e}_j + \mathbf{e}_j \mathbf{e}_i = -2\delta_{ij}, \quad \forall i, j \in \{1, \dots, n\} \quad (7.1)$$

In the following the notation $\mathbf{e}_i \mathbf{e}_j = \mathbf{e}_{ij}$ is used. For sets $A = \{i_1, \dots, i_h\} \subseteq \{1, \dots, n\}$ with $1 < i_1 < \dots < i_h < n$ and $\mathbf{e}_A = \mathbf{e}_{i_1} \mathbf{e}_{i_2} \dots \mathbf{e}_{i_h}$ the basis $(\mathbf{e}_A : A \subseteq \{1, \dots, n\})$ forms a basis for the Clifford algebra $\mathbb{R}_{0,n}$. Therefore an element $\mathbf{a} \in \mathbb{R}_{0,n}$ allows the representation $\mathbf{a} = \sum_{k=0}^n [\mathbf{a}]_k$ where $[\mathbf{a}]_k = \sum_{|A|=k} \mathbf{a}_A \mathbf{e}_A$ is called k -vector. Vectors $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ are identified with one-vectors $\mathbf{x} = \sum_{j=1}^n x_j \mathbf{e}_j$. The product of two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}_{0,n}$ is then defined by

$$\mathbf{x}\mathbf{y} = -\langle \mathbf{x}, \mathbf{y} \rangle + \mathbf{x} \wedge \mathbf{y} \quad (7.2)$$

where the inner product

$$\mathbf{x} \bullet \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n x_i y_i = -\frac{1}{2}(\mathbf{x}\mathbf{y} + \mathbf{y}\mathbf{x}) \quad (7.3)$$

results in a scalar and the wedge product or outer product

$$\mathbf{x} \wedge \mathbf{y} = \sum_{i < j} \mathbf{e}_i \mathbf{e}_j (x_i y_j - x_j y_i) = \frac{1}{2}(\mathbf{x}\mathbf{y} - \mathbf{y}\mathbf{x}). \quad (7.4)$$

results in a two-vector, which is also known as a bivector. The conjugation in $\mathbb{R}_{0,n}$ is given by $\overline{\mathbf{e}_i} = -\mathbf{e}_i$ and therefore the conjugation of a vector \mathbf{x} results in $\overline{\mathbf{x}} = -\mathbf{x}$.

In the following \mathbb{S}^2 denotes the unit sphere in \mathbb{R}^3 and \mathbb{B}^2 the unit ball, respectively. $\boldsymbol{\xi}, \boldsymbol{\omega}$, will be elements of \mathbb{S}^2 described by the angles (θ, φ) and (α, β) where θ, α are zenithal angles and φ, β are azimuthal angles. In Cartesian coordinates they are written as $\boldsymbol{\xi} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)^\top$ and $\boldsymbol{\omega} = (\sin \alpha \cos \beta, \sin \alpha \sin \beta, \cos \alpha)^\top$. Elements in \mathbb{B}^2 are denoted by \mathbf{x}, \mathbf{y} with $\mathbf{x} = r\boldsymbol{\xi}$, $\mathbf{y} = r\boldsymbol{\omega}$ and $0 < r < 1$ such that $\mathbf{x} = (r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta)^\top$ and $\mathbf{y} = (r \sin \alpha \cos \beta, r \sin \alpha \sin \beta, r \cos \alpha)^\top$. We identify the north pole $\boldsymbol{\eta}$ with $(0, 0, 1)^\top$. Furthermore dS will denote the surface measure on \mathbb{S}^2 and A_3 denotes the surface area of \mathbb{S}^2 .

7.2. Basic Fourier analysis on the sphere

7.2.1. Spherical harmonics

In order to expand a function in $L^2(\mathbb{S}^2)$ into a series, we introduce an orthonormal basis for functions in $L^2(\mathbb{S}^2)$:

Definition 7.1 (Standard spherical harmonics, Wiaux et al. (2006)).
The standard spherical harmonics are defined as

Spherical
harmonics

$$Y_{l,m}(\theta, \varphi) = \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos \theta) e^{im\varphi} \quad (7.5)$$

where $P_l^m(x)$ are the associated Legendre functions

$$P_l^{(m)}(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} (P_l(x)) \quad (7.6)$$

and $P_l(x)$ are the Legendre polynomials given by the Rodrigues' formula

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} [(x^2-1)^l]. \quad (7.7)$$

The spherical harmonics form an orthonormal basis for the functions in $L^2(\mathbb{S}^2)$. Every function $f \in L^2(\mathbb{S}^2)$ can be expanded into a Fourier series as

$$f(\boldsymbol{\omega}) = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \hat{f}_{l,m} Y_{l,m}(\theta, \varphi) \quad (7.8)$$

where the Fourier coefficients $\hat{f}_{l,m}$ are obtained as

$$\hat{f}_{l,m} = \int_0^{2\pi} \int_0^\pi f(\theta, \varphi) \overline{Y_{l,m}(\theta, \varphi)} \sin \theta d\theta d\varphi. \quad (7.9)$$

We will call the index l the frequency of a spherical harmonic and the index m the order.

In the Euclidean plane, the convolution operation is based on the multiplication of a translated filter kernel with the target function. On the sphere a translation corresponds to a rotation described by two angles. To expand convolution operations on the sphere into a series, we need an orthonormal basis for rotations which are representations of the group $SO(3)$.

7.2.2. Wigner-D functions

Definition 7.2 (Wigner-D functions, Wiaux et al. (2006)).

Wigner-D
functions

While the spherical harmonics form an orthonormal basis for square integrable functions on the sphere, the Wigner-D functions

$$D_{m,n}^l(\boldsymbol{\rho}) = D_{m,n}^l(\varphi, \theta, \psi) = e^{-im\varphi} d_{m,n}^l(\theta) e^{-in\psi} \quad (7.10)$$

with

$$d_{m,n}^l(\theta) = \sum_{k=L_1}^{L_2} \frac{(-1)^k ((l+m)!(l-m)!(l+n)!(l-n)!)^{(1/2)}}{(l+m-k)!(l-n-k)!k!(k+n-m)!} \cos(\theta/2)^{2l+m-n-2k} \sin(\theta/2)^{2k+n-m} \quad (7.11)$$

and $L_1 = \max(0, m-n)$, $L_2 = \min(l+m, l-n)$ form an orthonormal basis for square integrable functions $f \in L^2(SO(3))$ on the rotation group $SO(3)$ with respect to the standard Haar measure. Hence, the Fourier series expansion of a function $f \in L^2(SO(3))$ with $\boldsymbol{\rho} \in SO(3)$ reads (see e.g. (Wiaux et al. (2006)))

$$f(\boldsymbol{\rho}) = \sum_{l \in \mathbb{N}} \frac{2l+1}{8\pi^2} \sum_{m=-l}^l \sum_{n=-l}^l \hat{f}_{m,n}^l \overline{D_{m,n}^l(\boldsymbol{\rho})}. \quad (7.12)$$

with the Fourier coefficients $\hat{f}_{m,n}^l$ given by

$$\hat{f}_{m,n}^l = \int_{SO(3)} f(\boldsymbol{\rho}) D_{m,n}^l(\boldsymbol{\rho}) d\boldsymbol{\rho}. \quad (7.13)$$

7.3. The Cauchy and the Hilbert transform for subsets of \mathbb{R}^n

The generalized Hilbert transform in \mathbb{R}^n has already been defined and used in the previous chapters. It can be derived as the non-tangential boundary value of the *Cauchy transform*. The classical Cauchy integral formula is well known from complex analysis (see e.g. Needham (1997)). Its generalization in the sense of Clifford analysis working in the real Clifford algebra $\mathbb{R}_{0,n}$ is given by

Cauchy transform **Definition 7.3** (Cauchy transform in $\mathbb{R}_{0,n}$, Delanghe (2004)). Let $G \subseteq \mathbb{R}^n$ with smooth boundary ∂G , $x \in G$ and $f \in L^2(\partial G)$. Then the Cauchy transform of f in $\mathbb{R}_{0,n}$ is defined as

$$\begin{aligned}\mathcal{C}[f](\mathbf{x}) &= \frac{2}{A_n} \int_{\partial G} E(\mathbf{x} - \mathbf{y}) n(\mathbf{y}) f(\mathbf{y}) dS(\mathbf{y}) \\ &= \frac{2}{A_n} \int_{\partial G} \frac{\mathbf{x} - \mathbf{y}}{\|\mathbf{x} - \mathbf{y}\|^n} n(\mathbf{y}) f(\mathbf{y}) dS(\mathbf{y})\end{aligned}\quad (7.14)$$

with dS as the surface element of ∂G , $E(\cdot)$ the Cauchy kernel, A_n the surface area of ∂G and $n(\mathbf{y})$ the outward pointing unit normal at \mathbf{y} . The Cauchy transform generates a monogenic function in G which solves the differential equation $\partial \mathcal{C}[f](\mathbf{x}) = 0$, where ∂ is the Dirac operator. Additionally the Cauchy transform splits into the Poisson transform and conjugate Poisson transform of f (Brackx et al. (2006)):

$$\mathcal{C}[f](\mathbf{x}) = \frac{1}{2} \mathcal{P}[f](\mathbf{x}) + \frac{1}{2} \mathcal{Q}[f](\mathbf{x}) \quad (7.15)$$

with non-tangential boundary values of $\mathcal{P}[f](\mathbf{x})$ and $\mathcal{Q}[f](\mathbf{x})$ given by

$$\lim_{\text{n.t. } \mathbf{x} \rightarrow \boldsymbol{\xi}} \mathcal{P}[f](\mathbf{x}) = f(\boldsymbol{\xi}), \quad \lim_{\text{n.t. } \mathbf{x} \rightarrow \boldsymbol{\xi}} \mathcal{Q}[f](\mathbf{x}) = \mathcal{H}[f](\boldsymbol{\xi}). \quad (7.16)$$

We can therefore define monogenic signals for any closed subset $G \subseteq \mathbb{R}^n$ with smooth boundary in \mathbb{R}^n as the non-tangential boundary value of the Cauchy transform on that surface. Furthermore, the Cauchy transform provides a natural embedding in a scale-space concept, the Poisson scale space. To emphasize the relationship between the Cauchy transform and the analytic and monogenic signal representations we give some examples:

Example 7.4 ($G = \mathbb{R}_+^2$, $\partial G = \mathbb{R}$). In this case we obtained the classical analytic signal on the real line (Gabor (1946)) with $\mathbf{x} = (x_0, x_1)$, $x_0 > 0$:

$$\lim_{\text{n.t. } \mathbf{x} \rightarrow (0, x_1)} \mathcal{C}[f](\mathbf{x}) = f(\mathbf{x}) + i\mathcal{H}[f](\mathbf{x}) = f_a(\mathbf{x}). \quad (7.17)$$

Example 7.5 ($G = \mathbb{R}_+^3$, $\partial G = \mathbb{R}^2$). This case represents the monogenic signal (Felsberg and Sommer (2001)) with $\mathbf{x} = (x_0, x_1, x_2)$, $x_0 > 0$ and the classical Riesz transforms (Stein (1971)) R_{x_i} :

$$\lim_{\text{n.t. } \mathbf{x} \rightarrow (0, x_1, x_2)} \mathcal{C}[f](\mathbf{x}) = f(\mathbf{x}) + R_{x_1}[f](\mathbf{x})e_1 + R_{x_2}[f](\mathbf{x})e_2 = f_M(\mathbf{x}). \quad (7.18)$$

7.4. The Cauchy and the Hilbert transform on the sphere

We have seen that the Cauchy transform is well defined for subsets $G \subseteq \mathbb{R}^n$ with smooth boundary ∂G . The unit sphere \mathbb{S}^2 in \mathbb{R}^3 is one of these surfaces and therefore it allows the construction of a monogenic signal with the help of the Cauchy transform. In the following we consider functions $f \in L^2(\mathbb{S}^2)$. We call such functions signals, which are band-limited such that there exists a maximum frequency $L \in \mathbb{N}$ with $\hat{f}_{l,m} = 0$ for $l \geq L$. In order to construct an analogue to the analytic and monogenic signal on \mathbb{S}^2 , we proceed in accordance to (7.15). Using the Cauchy transform on \mathbb{S}^2 we obtain a splitting into the Poisson and the conjugate Poisson transform, which is the natural embedding of a monogenic signal on \mathbb{S}^2 in the Poisson scale-space. Taking the non-tangential boundary values of the Cauchy transform and therefore of the Poisson and conjugate Poisson transform, we obtain the original function f and its Hilbert transform $\mathcal{H}[f]$ on \mathbb{S}^2 .

