INSTITUT FÜR INFORMATIK

Signal Modeling for Two-Dimensional Image Structures and Scale-Space Based Image Analysis

Di Zang

Bericht Nr. 0705
May 2007

CHRISTIAN-ALBRECHTS-UNIVERSITÄT

KIEL

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

Signal Modeling for Two-Dimensional Image Structures and Scale-Space Based Image Analysis

Di Zang

Bericht Nr. 0705
May 2007

e-mail: zd@ks.informatik.uni-kiel.de

Dieser Bericht gibt den Inhalt der Dissertation wieder, die die Verfasserin
im Januar 2007 bei der Technischen Fakultät der
Christian-Albrechts-Universität zu Kiel eingereicht hat.
Datum der Disputation: 19. April 2007.

1. Gutachter Prof. Dr. Gerald Sommer (Kiel)

2. Gutachter Prof. Dr. Reinhard Koch (Kiel)

Datum der mündlichen Prüfung: 19.04.2007

ABSTRACT

Model based image representation plays an important role in many computer vision tasks. Consequently, it is of high significance to model image structures with more powerful representation capabilities. In the literature, there exist bulk of researches for intensity based modeling. However, most of them suffer from the illumination variation. On the other hand, phase information, which carries most essential structural information of the original signal, has the advantage of being invariant to the brightness change. Therefore, phase based image analysis is advantageous when compared to purely intensity based approaches. This thesis aims to propose novel image representations for 2D image structures, from which useful local features can be extracted, which are useful for phase based image analysis. The first approach presents a 2D rotationally invariant quadrature filter. This model is able to handle superimposed intrinsically two-dimensional (i2D) patterns with flexible angles of intersection. Hence, it can be regarded as an extension of the structure multivector. The second approach is the monogenic curvature tensor. Coupling methods of differential geometry, tensor algebra, monogenic signal and quadrature filter, we can design a general model for 2D structures as the monogenic extension of a curvature tensor. Based on it, local representations for the intrinsically one-dimensional (i1D) and i2D structures are derived as the monogenic signal and the generalized monogenic curvature signal, respectively. From them, independent features of local amplitude, phase and orientation are simultaneously extracted. Besides, a generalized monogenic curvature scale-space can be built by applying a Poisson kernel to the monogenic curvature tensor. Compared with other related work, the remarkable advantage of our approach lies in the rotationally invariant phase evaluation of 2D structures in a multi-scale framework, which delivers access to phase-based processing in many computer vision tasks. To demonstrate the efficiency and power of the theoretic framework, some computer vision applications are presented, which include the phase based image reconstruction, detecting i2D image structures using local phase and monogenic curvature tensor for optical flow estimation.

ACKNOWLEDGMENTS

The making of this thesis would not have been possible without the support of many helpful persons. Hence, it is my great pleasure to thank all of them here.

First of all, I would like to thank my supervisor Prof. Gerald Sommer, who has given me the valuable opportunity to finish the research work in the Kiel cognitive systems group. His scientific vision, high criterion, detail comments and discussions has been a major source of inspiration and motivation of my work. Besides, I appreciate his kind encouragements and financial supports during my hard time. I am also grateful to him for having provided me chances to attend many conferences. All these valuable experiences have inspired and stimulated me to conduct my research work.

Furthermore, my special thanks go to Prof. Fred Brackx, Bram De Knock and Hennie De Schepper at the Ghent university of Belgium. Their extensive discussions and theoretic analysis are very helpful for my work.

I thank also Antti Tolvanen, Stephan Al-Zubi and Lennart Wietzke for proofreading different chapters of this thesis. In addition to this, I am grateful for all the discussions with Michael Felsberg, Martin Krause, Lennart Wietzke, Christian Perwass and Oliver Granert. I further thank my colleagues Florian Hoppe, Herward Prehn, Marco Chavarria, Yohannes Kassahun, Christian Gebken, Sven Buchholz, Nils T Siebel for their kind help during my study in the cognitive systems group.

I appreciate the technical support from Gerd Diesner and Henrik Schmidt and also the friendly help in administrative matters from Françoise Maillard and Armgard Kruse.

I acknowledge the DFG Graduiertenkolleg No. 357 for giving me the financial support during my PhD studies.

Finally, I thank all members of my family for their full support and encouragement during these years.

CONTENTS

1. INTRODUCTION

Computer vision research has gained lots of attention over the past decades. In order to handle the dramatically increased low-level computer vision tasks, the methodology of image analysis has also been considerably progressed. As a consequence, image information has been represented in different ways as different types of features for further analysis. Among these, model-based image representation plays a significant role. It combines several important image features into a framework by a parametric way, which contributes a lot for the analysis of image information.

This thesis aims to introduce such novel model-based image representations for analyzing images in a multi-scale way. These novel image representations are derived by combining the knowledge from the mathematics, engineering and computer science fields. It turns out that the presented framework delivers access to many low-level computer vision tasks.

1.1 Motivation

Model based image representation plays an important role in many computer vision tasks such as object recognition, motion estimation, image retrieval, etc. Therefore, signal modeling for local image structures is of high significance in image processing. There are bulk of researches for intensity-based modeling, see [43, 65, 71, 34, 35]. However, most of those approaches suffer from non-stable performances when the illumination varies. Therefore, that intensively investigated area of research is not adequate to model local structures. On the other hand, phase information carries most essential structure information of the original signal [85]. It is invariant with respect to illumination changes. Consequently, modeling of local structures should take both the intensity and phase information into consideration.

In 1D signal processing, the analytic signal [45] is an important complex valued model which can be used for speech recognition, seismic data analysis, airfoil design and so on. The polar representation of the analytic signal yields the local amplitude and local phase, which are measures of quantitative and qualitative information of a signal, respectively. In 1D case, there exist four types of structures, they are the peak, pit, decreasing slope and increasing slope. The local amplitude is invariant with respect to local structures and it indicates the energetic information of the signal. The local phase allows to distinguish structures and it is invariant with respect to the local amplitude. If the local structure varies, the local phase will correspondingly change. Local amplitude and local phase are independent of each other and they fulfill the properties of invariance and equivariance. Invariance means that a feature value is not changed by a certain group acting on a signal. Opposite to invariance, equivariance means there is a monotonic dependency of a feature value on the parameter of the group action. If a set of features includes only invariant and equivariant features, it thus has the property of invariance-equivariance. In addition to satisfying the requirement of invariance and equivariance, if a set of features is at the same time a unique description of the signal, it then performs a split of identity [47]. The split of identity indicates that different features represent mutually different properties of the signal and the whole set of features describes completely the signal. Hence, the analytic signal performs a split of identity.

For 2D images, there exist infinite many types of structures. These can be classified with different features such as their intrinsic dimensions, the number and shape of junctions, or the type of curvature in a differential geometric setting. According to their intrinsic dimensionality, 2D images can locally belong to the intrinsically zero dimensional (i0D) signals which are constant signals, intrinsically one dimensional (i1D) signals representing straight lines and edges and intrinsically two dimensional (i2D) signals which do not belong to the above two cases. The i2D structures are composed of curved edges and lines, junctions, corners and line ends, etc. Intrinsic dimensionality [121] is a local property of a multidimensional signal, which expresses the number of degrees of freedom necessary to describe local structure. The term intrinsic dimension used in image processing corresponds to the term codimension in mathematics. In [121], a discrete concept of the intrinsic dimensionality has been proposed and it was later extended to a continuous one by Krüger and Felsberg [72]. The i1D and i2D structures carry most of the important information of the image, therefore, correct characterization of them has great significance for many computer vision applications.

1.2 Related Work

Many approaches have been proposed for the signal representation of local image structures. The structure tensor [43] and the boundary tensor [65] estimate the main orientation and the energy of the i2D signal. However, the split of identity is lost, because the phase is neglected. In [71], a nonlinear image operator for the detection of locally i2D signals was proposed, but it captures no information about the phase. There are lots of papers concerned with applications of the analytic signal for image analysis. But they have serious problems in transferring that concept from 1D to 2D in a rotation-invariant way. The partial Hilbert transform and the total Hilbert transform [48] provide some representations of the phase in 2D. Unfortunately, they lack the property of rotation invariance and are not adequate for detecting i2D features. Bülow and Sommer [20] proposed the quaternionic analytic signal, which enables the evaluation of the i2D signal phase, however, this approach also has the drawback of being not rotationally invariant.

For i1D signals, Felsberg and Sommer [36] proposed the monogenic signal as a novel model. It is a rotationally invariant generalization of the analytic signal in 2D and higher dimensions. In [70], the 3D monogenic signal has been used for image sequence analysis. From the monogenic signal, the local amplitude, local phase and local orientation can be simultaneously extracted. They deliver an orthogonal decomposition of the original signal into amplitude, phase and orientation. Thus, the monogenic signal has the property of split of identity [36]. However, it captures no information of the i2D part. A 2D phase model was proposed in [31], where the i2D signal is split into two perpendicularly superposed i1D signals and the corresponding two phases are evaluated. The operator derived from that signal model takes advantage of spherical harmonics up to order three. It delivers a new description of i2D structure by a so-called structure multivector. Unfortunately, steering is needed and only i2D patterns superimposed by two perpendicular i1D signals can be correctly handled.

Quite another approach of local signal analysis is based on differential geometry of curves and surfaces [63, 64]. The main points of concern are some invariance properties of signal analysis and regional symmetry with respect to certain combinations of Gaussian and mean curvatures of local surface patterns in a Gaussian multi-scale framework [42, 41]. We will pick up the differential geometry model of surfaces. But instead of a Gaussian blurring operator, we will apply a Poisson blurring operator as a consequence of the algebraic embedding we use.

Our purpose is to build a general model for all 2D structures. This model should contain both the amplitude and phase information of 2D structures in a rotation-invariant manner. In other words, the new model should be an extension of the analytic signal to the 2D case.

1.3 Contributions

In this thesis, we first present a rotationally invariant 2D quadrature filter. It is designed by combining the rotation invariance of the monogenic signal and the symmetry decomposition of the quaternionic analytic signal. The assumed signal model is the superimposed i2D pattern with a flexible angle of intersection. This approach also enables the extraction of local amplitude and local phase in a rotation invariant way. Moreover, it is capable to correctly treat superimposed i2D patterns with flexible angles of intersection. Hence, it is regarded as an extension of the structure multivector [37].

Secondly, we present a novel signal model which covers 2D structures of all intrinsic dimensionalities. By embedding our problem into a certain geometric algebra, more degrees of freedom can be obtained to derive a complete representation for the 2D structure. Based on the differential geometry, we are able to design that general model for 2D structures in a rotation-invariant manner by coupling the methods of tensor algebra, monogenic signal and quadrature filter. The proposed model can be considered as the monogenic extension of a curvature tensor. From this model, a local signal representation for i1D structures is obtained. It is exactly the monogenic signal [36] as a special case of this general model. The local representation for i2D structures, referred as the generalized monogenic curvature signal, can also be derived based on the proposed model. The odd part of the generalized monogenic curvature signal results from the even part by convolving with a second order 2D spherical harmonic, which can be interpreted as a generalized Hilbert transform in the case of i2D signals.

From the generalized monogenic curvature signal, three independent local features can be extracted. They are the amplitude, phase and orientation just like that in the case of the monogenic signal. Hence, the generalized monogenic curvature signal also performs the split of identity, i.e., the invariance-equivariance property of signal decomposition. The energy output (square of the amplitude) can be regarded as a junction strength for detecting the points of interest. The estimated orientation represents the local main orientation of the structure and the phase feature can be used to classify some specific i2D structures.

The amplitude, orientation and phase are coupled to a scale concept in one single framework which is called the generalized monogenic curvature scale-space. The generalized monogenic curvature scale-space is the scale-space of the generalized monogenic curvature signal. But instead of considering that unifying framework, the orthogonal scale-spaces for amplitude, phase and orientation can also be studied separately for local signal analysis. These features are scale dependent, but they are still independent of each other at each scale. The generalized monogenic curvature scale-space is completely different to the well-known Gaussian curvature scale-space [79, 78]. Both the definition of curvature and the scale generating operator are different. Compared with the related research work, our main contribution is the derivation of a general signal model for 2D structures, which enables us to simultaneously extract local features in a multi-scale way. The remarkable advantage lies in the possibility of evaluating the 2D structure phase information in a rotation-invariant manner, which gives access to many phase-based processing in computer vision tasks.

Thirdly, some scale-space based image analysis are presented as applications of the proposed image representation framework. They include the image reconstruction, detection of i2D image structures and optical flow estimation. Parts of this work have already been published in [112, 113, 106, 115, 114, 116, 119, 117, 118, 105].

1.4 Outline of the Thesis

The rest of this thesis is organized as follows. The second chapter introduces the fundamental mathematical background of this thesis. A brief overview of geometric algebra over 2D and 3D Euclidean space is given. We focus on only the group representation and Clifford analysis on the introduced algebra, which are related to the remaining part of this thesis.

Some signal processing techniques are very crucial for the understanding of the proposed new framework. Hence, in chapter three, signal processing basics represented in terms of the geometric algebra are introduced. Besides, the monogenic signal and the monogenic scale-space, as the newly developed low-level image processing framework, are briefly described for better understanding of the following chapters.

The fourth chapter describes a novel approach to design the rotationally invariant 2D quadrature filter. The assumed signal model is the superimposed i2D pattern with a flexible angle of intersection. Combining the rotation invariance of the monogenic signal with the symmetry decomposition of the quaternionic analytic signal, this new quadrature filter can be derived. From it, local amplitudes and phases of two i1D signals are easily to be obtained. Hence, it can be considered as the extension of the structure multivector [37, 31], since the superimposed two i1D signals are not necessarily required to be perpendicular to each other.

The most important contribution of this thesis is given in chapter five. Where we propose a novel image representation to model 2D image structures. This new model is called the monogenic curvature tensor, which consists of the curvature tensor and its harmonic conjugate part. Motivated from the differential geometry, the monogenic curvature tensor can be derived by coupling the methods of tensor algebra, monogenic signal and quadrature filter. It is unified within a scale-space framework by employing 2D damped spherical harmonics as basis functions. From this novel image representation, local amplitudes, main orientations and phases of 2D image structures can be extracted, which delivers access to many applications for the computer vision tasks.

The sixth chapter mainly focuses on the scale-space based image analysis. Three applications based on the presented theoretical framework are introduced, which include image reconstruction, detecting i2D image structures using local phase and optical flow estimation. All these applications demonstrate much advantages over the related work in the literature.

The summary and outlook of this thesis are given in the last chapter. The novel theoretical framework is summarized and important contributions are listed. Open problems and future work are also discussed. This thesis is completed by an appendix and a list of references.

2. GEOMETRIC ALGEBRA

2.1 General Overview

Geometric algebra is one of the most powerful available mathematical systems, which can be seen as the Clifford algebra mainly focusing on a suited geometric interpretation.

In the 19th century, Hamilton (1805-1865) invented the quaternions to generalize complex numbers and rotations to 3D. In a separate development, Grassmann (1809-1877) pioneered the introduction of the exterior or outer algebra. Based on the former work of Hamilton and Grassmann, Clifford (1845-1879) made a crucial step for the mathematical systems by introducing his Clifford algebra. This algebra combined all the advantages of quaternions with those of the vector geometry. The term geometric algebra was introduced by Hestenes in the 1960's, who made further development of Clifford algebra in classical geometry and mechanics.

Geometric algebras constitute a rich family of algebras as generalizations of vector algebra. Compared with the classical framework of vector algebra, the geometric algebra has a more powerful subspace structure. Compared with subspaces of a vector space, subspaces of the geometric algebra are not restricted to interpretations within set theory, instead, they contain group generators and have more powerful geometric meanings. Solving problems in the framework of the geometric algebra contains more degrees of freedom, which makes a tremendous extension of modeling capabilities available.

Geometric algebras have the properties of compact symbolic representations of higher order entities and of linear operations acting on them. However, in the classic framework, different representations are mixed which makes it very hard to formulate coherent derivations. One has to switch between different representations in order to perform necessary calculations or to interpret the results. For example, when one deals with vectors and rotations, the former ones are commonly denoted by column vectors and the latter ones are represented in terms of matrix algebra. In addition to this, complex numbers can also be a representation of 2D vectors, 2D rotations and scale transformations, respectively. In the classical way to formulate multi-dimensional signal theory, those different mathematical approaches for denoting the same entities are mixed. Moreover, rotations by means of matrix algebra are not intuitive with respect to their geometric meanings. The rotation angle and axis are not obviously shown in the corresponding matrix. Therefore, working in the geometric algebra is much easier than using the classic vector algebra since representations in geometric algebra are more simple and elegant. In this thesis, all derivations and image processing approaches are formulated in the framework of geometric algebra.

For a more general introduction to geometric algebra, one may refer to [2, 54, 53, 26, 97, 12, 76]. Much easier to understand the geometric algebra are those in [73, 91, 27, 96]. The GABLE tutorial [27] is combined with a Matlab-package for numerical computations in geometric algebra. There exist also some other softwares which are helpful for computations with geometric algebras. CLICAL [86] is a calculator type computer program for vectors, complex numbers, quaternions, bivectors, spinors, and multivectors in Clifford algebras. Under OpenGL environment, Perwass has developed a software library, called CluCalc, for geometric algebra computations and visualization [90]. The online geometric calculator [60] is an ordinary desk calculator that uses the Clifford numbers over a three dimensional Euclidean Space. Other useful packages for geometric algebras can also be found in [4, 55, 3, 94].

In general, the geometric algebra $\mathbb{R}_{p,q,r}$ is a linear space of multivectors of dimension 2^n, $n = p + q + r$, which results from a vector space $\mathbb{R}^{p,q,r}$. These indices mark the signature of the vector space. Hence, (p, q, r) means we have $p/q/r$ basis vectors which square to $+1/-1/0$, respectively. By choosing the right signature, the geometric algebra will take on certain geometric properties which enables adaption of the embedding framework to the task at hand [104]. For the problems we concentrate on, only the geometric algebras $\mathbb{R}_2$ and $\mathbb{R}_3$ over the Euclidean spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ are needed. In such cases, n is equal to p. The chosen dimensions of the vector spaces come from the fact that an n-dimensional signal will be embedded into an $(n + 1)$-dimensional vector space.

In this chapter, we will give brief introduction to the geometric algebras $\mathbb{R}_2$ and $\mathbb{R}_3$ over the Euclidean spaces $\mathbb{R}^2$ and $\mathbb{R}^3$.

2.2 Multiplying Vectors

In the linear space, a vector has the geometric concept of directed line segment. The well-known products for vectors are the inner product and the cross product, which enable algebraically the geometric ideas of the results.

The inner product is also known as the dot or scalar product. For two

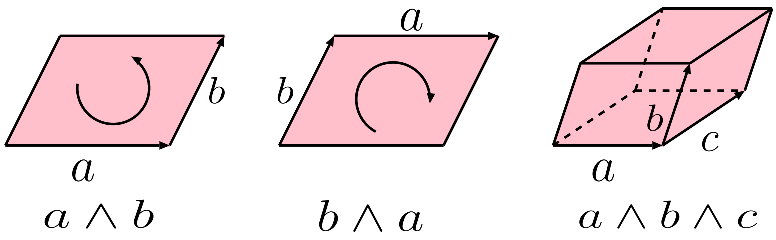

Fig. 2.1: Outer products of vectors.

vectors $\mathbf{x}$ and $\mathbf{y}$, their inner product is defined as

$$\mathbf{x} \cdot \mathbf{y} = |\mathbf{x}||\mathbf{y}| \cos(\theta) \ . \tag{2.1}$$

The result is a scalar with magnitude $|\mathbf{x}||\mathbf{y}| \cos(\theta)$, where $|\mathbf{x}|$ and $|\mathbf{y}|$ are the lengths of $\mathbf{x}$ and $\mathbf{y}$, and θ is the angle between them. Therefore, the inner product of two vectors is effectively an algebraic definition of $\cos(\theta)$. The inner product of two vectors vanishes if they are perpendicular to each other. Hence, this product encompasses partial information of the relative direction of two vectors. To capture the direction information, the vector cross product is introduced. The cross product of vectors $\mathbf{x}$ and $\mathbf{y}$ is $\mathbf{x} \times \mathbf{y}$. The result is a vector of magnitude $|\mathbf{x}||\mathbf{y}| \sin(\theta)$ in the direction perpendicular to the plane defined by $\mathbf{x}$ and $\mathbf{y}$. In this case, $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{x} \times \mathbf{y}$ forms a right-handed set. The cross product exists only in 3D space. In 2D, their exists no such a direction perpendicular to both $\mathbf{x}$ and $\mathbf{y}$ and this direction is ambiguous in four or more higher dimensions. Therefore, a more general concept is required such that full information about relative directions can still be encoded in all dimensions. The outer or exterior product, proposed by Grassmann, provides a way of encoding a plane geometrically without relying on the notion of a vector perpendicular to it.

The outer product of two vectors $\mathbf{x}$ and $\mathbf{y}$, denoted as $\mathbf{x} \wedge \mathbf{y}$, is a bivector. Same as the cross product, the magnitude of $\mathbf{x} \wedge \mathbf{y}$ is $|\mathbf{x}||\mathbf{y}| \sin(\theta)$. The outer product of vectors is anti-commutative and associative

$$\mathbf{x} \wedge \mathbf{y} = -\mathbf{y} \wedge \mathbf{x} \tag{2.2}$$

$$(\mathbf{x} \wedge \mathbf{y}) \wedge \mathbf{z} = \mathbf{x} \wedge (\mathbf{y} \wedge \mathbf{z}) = \mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z} \ . \tag{2.3}$$

Fig. 2.1 visualizes the outer products of vectors. The outer product of two vectors is not a scalar or a vector, but an oriented area in the plane containing $\mathbf{x}$ and $\mathbf{y}$. The outer product of three vectors $\mathbf{x} \wedge \mathbf{y} \wedge \mathbf{z}$ represents an oriented volume.

Clifford united the inner product and outer product into a single framework by introducing his geometric product. The geometric product of two

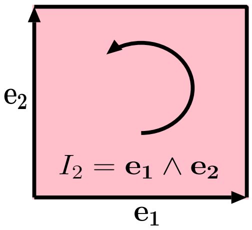

Fig. 2.2: Basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and the unit pseudoscalar $I_2 = \mathbf{e}_{12}$.

vectors is defined as

$$\mathbf{xy} = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \wedge \mathbf{y} \ . \tag{2.4}$$

The right-hand side is a sum of two distinct objects, they are carried into a single entity, which enables many mathematical simplifications.

2.3 Geometric Algebra in 2D Euclidean Space $\mathbb{R}^2$

2.3.1 Definition

The 2D Euclidean space $\mathbb{R}^2$ is spanned by the orthonormal basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. The geometric algebra of the 2D Euclidean space $\mathbb{R}_2$ consists of $2^2 = 4$ elements,

$$\mathbb{R}_2 = span\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{12} = I_2\} \ . \tag{2.5}$$

Here $\mathbf{e}_{12}$ is the bivector or unit pseudoscalar which is formed by the outer product of $\mathbf{e}_1$ and $\mathbf{e}_2$

$$\mathbf{e}_{12} = \mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_2 \wedge \mathbf{e}_1 = I_2 \ . \tag{2.6}$$

All the elements of $\mathbb{R}_2$ are shown in Fig. 2.2, which represents an oriented plane. For every two vectors $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2$, their inner product reads

$$\mathbf{x} \cdot \mathbf{y} = x_1y_1\mathbf{e}_1 \cdot \mathbf{e}_1 + x_1y_2\mathbf{e}_1 \cdot \mathbf{e}_2 + x_2y_1\mathbf{e}_2 \cdot \mathbf{e}_1 + x_2y_2\mathbf{e}_2 \cdot \mathbf{e}_2 = x_1y_1 + x_2y_2 \ . \tag{2.7}$$

Note that in geometric algebra, vectors are not denoted in terms of column or row vectors, but solely as linear combinations of basis elements. The outer product of two vectors is defined as

$$\mathbf{x} \wedge \mathbf{y} = (x_1y_2 - x_2y_1)\mathbf{e}_{12} \ . \tag{2.8}$$

The basic product of geometric algebra is the geometric product, indicated by juxtaposition of the operands. For orthogonal vectors, the geometric product is a pure bivector

$$\mathbf{e}_1\mathbf{e}_2 = \mathbf{e}_{12} = \mathbf{e}_1 \cdot \mathbf{e}_2 + \mathbf{e}_1 \wedge \mathbf{e}_2 = \mathbf{e}_1 \wedge \mathbf{e}_2 = -\mathbf{e}_2 \wedge \mathbf{e}_1 = -\mathbf{e}_{21} \ . \tag{2.9}$$

Tab. 2.1: The geometric product of basis elements $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{12}\}$.