Cauchy
transform on \mathbb{S}^2

Definition 7.6 (Cauchy transform on \mathbb{S}^2 , Delanghe (2004)). Let $\mathbf{x} \in \mathbb{B}^2 \setminus \mathbb{S}^2$ with $\mathbf{x} = r\xi$, $\xi \in \mathbb{S}^2$, $0 < r < 1$ and $f \in L^2(\mathbb{S}^2)$. Then the Cauchy transform of f on \mathbb{S}^2 is defined as

$$\mathcal{C}[f](\mathbf{x}) = \frac{2}{\mathcal{A}_3} \int_{\mathbb{S}^2} E(\mathbf{x} - \boldsymbol{\omega}) \boldsymbol{\omega} f(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) = \frac{2}{\mathcal{A}_3} \int_{\mathbb{S}^2} \frac{\mathbf{x} - \boldsymbol{\omega}}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} \boldsymbol{\omega} f(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) \quad (7.19)$$

where $E(\cdot)$ is the Cauchy kernel on \mathbb{S}^2 .

According the Brackx et al. (2006) the Cauchy transform on \mathbb{S}^2 splits into the Poisson and conjugate Poisson transform as

$$\begin{aligned} \mathcal{C}[f](\mathbf{x}) &= \frac{1}{2} \mathcal{P}[f](\mathbf{x}) + \frac{1}{2} \mathcal{Q}[f](\mathbf{x}) \\ &= \frac{1}{2\mathcal{A}_3} \int_{\mathbb{S}^2} \frac{1 - \|\mathbf{x}\|^2}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} f(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) + \frac{1}{2\mathcal{A}_3} \int_{\mathbb{S}^2} \frac{1 + \|\mathbf{x}\|^2 + 2\mathbf{x}\boldsymbol{\omega}}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} f(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) \end{aligned} \quad (7.20)$$

In our analysis we want to be able to select certain frequency bands in the spectral spherical harmonic domain. Therefore, we consider lowpass filtered versions of our signal in a linear scale space in \mathbb{B}^2 . Given two lowpass filters we construct a bandpass filter as the difference of the two lowpass filters. In the case of the monogenic signal the scale space concept which arises naturally is the Poisson scale space in the upper half space \mathbb{R}_+^3 (Felsberg and Sommer (2004)). Analogously the scale space in this context is the Poisson scale space in \mathbb{B}^2 which arises naturally from the definition of the Hilbert transform being the non tangential boundary value of the Cauchy transform (Delanghe et al. (1992)). Therefore the Cauchy transform encompasses the

Hilbert transform in the light of the Poisson scale space in the unit ball. The parameter r acts as the scale parameter. The more r tends towards the origin the more blurred or smoothed versions of the original signal are obtained. As r moves towards 1 the original signal and its Hilbert transform are obtained. This leads to the following definition of a Hilbert transform on \mathbb{S}^2 :

Definition 7.7 (Hilbert transform on \mathbb{S}^2 , Delanghe (2004)). Let $\xi \in \mathbb{S}^2$ and $f \in L^2(\mathbb{S}^2)$. Then the Hilbert transform of f on \mathbb{S}^2 is defined as the Cauchy Principal Value integral Hilbert transform on \mathbb{S}^2

$$\mathcal{H}[f](\xi) = \frac{2}{A_3} \text{P.V.} \int_{\mathbb{S}^2} \frac{\xi - \omega}{\|\xi - \omega\|^3} \omega f(\omega) dS(\omega). \quad (7.21)$$

Due to (7.15) we can analyze the Cauchy transform in two steps by analyzing the Poisson and the conjugate Poisson transform separately. Since the Cauchy transform generates a monogenic function, its component functions are harmonic in \mathbb{S}^2 with

$$\Delta \mathcal{P}[f](\mathbf{x}) = 0, \quad \Delta \mathcal{Q}[f](\mathbf{x}) = 0 \quad (7.22)$$

where Δ is the Laplace operator in \mathbb{B}^2 .

7.4.1. The Cauchy transform on \mathbb{S}^2 as a directional correlation

It is desirable from a signal processing viewpoint to interpret the Cauchy transform as a filtering operation. In \mathbb{R}^3 linear filters are applied to signals by convolution. The convolution on the sphere is the integration over the group $SO(3)$ of the input signal and the filter, as proposed for example by Driscoll and Healy (1994). Since the Cauchy transform integrates over \mathbb{S}^2 it cannot immediately be identified with a spherical convolution operation. Nonetheless, there exists an operation called directional correlation introduced in Wiaux et al. (2006):

Definition 7.8 (Directional correlation). Let $\rho \in SO(3)$, $\mathcal{R}(\rho)$ the rotation operator associated with ρ and $h \in L^2(\mathbb{S}^2)$, $f \in L^2(\mathbb{S}^2)$. The directional correlation of h and f in the direction ρ is given by Directional correlation

$$\mathcal{R}(\rho)[h] \star f = \int_{\mathbb{S}^2} \overline{h(\mathcal{R}^{-1}(\rho) \omega)} f(\omega) dS(\omega). \quad (7.23)$$

The directional correlation turns out to be a function in $L^2(SO(3))$. We want to analyze signals on the sphere locally. Therefore we are not interested in the global Cauchy transform of a certain signal. Instead every time we evaluate the Cauchy transform at some point on the sphere, we treat this point as the north pole η of a sphere,

rotated to the evaluation point. The spherical coordinates (θ, φ) at which we evaluate the Cauchy transform describe a rotation $\rho = (\varphi, \theta, 0)$. So with respect to original coordinate system we always evaluate the Cauchy transform at the north pole of a sphere in a rotated coordinate system. In that sense we obtain a local analysis of our signal for every point. Furthermore, this is the key to the expression of the Cauchy transform as a directional correlation. Suppose we want to evaluate the Cauchy transform at some fixed point $\mathbf{x} = r\xi$. Now we choose this point \mathbf{x} as the north pole $\boldsymbol{\eta}$ of some rotated coordinate system. Then the Cauchy kernel $E(\mathbf{x} - \boldsymbol{\omega})$ is just a rotated version of the kernel $E(r\boldsymbol{\eta} - \boldsymbol{\omega})$ centered at the north pole $\boldsymbol{\eta}$. The evaluation of the Cauchy transform at a point $\mathbf{x} = r\xi \in \mathbb{B}^2$ with spherical coordinates (r, θ, φ) reduces to the evaluation of the Cauchy transform at $\boldsymbol{\eta}$ rotated by $\rho = (\varphi, \theta, 0)$. Thinking of the Cauchy kernel as a filter mask we always use the same filter mask centered at the north pole, but at every point we apply it to a rotated version of the original function. So for every point $\mathbf{x} = r\xi$ we do not actually evaluate the transform with the kernel $E(r\xi - \boldsymbol{\omega})$ but we evaluate the transform as

$$\mathcal{C}[f](r, \theta, \varphi) = \int_{\mathbb{S}^2} E(r\boldsymbol{\eta} - \boldsymbol{\omega}) \boldsymbol{\omega} f(\mathcal{R}^{-1}(\varphi, \theta, 0)\boldsymbol{\omega}) dS(\boldsymbol{\omega}) \quad (7.24)$$

$$= \mathcal{R}(\varphi, \theta, 0)[h] \star f. \quad (7.25)$$

7.4.2. Fourier series expansion of the Cauchy kernel

The next step towards an intuitive interpretation of the Hilbert transform of the sphere is the series expansion of the directional correlation. As we have noticed, the directional correlation is a function in $L^2(SO(3))$. The Cauchy transform has been interpreted as a directional correlation evaluated for rotations corresponding to the Euler angle representations $(\varphi, \theta, 0)$. In this case the directional correlation is also known as a standard correlation on \mathbb{S}^2 according to Wiaux et al. (2006). We introduced the Wigner-D functions as an orthonormal basis for $L^2(SO(3))$. As a consequence, the directional correlation admits an expansion into a Fourier series on $SO(3)$ in terms of the Wigner-D basis functions. For a series expansion of the Cauchy transform we are in the need of the series expansion coefficients. It has been shown in Wiaux et al. (2006) that the coefficients of the directional correlation can be evaluated in terms of the spherical harmonic coefficients of the function and the filter respectively as

$$[\widehat{\mathcal{R}[h] \star f}]_{m,n}^l = \overline{\widehat{h}_{l,n}} \widehat{f}_{l,m} \quad (7.26)$$

such that the series expansion of the directional correlation reads

$$\mathcal{R}(\boldsymbol{\rho})[\mathbf{h}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \sum_{n=-l}^l [\widehat{\mathcal{R}[\mathbf{h}] \star f}]_{m,n}^l \overline{D_{m,n}^l(\boldsymbol{\rho})}. \quad (7.27)$$

Since we want to provide a spectral characterization of the Poisson and conjugate Poisson filter kernels, we have to expand them into a spherical harmonic series. The directional correlation in terms of the spherical harmonic coefficients is now used to evaluate the Cauchy transform. We seek for the spherical harmonic coefficients of the Cauchy kernel. Since the Cauchy transform splits into the Poisson and conjugate Poisson transform we evaluate the coefficients of the Poisson and conjugate Poisson kernel in (7.20) separately. The Poisson kernel belongs to the set of zonal functions on \mathbb{S}^2 which only depend on the inner product between their argument $\boldsymbol{\omega} \in \mathbb{S}^2$ and a unit vector $\boldsymbol{\xi} \in \mathbb{S}^2$. The spectral characterization of these zonal functions is obtained by the Funk-Hecke theorem:

Theorem 7.9 (Funk-Hecke formula on \mathbb{S}^2 , Freeden et al. (1998)). *Let f be a zonal function on \mathbb{S}^n and $\boldsymbol{\xi} \in \mathbb{S}^n$. Then*

$$\int_{\mathbb{S}^n} f(\langle \boldsymbol{\xi}, \boldsymbol{\omega} \rangle) Y_{l,m}(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) = \lambda_l Y_{l,m}(\boldsymbol{\xi}) \quad (7.28)$$

with the eigenvalue

$$\lambda_l = A_{n-1} \frac{1}{C_l^{(n-1)/2}(1)} \int_{-1}^1 f(t) C_l^{(n-1)/2}(t) (1-t^2)^{(n-2)/2} dt \quad (7.29)$$

where $C_l^{(n-1)/2}$ is the l -th Gegenbauer polynomial of order $(n-1)/2$. For $n=2$ the Gegenbauer Polynomials reduce to the Legendre polynomials. In that case the above transform used to obtain the eigenvalues λ_l reduces to the Legendre transform. Using the Funk-Hecke theorem, the zonal Poisson kernel in \mathbb{B}^2 , also known as the Abel-Poisson kernel, results in the series expansion (see e.g. Freeden et al. (1998))

$$\mathcal{P}_{x_0}(\mathbf{x}, \boldsymbol{\omega}) = \frac{1 - \|\mathbf{x}\|^2}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} = \frac{1 - r^2}{(1 - 2r \cos \theta + r^2)^{(3/2)}} = \sum_{k=0}^{\infty} (2k+1) r^k P_n(\langle \boldsymbol{\xi}, \boldsymbol{\omega} \rangle). \quad (7.30)$$

Remembering that we always evaluate the transforms at the north pole $\boldsymbol{\eta}$ with $\mathbf{x} = x_3 \mathbf{e}_3 = r \boldsymbol{\eta} = r \mathbf{e}_3, 0 < r < 1$ the above series leads to the spherical harmonic coefficients

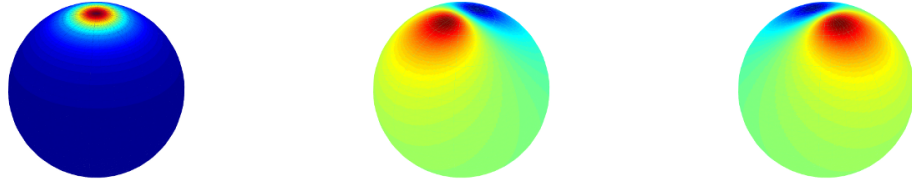


Figure 7.1.: From left to right: The Poisson kernel \mathcal{P}_{x_0} and the bivector parts of the conjugate Poisson kernel $\mathcal{Q}_{x_0}^{(1)}$ and $\mathcal{Q}_{x_0}^{(2)}$ for $r = 0.9$.

$$[\widehat{\mathcal{P}_{x_0}}]_{l,m} = \int_{\mathbb{S}^2} \mathcal{P}_{x_0}(r\boldsymbol{\eta}, \boldsymbol{\omega}) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) = \begin{cases} r^l & \text{for } m = 0 \\ 0 & \text{else} \end{cases}. \quad (7.31)$$

For the expansion of conjugate Poisson kernel we first rewrite the kernel as

$$\mathcal{Q}_{x_0}(\mathbf{x}, \boldsymbol{\omega}) = \frac{1 + \|\mathbf{x}\|^2 + 2\mathbf{x}\boldsymbol{\omega}}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} = \frac{1 + \|\mathbf{x}\|^2 - 2r\omega_3}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} - \frac{2r\omega_1 e_{13}}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} - \frac{2r\omega_2 e_{23}}{\|\mathbf{x} - \boldsymbol{\omega}\|^3}. \quad (7.32)$$

with $\boldsymbol{\omega} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)^\top$ and $\mathbf{x} = r\boldsymbol{\eta}$ such that

$$\langle \mathbf{x}, \boldsymbol{\omega} \rangle = \langle r\boldsymbol{\eta}, \boldsymbol{\omega} \rangle = r \cos \theta. \quad (7.33)$$

Since $\mathbf{x} = r\boldsymbol{\eta} = (0, 0, r)^\top$ we have

$$\|\mathbf{x} - \boldsymbol{\omega}\|^3 = ((-\omega_1)^2 + (-\omega_2)^2 + (r - \omega_3)^2)^{3/2} \quad (7.34)$$

$$= (1 + r^2 - 2r\omega_3)^{3/2} \quad (7.35)$$

$$= (1 + \|\mathbf{x}\|^2 - 2r\omega_3) (1 + \|\mathbf{x}\|^2 - 2r\omega_3)^{1/2} \quad (7.36)$$

yielding

$$\frac{1 + \|\mathbf{x}\|^2 - 2r\omega_3}{\|\mathbf{x} - \boldsymbol{\omega}\|^3} = \frac{1}{(1 + r^2 - 2r\omega_3)^{1/2}}. \quad (7.37)$$

According to Jeffrey and Zwillinger (2007) pp. 988, the generating function

$(1 + r^2 - 2r\omega_3)^{-1/2}$ has the following series expansion in terms of the Legendre polynomials:

$$\frac{1}{(1 + r^2 - 2r\omega_3)^{1/2}} = \sum_{k=0}^{\infty} r^k P_n(\langle \xi, \omega \rangle) = \sum_{k=0}^{\infty} r^k P_n(\cos \theta). \quad (7.38)$$

Accordingly, the generating function $(1 + \|\mathbf{x}\|^2 - 2r\omega_3)^{-3/2}$ expands into a series of the Gegenbauer polynomials (see also Jeffrey and Zwillinger (2007))

$$\frac{1}{(1 + \|\mathbf{x}\|^2 - 2r\omega_3)^{3/2}} = \sum_{k=0}^{\infty} r^k C_n^{3/2}(\langle \xi, \omega \rangle) = \sum_{k=0}^{\infty} r^k C_n^{3/2}(\cos \theta). \quad (7.39)$$

Therefore, we obtain a series expansion of the conjugate Poisson kernel in terms of Legendre and Gegenbauer polynomials as

$$Q_{x_0}(\mathbf{x}, \omega) = \sum_{k=0}^{\infty} r^k P_n(\cos \theta) \quad (7.40)$$

$$- 2r\omega_1 \sum_{k=0}^{\infty} r^k C_k^{3/2}(\cos \theta) e_{13} - 2r\omega_2 \sum_{k=0}^{\infty} r^k C_k^{3/2}(\cos \theta) e_{23} \quad (7.41)$$

$$= Q_{x_0}^{(0)}(\mathbf{x}, \omega) - 2rQ_{x_0}^{(1)}(\mathbf{x}, \omega) e_{13} - 2rQ_{x_0}^{(2)}(\mathbf{x}, \omega) e_{23} \quad (7.42)$$

While the Poisson kernel only consists of a scalar part, the conjugate Poisson kernel splits into a scalar and two bivector parts.

The scalar part of the conjugate Poisson kernel is equal to the *singularity kernel* (cmp. Freeden et al. (1995)). Both, the (Abel-)Poisson and the singularity kernel belong to the same class of reproducing kernels and can be used for a zonal basis function approximation on the sphere (see Keiner et al. (2006)). Figure 7.1 shows the Poisson kernel and the two bivector parts of the conjugate kernel at a certain scale. We treat the scalar and bivector parts of the conjugate Poisson kernel as three single kernels. For every single kernel, the Fourier coefficients are supposed to be determined, whereas the coefficients of the Poisson kernel are already given by (7.31) and the coefficients of the scalar part of the conjugate Poisson kernel (singularity kernel) are given by (Freeden and Michel (2004)):

$$\widehat{[Q_{x_0}^{(0)}]}_{l,m} = \int_{\mathbb{S}^2} Q_{x_0}^{(0)}(r\boldsymbol{\eta}, \omega) \overline{Y_{l,m}(\omega)} dS(\omega) = \begin{cases} 2r^l / (2l + 1) & \text{for } m = 0 \\ 0 & \text{else.} \end{cases} \quad (7.43)$$

In order to obtain the Fourier coefficients of the kernels, we express all parts of the kernels as a series depending on the standard spherical harmonics. We introduce the addition theorem for Gegenbauer polynomials with spherical coordinates (r_1, θ, ϕ) and (r_2, α, β) , subtended by the angle γ at the origin, as (see Srinivasan et al. (2005)):

$$C_n^\lambda(\cos \gamma) = 4\pi \sum_{m=0}^{\lfloor n/2 \rfloor} \langle \mathbf{Y}_{n,m}(\theta, \phi), \overline{\mathbf{Y}_{n,m}}(\alpha, \beta) \rangle \quad (7.44)$$

with the vector of scalar spherical harmonics

$$\mathbf{Y}_{n,m}(\mathbf{x}, \mathbf{y}) = \begin{pmatrix} Y_{l,-l} \\ Y_{l,-l+1} \\ \dots \\ Y_{l,l} \end{pmatrix} \quad (7.45)$$

where $l = n - 2m$ resulting in

$$C_n^\lambda(\cos \gamma) = 4\pi \sum_{m=0}^{\lfloor n/2 \rfloor} \sum_{j=-l}^l Y_{l,j}(\theta, \phi) \overline{Y_{l,j}}(\alpha, \beta). \quad (7.46)$$

In addition we expand ω_1 and ω_2 into a Fourier series with the Fourier coefficients

$$\widehat{[\omega_1]}_{l,m} = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} \sin \theta \cos \varphi \overline{Y_{l,m}}(\theta, \varphi) dS(\boldsymbol{\omega}) = \begin{cases} \pm \sqrt{\frac{2\pi}{3}} & \text{for } l = 1, m = \mp 1 \\ 0 & \text{else} \end{cases} \quad (7.47)$$

$$\widehat{[\omega_2]}_{l,m} = \int_{\varphi=0}^{2\pi} \int_{\theta=0}^{\pi} \sin \theta \sin \varphi \overline{Y_{l,m}}(\theta, \varphi) dS(\boldsymbol{\omega}) = \begin{cases} \mathbf{i} \sqrt{\frac{2\pi}{3}} & \text{for } l = 1, m = \pm 1 \\ 0 & \text{else} \end{cases} \quad (7.48)$$

such that

$$\omega_1 = \sum_{l=0}^{\infty} \sum_{m=-l}^l \widehat{[\omega_1]}_{l,m} Y_{l,m}(\theta, \varphi) = \sqrt{\frac{2\pi}{3}} Y_{1,-1}(\theta, \varphi) - \sqrt{\frac{2\pi}{3}} Y_{1,1}(\theta, \varphi) \quad (7.49)$$

$$\omega_2 = \sum_{l=0}^{\infty} \sum_{m=-l}^l \widehat{[\omega_2]}_{l,m} Y_{l,m}(\theta, \varphi) = \mathbf{i} \sqrt{\frac{2\pi}{3}} Y_{1,-1}(\theta, \varphi) + \mathbf{i} \sqrt{\frac{2\pi}{3}} Y_{1,1}(\theta, \varphi). \quad (7.50)$$