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_{12}$
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_{12}$
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$\mathbf{e}_2$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$-\mathbf{e}_1$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	1

The geometric product of two arbitrary vectors $\mathbf{x}$ and $\mathbf{y}$ can be decomposed into their inner product and outer product

$$\mathbf{x}\mathbf{y} = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \wedge \mathbf{y} = x_1 y_1 + x_2 y_2 + (x_1 y_2 - x_2 y_1)\mathbf{e}_{12} \ . \tag{2.10}$$

The inner and outer product of vectors can also be defined according to the commutator and the anti-commutator of the geometric product

$$\mathbf{x} \cdot \mathbf{y} = \frac{1}{2}(\mathbf{x}\mathbf{y} + \mathbf{y}\mathbf{x}) \tag{2.11}$$

$$\mathbf{x} \wedge \mathbf{y} = \frac{1}{2}(\mathbf{x}\mathbf{y} - \mathbf{y}\mathbf{x}) \ . \tag{2.12}$$

The multiplication table of the basis elements is shown in Tab. 2.1. The bivector $\mathbf{e}_{12} = I_2$ squares to -1 and it represents a rotation by 90^0 from $\mathbf{e}_1$ to $\mathbf{e}_2$. Therefore, the imaginary unit i of the complex numbers can be identified with I_2, yielding an algebra isomorphism. That means every calculation in the complex domain can be performed by a calculation in the algebra of the space spanned by $\{1, I_2\}$. Applying the algebra isomorphism, a rotation by angle θ can thus be represented by $\cos(\theta) + \sin(\theta)I_2 = \exp(\theta I_2)$ according to $\cos(\theta) + \sin(\theta)i = \exp(\theta i)$ in the complex domain.

For the geometric algebra in 2D Euclidean space, we have four basis elements, a general combination of these elements is called a multivector

$$M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_{12} \ . \tag{2.13}$$

The k-grade part of a multivector is obtained from the grade operator $\langle M \rangle_k$. Hence, $\langle M \rangle_0$ is the scalar part of M, $\langle M \rangle_1$ represents the vector part, $\langle M \rangle_2$ indicates the bivector part. Given two multivectors M_1 and M_2, their sum reads

$$\begin{aligned} M_1 + M_2 &= (a_1 + b_1\mathbf{e}_1 + c_1\mathbf{e}_2 + d_1\mathbf{e}_{12}) + (a_2 + b_2\mathbf{e}_1 + c_2\mathbf{e}_2 + d_2\mathbf{e}_{12}) \\ &= (a_1 + a_2) + (b_1 + b_2)\mathbf{e}_1 + (c_1 + c_2)\mathbf{e}_2 + (d_1 + d_2)\mathbf{e}_{12} \ . \end{aligned} \tag{2.14}$$

The geometric product of two multivectors takes the form

$$\begin{aligned} M_1M_2 &= (a_1+b_1\mathbf{e}_1+c_1\mathbf{e}_2+d_1\mathbf{e}_{12})(a_2+b_2\mathbf{e}_1+c_2\mathbf{e}_2+d_2\mathbf{e}_{12}) \\ &= (a_1a_2+b_1b_2+c_1c_2-d_1d_2)+(a_1b_2+b_1a_2-c_1d_2+d_1c_2)\mathbf{e}_1 \\ &+ (a_1c_1+b_1d_2+c_1a_2-d_1b_2)\mathbf{e}_2+(a_1d_2+b_1c_2-c_1b_2+d_1a_2)\mathbf{e}_{12}\ . \end{aligned} \tag{2.15}$$

Multivector addition and multiplication obey the associative and distributive laws. The geometric product of a multivector M by I_2^{-1} from the right is called the dual of M

$$M^* = MI_2^{-1} = -M\mathbf{e}_{12}\ . \tag{2.16}$$

The reverse of a multivector is denoted by a tilde and it takes the form

$$\widetilde{M} = \langle M\rangle_0 + \langle M\rangle_1 - \langle M\rangle_2\ . \tag{2.17}$$

The reversion describes the result when basis vectors are multiplied in the reverse order. Thereby, scalars and vectors remain unchanged, but bivectors are negated. The role that the reversion plays is just like the complex conjugate in many cases. The reversion is involutive and anti-automorphic, this means

$$\widetilde{\widetilde{M}} = M \quad \text{and} \quad \widetilde{MN} = \widetilde{N}\widetilde{M}\ . \tag{2.18}$$

The modulus of a multivector can be obtained from its reverse, that is

$$|M| = \sqrt{\langle M\widetilde{M}\rangle_0}\ . \tag{2.19}$$

As an alternative way, the modulus of a multivector $M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_{12}$ is also given by its Euclidean norm

$$|M| = \sqrt{a^2+b^2+c^2+d^2}\ . \tag{2.20}$$

Hence, for a vector $\mathbf{x}$, its modulus can be obtained from the square root of the product with itself. Its inverse vector is correspondingly given by $\mathbf{x}^{-1} = \frac{\mathbf{x}}{\mathbf{x}\cdot\mathbf{x}}$.

2.3.2 Reflection and Rotation

A special case of the multivector is the spinor, which is combined by a scalar and a bivector, e.g. $S = a + dI_2 = r\exp(\theta I_2)$. Here, r represents a scaling and the angle θ indicates a rotation. Therefore, the spinor is equivalent to a scaling-rotation. The inverse of a spinor reads

$$S^{-1} = r^{-1}\exp(-\theta I_2) = \frac{\widetilde{S}}{|S|^2}\ . \tag{2.21}$$

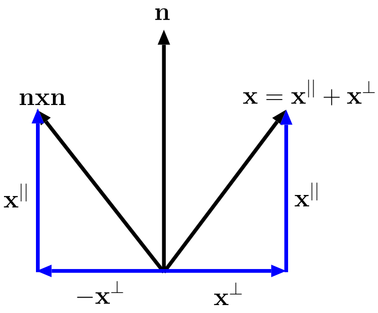

Fig. 2.3: Reflection of vector $\mathbf{x}$ with respect to vector $\mathbf{n}$.

All spinors in $\mathbb{R}_2$ form an even-grade subalgebra $\mathbb{R}_2^+$, which is isomorphic to the complex algebra, i.e $\mathbb{R}_2^+ \simeq \mathbb{C}$. Hence, complex numbers which are interpreted as scaling-rotations, are now replaced with spinors. This means $z = x + iy = r\exp(\theta i)$ can be substituted by $S = x + yI_2 = r\exp(\theta I_2)$. However, complex numbers which interpret vectors are then represented as the action of spinors on the basis vector $\mathbf{e}_1$, i.e $\mathbf{x} = \mathbf{e}_1 S = x\mathbf{e}_1 + y\mathbf{e}_2$.

By interpreting the scaling-rotation in terms of Lie groups [99, 9], the unit pseudoscalar I_2 is considered as the generator of the rotation. The group of scaling is generated by the scalar 1. The logarithm of the spinor plays an important role for the definition of the local phase. For any spinor $s \in \mathbb{R}_2^+$, its logarithm can be defined as [52]

$$\log(S) = \log(|S|) + \frac{\langle S\rangle_2}{|\langle S\rangle_2|}\text{atan}\left(\frac{|\langle S\rangle_2|}{\langle S\rangle_0}\right) , \tag{2.22}$$

where $\log(|S|)$ is the ordinary real logarithm of the modulus of S, atan denotes the inverse tangent mapping for the interval $[0, \pi)$.

The interpretation of reflection, same as the scaling-rotation, can also be represented in an algebraic way (see Fig. 2.3). Let $\mathbf{n}$ be a unit vector describing the orientation of an arbitrary line, any vector $\mathbf{x}$ which is reflected with respect to this arbitrary line can be split into two parts, i.e. $\mathbf{x} = \mathbf{x}^{||} + \mathbf{x}^{\perp}$. Here, $\mathbf{x}^{||}$ is the component of $\mathbf{x}$ parallel to $\mathbf{n}$ and $\mathbf{x}^{\perp}$ refers to the part of $\mathbf{x}$ which is perpendicular to $\mathbf{n}$. The part $\mathbf{x}^{||}$, called the projection of $\mathbf{x}$ on $\mathbf{n}$, is given by projecting $\mathbf{x}$ onto $\mathbf{n}$

$$\mathbf{x}^{||} = (\mathbf{x} \cdot \mathbf{n})\mathbf{n} . \tag{2.23}$$

The other component $\mathbf{x}^{\perp}$ is called the rejection of $\mathbf{x}$ with respect to $\mathbf{n}$ and it takes the following form

$$\mathbf{x}^{\perp} = \mathbf{x} - \mathbf{x}^{||} = \mathbf{x} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n} = (\mathbf{x}\mathbf{n})\mathbf{n} - (\mathbf{x} \cdot \mathbf{n})\mathbf{n} = (\mathbf{x}\mathbf{n} - \mathbf{x} \cdot \mathbf{n})\mathbf{n} = (\mathbf{x} \wedge \mathbf{n})\mathbf{n} . \tag{2.24}$$

From Fig. 2.3, it is obviously shown that the reflection of $\mathbf{x}$ can be obtained by $\mathbf{x}^{\perp} - \mathbf{x}^{||}$. Considering equations (2.23) and (2.24), we are able to have the reflected result

$$\begin{aligned} \mathbf{x}^{||} - \mathbf{x}^{\perp} &= (\mathbf{x} \cdot \mathbf{n})\mathbf{n} - (\mathbf{x} \wedge \mathbf{n})\mathbf{n} \\ &= (\mathbf{n} \cdot \mathbf{x})\mathbf{n} + (\mathbf{n} \wedge \mathbf{x})\mathbf{n} \\ &= (\mathbf{n} \cdot \mathbf{x} + \mathbf{n} \wedge \mathbf{x})\mathbf{n} \\ &= (\mathbf{n}\mathbf{x})\mathbf{n} = \mathbf{n}\mathbf{x}\mathbf{n} \ . \end{aligned} \tag{2.25}$$

Therefore, the reflected result of any vector $\mathbf{x}$ with respect to the orientation vector $\mathbf{n}$ is given by the two-sided multiplication with the orientation vector.

The rotation of a vector $\mathbf{x}$ can be implemented based on successive reflections, i.e. $\mathbf{n}_2\mathbf{n}_1\mathbf{x}\mathbf{n}_1\mathbf{n}_2$, where $\mathbf{n}_1$ and $\mathbf{n}_2$ denote two orientation vectors. In this way, the rotation angle is two times the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$, this means

$$\begin{aligned} \mathbf{x}' &= \mathbf{n}_2\mathbf{n}_1\mathbf{x}\mathbf{n}_1\mathbf{n}_2 = (\mathbf{n}_2\mathbf{n}_1)\mathbf{x}(\mathbf{n}_1\mathbf{n}_2) \\ &= \mathbf{x}(\mathbf{n}_1\mathbf{n}_2)^2 = \mathbf{x}(\cos(\theta) + \sin(\theta)\mathbf{e}_{12})^2 \\ &= \mathbf{x}\exp(2\theta\mathbf{e}_{12}) \ , \end{aligned} \tag{2.26}$$

where θ is the angle between the orientation vectors $\mathbf{n}_1$ and $\mathbf{n}_2$.

Up to now, we have understood some basic definitions of the geometric algebra in 2D Euclidean space. Some important features of $\mathbb{R}_2$ are in accordance emphasized as below

- The geometric product of two parallel vectors is a scalar number, which is the product of their lengths.
- The geometric product of two perpendicular vectors is a bivector, it represents the oriented area formed by the two vectors.
- Parallel vectors commute under the geometric product, however, perpendicular vectors are anti-commutative.
- The bivector squares to -1. It has the geometric effect of rotating $\mathbf{e}_1$ basis vector to $\mathbf{e}_2$ by the angle of 90^0.
- Complex numbers are contained in $\mathbb{R}_2$ as a subalgebra which is isomorphic to the even-grade subalgebra $\mathbb{R}_2^+$.
- The subalgebra of spinors represent the group action of scaling-rotation over $\mathbb{R}_2$.

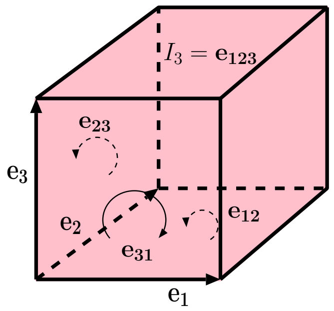

Fig. 2.4: Basis elements of $\mathbb{R}_3$.

Tab. 2.2: The geometric product of basis elements in $\mathbb{R}_3$.

	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
1	1	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	I_3
$\mathbf{e}_1$	$\mathbf{e}_1$	1	$\mathbf{e}_{12}$	$-\mathbf{e}_{31}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	$\mathbf{e}_{23}$
$\mathbf{e}_2$	$\mathbf{e}_2$	$-\mathbf{e}_{12}$	1	$\mathbf{e}_{23}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{31}$
$\mathbf{e}_3$	$\mathbf{e}_3$	$\mathbf{e}_{31}$	$-\mathbf{e}_{23}$	1	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$\mathbf{e}_{12}$
$\mathbf{e}_{23}$	$\mathbf{e}_{23}$	I_3	$-\mathbf{e}_3$	$\mathbf{e}_2$	-1	$-\mathbf{e}_{12}$	$\mathbf{e}_{31}$	$-\mathbf{e}_1$
$\mathbf{e}_{31}$	$\mathbf{e}_{31}$	$\mathbf{e}_3$	I_3	$-\mathbf{e}_1$	$\mathbf{e}_{12}$	-1	$-\mathbf{e}_{23}$	$-\mathbf{e}_2$
$\mathbf{e}_{12}$	$\mathbf{e}_{12}$	$-\mathbf{e}_2$	$\mathbf{e}_1$	I_3	$-\mathbf{e}_{31}$	$\mathbf{e}_{23}$	-1	$-\mathbf{e}_3$
I_3	I_3	$\mathbf{e}_{23}$	$\mathbf{e}_{31}$	$\mathbf{e}_{12}$	$-\mathbf{e}_1$	$-\mathbf{e}_2$	$-\mathbf{e}_3$	-1

2.4 Geometric Algebra in 3D Euclidean Space $\mathbb{R}^3$

2.4.1 Basic Definitions

The Euclidean space $\mathbb{R}^3$ is spanned by the orthonormal basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The geometric algebra of the 3D Euclidean space $\mathbb{R}_3$ consists of $2^3 = 8$ elements,

$$\mathbb{R}_3 = span\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_{23}, \mathbf{e}_{31}, \mathbf{e}_{12}, \mathbf{e}_{123} = I_3\} \ . \tag{2.27}$$

All the elements of $\mathbb{R}_3$ are illustrated in Fig. 2.4. Here $\mathbf{e}_{23}$, $\mathbf{e}_{31}$ and $\mathbf{e}_{12}$ are the unit bivectors and the element $\mathbf{e}_{123} = I_3$ is a trivector or unit pseudoscalar. These three bivectors are independent of each other, each encodes a distinct plane. Sweeping the bivector $\mathbf{e}_1\mathbf{e}_2$ along the vector $\mathbf{e}_3$ results in the trivector I_3, which is a 3D volume element. The squares of bivectors and the trivector equal one, and the trivector commutates with all elements in $\mathbb{R}_3$. The geometric products of these basis elements are shown in Tab. 2.2. According to the habits of geometric calculus [54], the differential operator is always combined with the corresponding basis. Hence, the vector-valued 3D nabla operator is defined as

$$\nabla_3 = \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2} + \mathbf{e}_3 \frac{\partial}{\partial x_3} \ . \tag{2.28}$$

The Laplacian is a scalar operator, which has the following form

$$\Delta_3 = \nabla_3 \nabla_3 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \ . \tag{2.29}$$

Therefore, for a vector $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ in the 3D Euclidean space, its derivative with respect to the basis vectors is given by

$$\nabla_3 \mathbf{x} = \mathbf{e}_1 \frac{\partial}{\partial x_1}\mathbf{x} + \mathbf{e}_2 \frac{\partial}{\partial x_2}\mathbf{x} + \mathbf{e}_3 \frac{\partial}{\partial x_3}\mathbf{x} \ . \tag{2.30}$$

In accordance, its Laplacian is shown as below

$$\Delta_3 \mathbf{x} = \frac{\partial^2}{\partial x_1^2}\mathbf{x} + \frac{\partial^2}{\partial x_2^2}\mathbf{x} + \frac{\partial^2}{\partial x_3^2}\mathbf{x} \ . \tag{2.31}$$

Any two vectors in the 3D Euclidean space, e.g. $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3$, can be multiplied by using the geometric product, and we still have

$$\mathbf{x}\mathbf{y} = \mathbf{x} \cdot \mathbf{y} + \mathbf{x} \wedge \mathbf{y} \ . \tag{2.32}$$

The inner product of two vectors is defined as

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + x_3 y_3 \ . \tag{2.33}$$

The bivector $\mathbf{x} \wedge \mathbf{y}$ belongs to the 3D space, spanned by the $\mathbf{e}_i \wedge \mathbf{e}_j$ (for all i, j). The outer product of two 3D vectors is given by

$$\mathbf{x} \wedge \mathbf{y} = (x_2y_3 - x_3y_2)\mathbf{e}_2 \wedge \mathbf{e}_3 + (x_3y_1 - x_1y_3)\mathbf{e}_3 \wedge \mathbf{e}_1 + (x_1y_2 - x_2y_1)\mathbf{e}_1 \wedge \mathbf{e}_2 \ . \tag{2.34}$$

The components are the same as those of the cross product, but the result is a bivector rather than a vector. The bivector $\mathbf{x} \wedge \mathbf{y}$ represents a plane spanned by $\mathbf{x}, \mathbf{y} \in \mathbf{R}^3$. The dual of this plane, given by $(\mathbf{x} \wedge \mathbf{y})^*$, is shown in Fig. 2.5. It indicates the normal of this plane. The side of the plane from which the normal $(\mathbf{x} \wedge \mathbf{y})^*$ sticks out from is usually regarded as the "front"-side of the plane. Thus, a bivector represents a sided plane. Accordingly, the dual of the sided plane $\mathbf{y} \wedge \mathbf{x}$ is $(\mathbf{y} \wedge \mathbf{x})^*$, which points in an opposite direction. The idea of a plane normal vector does only work in $\mathbb{R}^3$. In any dimension higher than three, the set of vectors perpendicular to one vector spans a higher dimensional space than a plane.

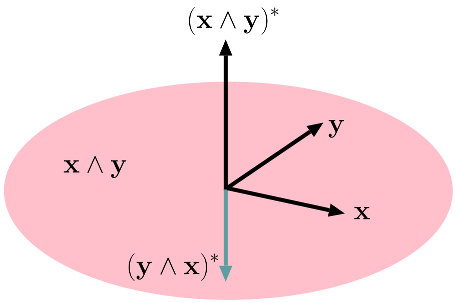

Fig. 2.5: Dual of plane represented by bivector $\mathbf{x}\wedge\mathbf{y}$.

Due to the orthogonality of basis vectors, their outer products are equivalent to their geometric products

$$\mathbf{e}_1\wedge\mathbf{e}_2=\mathbf{e}_1\mathbf{e}_2=\mathbf{e}_{12} \tag{2.35}$$

$$\mathbf{e}_2\wedge\mathbf{e}_3=\mathbf{e}_2\mathbf{e}_3=\mathbf{e}_{23} \tag{2.36}$$

$$\mathbf{e}_3\wedge\mathbf{e}_1=\mathbf{e}_3\mathbf{e}_1=\mathbf{e}_{31}\ . \tag{2.37}$$

An arbitrary bivector is a linear combination of these basis bivectors,

$$B=a\mathbf{e}_{12}+b\mathbf{e}_{23}+c\mathbf{e}_{23}\ . \tag{2.38}$$

The interpretation of bivector addition is most easily seen when the bivectors are expressed in terms of the outer product with a common vector factor. In 3D space, this is always possible because any two planes will either be parallel or intersect along a common line. Given two bivectors $B_1=\mathbf{x}\wedge\mathbf{z}$ and $B_2=\mathbf{y}\wedge\mathbf{z}$, the bivector addition is defined as

$$B_1+B_2=\mathbf{x}\wedge\mathbf{z}+\mathbf{y}\wedge\mathbf{z}=(\mathbf{x}+\mathbf{y})\wedge\mathbf{z}\ . \tag{2.39}$$

The geometric significance of this can be depicted in Fig. 2.6. By decomposing the vector $\mathbf{x}$ and $\mathbf{y}$ into components parallel and perpendicular to $\mathbf{x}+\mathbf{y}$, we have

$$\mathbf{x}=\mathbf{x}^{||}+\mathbf{x}^{\perp},\quad \mathbf{y}=\mathbf{y}^{||}+\mathbf{y}^{\perp}\ , \tag{2.40}$$

where $\mathbf{y}^{\perp}=-\mathbf{x}^{\perp}$. In this case, the addition of bivectors in 3D is reduced to the addition of coplanar bivectors. This is evidently shown as follow

$$\mathbf{x}\wedge\mathbf{z}+\mathbf{y}\wedge\mathbf{z}=(\mathbf{x}+\mathbf{y})\wedge\mathbf{z}=(\mathbf{x}^{||}+\mathbf{y}^{||})\wedge\mathbf{z}=\mathbf{x}^{||}\wedge\mathbf{z}+\mathbf{y}^{||}\wedge\mathbf{z}\ . \tag{2.41}$$

The geometric product of two basis bivectors results in a third basis bivector, for example,

$$(\mathbf{e}_2\wedge\mathbf{e}_3)(\mathbf{e}_1\wedge\mathbf{e}_2)=\mathbf{e}_3\mathbf{e}_2\mathbf{e}_2\mathbf{e}_1=\mathbf{e}_3\mathbf{e}_1 \tag{2.42}$$

$$(\mathbf{e}_1\wedge\mathbf{e}_2)(\mathbf{e}_2\wedge\mathbf{e}_3)=\mathbf{e}_1\mathbf{e}_2\mathbf{e}_2\mathbf{e}_3=\mathbf{e}_1\mathbf{e}_3=-\mathbf{e}_3\mathbf{e}_1\ . \tag{2.43}$$

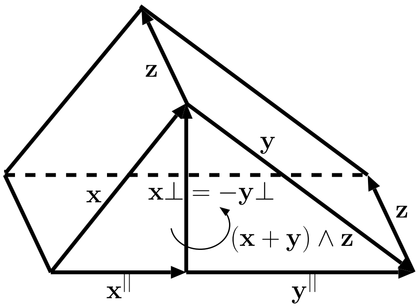

Fig. 2.6: Addition of two bivectors.

It is obvious that the geometric products of bivectors are anti-commutative. Furthermore, the basis bivectors all square to -1. These are the properties of the generators of the quaternion algebra. Quaternions are generalized complex numbers of the form $q = a + ib + jc + kd$, where the generalized imaginary units i, j and k satisfy the following multiplication rules:

$$i^2 = j^2 = k^2 = ijk = -1 \tag{2.44}$$

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik \; . \tag{2.45}$$

Therefore, we can identify the imaginary units i, j and k of quaternions with the following bivectors in $\mathbb{R}_3$.

$$i \rightarrow \mathbf{e}_{23}, \quad j \rightarrow \mathbf{e}_{12}, \quad k \rightarrow \mathbf{e}_{31} \; . \tag{2.46}$$

The property of the quaternion imaginary units based on the chosen bivectors can be checked as follow

$$i^2 = (\mathbf{e}_{23})^2 = -1, \quad j^2 = (\mathbf{e}_{12})^2 = -1, \quad k^2 = (\mathbf{e}_{31})^2 = -1 \tag{2.47}$$

$$ijk = \mathbf{e}_{23}\mathbf{e}_{12}\mathbf{e}_{31} = -1 \tag{2.48}$$

$$ij = \mathbf{e}_{23}\mathbf{e}_{12} = \mathbf{e}_{31} = k \tag{2.49}$$

$$jk = \mathbf{e}_{12}\mathbf{e}_{31} = \mathbf{e}_{23} = i \tag{2.50}$$

$$ki = \mathbf{e}_{31}\mathbf{e}_{23} = \mathbf{e}_{12} = j \; . \tag{2.51}$$

Hence, the subalgebra with basis $\{1, \mathbf{e}_{23}, \mathbf{e}_{12}, \mathbf{e}_{31}\}$ is isomorphic to the quaternions $\mathbb{H}$. However, the above mentioned identification is only one possible isomorphism.