We determine the Fourier coefficients of the conjugate Poisson kernel bivector parts as:

$$\widehat{[Q_{x_0}^{(1)}]_{l,m}} = \int_{\mathbb{S}^2} Q_{x_0}^{(1)}(\boldsymbol{\omega}) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.51)$$

$$= \int_{\mathbb{S}^2} \omega_1 \sum_{k=0}^{\infty} r^k C_k^{3/2}(\cos \theta) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.52)$$

$$\stackrel{l'=k-2m'}{=} 4\pi \sum_{k=0}^{\infty} r^k \int_{\mathbb{S}^2} \omega_1 \sum_{m'=0}^{[k/2]} \sum_{j=-l'}^{l'} Y_{l',j}(\boldsymbol{\omega}) \overline{Y_{l',j}(0)} \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.53)$$

$$= 4\pi \sqrt{\frac{2\pi}{3}} \sum_{k=0}^{\infty} r^k \quad (7.54)$$

$$\times \left(\int_{\mathbb{S}^2} Y_{1,-1}(\boldsymbol{\omega}) \sum_{m'=0}^{[k/2]} \sum_{j=-l'}^{l'} Y_{l',j}(\boldsymbol{\omega}) \overline{Y_{l',j}(0)} \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.55)$$

$$- \int_{\mathbb{S}^2} Y_{1,1}(\boldsymbol{\omega}) \sum_{m'=0}^{[k/2]} \sum_{j=-l'}^{l'} Y_{l',j}(\boldsymbol{\omega}) \overline{Y_{l',j}(0)} \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \right) \quad (7.56)$$

$$= 4\pi \sqrt{\frac{2\pi}{3}} \sum_{k=0}^{\infty} r^k \quad (7.57)$$

$$\times \left(\sum_{m'=0}^{[k/2]} \overline{Y_{l',0}(0)} \int_{\mathbb{S}^2} Y_{1,-1}(\boldsymbol{\omega}) Y_{l',0}(\boldsymbol{\omega}) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.58)$$

$$- \sum_{m'=0}^{[k/2]} \overline{Y_{l',0}(0)} \int_{\mathbb{S}^2} Y_{1,1}(\boldsymbol{\omega}) Y_{l',0}(\boldsymbol{\omega}) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \right) \quad (7.59)$$

Since $\overline{Y_{l,m}(\boldsymbol{\omega})} = (-1)^m Y_{l,-m}(\boldsymbol{\omega})$ the triple integrals evaluate to (Thompson (1994))

$$\int_{\mathbb{S}^2} Y_{1,1}(\boldsymbol{\omega}) Y_{l',0}(\boldsymbol{\omega}) \overline{Y_{l,m}(\boldsymbol{\omega})} dS(\boldsymbol{\omega}) \quad (7.60)$$

$$= (-1)^m \sqrt{\frac{(2+1)(2l'+1)(2l+1)}{4\pi}} \begin{pmatrix} 1 & l' & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & l' & l \\ 1 & 0 & (-m) \end{pmatrix} \quad (7.61)$$

and

$$\int_{\mathbb{S}^2} Y_{1,-1}(\boldsymbol{\omega}) Y_{l',0}(\boldsymbol{\omega}) \overline{Y_{l,m}}(\boldsymbol{\omega}) dS(\boldsymbol{\omega}) \quad (7.62)$$

$$= (-1)^m \sqrt{\frac{(2+1)(2l'+1)(2l+1)}{4\pi}} \begin{pmatrix} 1 & l' & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & l' & l \\ -1 & 0 & (-m) \end{pmatrix} \quad (7.63)$$

where $\begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix}$ are the Wigner 3j-symbols. The Wigner 3j-symbols evaluate to 0 unless they satisfy

Wigner-3j
symbols

$$(i) \quad |l_1 - l_3| \leq l_2 \leq l_1 + l_3 \quad (7.64)$$

$$(ii) \quad m_1 + m_2 = -m_3. \quad (7.65)$$

Therefore, it holds that

$$\begin{pmatrix} 1 & l' & l \\ 0 & 0 & 0 \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & l' & l \\ 0 & 0 & 0 \end{pmatrix} & \text{if } l' = l - 1 \vee l' = l \vee l' = l + 1 \\ 0 & \text{else} \end{cases} \quad (7.66)$$

$$\begin{pmatrix} 1 & l' & l \\ \pm 1 & 0 & (-m) \end{pmatrix} = \begin{cases} \begin{pmatrix} 1 & l' & l \\ \pm 1 & 0 & \pm 1 \end{pmatrix} & \text{if } m = \mp 1 \\ 0 & \text{else} \end{cases}. \quad (7.67)$$

Closed expressions for the Wigner 3j-symbols used here are given in (Baylis (2004)):

$$\begin{pmatrix} k & k & 1 \\ q & -q & 0 \end{pmatrix} = (-1)^{k-q} \frac{q}{[k(k+1)(2k+1)]^{1/2}} \quad (7.68)$$

$$\begin{pmatrix} k & k+1 & 1 \\ q & -q-1 & 1 \end{pmatrix} = (-1)^{k-q} \left[\frac{(k+q+1)(k+q+2)}{2(k+1)(2k+1)(2k+3)} \right]^{1/2} \quad (7.69)$$

$$\begin{pmatrix} k & k+1 & 1 \\ q & -q & 0 \end{pmatrix} = (-1)^{k-q} \left[\frac{(k-q-1)(k+q+1)}{(k+1)(2k+1)(2k+3)} \right]^{1/2}. \quad (7.70)$$

With (7.68) it follows that that $\begin{pmatrix} 1 & l & l \\ 0 & 0 & 0 \end{pmatrix} = 0$ such that $l' = l - 1 \vee l + 1$. For $l' = l - 1$

we use (7.70) with $k = l - 1, q = 0$ and (7.69) with $k = l - 1, q = 0$ to obtain

$$(-1)^m Y_{l-1,0}(0) \sqrt{\frac{(2+1)(2(l-1)+1)(2l+1)}{4\pi}} \begin{pmatrix} 1 & l-1 & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & l-1 & l \\ 1 & 0 & -1 \end{pmatrix} \quad (7.71)$$

$$= (-1)^{l-1} \sqrt{\frac{2l-1}{4\pi}} \sqrt{\frac{3(2l-1)(2l+1)}{4\pi}} \quad (7.72)$$

$$\times \sqrt{\frac{l^2}{l(2l-1)(2l+1)}} \sqrt{\frac{l(l+1)}{2l(2l-1)(2l+1)}} \quad (7.73)$$

$$= (-1)^{l-1} \frac{1}{4\pi} \sqrt{\frac{3l(l+1)}{2(2l+1)}} \quad (7.74)$$

$$(-1)^m Y_{l+1,0}(0) \sqrt{\frac{(2+1)(2(l+1)+1)(2l+1)}{4\pi}} \begin{pmatrix} 1 & l+1 & l \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & l+1 & l \\ 1 & 0 & -1 \end{pmatrix} \quad (7.75)$$

$$= (-1)^l \sqrt{\frac{2l+3}{4\pi}} \sqrt{\frac{3(2l+3)(2l+1)}{4\pi}} \quad (7.76)$$

$$\times \sqrt{\frac{(l+1)^2}{(l+1)(2l+1)(2l+3)}} \sqrt{\frac{l(l+1)}{2(l+1)(2l+1)(2l+3)}} \quad (7.77)$$

$$= (-1)^l \frac{1}{4\pi} \sqrt{\frac{3l(l+1)}{2(2l+1)}}. \quad (7.78)$$

The only contributing terms in the sums are the expressions with $l' = l - 1 \vee l + 1$. Since $l' = k - 2m$, this is only the case for $k = (l - 1) + 2n, n \geq 0$. For $k = l - 1$ this is only true for $n = 0$. For all other k there are always two solutions but as we have seen, they only differ in their sign. Therefore the sum for all $k > l - 1$ evaluates to zero such that

$$[\widehat{Q_{x_0}^{(1)}}]_{l,\pm 1} = \mp 2r 4\pi \sqrt{\frac{2\pi}{3}} \frac{1}{4\pi} \sqrt{\frac{3l(l+1)}{2(2l+1)}} r^{l-1} = \mp \sqrt{\frac{4\pi l(l+1)}{(2l+1)}} r^l \quad (7.79)$$

$$[\widehat{Q_{x_0}^{(2)}}]_{l,\pm 1} = \pm 1 2r 4\pi \sqrt{\frac{2\pi}{3}} \frac{1}{4\pi} \sqrt{\frac{3l(l+1)}{2(2l+1)}} r^{l-1} = \pm 1 \sqrt{\frac{4\pi l(l+1)}{(2l+1)}} r^l. \quad (7.80)$$

The spherical harmonic coefficients characterize the filter kernels in the frequency domain. As one notices the coefficients of the scalar part are zero for $m \neq 0$ whereas the coefficients of the bivector parts are zero for $m \neq \pm 1$. The factor r^l indicates the lowpass behaviour of all filter parts since $0 < r < 1$. Using the spherical harmonic

coefficients, the Poisson and conjugate Poisson transforms can now be expressed as a directional correlation in the spherical harmonic domain according to (7.27) as

$$\mathcal{R}(\boldsymbol{\rho})[\mathcal{P}_{x_0}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l \widehat{f}_{l,m} \overline{D_{m,0}^l(\boldsymbol{\rho})} = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l \widehat{f}_{l,m} \overline{Y_{l,m}(\theta, \varphi)} \quad (7.81)$$

$$\mathcal{R}(\boldsymbol{\rho})[\mathcal{Q}_{x_0}^{(1)}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \sqrt{\frac{4\pi l(l+1)}{(2l+1)}} r^l \widehat{f}_{l,m} \left(\overline{D_{m,-1}^l(\boldsymbol{\rho})} - \overline{D_{m,1}^l(\boldsymbol{\rho})} \right) \quad (7.82)$$

$$\mathcal{R}(\boldsymbol{\rho})[\mathcal{Q}_{x_0}^{(2)}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \imath \sqrt{\frac{4\pi l(l+1)}{(2l+1)}} r^l \widehat{f}_{l,m} \left(\overline{D_{m,-1}^l(\boldsymbol{\rho})} + \overline{D_{m,1}^l(\boldsymbol{\rho})} \right). \quad (7.83)$$

Since we evaluate the correlation for rotations $\boldsymbol{\rho} = (\varphi, \theta, 0)$ the Wigner-D functions reduce to spin-weighted spherical harmonics Wiaux et al. (2006)

Spin-weighted
spherical
harmonics

$${}_n Y_{l,m}(\theta, \varphi) = (-1)^n \sqrt{\frac{2l+1}{4\pi}} \overline{D_{m,n}^l(\varphi, \theta, 0)}. \quad (7.84)$$

We define the spin raising and lowering operators as in Newman and Penrose (1966)

Spin raising and
lowering
operators

$$\begin{aligned} [\delta_n G](\theta, \varphi) &= \left[-\sin^n \theta \left(\frac{\partial}{\partial \theta} + \frac{\imath}{\sin \theta} \frac{\partial}{\partial \varphi} \right) \sin^{-n} \theta {}_n G \right](\theta, \varphi) \\ [\bar{\delta}_n G](\theta, \varphi) &= \left[-\sin^n \theta \left(\frac{\partial}{\partial \theta} - \frac{\imath}{\sin \theta} \frac{\partial}{\partial \varphi} \right) \sin^{-n} \theta {}_n G \right](\theta, \varphi). \end{aligned} \quad (7.85)$$

With these operators the spin-weighted spherical harmonics are constructed from the scalar spherical harmonics as

$$\begin{aligned} {}_n Y_{l,m}(\theta, \varphi) &= \left[\frac{(l-n)!}{(l+n)!} \right]^{1/2} [\delta^n Y_{l,m}](\theta, \varphi) && \text{for } 0 \leq n \leq l \\ {}_n Y_{l,m}(\theta, \varphi) &= \left[\frac{(l-n)!}{(l+n)!} \right]^{1/2} (-1)^n [\bar{\delta}^n Y_{l,m}](\theta, \varphi) && \text{for } -l \leq n \leq 0. \end{aligned} \quad (7.86)$$