In 3D space, the geometric product of a vector and a bivector is expressed as $\mathbf{x}B$. To understand the properties of this result, the vector $\mathbf{x}$ is decomposed into two terms

$$\mathbf{x} = \mathbf{x}^{||} + \mathbf{x}^{\perp} \; , \tag{2.52}$$

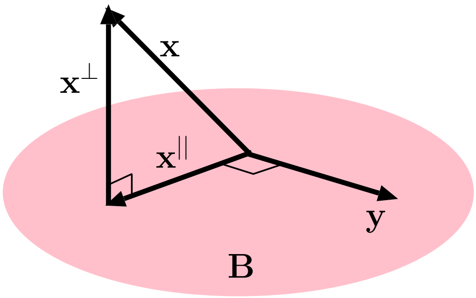

Fig. 2.7: Inner product of a vector and a bivector.

where $\mathbf{x}^{||}$ indicates the part which is parallel to the plane B and $\mathbf{x}^{\perp}$ is the component perpendicular to the plane. This plane can thus be represented as $B = \mathbf{x}^{||} \wedge \mathbf{y}$ with $\mathbf{y}$ being a vector that is orthogonal to $\mathbf{x}^{||}$ in this plane (see Fig. 2.7). Hence, the product $\mathbf{x}B$ takes the following form

$$\mathbf{x}B = (\mathbf{x}^{||} + \mathbf{x}^{\perp})(\mathbf{x}^{||} \wedge \mathbf{y}) = \mathbf{x}^{||}(\mathbf{x}^{||} \wedge \mathbf{y}) + \mathbf{x}^{\perp}(\mathbf{x}^{||} \wedge \mathbf{y}) \ . \tag{2.53}$$

The first component of equation (2.53) is further written as

$$\mathbf{x}^{||}(\mathbf{x}^{||} \wedge \mathbf{y}) = \mathbf{x}^{||}(\mathbf{x}^{||}\mathbf{y}) = (\mathbf{x}^{||})^2\mathbf{y} \ , \tag{2.54}$$

which is a vector in the $\mathbf{y}$ direction. The second component of equation (2.53) reads

$$\mathbf{x}^{\perp}(\mathbf{x}^{||} \wedge \mathbf{y}) = \mathbf{x}^{\perp} \wedge \mathbf{x}^{||} \wedge \mathbf{y} = \mathbf{x}^{\perp}\mathbf{x}^{||}\mathbf{y} \ . \tag{2.55}$$

It is the outer product of three orthogonal vectors, namely, a trivector. Therefore, the geometric product of a vector and a bivector results in two terms, i.e. a vector and a trivector. Hence, it can be obtained as

$$\mathbf{x}B = \mathbf{x} \cdot B + \mathbf{x} \wedge B \ . \tag{2.56}$$

The inner product $\mathbf{x} \cdot B = \mathbf{x}^{||} \cdot B$ now subtracts the subspace represented by $\mathbf{x}^{||}$ from the subspace indicated by B. Opposite to this, the outer product adds the subspace represented by $\mathbf{x}^{\perp}$ to the subspace indicated by $\mathbf{x}^{||} \wedge \mathbf{y}$.

The inner product between a vector and a bivector is anti-symmetric,

$$\mathbf{x} \cdot B = (\mathbf{x}^{||})^2\mathbf{y} = -(\mathbf{x}^{||}b)\mathbf{x}^{||} = -B \cdot \mathbf{x} \ . \tag{2.57}$$

Thereby, as an alternative, it can also be obtained as

$$\mathbf{x} \cdot B = \frac{1}{2}(\mathbf{x}B - B\mathbf{x}) \ . \tag{2.58}$$

The outer product $\mathbf{x} \wedge B$ returns a trivector which is symmetric

$$\mathbf{x} \wedge B = \mathbf{x}^{\perp}\mathbf{x}^{||}\mathbf{y} = \mathbf{x}^{||}\mathbf{y}\mathbf{x}^{\perp} = B \wedge \mathbf{x} \ . \tag{2.59}$$

Hence, it can be alternatively written as

$$\mathbf{x} \wedge B = \frac{1}{2}(\mathbf{x}B + B\mathbf{x}) \ . \tag{2.60}$$

A general combination of the elements in $\mathbb{R}_3$ is called a multivector,

$$M = a + b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_3 + e\mathbf{e}_{23} + f\mathbf{e}_{31} + g\mathbf{e}_{12} + hI_3 \ . \tag{2.61}$$

The geometric product of two multivectors M_1 and M_2 is indicated by juxtaposition of M_1 and M_2, i.e. M_1M_2. The k-grade part of a multivector is obtained from the grade operator $\langle M\rangle_k$. A blade of grade k, i.e. a k-blade B_k, is the outer product ($\wedge$) of k independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k \in \mathbb{R}^3$

$$B_k = \mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k = \langle \mathbf{x}_1 ... \mathbf{x}_k \rangle_k \ . \tag{2.62}$$

Hence, $\langle M\rangle_0$ is the scalar part of M, $\langle M\rangle_1$ represents the vector part, $\langle M\rangle_2$ indicates the bivector part and $\langle M\rangle_3$ is the trivector part, which commutes with every element of the $\mathbb{R}_3$.

2.4.2 Multiplication of Multivectors

The inner product of two multivectors is defined as

$$M_1 \cdot M_2 = \langle M_1M_2\rangle_{|r-s|} \ , \tag{2.63}$$

where $M_1 = \langle M_1\rangle_r$, $M_2 = \langle M_2\rangle_s$ and $r, s > 0$. For some specific cases, the following rules occur:

- If M_1 and M_2 are scalars, which means $r = s = 0$, their inner product is zero.
- For two vectors, their grades satisfy $r = s = 1$. Hence, the inner product is their scalar product.
- The inner product of a vector $\mathbf{x}$ ($r = 1$) and a bivector B ($s = 2$) is given by the vector part of their geometric product

$$\mathbf{x} \cdot B = \langle \mathbf{x}B\rangle_1 = \frac{1}{2}(\mathbf{x}B - B\mathbf{x}) = -B \cdot \mathbf{x} \ . \tag{2.64}$$

- For a vector $\mathbf{x}$ $(r = 1)$ and a pseudoscalar λI_3 $(s = 3)$, their inner product is identical to their geometric product

$$\mathbf{x} \cdot \lambda I_3 = \lambda \mathbf{x} I_3 = (\lambda I_3) \cdot \mathbf{x} \; . \tag{2.65}$$

- The inner product of two bivectors B_1 and B_2 $(r = s = 2)$ reads

$$B_1 \cdot B_2 = \langle B_1 B_2 \rangle_0 = \frac{1}{2}(B_1 B_2 + B_2 B_1) = B_2 \cdot B_1 \; . \tag{2.66}$$

- For a bivector B $(r = 2)$ and a pseudoscalar λI_3 $(s = 3)$, their inner product is given by their geometric product

$$B \cdot \lambda I_3 = \lambda B I_3 = (\lambda I_3) \cdot B \; . \tag{2.67}$$

- Given two pseudoscalars $\lambda_1 I_3$ and $\lambda_2 I_3$ $(r = s = 3)$, their inner product is just their geometric product

$$(\lambda_1 I_3) \cdot (\lambda_2 I_3) = -\lambda_1 \lambda_2 = (\lambda_2 I_3) \cdot (\lambda_1 I_3) \; . \tag{2.68}$$

The outer product of two multivectors takes this form

$$M_1 \wedge M_2 = \langle M_1 M_2 \rangle_{r+s} \; , \tag{2.69}$$

where $M_1 = \langle M_1 \rangle_r$, $M_2 = \langle M_2 \rangle_s$ and $r + s < 3$. To be more specific, we have the following cases:

- If either M_1 or M_2 is a scalar, that is $r = 0$ or $s = 0$, their outer product is identical to the scalar multiplication.
- Given two vectors $\mathbf{x}$ and $\mathbf{y}$ $(r = s = 1)$, the outer product reads

$$\mathbf{x} \wedge \mathbf{y} = (x_2 y_3 - x_3 y_2)\mathbf{e}_2\mathbf{e}_3 + (x_3 y_1 - x_1 y_3)\mathbf{e}_3\mathbf{e}_1 + (x_1 y_2 - x_2 y_1)\mathbf{e}_1\mathbf{e}_2 \; . \tag{2.70}$$

- For a vector $\mathbf{x}$ $(r = 1)$ and a bivector B $(s = 2)$, their outer product is given by the pseudoscalar part of their geometric product

$$\mathbf{x} \wedge B = \langle \mathbf{x} B \rangle_3 = \frac{1}{2}(\mathbf{x} B + B \mathbf{x}) = B \wedge \mathbf{x} \; . \tag{2.71}$$

Except for the case $(r = s = 2)$, the geometric product of two multivectors is always the combination of their inner product and their outer product

$$M_1 M_2 = M_1 \cdot M_2 + M_1 \wedge M_2 \; . \tag{2.72}$$

The geometric product of two bivectors B_1 and B_2 is the combination of their inner product and their cross product

$$B_1 B_2 = B_1 \cdot B_2 + B_1 \times B_2 \ , \tag{2.73}$$

where the cross product can be written as

$$B_1 \times B_2 = -B_2 \times B_1 = \frac{1}{2}(B_1 B_2 - B_2 B_1) \ . \tag{2.74}$$

The relationship between this cross product and the Gibbs cross product can be established through the duality. For a multivector M, its dual is obtained from the product of M with the inverse of the unit pseudoscalar

$$M^* = M I_3^{-1} = -M I_3 \ . \tag{2.75}$$

Hence, in the Euclidean 3D space, the dual of a vector is a bivector. This bivector is formed by two linear independent vectors which are orthogonal to the vector. Analogously, the dual of a bivector is a vector which is perpendicular to the plane represented by the bivector. Therefore, the duality maps the 3D bivector space into the 3D vector space and vice versa.

In terms of duality, the Gibbs cross product equals the cross product of geometric algebra. Given two vectors $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ and $\mathbf{y} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3$, their Gibbs cross product can be obtained as

$$\mathbf{x} \times_G \mathbf{y} = (x_2y_3 - x_3y_2)\mathbf{e}_1 + (x_1y_3 - x_3y_1)\mathbf{e}_2 + (x_2y_1 - x_1y_2)\mathbf{e}_3 \ . \tag{2.76}$$

The cross product of the dual of $\mathbf{x}$ and the dual of $\mathbf{y}$ is given by

$$\begin{aligned} \mathbf{x}^* \times \mathbf{y}^* &= (x_1\mathbf{e}_{23} + x_2\mathbf{e}_{31} + x_3\mathbf{e}_{12}) \times (y_1\mathbf{e}_{23} + y_2\mathbf{e}_{31} + y_3\mathbf{e}_{12}) \qquad (2.77) \\ &= (x_3y_2 - x_2y_3)\mathbf{e}_{23} + (x_3y_1 - x_1y_3)\mathbf{e}_{31} + (x_1y_2 - x_2y_1)\mathbf{e}_{12} \ . \end{aligned}$$

Hence, the following form can be derived

$$(\mathbf{x} \times_G \mathbf{y})^* = \mathbf{x}^* \times \mathbf{y}^* = \mathbf{y} \wedge \mathbf{x} \ . \tag{2.78}$$

Having different products occurring in $\mathbb{R}_3$, some precedence rules for the products should be taken into consideration. The wedge product binds strongest, the inner product binds stronger than the geometric product.

The modulus of a multivector in $\mathbb{R}_3$, same as that in $\mathbb{R}_2$, can be defined as

$$|M| = \sqrt{\langle M\widetilde{M}\rangle_0} \ , \tag{2.79}$$

where $\widetilde{M}$ refers to the reverse of M and it takes the following form

$$\widetilde{M} = \langle M\rangle_0 + \langle M\rangle_1 - \langle M\rangle_2 - \langle M\rangle_3 \ . \tag{2.80}$$

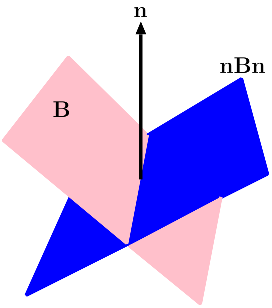

Fig. 2.8: Reflection of a bivector with respect to a vector.

The reverse of a multivector is obtained by reversing the order of basis elements. The scalar and vector components remain the same, but the bivector and the trivector parts are negated. As for $\mathbb{R}_2$, the reversion is an involutive anti-automorphism. The main automorphism of geometric algebra is the grade involution

$$\widehat{M} = \langle M \rangle_0 - \langle M \rangle_1 + \langle M \rangle_2 - \langle M \rangle_3 \ . \tag{2.81}$$

The grade involution satisfies the rule

$$(\widehat{M_1 M_2}) = \widehat{M_1}\widehat{M_2} \ . \tag{2.82}$$

The conjugation of a multivector [73] results from the combination of the grade involution and reversion

$$\overline{M} = \widehat{\widetilde{M}} = \langle M \rangle_0 - \langle M \rangle_1 - \langle M \rangle_2 + \langle M \rangle_3 \ . \tag{2.83}$$

2.4.3 Reflection and Rotation

In contrast to $\mathbb{R}_2$, the reflection in $\mathbb{R}_3$ has more degrees of freedom. Reflection of a vector with respect to a vector is very similar to that of $\mathbb{R}_2$, the only difference is that the vectors are now in 3D space. In addition to it, bivectors can also be reflected with respect to a vector (see Fig. 2.8). Let B be a bivector which is formed by the outer product of two vectors $\mathbf{x}$ and $\mathbf{y}$, i.e. $B = \mathbf{x} \wedge \mathbf{y}$. The reflection of the bivector B with respect to a vector $\mathbf{n}$ is obtained as

$$\mathbf{n} B \mathbf{n} = (\mathbf{nxn}) \wedge (\mathbf{nyn}) \ . \tag{2.84}$$

This means the reflection of a bivector with respect to a vector is the outer product of the separately reflected vectors. This property is also called the outer-morphism. Analogously, a vector $\mathbf{x}$ can also be reflected with respect to a bivector B as shown in Fig. 2.9. The reflected result is given by $B\mathbf{x}B$.

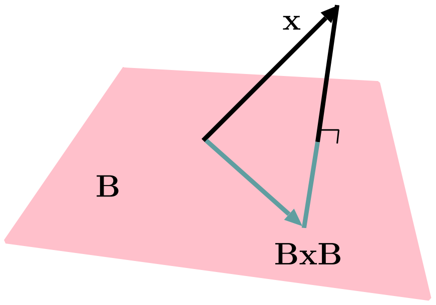

Fig. 2.9: Reflection of a vector with respect to a bivector.

Besides its interpreting algebraic properties, the geometric algebra over 3D Euclidean space is also able to represent the group of rotations and isotropic scale transformations by spinors. The spinor can be regarded as a special case of a multivector, i.e. the even graded multivector. A general spinor in $\mathbb{R}_3$ reads

$$S = a + b\mathbf{e}_{23} + c\mathbf{e}_{31} + d\mathbf{e}_{12} \ . \tag{2.85}$$

All spinors form a proper subalgebra of $\mathbb{R}_3$, i.e. $\mathbb{R}_3^+$, which is isomorphic to the algebra of quaternions, $\mathbb{R}_3^+ \simeq \mathbb{H}$ [73]. A spinor represents a scaling-rotation, i.e. $S = r\exp(\theta B)$, where B is a bivector indicating the rotation plane, θ is the rotation angle within that plane and r refers to the scaling factor. There exists the isomorphism between the algebra of complex numbers and the subalgebra of $\mathbb{R}_3$, which is generated by $\{1, I_3\}$ or $\{1, \frac{B}{|B|}\}$ with $\frac{B}{|B|}$ being a normalized bivector. The idea behind spinors is to generalize the interpretation of a unit complex number as a rotation. It is shown in Tab. 2.2 that the square of the basis bivector or trivector equals -1, therefore, the imaginary unit i of the complex numbers can be substituted by a bivector or a trivector, yielding an algebra isomorphism. A vector-valued signal $\mathbf{f}$ in $\mathbb{R}_3$ can be considered as the result of a spinor acting on the $\mathbf{e}_3$ basis, i.e. $\mathbf{f} = b\mathbf{e}_1 + c\mathbf{e}_2 + d\mathbf{e}_3 = \mathbf{e}_3 S$. The transformation performed under the action of the spinor delivers access to both the amplitude and phase information of the vector-valued signal $\mathbf{f}$ [106]. From the logarithm of the spinor representation, two parts can be obtained. They are the scaling which corresponds to the local amplitude and the rotation which corresponds to the local phase representation. The $\mathbb{R}_3$-logarithm of a spinor $S \in \mathbb{R}_3^+$ takes the following form

$$\log(S) = \langle \log(S) \rangle_0 + \langle \log(S) \rangle_2 = \log(|S|) + \frac{\langle S \rangle_2}{|\langle S \rangle_2|} \operatorname{atan}\left(\frac{|\langle S \rangle_2|}{\langle S \rangle_0} \right) \ , \tag{2.86}$$

where atan is the inverse tangent mapping for the interval $[0, \pi)$. The scalar

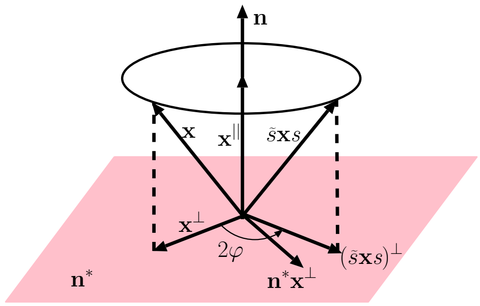

Fig. 2.10: Rotation in 3D space. The rotation axis is denoted by $\mathbf{n}$ and the rotation angle is 2φ.

part $\langle \log(S)\rangle_0 = \log(|S|)$ illustrates the logarithm of the local amplitude, hence, local amplitude is obtained as the exponential of it

$$|S| = \exp(\log|S|) = \exp(\langle \log(S)\rangle_0)\ . \tag{2.87}$$

The bivector part of $\log(S)$ indicates the local phase representation

$$\langle \log(S)\rangle_2 = \frac{\langle S\rangle_2}{|\langle S\rangle_2|}\operatorname{atan}\left(\frac{|\langle S\rangle_2|}{\langle S\rangle_0}\right)\ . \tag{2.88}$$

A spinor with a unit magnitude is called a rotor, which means $a^2+b^2+c^2+d^2=1$. In order to avoid pseudoscalar components, the action of a spinor is represented by means of a two-sided product. For example, the action of a vector $\mathbf{x}$ is denoted as $S\mathbf{x}\widetilde{S}$, where $\widetilde{S} = a - b\mathbf{e}_{23} - c\mathbf{e}_{31} - d\mathbf{e}_{12}$ is the reverse of the spinor S. In the case that the spinor is a rotor, $S\widetilde{S} = a^2+b^2+c^2+d^2 = 1$. Furthermore, a rotor S is possible to be written as $S = \cos\varphi - \mathbf{n}^*\sin\varphi$, where $\mathbf{n}$ is a unit vector and φ indicates a rotation angle. Hence, the two-sided action of the rotor on a vector $\mathbf{x}$ represents a rotation of the vector into 3D space. This 3D rotation can be visualized in Fig. 2.10. In this case, the vector is split into two parts

$$\mathbf{x} = \mathbf{x}^{||} + \mathbf{x}^{\perp} = (\mathbf{x}\cdot\mathbf{n})\mathbf{n} + (\mathbf{x}\wedge\mathbf{n})\mathbf{n}\ , \tag{2.89}$$

where $\mathbf{x}^{||}$ is the component which is parallel to the unit vector $\mathbf{n}$ and $\mathbf{x}^{\perp}$ indicates the part which is perpendicular to $\mathbf{n}$. Evaluating the action of the rotor yields

$$\begin{aligned} \widetilde{S}\mathbf{x}S &= (\cos\varphi + \mathbf{n}^*\sin\varphi)\mathbf{x}(\cos\varphi - \mathbf{n}^*\sin\varphi) \qquad (2.90)\\ &= (\cos\varphi + \mathbf{n}^*\sin\varphi)(\cos\varphi - \mathbf{n}^*\sin\varphi)\mathbf{x}^{||} + (\cos\varphi + \mathbf{n}^*\sin\varphi)^2\mathbf{x}^{\perp}\\ &= \mathbf{x}^{||}(\cos^2\varphi - \sin^2\varphi + 2\mathbf{n}^*\cos\varphi\sin\varphi)\\ &= \mathbf{x}^{||} + (\cos(2\varphi) + \mathbf{n}^*\sin(2\varphi))\mathbf{x}^{\perp}\ . \end{aligned}$$

Hence, the rotation of a vector in 3D space only rotates the part that is perpendicular to the unit vector, however, the part parallel to it remains unchanged. The double sided operation results in anti-clockwise rotation with an angle of 2φ.

Since there is the isomorphism between the complex numbers and the subalgebra of $\mathbb{R}_3$ generated by $\{1, \frac{B}{|B|}\}$, the exponential map of a bivector can be written as

$$\exp B = \cos|B| + \frac{B}{|B|}\sin|B| \ . \tag{2.91}$$

It is decomposed into the scalar and the bivector parts which establishes a generalized Euler equation. Correspondingly, the exponential map of the dual rotation vector is directly related with the rotor

$$\exp(\varphi\mathbf{n}^*) = \cos\varphi + \mathbf{n}^*\sin\varphi = \widetilde{S} \ . \tag{2.92}$$

Therefore, in terms of Lie group theory, the generation of an arbitrary 3D rotation is given by the bivector which is dual to the unit rotation vector. In contrast to 2D, an angle in 3D has the directional information. Due to the representation of a rotation by the dual of the rotation vector, angles are better considered as bivectors.

Reflections with respect to a normalized vector are always reflections on a line passing through the origin. Just like the 2D case, two consecutive reflections on different normalized vectors are equivalent to twice the angle between those two normalized vectors. Fig. 2.11 shows such a setup in 3D space. The normalized vectors $\mathbf{n}, \mathbf{m} \in \mathbb{R}^3$ enclose an angle θ and define a rotation plane through their outer product $\mathbf{n} \wedge \mathbf{m}$. Reflecting a vector $\mathbf{x} \in \mathbb{R}^3$ first on $\mathbf{n}$ and then on $\mathbf{m}$, rotates the component of $\mathbf{x}$ which lies in the rotation plane by a angle of 2θ. The component of a perpendicular to the rotation plane remains unchanged.

The rotation of the vector $\mathbf{x}$ in the plane $\mathbf{n} \wedge \mathbf{m}$ can be obtained as $\mathbf{mnxnm}$. From the definition of the geometric product, we can find that

$$\mathbf{mn} = \mathbf{m} \cdot \mathbf{n} + \mathbf{m} \wedge \mathbf{n} \tag{2.93}$$

and also

$$\mathbf{nm} = \mathbf{n} \cdot \mathbf{m} + \mathbf{n} \wedge \mathbf{m} = \mathbf{n} \cdot \mathbf{m} + \widetilde{(\mathbf{m} \wedge \mathbf{n})} \ . \tag{2.94}$$

Since the reverse of a scalar is still itself, it follows

$$\mathbf{mn} = \widetilde{\mathbf{nm}} \ . \tag{2.95}$$

Hence, consecutive reflections can be rewritten as

$$R\mathbf{x}\widetilde{R} \ , \tag{2.96}$$

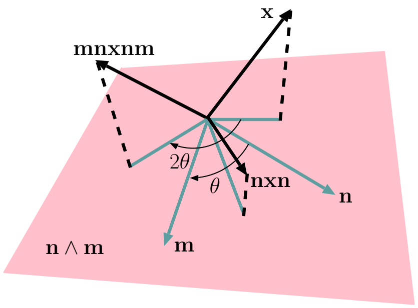

Fig. 2.11: Rotation of a vector by consecutive reflections on two different normalized vectors.

where $R = \mathbf{mn}$ refers to a rotor and satisfies $R\widetilde{R} = 1$. Expanding the rotor R, we have

$$R = \mathbf{mn} = \mathbf{m} \cdot \mathbf{n} + \mathbf{m} \wedge \mathbf{n} = \cos\theta + B \sin\theta = \exp(B\theta) \ , \tag{2.97}$$

where B indicates the normalized version of $\mathbf{m} \wedge \mathbf{n}$, that is

$$B = \frac{\mathbf{m} \wedge \mathbf{n}}{|\mathbf{m} \wedge \mathbf{n}|} \ . \tag{2.98}$$

Hence, it turns out that R actually represents a rotation by angle 2θ in the plane spanned by $\mathbf{n}$ and $\mathbf{m}$. Just as for reflections, with the same rotor, we can rotate vectors and bivectors. It turns out that a rotor also has the property of outer-morphism. This means that given a rotor and a blade $B_k =$ which is the outer product of k independent vectors $\mathbf{x}_1, ..., \mathbf{x}_k \in \mathbb{R}^3$

$$B_k = \mathbf{x}_1 \wedge ... \wedge \mathbf{x}_k = \langle \mathbf{x}_1 ... \mathbf{x}_k \rangle_k \ . \tag{2.99}$$

We can expand the expression $RB_k\widetilde{R}$ as

$$RB_k\widetilde{R} = (R\mathbf{x}_1\widetilde{R}) \wedge (R\mathbf{x}_2\widetilde{R}) \wedge ... \wedge (R\mathbf{x}_k\widetilde{R}) \ . \tag{2.100}$$

Hence, the rotation of the outer product of a number of vectors is identical to the outer product of a number of rotated vectors.

As a conclusion, some basic results of $\mathbb{R}_3$ are summarized as follow:

- $\mathbb{R}_3$ contains the vector space $\mathbb{R}^3$ as a subspace.
- The geometric product of two parallel vectors is equivalent to their inner product. For two perpendicular vectors, their geometric product equals their outer product.

- A spinor represents a scaling-rotation. The geometric product of a vector and a spinor results in a scaled and rotated vector.
- There exist the isomorphisms between the complex numbers and the subalgebras generated by $\{1, I_3\}$ and $\{1, \frac{B}{|B|}\}$.
- All spinors form the even subalgebra of $\mathbb{R}_3$, i.e. $\mathbb{R}_3^+$, which is isomorphic to the algebra of quaternions, i.e. $\mathbb{R}_3^+ \simeq \mathbb{H}$.
- One can switch from spinors to rotation vectors and from plane to normal vectors by using the property of duality.

2.5 Summary

In this chapter, a powerful algebraic framework was introduced. Brief introduction to the geometric algebra over Euclidean 2D and 3D spaces was given as the mathematical background for the remaining part of this thesis. The 2D and 3D vector spaces are extended by the outer product of vectors, which denote directed planes. Therefore, $\mathbb{R}^3$ is a subspace of $\mathbb{R}_3$. Subspaces of the geometric algebra are not restricted to interpretations within set theory, instead, they have geometric meanings.

The algebra product of geometric algebra is the geometric product. The geometric product of two vectors is the combination of their inner product and outer product. It represents the geometric relation between them. For a vector and a spinor, their geometric product indicates the group action on this vector, which will result in a scaled and rotated vector. The three-fold outer product of vectors denotes a directed volume spanned by three vectors. The outer product of three basis vectors is a trivector or pseudoscalar, which enables us to switch between spinors and rotation vectors or between planes and normal vectors by means of duality. All spinors form even subalgebras $\mathbb{R}_2^+$ and $\mathbb{R}_3^+$, which provide representations of scaling-rotations in 2D and 3D spaces by means of spinors. The subalgebras of $\mathbb{R}_2^+$ and $\mathbb{R}_3^+$ are isomorphic to the complex numbers and quaternions, respectively.

Geometric algebra enables simple and elegant representations. Compared with the classical framework of vector algebra, the geometric algebra has a more powerful subspace structure with particular geometric meanings. Solving problems in the framework of the geometric algebra contains more degrees of freedom, which makes a tremendous extension of modeling capabilities available.

3. FUNDAMENTAL OF SIGNAL PROCESSING FOR COMPUTER VISION

Computer vision has progressed considerably in the past decades. From methods only applicable to simple images, it has developed to treat dramatically increasing complex scenes, volumes and time sequences. However, there still exist many fundamental questions of how to handle and represent vision information in an effective and meaningful way. For low-level computer vision, signal processing techniques play important roles for modeling and representing image information. Low-level image features are extracted and united as building blocks for high-level computer vision tasks. In this chapter, some basic concepts of signal processing for computer vision will be introduced. They are fundamental knowledge for understanding the following chapters in this thesis.

3.1 Local Structures and Intrinsic Dimensionality

Local image structures play an significant role in many computer vision tasks. For an 2D image, its local structures are tightly associated with the term intrinsic dimension [121, 71]. This term is important for multi-dimensional signal processing. Hence, it is necessary to distinguish the global embedding dimension n of a signal and its local intrinsic dimension d. In image processing, intrinsic dimension corresponds to the term codimension in mathematics. If $S \subset \mathbb{R}^n$ is a subspace, then the intrinsic dimension is given by

$$d(S) = \mathrm{codim}(S) = n - \dim(S) \ . \tag{3.1}$$

Intrinsic dimensionality is a local property of a multidimensional signal, which expresses the number of degrees of freedom necessary to describe local structures. In an 2D image, there exist three types of structures. The image $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^2$ can be locally classified as one of the following intrinsic

d-dimensional (idD) structures

$$d(f) \in \begin{cases} \{i0D\}, & \forall x, y \in \mathbb{R}^2 : \ f(x) = f(y) \\ \{i1D\}, & \exists\, \mathbf{n} \in \mathbb{R}^2 \backslash \{(0,0)\}, \forall x \in \mathbb{R}^2, \ \alpha \in \mathbb{R} : \ f(x) = f(x + \mathbf{n} \cdot \alpha) \\ \{i2D\}, & \text{else} \end{cases} . \tag{3.2}$$

The first class is the intrinsically zero dimensional (i0D) signal which is constant in the spatial domain. The frequency representation of the i0D signal is a delta function in the origin which indicates the DC component. The intrinsically one dimensional (i1D) signal belongs to the second class. It is constant in one orientation in the spatial domain. Its spectral domain representation is concentrated on a line which is orthogonal to the constant lines in the spatial domain. In an 2D image, line and edge structures, which are constant along one orientation, can be regarded as i1D signals. The third class is the intrinsically two dimensional (i2D) signal which varies in at least two orientations. The i2D structures are composed of curved edges and lines, junctions, corners and line ends, etc. Fig. 3.1 illustrates these three basic classes of intrinsic dimensionality. There exists a close relationship between the concept of intrinsic dimensionality and the statistical properties of natural images [28, 123, 122]. Local structures show a clear order in their probability of occurrence. The i0D signals, which correspond to homogeneous regions in an image, have the highest probability. The i1D signals, representing edge and line like structures, appear less than i0D signals. The most rare cases in nature images are i2D signals, which have higher cardinality than the other two types of structures. This implies that the amount of information carried by a local image structure increases systematically with its intrinsic dimensionality. The relevance of the concept of intrinsic dimensionality with respect to the processing of images is also supported by a variety of results from neurophysiological and psychophysical research on biological vision systems [39, 110, 109, 120].

3.2 Analytic Signal in 1D and 2D

Signal processing is a fast-growing segment of modern science and technology, which integrals various techniques into a framework. Analytic signal is an important concept in signal processing with wide applications. Since the topic of this thesis focuses on modeling 2D image structures in terms of the generalization of analytic signal in a rotation-invariant manner. Thus, a brief introduction to existing definitions of analytic signals in 1D and 2D is given in this section.

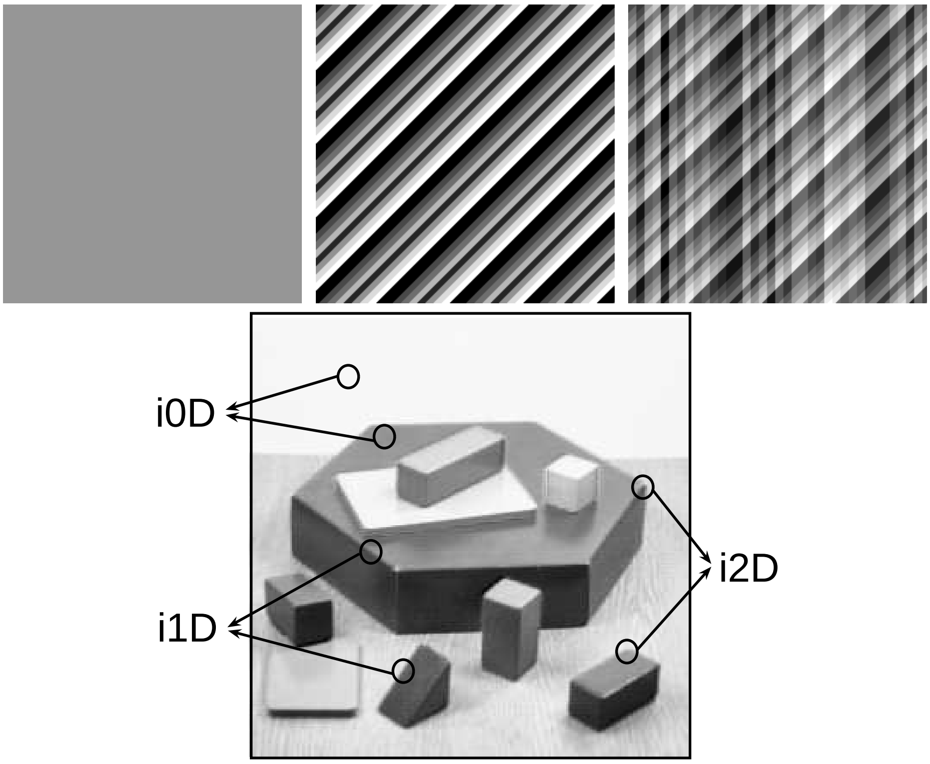

Fig. 3.1: Top row: from left to right are i0D, i1D and i2D signals. Bottom row: illustration of the three basic classes of intrinsic dimensionality as occurring in a natural image.

3.2.1 The 1D Analytic Signal

In one dimensional (1D) signal processing, the analytical signal [45] is an important complex valued model which can be used for narrow-band communication, speech recognition, seismic data analysis, airfoil design and so on, see [50, 29, 82, 83].

The analytic signal $f_A : \mathbb{R} \rightarrow \mathbb{C}$ of a real 1D signal f is defined as the sum of f and a version of f which is shifted in phase by $(-\frac{\pi}{2})$ multiplied by i. The shifted version of f is the Hilbert transform f_H of f. Thus, 1D analytic signal can be formulated in the spatial domain as

$$f_A(x) = f(x) + i f_H(x) = f(x) + i(f(x) * h(x)) \; , \tag{3.3}$$

where $h(x) = \frac{1}{\pi x}$ refers to the Hilbert transform in the spatial domain. In the frequency domain, the Hilbert transform is given by

$$H(u) = -\imath\text{sign}(u) \quad \text{with} \quad \text{sign}(u) = \begin{cases} 1 & \text{if } u > 0 \\ 0 & \text{if } u = 0 \\ -1 & \text{if } u < 0 \end{cases} \tag{3.4}$$

and the analytic signal takes the following form

$$F_A(u) = F(u) + iF_H(u) = (1 + \text{sign}(u))F(u) \ , \tag{3.5}$$

where $F(u)$ is the Fourier transform of the real signal f. The components f and f_H of the complex valued signal f_A are phase shifted by $-\frac{\pi}{2}$ and they are in quadrature phase relation.

The local amplitude and local phase of the 1D analytic signal are respectively defined as

$$a(x) = \sqrt{f^2(x) + f_H^2(x)} \tag{3.6}$$

$$\phi(x) = \arg f_A(x) \ , \tag{3.7}$$

where arg means the argument of the expression. The local energy of f_A is the square of its local amplitude. Only if the local energy exceeds a certain threshold of significance, the local phase feature then makes sense for analysis. The polar representation of the analytical signal yields the local amplitude and local phase, which are measures of quantitative and qualitative information of a signal, respectively. The local amplitude is invariant with respect to local structures and it indicates the energetic information of the original signal. The local phase is invariant with respect to the local amplitude. If the local structure varies, the local phase will correspondingly change. Local amplitude and local phase are independent of each other and they fulfill the properties of invariance and equivariance [47]. Hence, the analytic signal performs a split of identity.

Local phase of the analytic signal is tightly related with the concept of parity symmetry. Parity refers to the invariance of a process with respect to a reflection operation. A real function is of even symmetry (parity invariance) if $f(-\mathbf{x}) = f(\mathbf{x})$ and of odd symmetry (parity variance) if $f(-\mathbf{x}) = -f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$. Parity symmetry is understood as the structural feature of a signal. Hence, the parity symmetry of local phase enables a local structure analysis. The following ideal cases occur: peak: $\phi(x) = 0$, pit: $\phi(x) = \pi$, decreasing slope: $\phi(x) = \frac{\pi}{2}$, increasing slope: $\phi(x) = -\frac{\pi}{2}$. Both peak and pit indicate even symmetry, while slopes are of odd symmetry. Hence, the mapping $\mathcal{A}$: $f \rightarrow f_A$ indicates a rotation in the complex domain from the real axis to any location on the unit circle.

Because the Hilbert transform is an allpass filter, in practice, quadrature filter pairs [47] are preferred. Being bandpass filters, they have to guarantee the quadrature phase relation only within a passband. This is achieved by coupling two filters of even and odd symmetry according to

$$h_q(x) = h_e(x) + ih_o(x) \; . \tag{3.8}$$

The Gabor filter [45] is a widely used example. Gabor filters are used as standard quadrature filters in image analysis [47]. Due to their lack of rotation-invariance, they are applied as oriented filters. Convolution of f with each filter of the quadrature pair $\{h_e, h_o\}$ results in

$$g(x) = g_e(x) + ig_o(x) \; . \tag{3.9}$$

In that case, the real axis on the complex unit circle indicates pure even symmetry and the imaginary axis indicates pure odd symmetry.

3.2.2 Partial and Total Hilbert Transform

It is shown that the 1D analytic signal can be used to separate the local amplitude and local phase of a given real 1D signal. Such a separation would be of much use for multidimensional signals as well for feature extraction, classification and so on. There exist some approaches towards the generalization of the analytic signal to 2D.

The first one is the partial analytic signal [48]. It is defined as the combination of the original signal and its partial Hilbert transform. Let f be a real 2D signal, its partial Hilbert transform with respect to the orientation vector $\mathbf{n}$ in the spatial domain is given by

$$f_p(\mathbf{x}) = f(\mathbf{x}) * \left(\frac{1}{\pi(\mathbf{x}\cdot\mathbf{n})}\right) \; . \tag{3.10}$$

In the spectral domain, the partial Hilbert transform reads

$$F_p(\mathbf{u}) = -iF(\mathbf{u})\mathrm{sign}(\mathbf{u}\cdot\mathbf{n}) \; . \tag{3.11}$$

Hence, the partial analytic signal takes the following form

$$f_A^p(\mathbf{x}) = f(\mathbf{x}) * \left(\delta(\mathbf{x}) + \frac{1}{\pi(\mathbf{x}\cdot\mathbf{n})}\right) \; . \tag{3.12}$$

The symmetry of the partial Hilbert transform is illustrated in Fig. 3.2.

It is obvious that the partial analytic signal is only a generalization to i1D signal with a pre-defined orientation. Unfortunately, it cannot deliver information about i2D parts of a 2D signal.

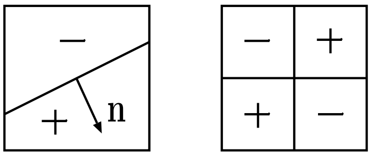

Fig. 3.2: From left to right: symmetry of the partial Hilbert transform with respect to an orientation vector **n** and the symmetry of the total Hilbert transform.

The second generalization of analytic signal to 2D is the total analytic signal [48], which is obtained by the successive Hilbert transforms with respect to both coordinate axes. The total Hilbert transform of a 2D signal in the spatial and frequency domains are respectively given by

$$f_t(\mathbf{x}) = f(\mathbf{x}) * \left(\frac{1}{\pi^2 xy}\right) \tag{3.13}$$

$$F_t(\mathbf{u}) = -F(\mathbf{u})\mathrm{sign}(u)\mathrm{sign}(u)\ . \tag{3.14}$$

Hence, the total analytic signal is defined as

$$f_A^t(\mathbf{x}) = f(\mathbf{x}) * \left(\delta(x) + \frac{i}{\pi^2 xy}\right)\ . \tag{3.15}$$

The resulting symmetry of the total Hilbert transform is even with respect to point symmetry and odd with respect to line-symmetry, see Fig. 3.2. Hence, the total analytic signal is not rotationally invariant and is not capable of detecting i1D features.

A further approach, proposed by Hahn [48] is obtained by combing the previous methods, that is

$$\begin{aligned} f_A^h(\mathbf{x}) &= f(\mathbf{x}) * \left[\left(\delta(\mathbf{x}) + \frac{i}{\pi x}\right)\left(\delta(\mathbf{x}) + \frac{i}{\pi y}\right)\right] \\ &= f(\mathbf{x}) - f_A^t(\mathbf{x}) + i(f_A^{p_1}(\mathbf{x}) + f_A^{p_2}(\mathbf{x}))\ , \end{aligned} \tag{3.16}$$

where $f_A^{p_1}(\mathbf{x})$ and $f_A^{p_2}(\mathbf{x})$ are partial Hilbert transforms with respect to x and y coordinates. This analytic signal is only non-zero in the first quadrant of the frequency domain which yields a loss of information. Hence, it is not a satisfactory generalization to 2D signal.

3.2.3 Quaternionic Analytic Signal

Another extension of the analytic signal to 2D case is the quaternionic analytic signal [20, 18], which has more powerful algebraic interpretations when compared with previous approaches. Derivation of the quaternionic analytic signal is based on the quaternionic Fourier transform (QFT), hence, a brief view of QFT is first given.

It is well known that any real function $f : \mathbb{R}^n \to \mathbb{R}$ at any location $\mathbf{x} \in \mathbb{R}^n$ may be decomposed into an even and an odd part [47] by reflection at the origin

$$f(\mathbf{x}) = f_e(\mathbf{x}) + f_o(\mathbf{x})\ . \tag{3.17}$$

Only in the case of a 1D function, i.e. $\mathbf{x} \in \mathbb{R}$, the Fourier transform preserves this symmetry decomposition in an integral manner due to the parity symmetry properties of its basis functions

$$\exp(-i2\pi ux) = \cos(2\pi ux) - i\sin(2\pi ux)\ . \tag{3.18}$$

Hence,

$$F(u) = F_e(u) + F_o(u) \tag{3.19}$$

for all $u \in \mathbb{R}$ with $F_e = \mathcal{F}\{f_e\}$ and $F_o = \mathcal{F}\{f_o\}$, $F_R = F_e$ is the real spectrum and $F_I = -iF_o$ denotes the imaginary spectrum. The amplitude spectrum $A(u) = |F(u)|$ is of even symmetry and the phase spectrum $\Phi(u) = \arg F(u)$ has the odd symmetry. For 2D case, basis functions of the Fourier transform are

$$\begin{aligned}\exp(-i2\pi\mathbf{u}\cdot\mathbf{x}) &= \cos(2\pi ux)\cos(2\pi vy) - \sin(2\pi ux)\sin(2\pi vy) \\ &+ i(\cos(2\pi ux)\sin(2\pi vy) + \sin(2\pi ux)\cos(2\pi vy))\ .\end{aligned} \tag{3.20}$$

Hence, for $\mathbf{x}, \mathbf{u} \in \mathbb{R}^2$, the following relation holds

$$F(\mathbf{u}) = F_{ee}(\mathbf{u}) + F_{oo}(\mathbf{u}) + F_{eo}(\mathbf{u}) + F_{oe}(\mathbf{u}) \tag{3.21}$$

and

$$f(\mathbf{x}) = f_{ee}(\mathbf{x}) + f_{oo}(\mathbf{x}) + f_{eo}(\mathbf{x}) + f_{oe}(\mathbf{x})\ , \tag{3.22}$$

where ee means even symmetry with respect to both x and y, and so forth. This relation corresponds to considering products of symmetries with respect to the coordinate axes. The limited degree of freedom in the algebraic structure of complex numbers results in a partial cover of symmetry with

$$F_R(\mathbf{u}) = F_{ee}(\mathbf{u}) + F_{oo}(\mathbf{u}) \tag{3.23}$$

$$F_I(\mathbf{u}) = -i(F_{eo}(\mathbf{u}) + F_{oe}(\mathbf{u}))\ . \tag{3.24}$$

The algebraic nature of the 2D Fourier transform in the complex domain causes a limited representation of symmetries. Hence, there exists no access to the single symmetries.

To overcome this problem, Bülow and Sommer [20] proposed the quaternionic Fourier transform, which is algebraically extended by employing the quaternion algebra. The QFT of a 2D function is given by

$$F^q(\mathbf{u}) = \int\int \exp(-i2\pi ux) f(\mathbf{x}) \exp(-j2\pi vy) dxdy \; . \tag{3.25}$$

The imaginary unit i in the second exponential of the original Fourier transform is now replaced by j. The units i and j are elements of the quaternion algebra and they obey the following relations

$$i^2 = j^2 = -1 \quad \text{and} \quad ij = -ji = k \; . \tag{3.26}$$

The algebraic decomposition of F^q reads

$$F^q(\mathbf{u}) = F^q_R(\mathbf{u}) + iF^q_I(\mathbf{u}) + jF^q_J(\mathbf{u}) + kF^q_K(\mathbf{u}) \; . \tag{3.27}$$

For real signal $f(\mathbf{x})$, this corresponds to the symmetry decomposition

$$f(\mathbf{x}) = f_{ee}(\mathbf{x}) + f_{oo}(\mathbf{x}) + f_{eo}(\mathbf{x}) + f_{oe}(\mathbf{x}) \; . \tag{3.28}$$

The extended polar representation of QFT gives access to a novel phase concept

$$F^q(\mathbf{u}) = |F^q(\mathbf{u})| \exp(i\phi(\mathbf{u})) \exp(k\psi(\mathbf{u})) \exp(j\theta(\mathbf{u})) \; . \tag{3.29}$$

Here the triple

$$(\phi, \theta, \psi) \in [-\pi, \pi[\times [-\frac{\pi}{2}, \frac{\pi}{2}[\times \left[-\frac{\pi}{4}, \frac{\pi}{4}\right] \tag{3.30}$$

of the quaternionic phase represents i1D phases in axes directions (ϕ, θ) and the i2D phase (ϕ), respectively. Basis functions of QFT are i2D, thus, they are able to represent explicitly i2D signals. However, QFT has the drawback of being not rotationally invariant.