With $\boldsymbol{\rho} = (\theta, \varphi, 0)$, $\boldsymbol{\omega} = (\theta, \varphi) \in \mathbb{S}^2$ the correlations (7.82), (7.83) reduce to

$$\mathcal{R}(\boldsymbol{\rho})[\mathcal{Q}_{x_0}^{(1)}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l \widehat{f}_{l,m} (-{}_1Y_{l,m}(\boldsymbol{\omega}) - {}_1Y_{l,m}(\boldsymbol{\omega})) \quad (7.87)$$

$$= 2 \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l \widehat{f}_{l,m} \frac{\partial}{\partial \theta} Y_{l,m}(\boldsymbol{\omega}) \quad (7.88)$$

$$\mathcal{R}(\boldsymbol{\rho})[\mathcal{Q}_{x_0}^{(2)}] \star f = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \imath r^l \widehat{f}_{l,m} (-{}_1Y_{l,m}(\boldsymbol{\omega}) + {}_1Y_{l,m}(\boldsymbol{\omega})) \quad (7.89)$$

$$= 2 \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l \widehat{f}_{l,m} \frac{1}{\sin \theta} \frac{\partial}{\partial \varphi} Y_{l,m}(\boldsymbol{\omega}). \quad (7.90)$$

The ∇ operator in a spherical coordinate system, which coincides with the spherical gradient operator in the case of a scalar valued function, with basis vectors e_r, e_θ, e_φ is defined as

$$\nabla = e_r \frac{\partial}{\partial r} + e_\theta \frac{1}{r} \frac{\partial}{\partial \theta} + e_\varphi \frac{1}{r \sin \theta} \frac{\partial}{\partial \varphi}. \quad (7.91)$$

It acts on a function $f \in L_2(\mathbb{S}^2)$ as

$$(\nabla f)(\boldsymbol{\omega}) = \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \widehat{f}_{l,m} (\nabla Y_{l,m})(\boldsymbol{\omega}) \quad (7.92)$$

$$= e_\theta \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \widehat{f}_{l,m} \frac{\partial}{\partial \theta} Y_{l,m}(\boldsymbol{\omega}) + e_\varphi \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \widehat{f}_{l,m} \frac{1}{\sin \theta} \frac{\partial}{\partial \varphi} Y_{l,m}(\boldsymbol{\omega}). \quad (7.93)$$

Comparing (7.93) with (7.88) and (7.90) one notices that the conjugate Poisson transform acts like a gradient operator but instead of acting on the original function it acts on the solution of the Laplace equation. Figure 7.2 demonstrates the application of the bivector parts of the conjugate Poisson transform at a certain scale.

7.5. Applications

7.5.1. Poisson scale space

The Poisson transform $\mathcal{P}[f]$ acts as a lowpass filter in the spherical harmonic domain on the frequencies l . Furthermore the Poisson transform generates a harmonic function and solves the Laplace equation in \mathbb{B}^2 with boundary values on \mathbb{S}^2 . It constitutes a linear scale space in \mathbb{B}^2 with scale parameter $r = e^{-p}$ for $p \in [0, \infty)$ such that $0 < r < 1$.

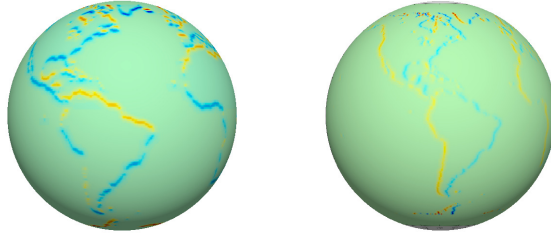


Figure 7.2.: Outputs of the filters $\mathcal{R}(\rho)[Q_{x_0}^{(1)}]$ and $\mathcal{R}(\rho)[Q_{x_0}^{(2)}]$ at the scale $r = 0.99$.

With this choice for r , for any $p_1, p_2 \in [0, \infty)$ and $f \in L^2(\mathbb{S}^2)$ we have

$$\mathcal{P}_{p_2}[\mathcal{P}_{p_1}[f]] = \sum_{l=0}^{\infty} \sum_{m=-l}^l e^{-lp_1} e^{-lp_2} \hat{f}_{l,m} Y_{l,m} \quad (7.94)$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^l e^{-l(p_1+p_2)} \hat{f}_{l,0} Y_{l,m} = \mathcal{P}_{p_1+p_2}[f] \quad (7.95)$$

since we can apply the spherical convolution theorem according to Driscoll and Healy (1994) to the convolution with the zonal Poisson kernel. The Poisson transform therefore fulfills the semigroup property. Furthermore it holds that $P[f] \geq 0$ for every $f \geq 0$ due to the positivity of the kernel. Since the Poisson transform solves the Laplace equation, which is the diffusion equation with constant source term, it is an diffusion semigroup operator on \mathbb{S}^2 . The maximum and minimum values of harmonic functions in \mathbb{B}^2 lie on the boundary \mathbb{S}^2 . In addition it holds for a harmonic function that every function value at some $x \in \mathbb{B}^2$ is equal to the average over an arbitrary sphere around x in \mathbb{B}^2 . It therefore generates no additional local extrema. These properties suggest the Poisson kernel as a smoothing filter. Figure 7.3 shows the Poisson transform at different scales on \mathbb{S}^2 .

7.5.2. Plane wave analysis

We want to determine the orientation of a plane wave $Ae^{i\langle x, k \mathbf{n} \rangle}$ and the phase with respect the θ, φ plane in our local spherical coordinate system (see Figure 7.4). The direction of the plane wave is given by $\mathbf{n} = (\sin \alpha \cos \beta, \sin \alpha \sin \beta, \cos \alpha)^T$ where k is the frequency. A denotes the amplitude of the plane wave. Without loss of generality we can always choose the point ξ at which we evaluate the filter operations as the north pole $\xi = (\theta, \varphi) = (0, 0)$. It is well known (see e.g. Baylis (2004)) that a plane wave in \mathbb{R}^3 can be expanded in terms of spherical harmonics as

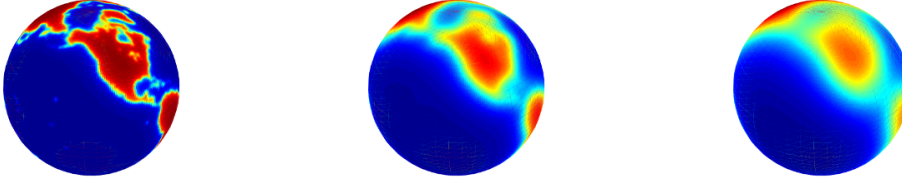


Figure 7.3.: Example for Poisson filtered functions on the sphere at decreasing scales.

$$f(\mathbf{x}) = A e^{i\langle \mathbf{x}, \mathbf{k} \mathbf{n} \rangle} = A \sum_{l \in \mathbb{N}} \sum_{m=-l}^l \hat{f}_{l,m} Y_{l,m}(\theta, \varphi) \quad (7.96)$$

with

$$\hat{f}_{l,m} = i^l j_l(rk) \overline{Y_{l,m}(\alpha, \beta)} \quad (7.97)$$

where $\mathbf{x} = r\xi$, $k\mathbf{n} = \boldsymbol{\omega}$ and j_l is the spherical Bessel function of the first kind and order l . The expansion splits the plane wave into an angular part represented by the spherical harmonics and a part depending on the frequency of the wave described by the spherical Bessel function. Due to the nature of our filter kernels, these act only on the angular portions of the plane wave. The frequency information k encoded as the argument of the Bessel function is left untouched by our filter set. Using (7.27) and the property $D_{m,0}^l = (\frac{4\pi}{2l+1})^{1/2} \overline{Y_{l,m}}$ in conjunction with the Wigner-D function addition theorem (Pendleton (2003))

$$\sum_{m=-l}^l D_{m,0}^l(\beta, \alpha, 0) \overline{D_{m,1}^l(\varphi, \theta, 0)} = \overline{D_{0,1}^l(\beta', \alpha', 0)} = D_{1,0}^l(\beta, \alpha, 0) \quad (7.98)$$

the conjugate Poisson transforms of f , assuming that they are evaluated at the north pole, read

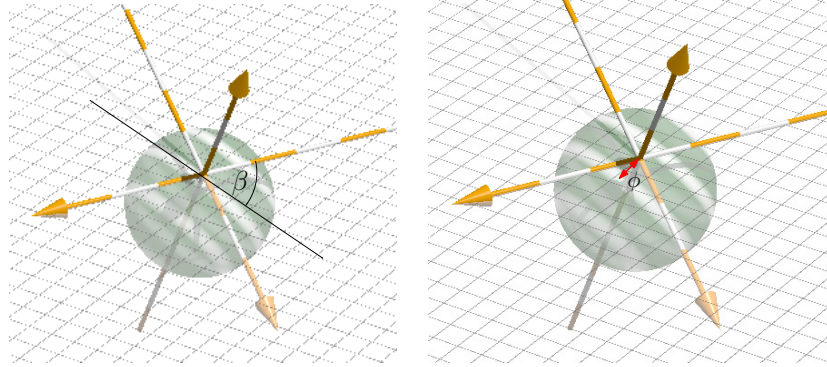


Figure 7.4.: From left to right: Orientation angle β of the plane wave in the θ, φ plane with respect to the local coordinate system. Phase ϕ of the plane wave with respect to the local coordinate system.

$$\mathcal{R}(\rho)[Q_{x_0}^{(1)}] \star f = A \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l i^l j_l(rk) P_l^1(\cos \alpha) \cos \beta \quad (7.99)$$

$$\mathcal{R}(\rho)[Q_{x_0}^{(2)}] \star f = A \sum_{l \in \mathbb{N}} \sum_{m=-l}^l r^l i^l j_l(rk) P_l^1(\cos \alpha) \sin \beta \quad (7.100)$$

where P_l^m are the associated Legendre functions. The angle β is obtained as

$$\beta = \arctan \frac{\mathcal{R}(\xi)[Q_{x_0}^{(2)}] \star f}{\mathcal{R}(\xi)[Q_{x_0}^{(1)}] \star f}. \quad (7.101)$$

Since

$$P_l^1(\cos \alpha) = -\sin \alpha \frac{d}{d \cos \alpha} P_l(\cos \alpha) \quad (7.102)$$

we notice that the bivector parts of the conjugate Poisson transform both act as differential operators with respect to $\cos \alpha$. We obtain the phase ϕ

$$\phi = \arctan \sqrt{\frac{(\mathcal{R}(0)[Q_{x_0}^{(1)}] \star f)^2 + (\mathcal{R}(0)[Q_{x_0}^{(2)}] \star f)^2}{(\mathcal{R}(0)[Q_{x_0}^{(0)}] \star f)^2}} \quad (7.103)$$

Note that the determination of the orientation β and the phase ϕ are invariant with respect the amplitude A .

Figure 7.5 shows the filter outputs of the amplitude, the local orientation and the phase obtained by our filter set. It acts as the analogue to the classical Riesz transform

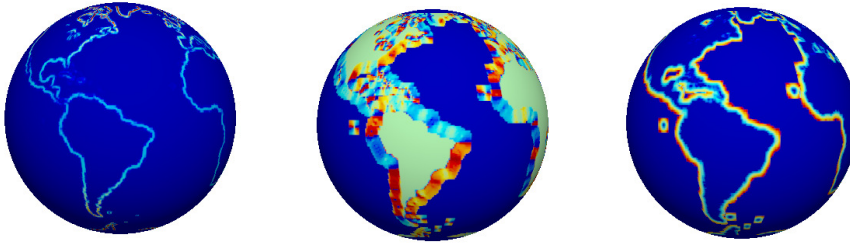


Figure 7.5.: From left to right: Amplitude , orientation and phase output of the filters with a filter mask size of 9 pixels and a scale of $r = 0.99$.

in the plane which has been used to obtain local amplitude, orientation and phase of plane waves in \mathbb{R}^2 .