The quaternionic analytic signal takes advantage of the additional degrees of freedom in the quaternionic domain. In the spectral domain, the quaternionic analytic signal reads

$$F^q_A(\mathbf{u}) = (1 + \mathrm{sign}(u))(1 + \mathrm{sign}(v)) F^q(\mathbf{u}) \; . \tag{3.31}$$

Its spatial domain representation is given by

$$f^q_A(\mathbf{x}) = f(\mathbf{x}) + \mathbf{n}^T \mathbf{f}^q_H(\mathbf{x}) \; , \tag{3.32}$$

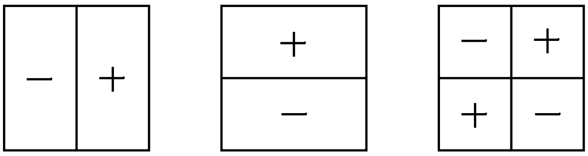

Fig. 3.3: Symmetries of the quaternionic analytic signal.

where $\mathbf{n} = (i, j, k)^T$ is the vector of quaternionic imaginary units and $\mathbf{f}_H^q$ is the vector of the Hilbert transformed signal

$$\mathbf{f}_H^q = (\mathbf{f}_{H1}, \mathbf{f}_{H2}, \mathbf{f}_{H3})^T = f(\mathbf{x}) * \left(\frac{\delta(y)}{\pi x}, \frac{\delta(x)}{\pi x}, \frac{1}{\pi^2 xy} \right)^T . \tag{3.33}$$

Local amplitude and local phase of the quaternionic analytic signal take the following forms, respectively

$$|\mathbf{f}_H^q| = \sqrt{f^2(\mathbf{x}) + \mathbf{f}_{H1}^2 + \mathbf{f}_{H2}^2 + \mathbf{f}_{H3}^2} \tag{3.34}$$

$$\Phi = \arg(f(\mathbf{x}) + \mathbf{n}^T \mathbf{f}_H^q) . \tag{3.35}$$

The resulting quaternionic analytic signal consists of four parts. Two parts correspond to the partial Hilbert transforms with respect to the coordinate axes and one corresponds to the total Hilbert transform. The corresponding symmetries of the quaternionic analytic signal is shown in Fig. 3.3. The quaternionic analytic signal is able to estimate three phases which correspond to two i1D phases and one i2D phase, just as in the case of QFT. Local phase of the quaternionic analytic signal will lose its direct relation to the local structure when the signal is not of narrow bandwidth. Hence, quaternionic Gabor filters [18] are used in real applications just like in the case of 1D analytic signal. The impulse response of a quaternionic Gabor filter is a Gaussian-windowed basis functions of the QFT

$$\begin{aligned} h^q(\mathbf{x}; \mathbf{u}_0, \sigma, \varepsilon) &= g(x; \sigma, \varepsilon)\mathrm{exp}(i2\pi u_0 x)\mathrm{exp}(j2\pi v_0 y) \\ &= (h_{ee}^q + ih_{oe}^q + jh_{eo}^q + kh_{oo}^q)(\mathbf{x}; \mathbf{u}_0, \sigma, \varepsilon) \end{aligned} \tag{3.36}$$

with

$$g(x; \sigma, \varepsilon) = N\mathrm{exp}\left(-\frac{x^2 + (\varepsilon y)^2}{\sigma^2} \right) , \tag{3.37}$$

where N is the normalization constant, $\mathbf{u}_0$ means the center frequency, σ indicates the variance of Gaussian and ε denotes the aspect ratio. The QFT of a quaternionic Gabor filter reads

$$H^q(\mathbf{u}; \mathbf{u}_0, \sigma, \varepsilon) = \exp(-2\pi^2\sigma^2[u - u_0]^2 + (v - v_0)^2/\varepsilon^2) . \tag{3.38}$$

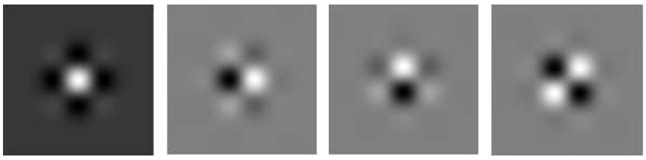

Fig. 3.4: The impulse response of a quaternionic Gabor filter.

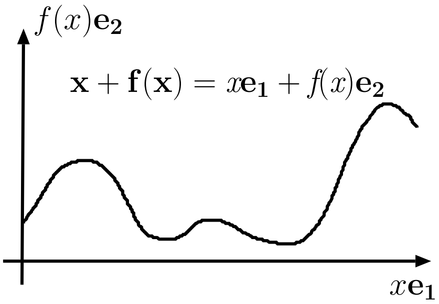

Fig. 3.5: A real scalar valued 1D signal $f(x)$ is embedded into 2D space as a curve $\mathbf{x}+\mathbf{f}(\mathbf{x}) = x\mathbf{e}_1 + f(x)\mathbf{e}_2$.

A quaternionic Gabor filter consists of four components with four different symmetries i.e. $h^q = h^q_{ee} + h^q_{oe} + h^q_{eo} + h^q_{oo}$. The impulse response of a quaternionic Gabor filter is illustrated in Fig. 3.4. Although the quaternionic analytic signal enables the evaluation of i2D signal phase information. Due to its symmetry decomposition, it still lacks the property of rotation invariance.

3.3 Fourier Transform and 2D Spherical Harmonics

Introducing geometric algebra to signal processing will result in some tiny changes with respect to algebraic aspects. Hence, it is necessary to investigate some fundamental theorems within the embedded algebraic framework.

3.3.1 Fourier Transform

For 1D case, a real scalar valued signal $f(x)$ is embedded into 2D space as a curve in the $\mathbf{e}_1 \wedge \mathbf{e}_2$ plane, see Fig. 3.5. According to this embedding, the scalar-valued signal becomes the $\mathbf{e}_2$-valued signal, that is

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1) = f(x)\mathbf{e}_2 \; . \tag{3.39}$$

The 1D Fourier transform of a multivector valued signal $g(\mathbf{x})$ [31] is given by

$$G(\mathbf{u}) = \mathcal{F}_1\{g(\mathbf{x})\}(\mathbf{u}) = \int_{x=-\infty}^{\infty} \exp(-\mathbf{e}_{12} 2\pi \mathbf{x}\mathbf{u}) g(\mathbf{x}) dx \; , \tag{3.40}$$

where $\mathbf{x} = x\mathbf{e}_1$, $\mathbf{u} = u\mathbf{e}_1$ and their geometric product equals their inner product, i.e. $\mathbf{xu} = \mathbf{x} \cdot \mathbf{u}$. The commonly used imaginary unit i in the classical Fourier transform is now replaced with $\mathbf{e}_{12}$ according to the isomorphism $\mathbb{R}_2^+ \simeq \mathbb{C}$. The inverse 1D Fourier transform can be obtained as the following

$$g(\mathbf{x}) = \mathcal{F}_1^{-1}\{G(\mathbf{u})\}(\mathbf{x}) = \int_{u=-\infty}^{\infty} \exp(\mathbf{e}_{12} 2\pi \mathbf{xu}) G(\mathbf{u}) du \ . \tag{3.41}$$

According to the Hermite symmetry of the complex Fourier transform [11], 1D Fourier transform can be split into an even and an odd part

$$G(\mathbf{u}) = G_e(\mathbf{u}) + \mathbf{e}_{12} G_o(\mathbf{u}) \ , \tag{3.42}$$

where $G_e(\mathbf{u})$ and $G_o(\mathbf{u})$ take the following forms, respectively

$$G_e(\mathbf{u}) = \int_{x=-\infty}^{\infty} g(\mathbf{x}) \cos(2\pi \mathbf{xu}) dx \tag{3.43}$$

$$G_o(\mathbf{u}) = \int_{x=-\infty}^{\infty} g(\mathbf{x}) \sin(2\pi \mathbf{xu}) dx \ . \tag{3.44}$$

Hence, for a signal $\mathbf{f}(\mathbf{x}) = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2$, its Fourier transform $\mathbf{F}(\mathbf{u}) = F_1\{\mathbf{f}(\mathbf{x})\}$ can be decomposed as

$$\mathbf{F}(\mathbf{u}) = F_1(\mathbf{u})\mathbf{e}_1 + F_2(\mathbf{u})\mathbf{e}_2 = (F_{e1}(\mathbf{u}) + F_{o2}(\mathbf{u}))\mathbf{e}_1 + (-F_{o1}(\mathbf{u}) + F_{e2}(\mathbf{u}))\mathbf{e}_2 \ , \tag{3.45}$$

where $F_1 = F_{e1} + F_{o1}\mathbf{e}_{12}$ and $F_2 = F_{e2} + F_{o2}\mathbf{e}_{12}$ are the Fourier transform of f_1 and f_2, respectively.

The derivative operator, or the nabla operator, is one of the most often used linear, shift invariant (LSI) operators in signal processing techniques. According to the habits of geometric algebra and Clifford analysis, the first derivative operator in 1D case is a vector valued operator. For a given multivector valued function $g(\mathbf{x})$, its first order derivative is given by

$$\nabla_1 g(\mathbf{x}) = \mathbf{e}_1 \frac{\partial}{\partial x} g(\mathbf{x}) \ . \tag{3.46}$$

Its Fourier transform reads

$$F_1\{\nabla_1 g(\mathbf{x})\} = -2\pi \mathbf{u}\mathbf{e}_{12} G(-\mathbf{u}) \ . \tag{3.47}$$

The second order derivative of a multivector valued function $g(\mathbf{x})$ takes the following form

$$\nabla_1^2 g(\mathbf{x}) = \Delta_1 g(\mathbf{x}) = \frac{\partial}{\partial x^2} g(\mathbf{x}) \ , \tag{3.48}$$

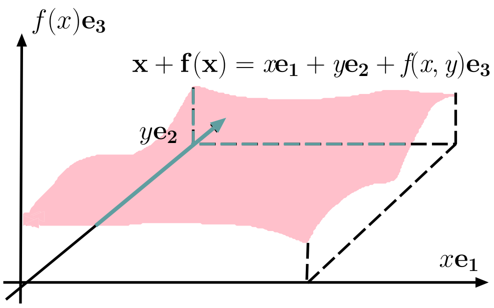

Fig. 3.6: A real scalar valued 2D signal $f(x, y)$ is embedded into 2D space as a surface $\mathbf{x} + \mathbf{f}(\mathbf{x}) = x\mathbf{e}_1 + y\mathbf{e}_2 + f(x, y)\mathbf{e}_3$.

where $\nabla_1^2 = \Delta_1$ is called the 1D Laplace operator. It is obvious that this operator is a scalar valued operator. The Fourier transform of the second derivative can be obtained as

$$\mathcal{F}_1\{\nabla_1^2 g(\mathbf{x})\} = \mathcal{F}_1\{\Delta_1 g(\mathbf{x})\} = -4\pi^2\mathbf{u}^2 G(\mathbf{u}) \ . \qquad (3.49)$$

Hence, the Fourier transform of the Laplace operator reads

$$\mathcal{F}_1\{\Delta_1\} = -4\pi^2\mathbf{u}^2 \ . \qquad (3.50)$$

For real 2D signals, they are considered as surfaces in 3D space as shown in Fig. 3.6. Hence, a real scalar valued signal $f(x, y)$ is embedded as $\mathbf{e}_3$-valued signal according to

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2) = f(x, y)\mathbf{e}_3 \ . \qquad (3.51)$$

The 2D Fourier transform of a multivector valued function $g(\mathbf{x})$ is defined as

$$G(\mathbf{u}) = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g(\mathbf{x})\exp(-I_3 2\pi\mathbf{x}\cdot\mathbf{u})dxdy \ . \qquad (3.52)$$

In this case, the pseudoscalar I_3 substitutes the imaginary unit i due to the isomorphism between the subalgebra $\{1, I_3\}$ and the complex numbers. Since I_3 commutes with every element of $\mathbb{R}_3$, the transform kernel can be placed on the left or on the right, whereas it has to be on the left for the 1D case. The 2D inverse Fourier transform reads

$$g(\mathbf{x}) = \int_{v=-\infty}^{\infty}\int_{u=-\infty}^{\infty} G(\mathbf{u})\exp(I_3 2\pi\mathbf{x}\cdot\mathbf{u})dudv \ . \qquad (3.53)$$

Similar as in the case of 1D, 2D Fourier transform can also be decomposed to even and odd parts

$$G(\mathbf{u}) = G_e(\mathbf{u}) + G_o(\mathbf{u})I_3 \ , \qquad (3.54)$$

where $G_e(\mathbf{u})$ and $G_o(\mathbf{u})$ are defined as the following forms, respectively

$$\begin{aligned} G_e(\mathbf{u}) &= \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g(\mathbf{x})\cos(2\pi\mathbf{x}\cdot\mathbf{u}) && (3.55)\\ G_o(\mathbf{u}) &= \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g(\mathbf{x})\sin(2\pi\mathbf{x}\cdot\mathbf{u}) \ . && (3.56)\end{aligned}$$

Hence, the spectral domain representation of a signal $\mathbf{f}(\mathbf{x}) = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2 + f_3(\mathbf{x})\mathbf{e}_3$ is given by

$$\begin{aligned} \mathbf{F}(\mathbf{u}) &= \mathcal{F}_2\{\mathbf{f}(\mathbf{x})\}(\mathbf{u}) = \mathbf{e}_1F_1(\mathbf{u}) + \mathbf{e}_2F_2(\mathbf{u}) + \mathbf{e}_3F_3(\mathbf{u}) && (3.57)\\ &= \mathbf{e}_1F_{e1}(\mathbf{u}) - \mathbf{e}_1^*F_{o1}(\mathbf{u}) + \mathbf{e}_2F_{e2}(\mathbf{u}) - \mathbf{e}_2^*F_{o2}(\mathbf{u}) + \mathbf{e}_3F_{e3}(\mathbf{u}) - \mathbf{e}_3^*F_{o3}(\mathbf{u}) \ , \end{aligned}$$

where $F_k = F_{ek} + F_{ok}I_3$, $k = 1, 2, 3$ is the Fourier transform of f_k.

In 2D case, the nabla operator still takes vector value. The first order derivative of a multivector valued function $g(\mathbf{x})$ is obtained by

$$\nabla_2 g(\mathbf{x}) = \mathbf{e}_1\frac{\partial}{\partial x}g(\mathbf{x}) + \mathbf{e}_2\frac{\partial}{\partial y}g(\mathbf{x}) \qquad (3.58)$$

and its frequency domain representation reads

$$\mathcal{F}_2\{\nabla_2 g(\mathbf{x})\}(\mathbf{u}) = 2\pi\mathbf{u}I_3G(\mathbf{u}) \ . \qquad (3.59)$$

The second order derivative of $g(\mathbf{x})$ in the spatial and Fourier domain are given as following, respectively

$$\nabla_2^2 g(\mathbf{x}) = \frac{\partial^2}{\partial x^2}g(\mathbf{x}) + \frac{\partial^2}{\partial y^2}g(\mathbf{x}) \qquad (3.60)$$

$$\mathcal{F}_2\{\nabla_2^2 g(\mathbf{x})\}(\mathbf{u}) = \mathcal{F}_2\{\Delta_2 g(\mathbf{x})\}(\mathbf{u}) = -4\pi^2\mathbf{u}^2G(\mathbf{u}) \ . \qquad (3.61)$$

As in the 1D case, two-fold of the first derivative operator results in the Laplace operator Δ_2.

3.3.2 2D Spherical Harmonics and Plancherel Theorem

For an arbitrary 2D spatial vector $\mathbf{x}$, it can be regarded as an oriented vector with an angle θ, i.e. $\mathbf{x} = |\mathbf{x}|(\cos\theta\mathbf{e}_1 + \sin\theta\mathbf{e}_2)$. According to the introduction in Section 2.2, this angle can be obtained as the bivector part of the logarithm of a spinor field $\mathbf{e}_1\mathbf{x}$

$$\theta\mathbf{e}_{12} = \langle\log(\mathbf{e}_1\mathbf{x})\rangle_2 \ . \qquad (3.62)$$

In this case, angles are considered as bivectors. The imaginary unit is directly given as $\mathbf{e}_{12}$. Consequently, the definition of the Fourier series of an angular, $\mathbb{R}_2^+$-valued function h can be represented as

$$h(\theta) = \sum_{n\in\mathbb{Z}} H_n \exp(n\theta\mathbf{e}_{12}) \ , \tag{3.63}$$

where H_n indicate the coefficients. If this Fourier series is truncated by $|n| < n_{max}$, it is an optimal approximation of $h(\theta)$ with respect to the L^2-norm [69]. The basis functions $\exp(n\theta\mathbf{e}_{12}) = \cos(n\theta) + \sin(n\theta)\mathbf{e}_{12}$ are called 2D spherical harmonics or circular harmonics of order n. Actually, spherical harmonic is a general term which denotes the harmonic oscillations on the unit sphere of a multi-dimensional Euclidean space. In this thesis, only the spherical harmonics of 2D space are used as basis functions for signal modeling.

In the classical case, the Plancherel theorem [108] is formulated for the complex function space. Considering the present geometric embedding, the scalar product of two multivectors $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$ is now defined as

$$\langle g_1, g_2\rangle = \int_{y=-\infty}^{\infty}\int_{x=-\infty}^{\infty} g_1(\mathbf{x})\widetilde{g}_2(\mathbf{x})dxdy \ . \tag{3.64}$$

In terms of scalar product, the Plancherel theorem in $\mathbb{R}_3$ reads [31]

$$\langle g_1, g_2\rangle = \langle G_1, G_2\rangle \ , \tag{3.65}$$

where G_1 and G_2 are the Fourier transforms of g_1 and g_2, respectively. From this definition, it can be naturally concluded that each component of the scalar product is equivalent

$$\langle\langle g_1, g_2\rangle\rangle_k = \langle\langle G_1, G_2\rangle\rangle_k \ , \quad k \in \{0,1,2,3\} \ . \tag{3.66}$$

Hence, for the special case with $k = 0$ and $g = g_1 = g_2$, the Parseval theorem is obtained

$$E_g = \langle\langle g_1, g_2\rangle\rangle_0 = \langle\langle G_1, G_2\rangle\rangle_0 = E_G \ , \tag{3.67}$$

where E_g and E_G indicate the energies of g and G, respectively.

3.4 Tensor Algebra

The main contribution of this thesis is a general signal model for 2D image structures, which consists of a curvature tensor and its conjugate part. As the fundamental knowledge, a brief overview of tensor algebra is given in this section.

Tensor analysis [15, 61, 51] is a generalization of the notions from vector analysis. The need for such a theory is motivated by the fact that there exist complicated quantities that cannot naturally be described or represented by scalars or vectors. These quantities can be represented adequately only by the more sophisticated mathematical entities called tensors. Scalars and vectors are special cases of tensors. Associated with a tensor is its order, which can be regarded as the complexity of the entity it describes. A zeroth order tensor is a scalar and a first order tensor indicates a vector. A second order tensor is a matrix which is particularly suited for the simultaneous analysis of i1D and i2D image structures.

3.4.1 Definition

An nth-rank tensor in m-dimensional space is a mathematical object that has n indices and m^n components and obeys certain transformation rules. Each index of a tensor ranges over the number of dimensions of space. However, the dimension of the space is largely irrelevant in most tensor equations (with the notable exception of the contracted Kronecker delta). Tensors are generalizations of scalars (that have no indices), vectors (that have exactly one index), and matrices (that have exactly two indices) to an arbitrary number of indices.

The notation for a tensor is similar to that of a matrix (i.e. $A = (a_{ij})$), except that a tensor $a_{ijk...}$, $a^{ijk...}$, $a_i^{jk}...$, etc., may have an arbitrary number of indices. In addition, a tensor with rank $r + s$ may be of mixed type (r, s), consisting of r so-called "contravariant" (upper) indices and s "covariant" (lower) indices. Note that the positions of the slots in which contravariant and covariant indices are placed are significant so, for example, $a_{\mu\nu}^{\lambda}$ is distinct from $a_{\mu}^{\nu\lambda}$. While the distinction between covariant and contravariant indices must be made for general tensors, the two are equivalent for tensors in three-dimensional Euclidean space, and such tensors are known as Cartesian tensors. Tensors may be operated on by other tensors (such as metric tensors, the permutation tensor, or the Kronecker delta) or by tensor operators (such as the covariant or semicolon derivatives).

In the computer vision community, there exist two well known tensor approaches, that is the structure tensor and the orientation tensor which will be briefly introduced in terms of classical matrix algebra in the following sections.

3.4.2 Structure Tensor

The structure tensor [43, 59] is one of the well known methods in multidimensional signal processing. It is a second order tensor which enables the detection of edge and corner information. The structure tensor T_s is constructed by the tensor product of a gradient vector $\nabla_2 \mathbf{f} = [\mathbf{f}_x, \mathbf{f}_y]^T$ with itself, that is

$$T_s = g * (\nabla_2 \mathbf{f})(\nabla_2 \mathbf{f})^T = g * \begin{bmatrix} \mathbf{f}_x^2 & \mathbf{f}_x\mathbf{f}_y \\ \mathbf{f}_x\mathbf{f}_y & \mathbf{f}_y^2 \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{f}_x^2} & \overline{\mathbf{f}_x\mathbf{f}_y} \\ \overline{\mathbf{f}_x\mathbf{f}_y} & \overline{\mathbf{f}_y^2} \end{bmatrix} , \tag{3.68}$$

where g represents the Gaussian filter, which is used for averaging the components, every component with a bar over it indicates the averaged result. The gradient itself encodes only step edge information, but the averaging distributes this information over a local neighborhood, and points that receive contributions from edges with different orientations are considered as corners or junctions. According to [59], the structure tensor can also be related to the tensor of inertia.

The structure tensor is a positive semi-definite symmetric tensor, it has two non-negative eigenvalues which are given by

$$\lambda_{1,2} = \frac{1}{2}\left(\overline{\mathbf{f}_x^2} + \overline{\mathbf{f}_y^2} \pm \sqrt{(\overline{\mathbf{f}_x^2} - \overline{\mathbf{f}_y^2})^2 + 4\overline{\mathbf{f}_x\mathbf{f}_y}}\right) . \tag{3.69}$$

They encode the magnitudes of the quantity of interest in the directions given by the corresponding orthogonal eigenvectors. Let θ_0 be the local main orientation, the first and second eigenvectors are $\mathbf{v}_1 = [\cos(\theta_0), \sin(\theta_0)]^T$ and $\mathbf{v}_2 = [-\sin(\theta_0), \cos(\theta_0)]^T$, respectively.

The local main orientation vector can also be obtained from a double angle representation [5, 46]

$$\mathbf{o} = |\mathbf{o}|[\cos(2\theta_0), \sin(2\theta_0)]^T = \begin{bmatrix} \overline{\mathbf{f}_x^2} - \overline{\mathbf{f}_y^2} \\ 2\overline{\mathbf{f}_x\mathbf{f}_y} \end{bmatrix} , \tag{3.70}$$

where $\mathbf{o}$ means the local main orientation vector and its amplitude equals the difference of eigenvalues, i.e. $|\mathbf{o}| = \lambda_1 - \lambda_2$. According to the eigenvalues, the coherence [59] of the local structure is obtained as

$$c = \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}\right) = \left(\frac{|\mathbf{o}|}{\mathrm{trace}(T)}\right) . \tag{3.71}$$

Assume the denominator is non-zero, the coherence is of value one if the underlying structure is i1D. It will decrease with raising intensity of i2D structures. If the energy of the i2D structure is distributed uniformly for all orientations, the coherence equals zero.

In order to detect i1D (edge) and i2D (corner, junction) structures, the structure tensor should be decomposed into two parts

$$T = T_{i1D} + T_{i2D} = (\lambda_1 - \lambda_2)\mathbf{v}_1\mathbf{v}_1^T + \lambda_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} , \tag{3.72}$$

where $\lambda_1 - \lambda_2$ can be interpreted as an edge strength and $2\lambda_2$ as a junction strength. Since the evaluations of i1D and i2D structures are tightly related with the eigenvalues, the structure tensor hence enables the rotationally invariant detection of local structures. From the structure tensor, the Harris detector [49] which is a well known approach to detect corners and junctions, can also be obtained. In this case, the corner strength measure reads

$$s = \det(T) - 0.04(\mathrm{trace}(T))^2 . \tag{3.73}$$

The eigenvalues of the structure tensor can also be considered as a $\cos^2$-decomposition of the local amplitude [31]. According to the definitions of eigenvectors, it follows that

$$\mathbf{v}_k^T T \mathbf{v}_k = \lambda_k, \quad k = 1, 2 . \tag{3.74}$$

Let T now be the tensor of a structure rotated by a small angle $\Delta\theta$, the two-sided multiplication by the first eigenvector $\mathbf{v}_1$ reads

$$\begin{aligned} &\mathbf{v}_1^T \begin{bmatrix} \cos^2(\theta_0 + \Delta\theta) & \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) \\ \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) & \sin^2(\theta_0 + \Delta\theta) \end{bmatrix} \mathbf{v}_1 \\ = &\mathbf{v}_1^T \begin{bmatrix} \cos(\theta_0 + \Delta\theta) \\ \sin(\theta_0 + \Delta\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta_0 + \Delta\theta) & \sin(\theta_0 + \Delta\theta) \end{bmatrix} \mathbf{v}_1 = \cos^2(\Delta\theta) . \end{aligned}$$

Analogously, for the two-sides multiplication by the second eigenvector $\mathbf{v}_2$, we have

$$\begin{aligned} &\mathbf{v}_2^T \begin{bmatrix} \cos^2(\theta_0 + \Delta\theta) & \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) \\ \cos(\theta_0 + \Delta\theta)\sin(\theta_0 + \Delta\theta) & \sin^2(\theta_0 + \Delta\theta) \end{bmatrix} \mathbf{v}_2 \\ = &\mathbf{v}_2^T \begin{bmatrix} \cos(\theta_0 + \Delta\theta) \\ \sin(\theta_0 + \Delta\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta_0 + \Delta\theta) & \sin(\theta_0 + \Delta\theta) \end{bmatrix} \mathbf{v}_2 = \sin^2(\Delta\theta) . \end{aligned}$$

Hence, eigenvalues of the structure tensor can be obtained from two oriented filters $\cos^2(\Delta\theta)$ and $\sin^2(\Delta\theta)$.

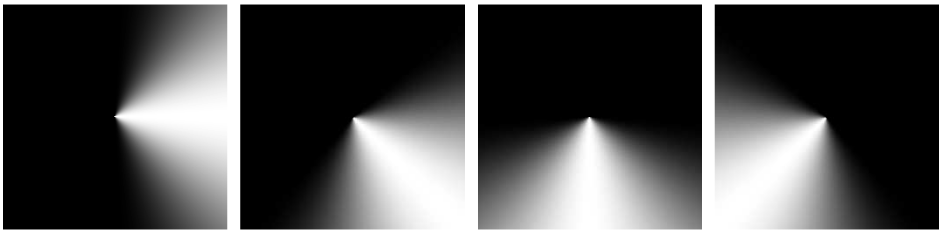

Fig. 3.7: Angular spectral responses of four oriented quadrature filters.

3.4.3 Orientation Tensor

The orientation tensor [47, 10], which encodes both step and roof edge information simultaneously, is derived by combining four oriented quadrature filter responses. These four filters, denoted as q_1, q_2, q_3, q_4, are oriented in the directions 0^0, 45^0, 90^0, 135^0, respectively. As basis functions to construct the orientation tensor, these four filters in the spectral domain are built by lognormal radial bandpass filters $B(|\mathbf{u}|)$ and the squared scalar product of frequency vector and directional vector in the positive half space of the directional vector

$$Q_k(\mathbf{u}) = \begin{cases} B(|\mathbf{u}|)(\mathbf{u}\cdot\mathbf{n}_k)^2, & \text{if}\,\mathbf{u}\cdot\mathbf{n}_k > 0 \\ 0, & \text{else} \end{cases} . \tag{3.75}$$

The resulting angular functions is a $\cos^2$-function which is centered at the angle of the directional vector, see Fig. 3.7. Thus, the orientation tensor is built as

$$T = \frac{1}{4}\begin{bmatrix} 3|q_1| + |q_2| - |q_3| + |q_4| & 2|q_2| - 2|q_4| \\ 2|q_2| - 2|q_4| & -|q_1| + |q_2| + 3|q_3| + |q_4| \end{bmatrix} . \tag{3.76}$$

The key idea is to analyze the angular behavior of the local spectrum. From the orientation tensor, local main orientation can be obtained in double angle representation as

$$\mathbf{o} = \begin{bmatrix} |q_1| - |q_3| \\ |q_2| - |q_4| \end{bmatrix} . \tag{3.77}$$

Returning to the eigenvalue decomposition of the orientation tensor, it is obvious that it has a rank of one, hence, the second eigenvalue is always zero. From this point of view, the orientation tensor can only handle the i1D structure. By employing a family of oriented quadrature filters as basis functions, the orientation tensor encodes step and roof edge information simultaneously. However, the suitability for corner or junction detection has apparently never been investigated.

3.5 Monogenic Signal and the Phase Concept

The main contribution of this thesis, i.e. the signal modeling for 2D image structures, is derived on the basis of the monogenic signal. Hence, a brief introduction to the monogenic signal and its phase concept is given.

3.5.1 Definition

The scalar-valued real 2D image signal $f(x, y)$ is embedded into $\mathbb{R}_3$ as a real-valued vector field, i.e. $f(x, y)\mathbf{e}_3 \equiv \mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3)$, $\mathbf{x} \in \mathbb{R}^3$ and $z = 0$. Thus, rotating the signal $\mathbf{f}$ out of the $\mathbf{e}_3$ axis results in a representation $\mathbf{f}_M$ which contains additional non-real components. Hence, $\mathbf{f}_M$ takes the following form

$$\mathbf{f}_M = f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2 + f_3(\mathbf{x})\mathbf{e}_3 \ , \tag{3.78}$$

where $f_3(\mathbf{x})\mathbf{e}_3$ indicates the real component which is identical to the original signal $\mathbf{f}(\mathbf{x})$, $f_1(\mathbf{x})\mathbf{e}_1$ and $f_2(\mathbf{x})\mathbf{e}_2$ are two non-real components that can be obtained from the Riesz transform of the original signal. This representation of a 2D signal is called a monogenic signal [36, 31], as the rotationally invariant generalization of the analytic signal.

Let $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2$ and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ be the Cartesian coordinates of the spatial domain and the Fourier domain, respectively. The convolution mask of the Riesz transform is given by

$$h_R(\mathbf{x}) = -\frac{x}{2\pi|\mathbf{x}|^3}\mathbf{e}_{31} + \frac{y}{2\pi|\mathbf{x}|^3}\mathbf{e}_{23} = \frac{\mathbf{x}\mathbf{e}_3}{2\pi|\mathbf{x}|^3} \tag{3.79}$$

and its frequency domain representation reads

$$H_R(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|}I_2^{-1} \ . \tag{3.80}$$

The Riesz kernel consists of two components and it is basically identical to the first order spherical harmonic. Combining the signal and its Riesz transformed result yields the monogenic signal

$$\mathbf{f}_M(\mathbf{x}) = \mathbf{f}(\mathbf{x}) + (h_R * \mathbf{f})(\mathbf{x}) \ . \tag{3.81}$$

The real part $\mathbf{f}(\mathbf{x})$ is also called the even component of the monogenic signal. And the odd component of the monogenic signal, called the figure flow, is defined as $(h_R * \mathbf{f})(\mathbf{x})$, which is in quadrature relation to the even part.

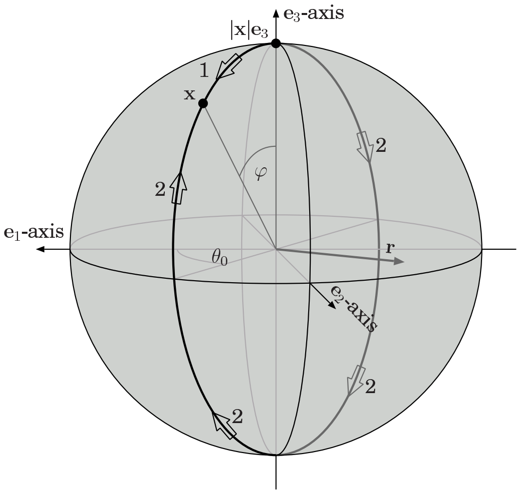

Fig. 3.8: Geometric model for the phase representation of the monogenic signal. Local phase angle is indicated by ϕ, θ_0 denotes the local orientation and $\mathbf{r}$ represents the local rotation vector [31].

3.5.2 Local Features

In the light of the discussion in Section 2.4.3, $\mathbf{f}_M(\mathbf{x})$ can be regarded as the impulse response of a spinor S acting on the $\mathbf{e}_3$ basis vector. Therefore, the logarithm of this spinor gives access to both the amplitude and phase information of the original signal. The spinor field which maps $\mathbf{e}_3$ basis vector to the vector valued signal $\mathbf{f}_M(\mathbf{x})$ is $\mathbf{e}_3\mathbf{f}_M(\mathbf{x})$. Hence, according to equation (2.87) and equation (2.88), local amplitude and local phase representation of the monogenic signal are obtained as follow

$$A_M(\mathbf{x}) = |\mathbf{f}_M(\mathbf{x})| = \exp(\log|\mathbf{e}_3\mathbf{f}_M(\mathbf{x})|) = \exp(\langle\log(\mathbf{e}_3\mathbf{f}_M(\mathbf{x}))\rangle_0) \tag{3.82}$$

$$\Phi_M(\mathbf{x}) = \langle\log(\mathbf{e}_3\mathbf{f}_M(\mathbf{x}))\rangle_2 = \frac{\langle\mathbf{e}_3\mathbf{f}_M(\mathbf{x})\rangle_2}{|\langle\mathbf{e}_3\mathbf{f}_M(\mathbf{x})\rangle_2|}\mathrm{atan}\left(\frac{|\langle\mathbf{e}_3\mathbf{f}_M(\mathbf{x})\rangle_2|}{\langle\mathbf{e}_3\mathbf{f}_M(\mathbf{x})\rangle_0}\right) . \tag{3.83}$$

According to the duality in 3D space, the dual of a bivector is a vector which is perpendicular to the plane represented by the bivector. Hence, the local rotation vector $\mathbf{r}$ of the monogenic signal can be defined as the dual of the local phase representation

$$\mathbf{r}(\mathbf{x}) = (\Phi_M(\mathbf{x}))^* = \langle\log(\mathbf{e}_3\mathbf{f}_M(\mathbf{x}))\rangle_2^* . \tag{3.84}$$

Through the above mentioned geometric embedding, a geometric model for the phase representation of the monogenic signal can be obtained as shown in Fig. 3.8.

The rotation vector always lies in the plane orthogonal to $\mathbf{e}_3$. At those positions where $\mathbf{f}_M(\mathbf{x}) = -|\mathbf{f}_M(\mathbf{x})|\mathbf{e}_3$, the local rotation vector is not well defined. Any vector lying in the plane $\mathbf{e}_1 \wedge \mathbf{e}_2$ with magnitude π is a correct solution. There exists a wrapping of the rotation vector of the 2D local phase. Once a vector in a certain direction exceeds the amplitude π, it is replaced by the vector minus 2π times the unit vector in that direction, i.e. it points in the opposite direction. The rotation vector $\mathbf{r}$ is orthogonal to the local orientation of the 2D signal and its magnitude equals the local phase angle ϕ, i.e. $\text{sign}(\mathbf{r} \cdot \mathbf{e}_1)|\mathbf{r}|$ represents the local i1D phase of the 2D signal.

Given the local amplitude and local phase representation, the monogenic signal can be reconstructed by

$$\mathbf{f}_M(\mathbf{x}) = A_M(\mathbf{x})\exp(\Phi_M(\mathbf{x})) \ . \tag{3.85}$$

Since the local amplitude and local phase representation are independent of each other, like in the case of 1D analytic signal, the monogenic signal fulfills the split of identity [47].

It is also possible to extend the monogenic signal to higher dimensions, see [106, 70]. Hence, the monogenic signal can be regarded as a generalization of the analytic signal in 2D and higher dimensions. The monogenic signal is a novel model for i1D signals, however, it delivers no information about the i2D parts of the 2D image.

3.6 Monogenic Scale-Space

Up to now, we only consider the 2D image signal $\mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3)$ with $z = 0$. If we investigate the case for the half space $z > 0$, this signal will become a smoothed version of the original signal ($z = 0$). Hence, z is regarded as a scale parameter.

3.6.1 Definition and Local Features

Applying a Poisson kernel to the original signal results in smoothed signals which form a Poisson scale-space

$$p(\mathbf{x}; s) = (\mathbf{f} * h_P)(\mathbf{x}) \ , \tag{3.86}$$

where s denotes the scale parameter, $p(\mathbf{x}; s)$ is the Poisson scale-space and h_P indicates the low-pass filter Poisson kernel, which is defined as

$$h_P(\mathbf{x}) = \frac{s}{2\pi(|\mathbf{x} + s\mathbf{e}_3|)^3} \tag{3.87}$$

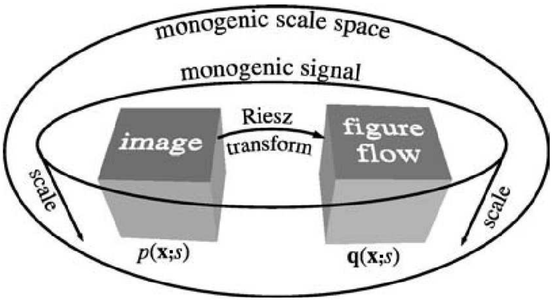

Fig. 3.9: The structure of the monogenic scale-space [38]

and

$$\mathcal{F}\{h_P\} = \exp(-2\pi|\mathbf{u}|s) \ , \tag{3.88}$$

where $\mathcal{F}$ refers to the Fourier transform. The harmonic conjugate of the Poisson scale-space reads

$$\mathbf{q}(\mathbf{x};s) = (\mathbf{f} * h_Q)(\mathbf{x}) \ , \tag{3.89}$$

where $\mathbf{q}(\mathbf{x};s)$ is the conjugate Poisson scale-space, h_Q denotes the conjugate Poisson kernel which takes the following forms

$$h_Q(\mathbf{x}) = h_{Q1}(\mathbf{x}) + h_{Q2}(\mathbf{x}) = -\frac{x\mathbf{e}_{31}}{2\pi(|\mathbf{x}+s\mathbf{e}_3|)^3} + \frac{y\mathbf{e}_{23}}{2\pi(|\mathbf{x}+s\mathbf{e}_3|)^3} = \frac{\mathbf{x}\mathbf{e}_3}{2\pi(|\mathbf{x}+s\mathbf{e}_3|)^3} \tag{3.90}$$

and

$$\mathcal{F}\{h_Q\} = \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1} \exp(-2\pi|\mathbf{u}|s) \ . \tag{3.91}$$

Combining the Poisson scale-space with its harmonic conjugate yields the monogenic scale-space [38], which can be visualized in Fig. 3.9. When the scale parameter is set to zero, the monogenic signal is obtained. From an alternative point of view, the monogenic scale-space can also be built by the monogenic signals at all scales, where the monogenic signals are formed by the smoothed image signals and their Riesz transformed results, i.e. the figure flows. In the monogenic scale-space, the figure flow and the smoothed signal are in quadrature phase relation at each scale.

In the monogenic scale-space, expressions for the local amplitude $A_M(\mathbf{x};s)$ and local phase $\Phi(\mathbf{x};s)$ have to be generalized accordingly as

$$A_M(\mathbf{x};s) = \frac{1}{2}\log\left(|p(\mathbf{x};s)|^2 + |\mathbf{q}(\mathbf{x};s)|^2\right) \tag{3.92}$$

$$\Phi_M(\mathbf{x};s) = \frac{\mathbf{q}(\mathbf{x};s)}{|\mathbf{q}(\mathbf{x};s)|} \arctan\left(\frac{|\mathbf{q}(\mathbf{x};s)|}{p(\mathbf{x};s)}\right) \ . \tag{3.93}$$

Accordingly, the local rotation vector in the monogenic scale-space is given by

$$\mathbf{r}(\mathbf{x}; s) = (\Phi_M(\mathbf{x}; s))^* \ . \tag{3.94}$$

3.6.2 DOP and DOCP Filters

As in the case of the well-known bandpass filter, difference of Gaussian (DOG), it is also possible to build up two bandpass filters. They are the difference of Poisson (DOP) and the difference of conjugate Poisson (DOCP). Combining two Poisson kernels with a fixed ratio of scale parameters yields a family of bandpasses with constant relative bandwidth, that is the DOP bandpass filter. Its expressions in the spatial and the Fourier domains are given as follows, respectively

$$b_{s_0,\lambda,k}(\mathbf{x}) = \frac{s_0\lambda^k}{2\pi(\mathbf{x}^2 + (s_0\lambda^k)^2)^{3/2}} - \frac{s_0\lambda^{k-1}}{2\pi(\mathbf{x}^2 + (s_0\lambda^{k-1})^2)^{3/2}} \tag{3.95}$$

$$B_{s_0,\lambda,k}(\mathbf{u}) = \exp(-2\pi|\mathbf{u}|s_0\lambda^k) - \exp(-2\pi|\mathbf{u}|s_0\lambda^{k-1}) \ , \tag{3.96}$$

where $\lambda \in (0, 1)$ indicates the relative bandwidth, s_0 denotes the coarsest scale, and $k \in \mathbb{N}$ is the bandpass number. When $\lambda = 0.5k$, it results in the octave sampling scheme.

The conjugate part of the DOP in the spatial domain is given by

$$c_{s_0,\lambda,k}(\mathbf{x}) = \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2 + (s_0\lambda^k)^2)^{3/2}} - \frac{\mathbf{x}\mathbf{e}_3}{2\pi(\mathbf{x}^2 + (s_0\lambda^{k-1})^2)^{3/2}} \ . \tag{3.97}$$

Its corresponding frequency representation reads

$$c_{s_0,\lambda,k}(\mathbf{u}) = \frac{\mathbf{u}}{|\mathbf{u}|} I_2^{-1} \exp(-2\pi|\mathbf{u}|s_0\lambda^k) - \exp(-2\pi|\mathbf{u}|s_0\lambda^{k-1}) \ . \tag{3.98}$$

The triple consisting of the DOP and DOCP forms a Riesz triple. According to the definition of quadrature filter, the combined filter DOP+DOCP yields also a generalized quadrature filter called the spherical quadrature filter (SQF), see Fig. 3.10.

When compared with the Gaussian kernel, see [31], the peak of the Poisson kernel is sharper than that of the Gaussian kernel and the extent is larger. The uncertainty of the Poisson kernel is only slightly worse than that of the Gaussian. Hence, the monogenic scale-space is an interesting alternative to the Gaussian scale-space. The unique advantage of the monogenic scale-space, compared with that of the Gaussian scale-space, is the figure flow being in quadrature phase relation to the image at each scale. Therefore, the monogenic scale-space is superior to the Gaussian scale-space if a quadrature relation concept is required [38].

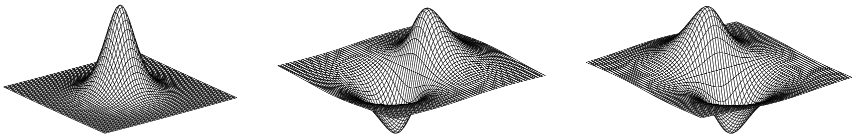

Fig. 3.10: The spherical quadrature filter in the spatial domain. From left to right are the DOP filter and the two components of the DOCP filter. Note that the two components of the DOCP filter constitute one single isotropic operator.

3.7 Summary

In this chapter, some fundamental knowledge of the signal processing for computer vision was introduced. Local structures, which contain most important information of 2D images, are tightly related with the term local intrinsic dimensionality. According to this local property of a signal, the 2D image structure can be classified into three categories, that is the DC component (i0D signals), lines or edges (i1D signals) and corners or junctions (i2D signals). Modeling of local structures is of high significance, since it is the building block of high-level computer vision tasks. Analytic signals are important models for local structures which enable the decomposition of the original signal into two orthogonal parts, i.e. the local amplitude and local phase. In contrast to the classical representation, 2D Fourier transform is reinterpreted in terms of geometric algebra.

In the following chapters, the main contribution of this thesis will be described. It is a general signal model for 2D image structures, which can be constructed based on tensor representations. Hence, a brief overview of two well known tensor approaches in the computer vision community was given. The monogenic signal and the monogenic scale-space play very important roles for the derivation of the general signal model. As a consequence, more detail descriptions about them were given.

4. ROTATIONALLY INVARIANT 2D QUADRATURE FILTER

4.1 Motivation and Background

Quadrature filters are well known issues in signal processing and low-level computer vision. They make it possible to estimate the local amplitudes and local phases of original signals. The local amplitude is a measure for the local dynamics of a structure and the local phase describes the structure or shape of the signal [85]. For 1D signals, the quadrature filter is obtained as a bandpass filter and its Hilbert transform, they together form a pair of an even and an odd filter. This Hilbert pair of filters can be used to detect peaks and jumps of signals. Furthermore, the local phase allows to distinguish detected structures, see the introduction in Chapter 3. The 2D generalization of quadrature filters is far from being trivial. A further degree of freedom for the features should be introduced, i.e. the orientation. For the detection and classification of features it is reasonable to have a rotation invariant approach since the orientation information neither affects the intensity nor the classification of a feature.

There exist also some approaches trying to generalize 2D quadrature filters. One way is just to project 1D quadrature filters onto 2D space. The partial Hilbert transform [48], obtained by projecting the frequency vector onto a preference direction and applying the frequency response of the 1D Hilbert transform to this scalar product, is used to create the odd filter. Special cases of the partial Hilbert transform are the Hilbert transforms with respect to the coordinate axes. Quadrature filters obtained from the partial Hilbert transform are obviously not rotation invariant and are not adequate for detecting i2D features. A second 2D extension of quadrature filters is obtained by means of the total Hilbert transform [48], which is just the successive application of the Hilbert transforms with respect to both coordinate axes. Quadrature filters obtained from the total Hilbert transforms also have the drawback of being not rotationally invariant and are not able to handle i1D features. A further quadrature approach is obtained by combining the previous two methods. The resulting quadrature filter is only non-zero in the first quadrant of the frequency domain [48]. Unfortunately, the reduction to one quadrant yields a loss of information. Therefore, Hahn suggests [48] to consider a second operator output which is non-zero either in the second or in the fourth quadrant. However, this representation in two complex signals is not totally satisfactory. Thus, Bülow and Sommer proposed to use the quaternionic Fourier transform [19, 98]. The resulting quaternionic analytic signal [20, 18] consists of four parts. Two parts correspond to the partial Hilbert transforms with respect to the coordinate axes and one corresponds to the total Hilbert transform. The phase approach of the quaternionic analytic signal also reflects the intrinsic dimension to some extent.

The steerable quadrature filters proposed in [44] are rotation variant since they are orientation adaptive. However, they are not capable to deal with i2D signals either. The only non-steered, rotation invariant approach to quadrature filters which occurred in the literature so far, is obtained from the monogenic signal [36]. It is adequate for treating i1D signals but delivers no information about the 2D part of a signal. The monogenic signal is based on the Riesz transform which is a 2D generalization of the Hilbert transform. The Riesz transform is antisymmetric with respect to the origin since its frequency response is basically given by the normalized frequency vectors [107]. Another generalization, called the structure multivector, is proposed in [37, 31], where an i2D signal is split into two perpendicular parts and the corresponding amplitudes and phases are evaluated. Quadrature filters obtained from this approach are rotation invariant and allow to analyze i2D signals. Unfortunately, only i2D patterns with 90^0 angle of intersection can be correctly handled.

In this chapter, a new approach for designing a 2D rotationally invariant quadrature filter will be presented. According to the previously discussed related work, it is desirable to combine the rotation invariance of the monogenic signal with the symmetry decomposition of the quaternionic analytic signal. The assumed signal model is a superimposed i2D pattern with flexible angle of intersection. This approach also enables the extraction of local amplitude and local phase in a rotation invariant way. Moreover, it is capable to correctly treat superimposed i2D patterns with flexible angle of intersection. The main idea is first to evaluate multi-orientation of local structures by analyzing the eigensystem of a generalized structure tensor. Based on estimated orientations, two i1D signals are separated and the corresponding Riesz transform are evaluated. In accordance, quadrature filters can be derived from this model. Hence, this novel approach can be regarded as an extension of the structure multivector [37].

This chapter is organized as follows. Section 4.2 describes a way to evaluate multiple orientations of local structures. Based on the estimation, a rotationally invariant 2D quadrature filter is designed in Section 4.3. Section 4.4 presents local features which can be obtained by using the newly developed filter. Some experimental results and comparisons are also demonstrated. The final section gives the summary of this chapter.