7.6. Summary

In this chapter we introduced a further generalization of the Hilbert transform in \mathbb{R}^n to arbitrary subsets with smooth boundary in \mathbb{R}^n . The generalization naturally arises as the non-tangential boundary value of the corresponding Cauchy transform on the subset. Both transforms have been introduced in the context of Clifford analysis. The introduced Hilbert transform from the previous chapters, turns out to be a special case of the new generalized transform. The new generalized Hilbert transform has been used to derive the Hilbert transform on the unit sphere in \mathbb{R}^n . A spectral characterization of the Cauchy kernel has been derived in terms of its spherical harmonic decomposition which led to a Fourier series of the Cauchy kernel on the unit sphere. Using the analogue of the convolution theorem of the Fourier transform in \mathbb{R}^n , the spherical harmonic coefficients of the Cauchy kernel allow an easy implementation of the transform. Further, the Cauchy transform naturally gives rise to a linear scale space on the sphere. Locally, intrinsically one-dimensional signals in \mathbb{R}^3 , which correspond to the real parts of complex plane waves in \mathbb{R}^3 can be analyzed in terms of their parameters orientation, amplitude and phase in a multiscale framework.

8. Conclusion and outlook

This thesis aimed to continue the work on generalized Hilbert transforms and their applications in computer and machine vision which was carried out in the Cognitive Systems Group at the University of Kiel by Thomas Bülow, Michael Felsberg, Di Zang and Lennart Wietzke and their respective theses Bülow (1999); Felsberg (2002); Zang (2007); Wietzke (2011). While Bülow and Felsberg provided the foundations in terms of hypercomplex signal representations using generalized Hilbert transforms in the plane, Bülow, Felsberg, Zang and Wietzke used the generalized Hilbert transform and its higher-order variants to analyze the curvature of intrinsically two-dimensional signals and the decomposition of two intrinsically one-dimensional signals. In this thesis, three model based approaches for the description of low-level signal structures have been studied:

1. Superpositions of an arbitrary number of (sinusoidal) intrinsically one-dimensional signals in the plane
2. Curved intrinsically one-dimensional signals in the plane
3. Intrinsically one-dimensional signals in \mathbb{R}^3 restricted to the unit sphere

The basic goal was an identification of the invariant and equivariant parameters of the signal models with respect to certain group actions, such as the general linear group in the plane, and the estimation of parameters in terms of certain linear shift-invariant operators. The investigations have been carried out in terms of the generalized Hilbert transforms and the partial derivatives as well as their higher-order variants.

It has been argued that the filtering with simple linear shift-invariant filters such as the Hilbert transform and the partial derivatives resembles the behavior of the first stages in the human visual system.

In the case of the superimposed signal model in chapter 5, the properties of the classical monogenic signal have been extended to an arbitrary number of superpositions. The decomposition relies on higher-order Hilbert transforms or partial derivatives. While both of these operators are able to obtain the orientations of the underlying signals, the Hilbert transform is mandatory for the phase estimation of sinusoidal intrinsically one-dimensional signals. The phases of the superpositional model constitute a set of invariants with respect to the general linear group in the plane, while the orientations transform equivariantly and allow to infer about certain transformations which have acted on the signals. The methods have been introduced in the context of a multiscale embedding which allows the analysis in a linear scale-space.

Further, a local averaging procedure of the local operator responses has been proposed to eliminate the problem of singularities in the orientation estimation process. Based on the orientation estimates, a novel affine equivariant regional feature has been constructed. It has been coupled by an estimation of the characteristic scale for the underlying intrinsically one-dimensional signals yielding the final parameter set of the affine equivariant regional feature, the affine characteristic scale. Experimentally the validity of the model has been confirmed in comparison to state of the art affine equivariant region detectors.

In chapter 6 curved signal structures have been modeled by curved intrinsically one-dimensional signals. Using a nonlinear embedding, the signals are projected to a sphere in the three-dimensional Euclidean space \mathbb{R}^3 . It has been shown that this projection locally approximates a intrinsically one-dimensional signal in \mathbb{R}^3 such that an analysis in terms of the generalized Hilbert transforms or partial derivatives in \mathbb{R}^3 is possible. The orientation vector of this intrinsically one-dimensional signal has turned out to encode, in addition to the classical orientation of the signal in the plane, its curvature. The validity of the model has experimentally been confirmed by comparisons with classical differential geometric curvature estimation approaches for curves and images.

Chapter 7, being of more theoretical nature, has studied signals arising on the unit sphere in \mathbb{R}^3 in terms of a further generalization of the Hilbert transform. To gain insights into this generalization, which is well known in the field of Clifford analysis, a spectral characterization in terms of its spherical harmonic decomposition has been provided. Not only does this decomposition allow an easy implementation of the transform using a theorem similar to the convolution theorem in Euclidean spaces, but it also provides insight to the spectral behavior of the operator.

The results from this thesis lead to new questions for future work, especially in the field of practical applications using the proposed methods. Using the results from chapter 5, interest point detectors should be investigated, which detect points of a certain phase, e.g. points, where the phases of all component signals are zero. Since the phase is invariant with respect to actions of the general linear group, such points should turn out to be useful affine invariant interest points.

Another question of more theoretical nature asks, how the results from chapter 5 can be generalized to obtain the orientations and phases for superpositions of intrinsically one-dimensional signals in dimensions higher than 2. Since a closed form solution for the corresponding symmetric tensor decomposition problem is not available in higher dimensions, solutions of algorithmic and approximating nature are required and have to be developed.

In the context of chapter 6 future work could investigate the application of the proposed method to three dimensional volume images and image sequences to study curvature properties in such signals.

Appendix A.

Algorithms

Algorithm 1: Pseudocode of the main algorithm to extract the angles and phases.

Data: Superimposed image signal I , number of superpositions d

Result: Angles $(\theta_1, \dots, \theta_d)$, Phases $(\varphi_1, \dots, \varphi_d)$

$m \leftarrow 2d - 1$ /* minimum operator order */

$\{k_0, \dots, k_m\} \leftarrow$ convolution kernels of \mathcal{H}^m (at a scale s)

$\{\bar{k}_0, \dots, \bar{k}_{m-1}\} \leftarrow$ convolution kernels of \mathcal{H}^{m-1} (at a scale s)

$\{I_0, \dots, I_m\} \leftarrow \{I * k_0, \dots, I * k_m\}$

$\{\bar{I}_0, \dots, \bar{I}_{m-1}\} \leftarrow \{I * \bar{k}_0, \dots, I * \bar{k}_m\}$

for $\mathbf{x} \in \text{Dom}(I)$ **do**

$(\theta_1, \dots, \theta_d) \leftarrow \text{getAngles}((I_0(\mathbf{x}), \dots, I_m(\mathbf{x})), d)$

$(\varphi_1, \dots, \varphi_d) \leftarrow \text{getPhases}((I_0(\mathbf{x}), \dots, I_m(\mathbf{x})), (\bar{I}_0(\mathbf{x}), \dots, \bar{I}_{m-1}(\mathbf{x})), (\theta_1, \dots, \theta_d))$

end

Algorithm 2: getAngles

Data: m -th order operator responses (c_0, \dots, c_m) at a pixel, number of superpositions d

Result: Angles $(\theta_1, \dots, \theta_d)$

$$\mathbf{C} \leftarrow \begin{bmatrix} c_0 & c_1 & \dots & c_d \\ \vdots & \vdots & \ddots & \vdots \\ c_{m-d} & \dots & c_{m-1} & c_m \end{bmatrix}$$

if $\text{rank}(\mathbf{C}) \neq d$ **then**
| return \emptyset

end

$(p_0, \dots, p_d) \leftarrow \arg \min_{\mathbf{p}} \|\mathbf{C}\mathbf{p}\|^2$ s.t. $\|\mathbf{p}\| = 1$ /* e.g. via SVD */

$\mathbf{R} \leftarrow$ roots of the polynomial $\tilde{Q}(x) = \sum_{i=0}^d p_i x^{d-i}$

$\mathbf{A} \leftarrow \emptyset$

if $\tilde{Q}(0) = 0$ **then**

| $\mathbf{A} \leftarrow \mathbf{A} \cup 0$ /* there exists a root at infinity */

end

for $r \in \mathbf{R}$ **do**

| **if** $r \notin \mathbb{R}$ **then**
| | continue

| **end**

| **if** $r = 0$ **then**

| | $\alpha \leftarrow \frac{\pi}{2}$

| **end**

| **else**

| | $\alpha \leftarrow \arctan(1.0/r)$

| **end**

| $\mathbf{A} \leftarrow \mathbf{A} \cup \alpha$

end

return \mathbf{A}

Algorithm 3: getPhases

Data: m -th and $(m-1)$ -th operator responses $\mathbf{c} = (c_0, \dots, c_m)$ and $\bar{\mathbf{c}} = (\bar{c}_0, \dots, \bar{c}_{m-1})$ at a pixel, angles $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$

Result: Phases $(\varphi_1, \dots, \varphi_d)$

$d \leftarrow |\boldsymbol{\theta}|$

$(\lambda_1, \dots, \lambda_d) \leftarrow \text{getCoefficients}((c_0(\mathbf{x}), \dots, c_m(\mathbf{x})), \boldsymbol{\theta})$ /* odd coefficients */

$(\bar{\lambda}_1, \dots, \bar{\lambda}_d) \leftarrow \text{getCoefficients}((\bar{c}_0(\mathbf{x}), \dots, \bar{c}_{m-1}(\mathbf{x})), \boldsymbol{\theta})$ /* even coefficients */

$(\varphi_1, \dots, \varphi_d) \leftarrow (0, \dots, 0)$

for $i \in \{1, \dots, d\}$ **do**

$\varphi_i \leftarrow \arctan 2(\lambda_i, \bar{\lambda}_i)$

end

return $(\varphi_1, \dots, \varphi_d)$

Algorithm 4: getCoefficients

Data: m -th operator responses $\mathbf{c} = (c_0, \dots, c_m)$ at a pixel, angles $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d)$

Result: Coefficients $(\lambda_1, \dots, \lambda_d)$

for $\theta_j \in (\theta_1, \dots, \theta_d)$ **do**

$\mathbf{n}_j \leftarrow (\cos(\theta_j), \sin(\theta_j))$

end

$$\mathbf{M} \leftarrow \begin{bmatrix} n_{1,1}^m & \dots & n_{d,1}^m \\ n_{1,1}^{m-1} n_{1,2} & \dots & n_{d,1}^{m-1} n_{d,2} \\ n_{1,1}^{m-2} n_{1,2}^2 & \dots & n_{d,1}^{m-2} n_{d,2}^2 \\ \vdots & \ddots & \vdots \\ n_{1,2}^m & \dots & n_{d,2}^m \end{bmatrix}$$