4.2 Estimation of Multiple Orientations

The design of the new quadrature filter is based first on the orientation estimation of local structures. Hence, detecting multiple orientations at a local neighborhood of the structure is introduced. Since the multiple orientation can be estimated on the basis of a generalized structure tensor, the classical matrix algebra is employed for better understanding of the estimation.

4.2.1 Orientation Estimation for I1D Signals

Assume $\mathbf{f} = \mathbf{f}(\mathbf{x}\cdot\mathbf{n})$ is an i1D signal, where $\mathbf{x} = [x\, y]^T$ and $\mathbf{n}$ is its orientation vector, i.e. $\mathbf{n} = [\cos\theta \; \sin\theta]^T$. The monogenic signal as the novel model of the i1D signal is able to capture simultaneously the local phase and local orientation (i.e., the orientation of the Dirac line in the Fourier domain). Hence, the constraint equation for this i1D signal then takes the form

$$\mathbf{n}^T\mathbf{f}_R = \mathbf{n}^T\left(h_R * \mathbf{f}\right) = 0 \;, \tag{4.1}$$

where $\mathbf{f}_R$ represents the Riesz transformed signal and h_R indicates the Riesz kernel

$$h_R = [h_{R1}\, h_{R2}]^T = \left[\frac{x}{2\pi(x^2+y^2)^{\frac{3}{2}}} \quad \frac{y}{2\pi(x^2+y^2)^{\frac{3}{2}}}\right]^T . \tag{4.2}$$

Hence, the orientation vector can be estimated within a local neighborhood Ω by minimizing the energy integral which is defined as

$$E_{i1D} = \int_\Omega \left(\mathbf{n}^T\left(h_R * \mathbf{f}\right)\right)^2 d\Omega = \mathbf{n}^T T_1 \mathbf{n} \quad \text{with} \quad \mathbf{n}^T\mathbf{n} = 1 \;, \tag{4.3}$$

where the constraint $\mathbf{n}^T\mathbf{n} = 1$ excludes the trivial solution $\mathbf{n} = 0$ and T_1 is a two by two symmetric tensor which is obtained as

$$\begin{aligned} T_1 &= \int_\Omega (h_R * \mathbf{f}) \otimes (h_R * \mathbf{f}) d\Omega = \int_\Omega (h_R * \mathbf{f})(h_R * \mathbf{f})^T d\Omega \qquad (4.4) \\ &= \int_\Omega \begin{bmatrix} (h_{R1} * \mathbf{f})^2 & (h_{R1} * \mathbf{f})(h_{R2} * \mathbf{f}) \\ (h_{R1} * \mathbf{f})(h_{R2} * \mathbf{f}) & (h_{R2} * \mathbf{f})^2 \end{bmatrix} d\Omega \;, \end{aligned}$$

where $\otimes$ indicates the tensor product. Minimizing the energy E_{i1D} is equivalent to find $\mathbf{n}$ such that

$$T_1\mathbf{n} = \lambda\mathbf{n} \quad \text{with} \quad \mathbf{n}^T\mathbf{n} = 1 \; , \tag{4.5}$$

where λ denotes the smallest eigenvalue of the tensor T_1. For i1D signals, the tensor T_1 has a rank of 1. This implies that $\mathbf{n}$ is the null-eigenspace of T_1 and its eigenvector corresponding to the zero eigenvalue gives the best fit of the orientation vector.

4.2.2 Multiple Orientation Estimation for I2D Signals

In contrast to i1D signals, i2D signals have a lot of degrees of freedoms. Due to the diversity of i2D signals, the designed quadrature filter takes only the double oriented pattern with flexible angle of intersection into consideration. The double oriented pattern is formed by the superimpose of two differently oriented i1D signals. Hence, it is necessary to detect exact orientations of these two i1D signals which form the i2D pattern.

For a double oriented i2D pattern, it can be modeled as the superimpose of two differently oriented i1D signals, that is

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}\cdot\mathbf{n}_1) + \mathbf{f}_2(\mathbf{x}\cdot\mathbf{n}_2) \; , \tag{4.6}$$

where $\mathbf{n}_1 = [\cos\theta_1 \; \sin\theta_1]^T$ and $\mathbf{n}_2 = [\cos\theta_2 \; \sin\theta_2]^T$ indicate two orientation vectors.

Normally, to estimate the orientation of such an i2D neighborhood, it is natural to consider the structure tensor [43, 59]. It can be considered in the context of an approximation of the autocorrelation of the signal. Hence, the two eigenvalues of the structure tensor are related to the principal curvatures of the autocorrelation function, whereas the eigenvectors indicate the corresponding coordinate system. Therefore, the structure tensor can be used to estimate the local main orientation of a structure by means of the eigenvector which corresponds to the largest eigenvalue. The eigenvalues themselves and their relation also provide a measure for the intrinsic dimension of the signal. However, to handle the i2D pattern with flexible angle of intersection, it is of great importance to evaluate exactly two orientations of the two i1D signals instead of the local main orientation. For this purpose, we take the idea which is similar to that in [1, 81]. First, the best fit for the mixture of two i1D signal orientations is estimated, then the mixed information is decomposed into two orientations of the i1D signals.

Similar as the i1D case, for two i1D signals of a double oriented pattern, the following constraint is satisfied

$$\mathbf{n}_1^T(h_R * \mathbf{f}_1) = \mathbf{n}_2^T(h_R * \mathbf{f}_2) = 0 \; . \tag{4.7}$$

Hence, for the i2D neighborhood, the following relation holds

$$\mathbf{n}_2^T \left(h_R * (\mathbf{n}_1^T h_R * \mathbf{f}) \right) = \mathbf{N}^T \mathbf{J} = 0 \ . \tag{4.8}$$

with

$$\begin{aligned} \mathbf{N} &= [\cos\theta_1 \cos\theta_2 \quad \sin(\theta_1 + \theta_2) \quad \sin\theta_1 \sin\theta_2]^T & (4.9) \\ \mathbf{J} &= [h_{R1} * (h_{R1} * \mathbf{f}) \quad h_{R1} * (h_{R2} * \mathbf{f}) \quad h_{R2} * (h_{R2} * \mathbf{f})]^T \ , & (4.10) \end{aligned}$$

where $\mathbf{N}$ represents the mixed information of $\mathbf{n}_1$ and $\mathbf{n}_2$.

In the spectral domain, this relation for i2D signals reads

$$(\mathbf{n}_1^T \omega)(\mathbf{n}_2^T \omega)\mathbf{F}(\omega) = 0 \ , \tag{4.11}$$

where ω denotes the frequency vector and $\mathbf{F}(\omega)$ is the Fourier transform of the original signal $\mathbf{f}$. Since the i1D signal is an oriented line in the Fourier domain, equation (4.11) indicates that the local spectrum of the i2D signal must be zero for both lines, i.e. $\mathbf{n}_1^T \omega = 0$ and $\mathbf{n}_2^T \omega = 0$.

Similar as the i1D signal case, the estimation of the mixed orientation information $\mathbf{N}$ can be realized within a local neighborhood by minimizing the following energy integral

$$E_{i2D} = \int_\Omega (\mathbf{N}^T \mathbf{J})^2 d\Omega = \mathbf{N}^T T_2 \mathbf{N} \ . \tag{4.12}$$

To exclude the trivial solution, the constraint $\mathbf{N}^T \mathbf{N} = a$, $a > 0$ must be imposed, where a works only as a scaling factor. Finding the best fit of the mixed orientation $\mathbf{N}$ is equivalent to analyze the eigensystem of the tensor T_2, that is

$$T_2 \mathbf{N} = \lambda \mathbf{N} \ , \tag{4.13}$$

where T_2 is defined as

$$\begin{aligned} T_2 &= \int_\Omega \mathbf{J} \otimes \mathbf{J} d\Omega & (4.14) \\ &= \int_\Omega \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} d\Omega \end{aligned}$$

with

$$\begin{aligned} T_{11} &= (h_{R1} * (h_{R1} * \mathbf{f}))^2 & (4.15) \\ T_{12} &= T_{21} = (h_{R1} * (h_{R1} * \mathbf{f}))(h_{R1} * (h_{R2} * \mathbf{f})) \\ T_{13} &= T_{31} = (h_{R1} * (h_{R1} * \mathbf{f}))(h_{R2} * (h_{R2} * \mathbf{f})) \\ T_{22} &= (h_{R2} * (h_{R1} * \mathbf{f}))^2 \\ T_{23} &= T_{32} = (h_{R2} * (h_{R1} * \mathbf{f}))(h_{R2} * (h_{R2} * \mathbf{f})) \\ T_{33} &= (h_{R2} * (h_{R2} * \mathbf{f}))^2 \ . \end{aligned}$$

In the ideal case of the i2D signal, T_2 has only a rank of two. Hence, its null-space indicates the best fit of the mixed orientation $\mathbf{N}$. The eigenvector which corresponds to the smallest eigenvalue of T_2 indicates the mixed orientation vector. In order to design a 2D quadrature filter which suits also double oriented patterns with flexible angle of intersections, two explicit orientation vectors $\mathbf{n}_1$ and $\mathbf{n}_2$ should be extracted from the mixed orientation $\mathbf{N}$. Let $\mathbf{N}$ be represented as

$$\mathbf{N} = [a \quad b \quad c]^T = [\cos\theta_1 \cos\theta_2 \quad \sin(\theta_1 + \theta_2) \quad \sin\theta_1 \sin\theta_2]^T \ , \tag{4.16}$$

a 2×2 matrix M can be constructed based on the mixed orientation with the following form

$$M = \begin{bmatrix} \cos\theta_1 \cos\theta_2 & \cos\theta_1 \sin\theta_2 \\ \sin\theta_1 \cos\theta_2 & \sin\theta_1 \sin\theta_2 \end{bmatrix} = \begin{bmatrix} a & z_1 \\ z_2 & c \end{bmatrix} . \tag{4.17}$$

It is obvious that $z_1 z_2 = ac$ and $z_1 + z_2 = b$. Thereby, z_1 and z_2 can be obtained as the roots of the following polynomial, that is

$$f(z) = (z - z_1)(z - z_2) = z^2 - bz + ac \ . \tag{4.18}$$

Hence, two explicit angles θ_1 and θ_2 are given as

$$\theta_1 = \arctan\left(\frac{z_1}{c}\right) = \arctan\left(\frac{a}{z_2}\right) \quad \theta_1 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{4.19}$$

$$\theta_2 = \arctan\left(\frac{z_1}{a}\right) = \arctan\left(\frac{c}{z_2}\right) \quad \theta_1 \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right] \ . \tag{4.20}$$

The corresponding angle of intersection θ thus reads

$$\theta = |\theta_1 - \theta_2| \ . \tag{4.21}$$

4.3 Quadrature Filter Design

Based on the multiple orientation estimation, a rotationally invariant 2D quadrature filter can be designed. In contrast to the structure multivector [31, 37] which is only suited for the patterns superimposed by two perpendicular i1D signals, this approach also fits for double oriented patterns with flexible angle of intersections. In this section, we will switch to the geometric algebra since it enables much simpler and compact representations.

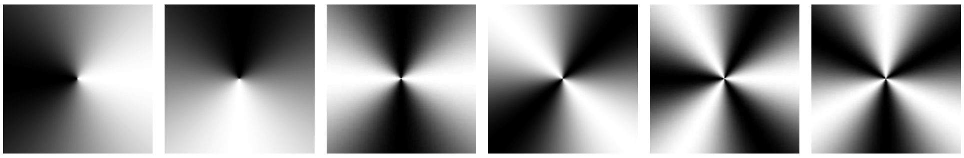

Fig. 4.1: From left to right are 2D spherical harmonics from order 1 to 3 in the spatial domain, except for the zero order, every spherical harmonic consists of two orthogonal components. White indicates positive one and black represents negative one.

4.3.1 Basis Functions

In [44], 2D spherical harmonics are used to design steerable filters which are orientation adaptive. Therefore, we take also 2D spherical harmonics as basis functions to design the 2D rotationally invariant quadrature filter. According to [31], only zero to three order spherical harmonics are employed in order to avoid aliasing. To design the filter, the main concern is the angular portions of spherical harmonics. Therefore, polar representation of spherical harmonics are used instead of the Cartesian form stated in [31].

In the frequency domain, 2D spherical harmonics have much simpler representations than those in the spatial domain. An *nth* order spherical harmonic H_n in the spectral domain takes the following form

$$H_n(\alpha) = \exp(n\alpha\mathbf{e}_{12}) = \cos(n\alpha) + \sin(n\alpha)\mathbf{e}_{12} \ , \tag{4.22}$$

where α denotes the angular part of the polar coordinate in the Fourier domain, n refers to the order of the spherical harmonic. Every spherical harmonic consists of two orthogonal components and the first order spherical harmonic is basically identical to the Riesz kernel [36]. In order to decompose the signal into its distinct frequency components, 2D spherical harmonics are combined with radial bandpass filters. In this thesis, the difference of Poisson (DOP) bandpass filter [31] is employed, see also Section 3.6.2. Fig. 4.1 illustrates spherical harmonics from order 1 to 3 in the spatial domain, they have the same angular portions as those in the spectral domain.

4.3.2 2D Quadrature Filter

Intrinsically 2D signals have a much greater variety than the i1D signals. This increase of possible signal realizations can be considered more formally by means of symmetries. It is reported in [20, 18] that a 2D signal can be split into even and odd parts along the x-axis and along the y-axis as well. Hence, every 2D signal can be written in the form of $\mathbf{f} = \mathbf{f}_{ee} + \mathbf{f}_{eo} + \mathbf{f}_{oe} + \mathbf{f}_{oo}$, see also Chapter 3. In this case, an arbitrary 2D signal, according to the quaternionic Fourier transform (see Section 3.2.3), can be represented by the following basis functions

$$\begin{aligned} b_1(\mathbf{x};\mathbf{u}) &= \cos(2\pi xu)\cos(2\pi yv) \\ b_2(\mathbf{x};\mathbf{u}) &= \cos(2\pi xu)\sin(2\pi yv) \\ b_3(\mathbf{x};\mathbf{u}) &= \sin(2\pi xu)\cos(2\pi yv) \\ b_4(\mathbf{x};\mathbf{u}) &= \sin(2\pi xu)\sin(2\pi yv)\ . \end{aligned} \tag{4.23}$$

These four basis functions are intrinsically 2D, which allows them to capture the i2D image structures. They indicate the even-even, even-odd, odd-even and odd-odd symmetries with respect to the x and y axes. Obviously, this representation lacks the property of rotation invariance. Hence, Felsberg [31, 37] developed the so-called structure multivector to cover this drawback. The signal model in his approach is the superimpose of two perpendicular i1D signals which can be given by

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}\cdot\mathbf{n}) + \mathbf{f}_2(\mathbf{x}\cdot\mathbf{n}^\perp)\ , \tag{4.24}$$

where $\mathbf{n}$ is a orientation vector, $\mathbf{n}^\perp$ is the vector obtained by rotating $\mathbf{n}$ by $\frac{\pi}{2}$ anticlockwise. The structure multivector is also based on the four symmetries, however, they are with respect to a new coordinate system $\mathbf{n}$ and $\mathbf{n}^\perp$ instead of the x and y axes. Unfortunately, this approach produces errors if the i2D structure is composed of two non-perpendicular i1D signals.

The proposed 2D quadrature filter can be regarded as an extension of the structure multivector since it can correctly handle the double oriented patterns with flexible angle of intersections. In this case, a 2D signal is split into two i1D signals which are not necessarily perpendicular to each other, the signal model thus reads

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}_1(\mathbf{x}\cdot\mathbf{n}_1) + \mathbf{f}_2(\mathbf{x}\cdot\mathbf{n}_2)\ , \tag{4.25}$$

where $\mathbf{n}_1 = \cos\theta_1\mathbf{e}_1 + \sin\theta_1\mathbf{e}_2$ and $\mathbf{n}_2 = \cos\theta_2\mathbf{e}_1 + \sin\theta_2\mathbf{e}_2$ represent orientation vectors of the two i1D signals which can be evaluated according to the multiple orientation estimation approach. The main idea for designing the 2D quadrature filter is to split an i2D signal as two i1D components according to their orientations and then to evaluate the corresponding Riesz transforms.

The starting point is still the four symmetries. However, the four symmetries in the current case should be attached to two new coordinate systems, respectively, for the sake of two i1D signals. The two new coordinate systems

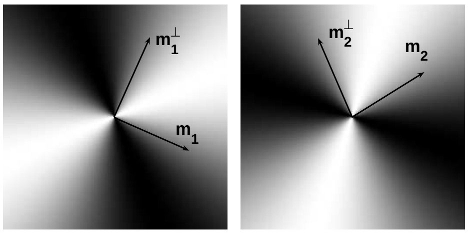

Fig. 4.2: Two odd-odd symmetries with respect to two coordinate systems. They are obtained from $\cos(2(\theta-\theta_1))$ and $\cos(2(\theta-\theta_2))$ (white:+1, black:-1).

are decided by $\mathbf{m}_1$, $\mathbf{m}_1^\perp$ and $\mathbf{m}_2$, $\mathbf{m}_2^\perp$ which can be obtained from $\mathbf{n}_1$ and $\mathbf{n}_2$ as follows

$$\begin{aligned}\mathbf{m}_1 &= \frac{1}{2}(\mathbf{n}_1+\mathbf{n}_1^*) \quad \text{and} \quad \mathbf{m}_1^\perp = \mathbf{m}_1 I_2 = \frac{1}{2}(\mathbf{n}_1-\mathbf{n}_1^*) \qquad (4.26)\\ \mathbf{m}_2 &= \frac{1}{2}(\mathbf{n}_2+\mathbf{n}_2^*) \quad \text{and} \quad \mathbf{m}_2^\perp = \mathbf{m}_2 I_2 = \frac{1}{2}(\mathbf{n}_2-\mathbf{n}_2^*)\ ,\end{aligned}$$

where $\mathbf{n}_1^* = -\mathbf{n}_1\mathbf{e}_{12}$ and $\mathbf{n}_2^* = -\mathbf{n}_2\mathbf{e}_{12}$ are dual vectors of $\mathbf{n}_1$ and $\mathbf{n}_2$, respectively.

Spherical harmonics can be used as basis functions to create the desired symmetry properties. The even-even symmetries for both coordinate systems are trivial and can simply be obtained from the zero order spherical harmonic. Hence, even-even symmetries in the coordinate systems $\mathbf{m}_1$, $\mathbf{m}_1^\perp$ and $\mathbf{m}_2$, $\mathbf{m}_2^\perp$ are represented as $p_{ee1} = p_{ee2} = H_0 = 1$. The odd-odd symmetries with respect to the new two coordinate systems are given by steering and projecting the second order spherical harmonic onto the double-angle orientation vectors, that is

$$\begin{aligned}p_{oo1} &= (\mathbf{n}_1\mathbf{e}_1\mathbf{n}_1)\cdot(\mathbf{e}_1 H_2) = \cos(2(\theta-\theta_1)) \qquad (4.27)\\ p_{oo2} &= (\mathbf{n}_2\mathbf{e}_1\mathbf{n}_2)\cdot(\mathbf{e}_1 H_2) = \cos(2(\theta-\theta_2))\ .\end{aligned}$$

Fig. 4.2 indicates the acquired two odd-odd symmetries with respect to two coordinate systems by using the second order spherical harmonic. Combing these even-even and odd-odd symmetries, two angular windowing functions W_1 and W_2 are obtained to yield two oriented i1D signals, see also Fig. 4.3

$$\begin{aligned}W_1 &= p_{ee1} + p_{oo1} = 1 + \cos(2(\theta-\theta_1)) = 2\cos^2(\theta-\theta_1) \qquad (4.28)\\ W_2 &= p_{ee2} + p_{oo2} = 1 + \cos(2(\theta-\theta_2)) = 2\cos^2(\theta-\theta_2)\ .\end{aligned}$$

After splitting the signal into its two i1D parts, the corresponding Riesz

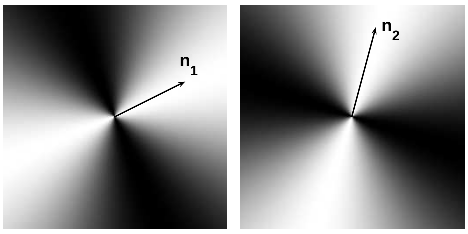

Fig. 4.3: Two angular windowing functions obtained from $\cos^2(\theta - \theta_1)$ and $\cos^2(\theta - \theta_2)$ (white:+1, black:0).

transforms [36] should be evaluated in order to obtain the phase information. Because orientations of these two i1D signals are already known, it is necessary to steer the Riesz transforms according to the two orientations. Hence, for the i1D signal with orientation vector $\mathbf{n}_1$, the Riesz kernel can be applied to the first angular windowing function W_1 which results in the following form

$$\begin{aligned} W_3 &= \mathbf{e}_1\mathbf{n}_1 H_1 W_1 \qquad (4.29)\\ &= \mathbf{e}_1(\cos\theta_1\mathbf{e}_1 + \sin\theta_1\mathbf{e}_2)(\cos\theta + \sin\theta\mathbf{e}_{12})2\cos^2(\theta-\theta_1)\\ &= \frac{1}{2}(3\cos(\theta-\theta_1) + \sin(\theta-\theta_1))\mathbf{e}_{12} + \cos(3(\theta-\theta_1)) + \sin(3(\theta-\theta_1))\mathbf{e}_{12}\ . \end{aligned}$$

Accordingly, the Riesz transform of the second angular windowing function is given by

$$\begin{aligned} W_4 &= \mathbf{e}_1\mathbf{n}_2 H_1 W_2 \qquad (4.30)\\ &= \mathbf{e}_1(\cos\theta_2\mathbf{e}_1 + \sin\theta_2\mathbf{e}_2)(\cos\theta + \sin\theta\mathbf{e}_{12})2\cos^2(\theta-\theta_2)\\ &= \frac{1}{2}(3\cos(\theta-\theta_2) + \sin(\theta-\theta_2))\mathbf{e}_{12} + \cos(3(\theta-\theta_2)) + \sin(3(\theta-\theta_2))\mathbf{e}_{12}\ . \end{aligned}$$

The Riesz transformed parts of these two angular windowing functions consist of odd order spherical harmonics. Hence, using spherical harmonics from order zero to three as basis functions enables the extraction of two i1D signals and their corresponding Riesz transformed results. This means that the proposed 2D rotationally invariant quadrature filter response is equivalent to two monogenic signals. As in the case of the monogenic signal, the original signal is encoded by $\mathbf{e}_3$ axis and its Riesz transformed part is encoded into $\mathbf{e}_1 \wedge \mathbf{e}_2$ plane. Hence, the angular windowing functions are multiplied by $\mathbf{e}_3$ basis vector. Because their Riesz transforms are spinor valued, they are multiplied by $\mathbf{e}_1$ basis vector from the left side. Thus, the 2D rotationally invariant quadrature filter Q is finally the combination of two i1D quadrature filters which consist of the two angular windowing functions and their Riesz transformed results, that is

$$Q = Q_1 + Q_2 = (\mathbf{e}_3 W_1 + \mathbf{e}_1 W_3) + (\mathbf{e}_3 W_2 + \mathbf{e}_1 W_4) \ . \tag{4.31}$$

4.4 Feature Extraction and Experiments

From the filter response of Q, a rich set of local features which fulfill the invariance-equivariance property can be extracted. For an $\mathbf{e}_3$ valued signal $\mathbf{f}$, the filter response $\mathbf{r}$ of the 2D quadrature filter reads

$$\begin{aligned} \mathbf{r} &= \mathcal{F}^{-1}\{Q\mathbf{F}\} \qquad (4.32) \\ &= \mathcal{F}^{-1}\{[(\mathbf{e}_3 W_1 + \mathbf{e}_1 W_3) + (\mathbf{e}_3 W_2 + \mathbf{e}_1 W_4)]\mathbf{F}\} \\ &= (\mathbf{r}_1 + \mathbf{r}_3) + (\mathbf{r}_2 + \mathbf{r}_4) \ , \end{aligned}$$

where $\mathbf{F}$ is the Fourier transform of $\mathbf{f}$. Based on this, local amplitudes (A_1, A_2) and local phases (φ_1, φ_2) for two i1D signals are obtained as follows

$$A_1 = \sqrt{\mathbf{r}_1^2 + \mathbf{r}_3^2} \tag{4.33}$$

$$A_2 = \sqrt{\mathbf{r}_2^2 + \mathbf{r}_4^2} \tag{4.34}$$

$$\varphi_1 = \arg(\mathbf{r}_1 + \mathbf{r}_3) \tag{4.35}$$

$$\varphi_2 = \arg(\mathbf{r}_2 + \mathbf{r}_4) \ , \tag{4.36}$$

where arg means the argument of a expression.