$\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_d) \leftarrow \arg \min_{\boldsymbol{\lambda}} \|\mathbf{M}\boldsymbol{\lambda} - \mathbf{c}\|^2$

return $\boldsymbol{\lambda}$

Bibliography

- Aach, T., C. Mota, I. Stuke, M. Mühlich, and E. Barth (2006). Analysis of Superimposed Oriented Patterns. *IEEE Transactions on Image Processing* 15(12), 3690–3700.
- Akhiezer, N., I. Glazman, and I. Glazman (1993). *Theory of Linear Operators in Hilbert Space*. Dover Publications.
- Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool (2008). Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding* 110(3), 346–359.
- Baylis, W. E. (2004). *Electrodynamics: A Modern Geometric Approach (Progress in Mathematical Physics)* (Corrected ed.). Birkhäuser Boston.
- Bernstein, S., R. Hielscher, and H. Schaeben (2009). The Generalized Totally Geodesic Radon Transform and its Application to Texture Analysis. *Mathematical Methods in the Applied Sciences* 32(4), 379–394.
- Boashash, B. (1992). Estimating and Interpreting the Instantaneous Frequency of a Signal I: Fundamentals. *Proceedings of the IEEE* 80(4), 520–538.
- Bogdanova, I., X. Bresson, J. P. Thiran, and P. Vandergheynst (2007). Scale Space Analysis and Active Contours for Omnidirectional Images. *Image Processing, IEEE Transactions on* 16(7), 1888–1901.
- Bracewell, R. (1978). *The Fourier Transform and its Applications*. McGraw-Hill.
- Bracewell, R. N. (2003). *Fourier Analysis and Imaging*. Kluwer Academic Plenum Publishers.
- Brachat, J., P. Comon, B. Mourrain, and E. Tsigaridas (2010). Symmetric Tensor Decomposition. *Linear Algebra and its Applications* 433(11-12), 1851–1872.
- Brackx, F., B. D. Knock, H. D. Schepper, and D. Eelbode (2006). On the interplay between the Hilbert transform and conjugate harmonic functions. *Mathematical Methods in the Applied Sciences* 29(12), 1435–1450.
- Bülow, T. (1999). *Hypercomplex Spectral Signal Representations for the Processing and Analysis of Images*. Ph. D. thesis, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik. Technical report 9903.

- Campbell, F. and J. Robson (1968). Application of Fourier Analysis to the Visibility of Gratings. *The Journal of Physiology* 197(3), 551.
- Chenouard, N. and M. Unser (2011). 3D Steerable Wavelets and Monogenic Analysis for Bioimaging. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 2132–2135. IEEE.
- Cohen, L. (1995). *Time-Frequency Analysis*. Prentice Hall PTR Englewood Cliffs, New Jersey.
- Comon, P., G. Golub, L.-H. Lim, and B. Mourrain (2008). Symmetric Tensors and Symmetric Tensor Rank. *SIAM J. Matrix Anal. Appl.* 30(3), 1254–1279.
- Comon, P. and B. Mourrain (1996). Decomposition of Quantics in Sums of Powers of Linear Forms. *Signal Processing* 53(2-3), 93–107.
- Coope, I. D. (1993). Circle Fitting by Linear and Nonlinear Least Squares. *Journal of Optimization Theory and Applications* 76(2), 381–388.
- Crotwell, P. (2000). Constructive Approximation on the Sphere. *Mathematical Geology* 32(2), 247–248.
- Dalal, N. and B. Triggs (2005). Histograms of Oriented Gradients for Human Detection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, Volume 1, pp. 886–893.
- Delanghe, R. (2004). On some Properties of the Hilbert Transform in Euclidean Space. *Bull. Belg. Math. Soc. Simon Stevin* 11(2), 163–180.
- Delanghe, R., F. Sommen, and V. Soucek (1992). *Clifford Algebra and Spinor-Valued Functions: A Function Theory for the Dirac Operator (Mathematics and Its Applications)* (1 ed.). Springer.
- do Carmo, M. P. (1976). *Differential Geometry of Curves and Surfaces*. Prentice-Hall.
- Driscoll, J. R. and D. M. Healy, Jr. (1994). Computing Fourier Transforms and Convolutions on the 2-sphere. *Adv. Appl. Math.* 15(2), 202–250.
- Felsberg, M. (2002). *Low-Level Image Processing with the Structure Multivector*. Ph. D. thesis, Institut für Informatik und Praktische Mathematik der Christian-Albrechts-Universität zu Kiel. Technical Report 0203.
- Felsberg, M. and G. Sommer (2001). The Monogenic Signal. *IEEE Transactions on Signal Processing* 49(12), 3136–3144.

- Felsberg, M. and G. Sommer (2004). The Monogenic Scale-Space: A Unifying Approach to Phase-Based Image Processing in Scale-Space. *Journal of Mathematical Imaging and Vision* 21(1), 5–26.
- Field, D. (1993). Scale-Invariance and Self-Similar ‘Wavelet’ Transforms: An Analysis of Natural Scenes and Mammalian Visual Systems. *Wavelets, Fractals, and Fourier Transforms: New Developments and New Applications*, 151–193.
- Fleet, D. J. and A. D. Jepson (1993). Stability of Phase Information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(12), 1253–1268.
- Fleet, D. J., A. D. Jepson, and M. R. M. Jenkin (1991). Phase-based Disparity Measurement. *CVGIP: Image Understanding* 53, 198–210.
- Fleischmann, O. and G. Sommer (2012). Automatic Scale Selection of Superimposed Signals. In A. Pinz, T. Pock, H. Bischof, and F. Leberl (Eds.), *Pattern Recognition*, Volume 7476 of LNCS, pp. 297–306. Springer Berlin / Heidelberg.
- Fleischmann, O., L. Wietzke, and G. Sommer (2009). A Spherical Harmonic Expansion for the Hilbert Transform on the Two-Sphere. In *International Conference on Numerical Analysis and Applied Mathematics*, Volume 1168 of AIP Conference Proceedings, pp. 777–780. Melville, New York.
- Fleischmann, O., L. Wietzke, and G. Sommer (2010a). A Novel Curvature Estimator for Digital Curves and Images. In M. Goesele, S. Roth, A. Kuijper, B. Schiele, and K. Schindler (Eds.), *Pattern Recognition*, Volume 6376 of LNCS, pp. 442–451. Springer Berlin / Heidelberg.
- Fleischmann, O., L. Wietzke, and G. Sommer (2010b). The Hilbert Transform on the Two-Sphere: A Spectral Characterization. *Mathematical Geosciences* 42(7), 857–876.
- Fleischmann, O., L. Wietzke, and G. Sommer (2011). Image Analysis by Conformal Embedding. *Journal of Mathematical Imaging and Vision* 40(3), 305–325.
- Förstner, W. and E. Gülch (1987). A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In *Proc. Intercommission Conference on Fast Processing of Photogrammetric Data (ISPRS)*, pp. 281–305.
- Freeden, W., T. Gervens, and M. Schreiner (1998). *Constructive approximation on the sphere*. Oxford: Clarendon Press.
- Freeden, W. and K. Hesse (2002). On the Multiscale Solution of Satellite Problems by use of Locally Supported Kernel Functions Corresponding to Equidistributed Data on Spherical Orbits. *Studia Scientiarum Mathematicarum Hungarica* 39(1), 37–74.

- Freeden, W. and V. Michel (2004). *Multiscale Potential Theory: With Applications to Geoscience*. Birkhäuser Boston.
- Freeden, W., M. Schreiner, and R. Franke (1995). A Survey on Spherical Spline Approximation. Technical Report 157, Fachbereich Mathematik.
- Freeman, W. and E. Adelson (1991). The Design and Use of Steerable Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(9), 891–906.
- Gabor, D. (1946). Theory of Communication. *Journal of the IEE (London)* 93, 429–457.
- Gander, W., G. H. Golub, and R. Strebler (1994). Least-Squares Fitting of Circles and Ellipses. *BIT* 34(4), 558–578.
- Gibson, J. (1979). *The Ecological Approach to Visual Perception*. Houghton, Mifflin and Company.
- Granlund, G. H. and H. Knutsson (1995). *Signal Processing for Computer Vision*. Kluwer Academic Publisher, Dordrecht.
- Haja, A., B. Jähne, and S. Abraham (2008). Localization Accuracy of Region Detectors. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, pp. 1–8. IEEE.
- Harris, C. and M. Stephens (1988). A Combined Corner and Edge Detector. In *Alvey Vision Conference*, pp. 147–152.
- Hel-Or, Y. and P. Teo (1998). Canonical Decomposition of Steerable Functions. *Journal of Mathematical Imaging and Vision* 9(1), 83–95.
- Hillar, C. J. and L.-H. Lim (2009). Most Tensor Problems are NP Hard. *Pre-Print abs/0911.1393*. arXiv:0911.1393v3 [cs.CC].
- Jeffrey, A. and D. Zwillinger (2007). *Table of Integrals, Series, and Products, Seventh Edition* (7 ed.). Academic Press.
- Kanatani, K. (1996). *Statistical Optimization for Geometric Computation: Theory and Practice*. New York, NY, USA: Elsevier Science Inc.
- Keiner, J., S. Kunis, and D. Potts (2006). Fast summation of radial functions on the sphere. *Computing* 78, 1–15.
- Knutsson, H. (1989). Representing Local Structure Using Tensors. In *Proceedings of the 6th Scandinavian Conference on Image Analysis*, Volume 1019 of *LiTH-ISY-I*, Linköping, pp. 244–251. Linköping University Electronic Press.

- Knutsson, H., C.-F. Westin, and M. Andersson (2011). Representing Local Structure Using Tensors II. In A. Heyden and F. Kahl (Eds.), *Image Analysis*, Volume 6688 of LNCS, pp. 545–556. Springer Berlin / Heidelberg.
- Koenderink, J. (1984). The Structure of Images. *Biological Cybernetics* 50(5), 363–370.
- Koenderink, J. (1993). What is a “Feature”? *Journal of Intelligent Systems* 3(1), 49–82.
- Kolda, T. G. and B. W. Bader (2009). Tensor Decompositions and Applications. *SIAM Rev.* 51(3), 455–500.
- Krieger, G. and C. Zetsche (1996). Nonlinear Image Operators for the Evaluation of Local Intrinsic Dimensionality. *IEEE Transactions on Image Processing* 5(6), 1026–1042.
- Langley, K., D. Fleet, and P. Hibbard (1996). Linear Filtering Precedes Nonlinear Processing in Early Vision. *Current Biology* 6(7), 891–896.
- Lichtenauer, J., E. A. Hendriks, and M. J. T. Reinders (2005). Isophote Properties as Features for Object Detection. *CVPR* (2), 649–654.
- Liew, A. and D. Nguyen (1995). Reconstruction from Wavelet Transform Modulus Maxima using Nonexpansive Projections. *Electronics Letters* 31(13), 1038–1039.
- Lindeberg, T. (1993). On Scale Selection for Differential Operators. In *Proceedings of the Scandinavian Conference on Image Analysis*, Volume 2, pp. 857–857.
- Lindeberg, T. (1994). *Scale-Space Theory in Computer Vision*. Springer.
- Lindeberg, T. (1997). *On The Axiomatic Foundations of Linear Scale-Space*, Volume 8 of *Computational Imaging and Vision*, Chapter 6, pp. 75–98. Kluwer Academic Publishers.
- Lindeberg, T. (1998). Feature Detection with Automatic Scale Selection. *International Journal of Computer Vision* 30(2), 79–116.
- Lowe, D. (1999). Object Recognition from Local Scale-Invariant Features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, Volume 2, pp. 1150–1157. IEEE.
- Lowe, D. G. (2004). Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision* 60, 91–110.
- Makadia, A., L. Sorigi, and K. Daniilidis (2004). Rotation Estimation from Spherical Images. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 3*, Washington, DC, USA, pp. 590–593. IEEE Computer Society.

- Mallat, S. (1999). *A Wavelet Tour of Signal Processing*. Academic Press.
- Mallat, S. and S. Zhong (1992). Characterization of Signals from Multiscale Edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(7), 710–732.
- Marr, D. (1976). Early Processing of Visual Information. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* 275(942), 483–519.
- Marr, D. (1982). *Vision: A Computational Approach*. Freeman & Co., San Francisco.
- Matas, J., O. Chum, M. Urban, and T. Pajdla (2004). Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions. *Image and Vision Computing* 22(10), 761–767.
- Mausfeld, R. (2005). Vom Sinn in den Sinnen. Wie kann ein biologisches System Bedeutung generieren? In „... sind eben auch nur Menschen – Verhalten zwischen Zwang, Freiheit und Verantwortung“, pp. 47–79. Wallstein, Göttingen.
- Mausfeld, R. and V. Knuesel (2002). The Physicalist Trap in Perception Theory. *Perception and the Physical World*, ed. D. Heyer and R. Mausfeld (Wiley, New York).
- Mcewen, J., P. Vielva, Y. Wiaux, R. Barreiro, L. Cayon, M. P. Hobson, A. N. Lasenby, E. Martinez-Gonzalez, and J. Sanz (2007). Cosmological Applications of a Wavelet Analysis on the Sphere. *Journal of Fourier Analysis and Applications* 13(4), 495–510.
- Mikolajczyk, K. and C. Schmid (2004). Scale & Affine Invariant Interest Point Detectors. *International Journal of Computer Vision* 60(1), 63–86.
- Needham, T. (1997). *Visual Complex Analysis*. Oxford University Press, Oxford.
- Newman, E. and R. Penrose (1966). Note on the Bondi-Metzner-Sachs Group. *Journal of Mathematical Physics* 7, 863.
- Nordberg, K., H. Knutsson, and G. Granlund (1993). On the Equivariance of the Orientation and the Tensor Field Representation. In *SCIA8*, pp. 57–63. NOBIM, Norwegian Society for Image Processing and Pattern Recognition.
- Olshausen, B. and D. Field (1997). Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1? *Vision research* 37(23), 3311–3325.
- Olver, P. (2000). *Applications of Lie Groups to Differential Equations*, Volume 107. Springer.
- Otsu, N. (1986). Recognition of Shape and Transformation – An Invariant-theoretical Foundation. *Science on Form*, 413–420.
- Pendleton, D. J. (2003). Euler Angle Geometry, Helicity Basis Vectors, and the Wigner D-function Addition Theorem. *American Journal of Physics* 71(12), 1280–1291.

- Perona, P. (1995). Deformable Kernels for Early Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17(5), 488–499.
- Romeny (Ed.), B. M. (1994). *Geometry-Driven Diffusion in Computer Vision*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Rudin, W. (1976). *Principles of Mathematical Analysis*, Volume 3. McGraw-Hill New York.
- Schaeben, H. (1996). A Unified View of Methods to Resolve the Inverse Problem of Texture Goniometry. *Textures and Microstructures* 25(2), 171–182.
- Schaeben, H. and K. van den Boogaart (2003). Spherical Harmonics in Texture Analysis. *Tectonophysics* 370(1-4), 253–268.
- Schmid, C., R. Mohr, and C. Bauckhage (2000). Evaluation of Interest Point Detectors. *International Journal of Computer Vision* 37(2), 151–172.
- Serre, J. (1977). *Linear Representations of Finite Groups*, Volume 42. Springer Verlag.
- Sommer, G., M. Michaelis, and R. Herpers (1998). The SVD Approach for Steerable Filter Design. In *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, Volume 5, pp. 349–353. IEEE.
- Srinivasan, K., H. Mahawar, and V. Sarin (2005). A Multipole Based Treecode Using Spherical Harmonics for Potentials of the Form r^λ . In *Computational Science - ICCS 2005*, Volume 3514 of LNCS, pp. 107–114. Springer Berlin / Heidelberg.
- Stein, E. M. (1971). *Singular Integrals and Differentiability Properties of Functions*. (PMS-30). Princeton University Press.
- Sylvester, J. (1886). Sur une Extension d'un Théorème de Clebsch Relatif aux Courbes du Quatrième Degré. *Comptes Rendus, Math. Acad. Sci. Paris* 102, 1532–1534.
- Teo, P. and Y. Hel-Or (1998). Design of Multi-Parameter Steerable Functions using Cascade Basis Reduction. In *Computer Vision, 1998. Sixth International Conference on*, pp. 187–192. IEEE.
- Thompson, W. (1994). *Angular Momentum: An Illustrated Guide to Rotational Symmetries for Physical Systems*. Wiley-Interscience.
- Unser, M., K. Balac, and D. Van De Ville (2008). The Monogenic Riesz-Laplace Wavelet Transform. In *Proceedings of the Sixteenth European Signal Processing Conference (EU-SIPCO)*.

- van de Weijer, J., L. J. van Vliet, P. W. Verbeek, and M. van Ginkel (2001). Curvature Estimation in Oriented Patterns Using Curvilinear Models Applied to Gradient Vector Fields. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 23, pp. 1035–1042.
- van Ginkel, M., J. van de Weijer, L. J. van Vliet, and P. W. Verbeek (1999). Curvature Estimation from Orientation Fields. In B. K. Ersboll (Ed.), *11th Scandinavian Conference on Image Analysis*, pp. 545–551. Pattern Recognition Society of Denmark.
- Weickert, J., S. Ishikawa, and A. Imiya (1999). Linear Scale-Space has first been proposed in Japan. *Journal of Mathematical Imaging and Vision* 10(3), 237–252.
- Wiaux, Y., L. Jacques, P. Vielva, , and P. Vandergheynst (2006). Fast Directional Correlation on the Sphere with Steerable Filters. *The Astrophysical Journal* 652(1), 820–832.
- Wietzke, L. (2011). *Algebraic Representation and Geometric Interpretation of Hilbert Transformed Signals*. Ph. D. thesis, Institut für Informatik der Christian-Albrechts-Universität zu Kiel.
- Wietzke, L., O. Fleischmann, A. Sedlazeck, and G. Sommer (2010). Local Structure Analysis by Isotropic Hilbert Transforms. In *Pattern Recognition*, Volume 6376 of LNCS, pp. 131–140. Springer Berlin / Heidelberg.
- Wietzke, L., O. Fleischmann, and G. Sommer (2008a). 2D Image Analysis by Generalized Hilbert Transforms in Conformal Space. In *ECCV (2)*, Volume 5303 of LNCS, pp. 638–649.
- Wietzke, L., O. Fleischmann, and G. Sommer (2008b). Signal Analysis by Generalized Hilbert Transforms on the Unit Sphere. In T. E. Simos (Ed.), *International Conference on Numerical Analysis and Applied Mathematics*, Volume 1048 of AIP Conference Proceedings, pp. 706–709. Melville, New York.
- Wietzke, L., O. Fleischmann, and G. Sommer (2009). The Solution of the Interference-Problem of 2D Waves by Means of Clifford Structures. In T. E. Simos, G. Psihoyios, and C. Tsitouras (Eds.), *International Conference on Numerical Analysis and Applied Mathematics*, Volume 1168 of AIP Conference Proceedings, pp. 817–820. Melville, New York.
- Wietzke, L. and G. Sommer (2010). The Signal Multi-Vector. *Journal of Mathematical Imaging and Vision* 37, 132–150.
- Wietzke, L., G. Sommer, and O. Fleischmann (2009). The Geometry of 2D Image Signals. In *CVPR*, pp. 1690–1697. IEEE Computer Society.

- Witkin, A. (1983). Scale-space filtering. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence - Volume 2*, pp. 1019–1022. Morgan Kaufmann Publishers Inc.
- Yu, W., K. Daniilidis, and G. Sommer (2000). A New 3D Orientation Steerable Filter. In G. Sommer, N. Krüger, and C. Perwass (Eds.), *22. Symposium für Mustererkennung, DAGM 2000, Kiel*, pp. 203–212. Springer-Verlag.
- Zang, D. (2007). *Signal Modeling for Two-Dimensional Image Structures and Scale-Space Based Image Analysis*. Ph. D. thesis, Institut für Informatik und Praktische Mathematik der Christian-Albrechts-Universität zu Kiel. Technical report 0705.
- Zhang, L., T. Qian, and Q. Zeng (2007). Radon Measure Formulation for Edge Detection using Rotational Wavelets. *Communication on Pure and Applied Analysis* 6(3), 899–915.

BIBLIOGRAPHY

Index

- affine characteristic scale, 74
- affine group, 9
- analytic signal, 24
- Cauchy transform
 - in \mathbb{R}^n , 116
 - on the two-sphere, 118
- characteristic scale, 72
- Clifford algebra, 114
- conformal monogenic signal
 - in \mathbb{R}^2 , 98
 - in \mathbb{R}^3 , 110
- conjugate Poisson kernel, 32
- conjugate Poisson transform, 32
- directional correlation, 119
- equivariant
 - function system, 14
 - map, 13
- evaluation functional, 10
- feature, 11
- Fourier series
 - on \mathbb{S}^2 , 115
 - on $SO(3)$, 115
- Fourier transform, 16
- functional, 10
- Gaussian kernel
 - in \mathbb{R}^n , 31
- general linear group, 8
- generalized Hilbert transform, 25
 - fourier multiplier, 25
 - multiscale, 32
 - on the two-sphere, 119
- generalized structure tensor, 68
- group, 7
 - action, 8
 - on functions, 10
 - homomorphism, 8
 - transformation, 8
- higher order
 - derivatives, 38
 - generalized Hilbert transform, 38
- Hilbert transform, 24
- invariant
 - map, 13
- isophote
 - classification, 92
 - curvature, 87, 92
 - of a signal, 86
- kernel
 - convolution, 11
- Legendre functions, 115
- linear feature, 11
- linear functional, 10
- local phase, 28
- matrix group, 9
- monogenic signal, 26
- multiindex notation, 36
- multiscale differential operator, 32
- Operator
 - linear, 10
 - linear shift-invariant, 11
- orbit, 12
- orthogonal group, 9
- outer product, 37

- parameter space
 - of i1D signals, 20
- partial derivatives, 29
- phase
 - local, 24
- Poisson kernel
 - in \mathbb{R}^n , 31
- Poisson transform, 31
- quotient, 12
- representation, 8
- Riesz transform, 25
- Rodrigues formula, 115
- rotation group, 9
- scale equivariance, 72
- scale-space
 - Gaussian, 31
 - linear, 30
 - Poisson, 31
- signal, 7
 - i0D, 18
 - i1D, 19
 - curved, 85
 - i2D, 20
 - sinusoidal i1D, 17
 - curved, 86
 - superimposed, 19
- signal multivector, 36
- singularity
 - of an i1D signal, 62
- special linear group, 9
- spherical harmonics, 115
- stereographic projection, 87
 - inverse, 88
- structure multivector, 35
- structure tensor, 68
- symmetric tensor, 37
- tensor, 36
- tensor field, 41
- Wigner-3j symbols, 126
- Wigner-D functions, 115