In the following, some experiments are conducted to evaluate the performance of the newly developed 2D quadrature filter. The first experiment uses a synthetic image shown in Fig. 4.4, it is superimposed by two i1D signals with orientations of $\frac{\pi}{4}$ and $\frac{\pi}{2}$. This pattern has obviously a non-90^0 angle of intersection. According to the novel 2D rotationally invariant quadrature filter, at those i2D points, double orientations will first be estimated, see also Fig. 4.4. Based on the signal model, two i1D signals along different orientations are extracted and the corresponding amplitudes and phases can be correctly evaluated. Fig. 4.5 illustrates the estimated amplitudes and phases. It can be shown that the proposed approach enables the correct evaluation of local features of double oriented patterns. As a comparison, feature extraction based on the structure multivector is also given in Fig. 4.6. As introduced in the background section, the signal model of the structure multivector is the superimpose of two perpendicular i1D signals. From the feature extraction result of the structure multivector, one can easily see

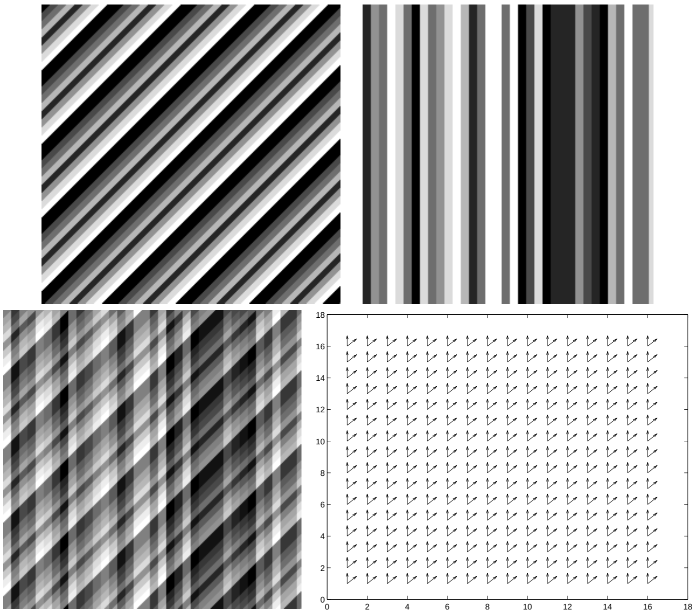

Fig. 4.4: Top row: two i1D patterns with different orientations. Bottom row: the i2D pattern superimposed by the two i1D patterns and the sampled orientation estimation by using the novel 2D rotationally invariant quadrature filter.

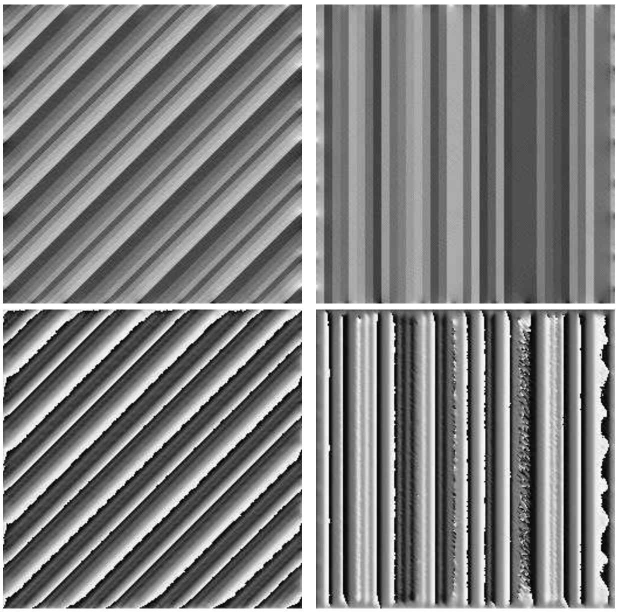

Fig. 4.5: Top row: local amplitudes of two i1D patterns. Bottom row: local phases of two i1D patterns.

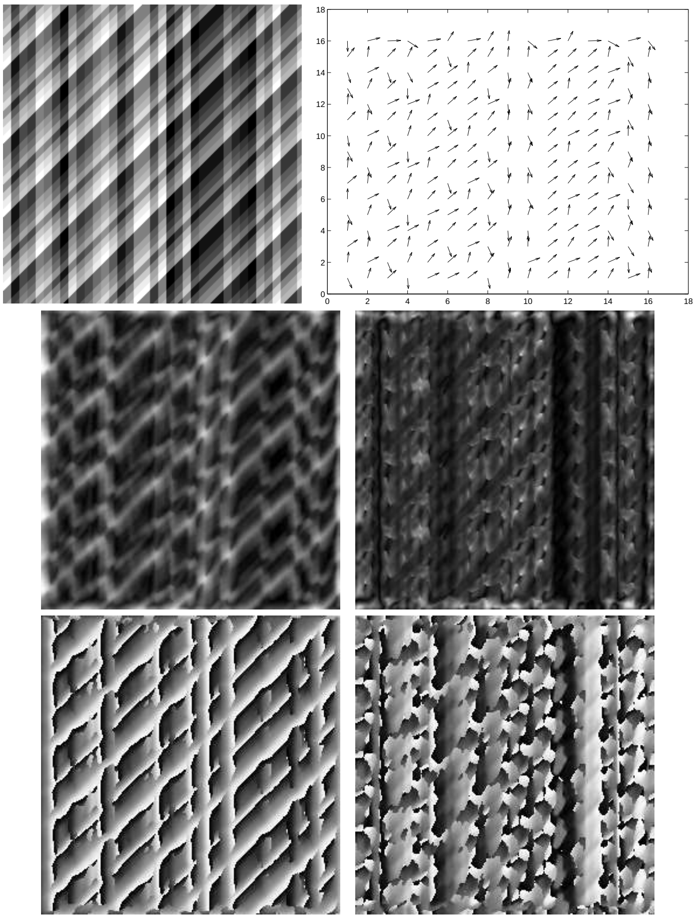

Fig. 4.6: Top row: the synthetic image and the estimated main orientation from the structure multivector. Middle row: local major and minor amplitudes extracted by using the structure multivector. Bottom row: local major and minor phases extracted by using the structure multivector.

that this approach fails to extract correct local features once its signal model is violated.

Another double oriented pattern is chosen for the second experiment, as visualized in Fig. 4.7. The two i1D signals which form the i2D pattern are not perpendicular to each other. Using the newly proposed 2D quadrature filter, two orientations can be correctly estimated. Based on the orientation information, corresponding local amplitudes and phases are extracted. From Fig. 4.7, one can see that the proposed approach is able to correctly handle those double oriented i2D patterns with flexible angle of intersections. In contrast to this, due to the violation of the assumed signal model, the struc-

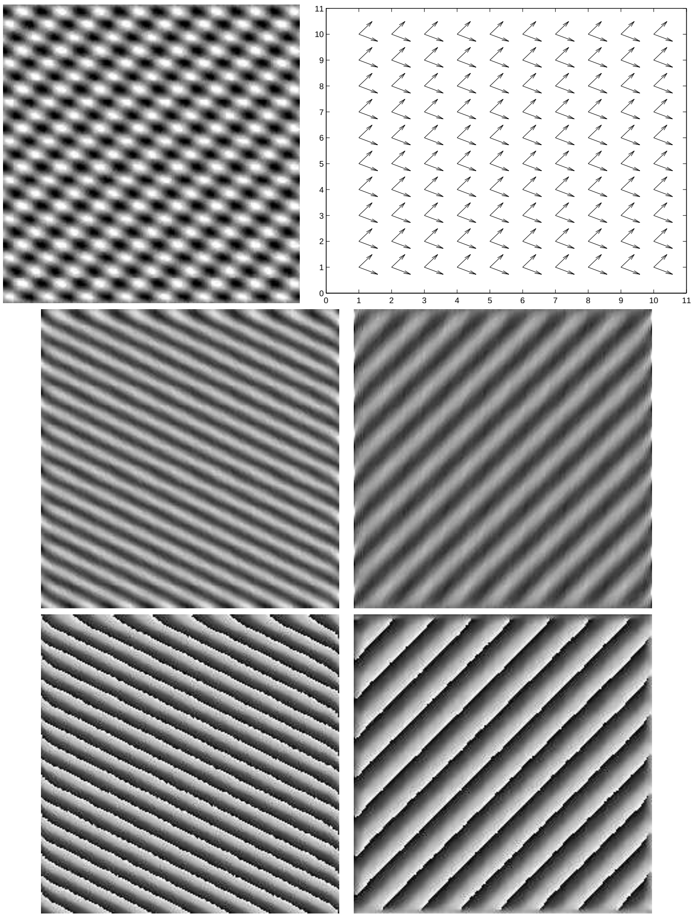

Fig. 4.7: Top row: the synthetic image and the estimated main orientation from the proposed 2D quadrature filter. Middle row: local amplitudes extracted by using the proposed 2D quadrature filter. Bottom row: local phases extracted by using the proposed 2D quadrature filter.

ture multivector approach produces errors at those i2D points, see Fig. 4.8. The third experiment aims to show that the newly developed 2D quadrature filter can be considered as an extension of the structure multivector since it is also able to handle i2D patterns superimposed by two perpendicular i1D patterns. Fig. 4.9 illustrates the corresponding local features.

4.5 Summary

In this chapter, a novel 2D rotationally invariant quadrature filter was presented. The motivation of designing such a filter is based on an assumed

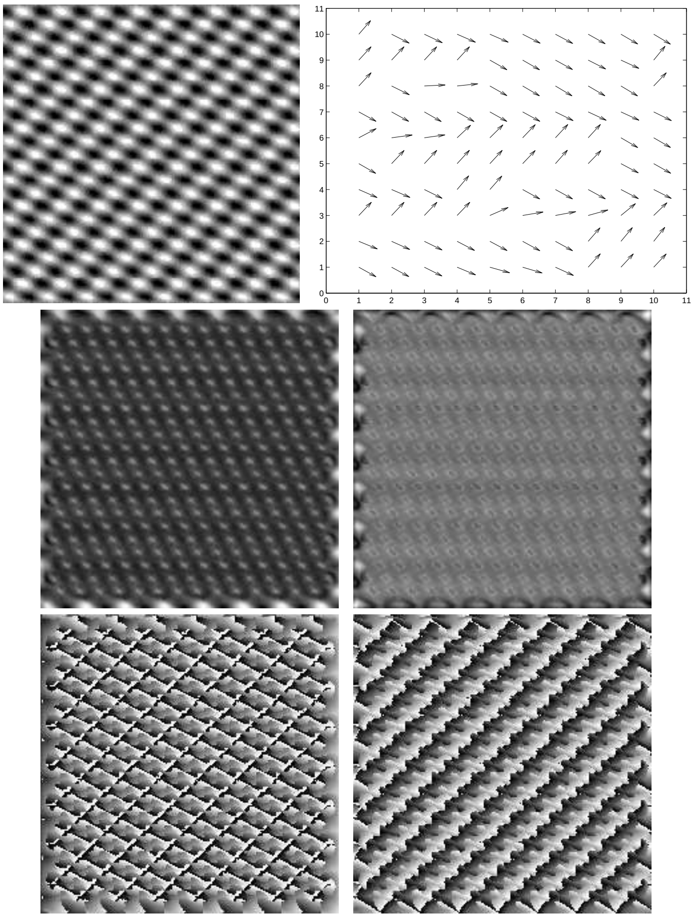

Fig. 4.8: Top row: the synthetic image and the estimated main orientation from the structure multivector. Middle row: local major and minor amplitudes extracted by using the structure multivector. Bottom row: local major and minor phases extracted by using the structure multivector.

signal model which is the superposition of two differently oriented i1D signals. Those two i1D signals are not necessarily to be perpendicular to each other. By constructing a tensor representation, two orientations of a i2D point can be correctly estimated. According to the double orientation information, at the i2D point, a signal is split into two oriented i1D signals and the corresponding Riesz transforms are employed for the phase evaluation. In contrast to this, the structure multivector approach works only for the i2D signal which is superimposed by two perpendicular i1D signals. Hence, the proposed 2D quadrature filter can be regarded as an extension of the structure multivector since it is able to handle double oriented i2D patterns

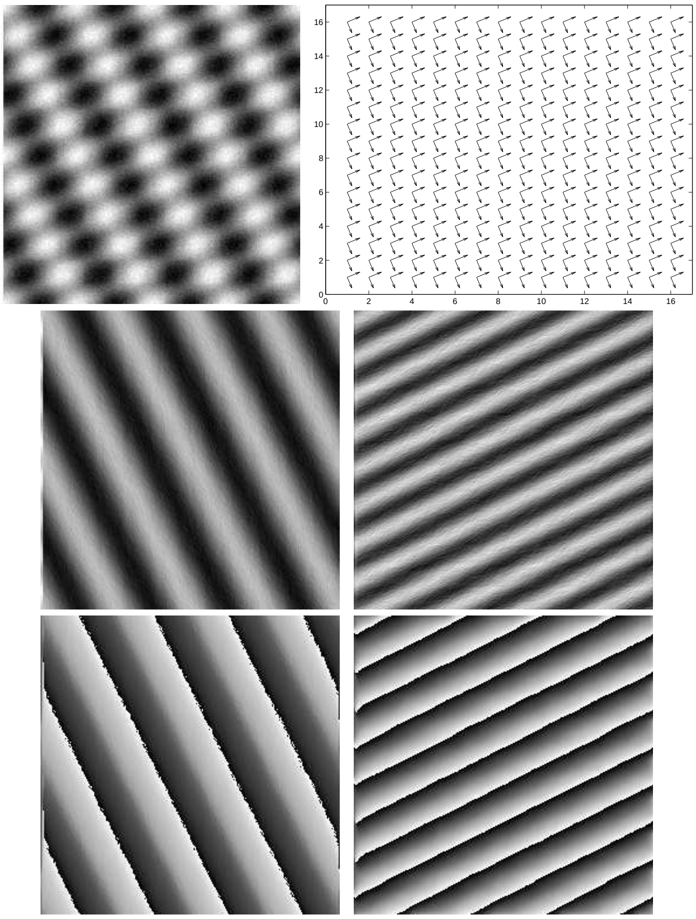

Fig. 4.9: Top row: the synthetic image and the estimated main orientation from the proposed 2D quadrature filter. Middle row: local amplitudes extracted by using the proposed 2D quadrature filter. Bottom row: local phases extracted by using the proposed 2D quadrature filter.

with flexible angle of intersections.

5. SIGNAL MODELING FOR TWO-DIMENSIONAL IMAGE STRUCTURES

Model based image representation plays an important role in many computer vision tasks such as object recognition, motion estimation, image retrieval, etc. Therefore, signal modeling for local structures is of high significance in image processing. In this chapter, a novel image model for two-dimensional image structures, called the monogenic curvature tensor, will be presented. The first section introduces some mathematical preliminaries which include the geometric embedding of the signal and basis functions employed to construct the novel model. Detail information to construct the monogenic curvature tensor is described in the second section. Section three focuses on the interpretation of the generalized monogenic curvature signal, which is derived from the monogenic curvature tensor as a local representation of i2D image structures. The fourth section illustrates the parity symmetry analysis of the generalized monogenic curvature signal. The generalized monogenic curvature scale-space, generated from the monogenic curvature tensor, is introduced in the fifth section. Summary of this chapter is given in the final section.

5.1 Mathematical Preliminaries

5.1.1 Geometric Embedding of the Signal

The way we intend to design a general model for 2D structures is a generalization of the analytic signal. It cannot be realized in the domain of complex numbers. Instead, a more powerful algebraic system should be taken into consideration. Geometric algebras [73, 2, 53] constitute a rich family of algebras as generalization of vector algebra [54]. Compared with the classical framework of vector algebra, the geometric algebra makes available a tremendous extension of modeling capabilities. By embedding our problem into a certain geometric algebra, more degrees of freedom can be obtained, which makes it possible to extract multiple features of 2D structure. For the problem we are concerned, the 2D signal will be algebraically embedded into the Euclidean 3D space $\mathbb{R}^3$. Hence, geometric algebra over 3D Euclidean space $\mathbb{R}^3$ (see Chapter 2 for details) is employed as the mathematical background for signal modeling.

The scalar-valued real 2D image signal $f(x, y)$ will be embedded into $\mathbb{R}_3$ as a real-valued vector field, i.e. $f(x, y)\mathbf{e}_3 \equiv \mathbf{f}(\mathbf{x}) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3)$, $\mathbf{x} \in \mathbb{R}^3$ and $z = 0$. Thus, rotating the signal $\mathbf{f}$ out of the $\mathbf{e}_3$ axis results in a monogenic signal $\mathbf{f}_M$ [36], which is composed of the original signal and its Riesz transformed part.

In terms of polar coordinates, the spatial and spectral domain representations of the Riesz kernel take the following forms

$$h_R(r, \beta) = \frac{1}{2\pi r^2}(-\cos(\beta)\mathbf{e}_{31} + \sin(\beta)\mathbf{e}_{23}) \tag{5.1}$$

$$H_R(\rho, \alpha) = -\cos(\alpha)\mathbf{e}_2 + \sin(\alpha)\mathbf{e}_1 \ . \tag{5.2}$$

The vector-valued signal $\mathbf{f}_M$ in $\mathbb{R}_3$ can be considered as the impulse response of a spinor S acting on the $\mathbf{e}_3$ basis vector, i.e. $\mathbf{f}_M = \mathbf{e}_3 S$. The transformation performed under the action of the spinor delivers access to both the amplitude and phase information of the vector-valued signal $\mathbf{f}_M$ [106]. To make this clear, we will consider the spinor more in depth. The spinor can be represented in polar coordinates as an exponential form. Therefore, from the logarithm of the spinor, two parts can be obtained. They are the scaling which corresponds to the local amplitude and the rotation which corresponds to the local phase representation. The $\mathbb{R}_3$-logarithm of a spinor $S \in \mathbb{R}_3^+$ takes the following form

$$\log(S) = \langle \log(S) \rangle_0 + \langle \log(S) \rangle_2 = \log(|S|) + \frac{\langle S \rangle_2}{|\langle S \rangle_2|} \mathrm{atan}\left(\frac{|\langle S \rangle_2|}{\langle S \rangle_0} \right) \ , \tag{5.3}$$

where atan is the arc tangent mapping for the interval $[0, \pi)$. The scalar part $\langle \log(S) \rangle_0 = \log(|S|)$ illustrates the attenuation [38] as the logarithm of the local amplitude. Hence, the local amplitude is obtained as the exponential of it

$$|S| = \exp(\log|S|) = \exp(\langle \log(S) \rangle_0) \ . \tag{5.4}$$

The bivector part of $\log(S)$ indicates the local phase representation

$$\arg(S) = \langle \log(S) \rangle_2 = \frac{\langle S \rangle_2}{|\langle S \rangle_2|} \mathrm{atan}\left(\frac{|\langle S \rangle_2|}{\langle S \rangle_0} \right) \ . \tag{5.5}$$

It is also possible to extend the monogenic signal to higher dimensions, see [106, 70]. Hence, the monogenic signal can be regarded as a generalization of the analytic signal in 2D and higher dimensions. The monogenic signal is a novel model for i1D signals, however, it delivers no information about the i2D parts of the 2D image.

If we investigate the case for the half space $z > 0$, this signal will become a smoothed version of the original signal ($z = 0$). Hence, z is regarded as a scale parameter. Applying a Poisson kernel to the original signal results in a smoothed signal which, for all scale parameters s, results in a Poisson scale-space $p(\mathbf{x}; s)$ [38]. The conjugate of the Poisson scale-space $\mathbf{q}(\mathbf{x}; s)$ is obtained by convolving the original signal with a conjugate Poisson kernel. The monogenic scale-space is thus formed by the combination of the Poisson scale-space and its conjugate part, see also Chapter 3 for details.

5.1.2 Basis Functions

In order to analyze 2D patterns, we choose 2D spherical harmonics as basis functions according to the proposal in [31]. Actually, spherical harmonic is a general term which denotes the harmonic oscillations on the unit sphere of a multi-dimensional Euclidean space. However, in this thesis, only the spherical harmonics of 2D space are employed, they are also called circular harmonics. Due to the theory of Fourier series, one can approximate any plane angular function (in L_2 sense) by using 2D spherical harmonics.

In Chapter 4, because the angular behavior of a signal can be regarded as band limited, only spherical harmonics of order zero to three are applied, otherwise, aliasing would occur on a discrete grid around a location $\mathbf{x}$. To construct the general signal model of this chapter, 2D spherical harmonics from zero to three orders are also employed. However, this choice is motivated from the differential geometry. The proposed signal model, called the monogenic curvature tensor, consists of a curvature tensor and its monogenic extension. The curvature tensor is tightly related with the Hessian matrix and its monogenic extension is obtained from the Riesz transform. Hence, only 2D spherical harmonics from zero to three orders are needed.

To build the signal model, we are more concerned of the angular portions. Therefore, we use the polar representation of spherical harmonics instead of the Cartesian form applied in [31].

In the frequency domain, an nth order spherical harmonic H_n takes the following form

$$H_n(\rho, \alpha) = H_n(\rho) H_n(\alpha) \tag{5.6}$$

with

$$H_n(\rho) = 1, \quad H_n(\alpha) = \exp(n\alpha \mathbf{e}_{12}) = \cos(n\alpha) + \sin(n\alpha)\mathbf{e}_{12} \ , \tag{5.7}$$

where ρ and α denote the polar coordinates in the Fourier domain and n refers to the order of the spherical harmonic. Every spherical harmonic consists of two orthogonal components and the first order spherical harmonic is basically identical to the Riesz kernel [36] which is well known in Clifford analysis as the multidimensional generalization of the Hilbert kernel. Since H_n is separable into radial part and angular variation, and its angular variation is harmonic, according to the theorem in [24] (page 262), its angular variation is preserved in the spatial domain while the radial function is the Hankel transform of the radial part in the frequency domain. Therefore, the spatial domain representation of an nth order spherical harmonic h_n reads

$$h_n(r,\beta) = c(I_3)^n h_n(r) h_n(\beta) = c(I_3)^n h_n(r)[\cos(n\beta) + \sin(n\beta)\mathbf{e}_{12}] \ , \tag{5.8}$$

where c indicates a constant, r and β are the polar coordinates in the spatial domain, $h_n(r)$ represents the radial part which is obtained by the nth order Hankel transform of $H_n(\rho)$.

In practice, 2D spherical harmonics are normally considered only as angular parts which should be combined with radial bandpass filters. In this thesis, the difference of Poisson (DOP) kernel [31] is employed as the radial bandpass filter. As a result, local signal analysis can be realized in a multi-scale approach in the monogenic scale-space [38]. The DOP is an isotropic bandpass filter which in spectral domain takes the form

$$H_{DOP}(\rho; s) = \exp(-2\pi\rho s_1) - \exp(-2\pi\rho s_2) \ , \tag{5.9}$$

where s_1 and s_2 represent the fine and coarse scales parameters, respectively. Therefore, an nth order bandpass bounded spherical harmonic reads

$$H_n(\rho, \alpha; s) = H_{DOP}(\rho; s) H_n(\alpha) \ . \tag{5.10}$$

Hence, $H_n(\rho, \alpha; s)$ is separable into radial variation $H_{DOP}(\rho; s)$ and angular variation $H_n(\alpha)$. Since the angular portion is harmonic, according to the theorem stated in [24], $H_n(\rho, \alpha; s)$ is separable both in the spatial and spectral domains. Thereby, we are able to obtain the spatial representation of the bandpass bounded spherical harmonic as follows

$$\mathcal{F}^{-1}\{H_n(\rho, \alpha; s)\} = c(I_3)^n h_n(r; s)(\cos(n\beta) + \sin(n\beta)\mathbf{e}_{12}) \ . \tag{5.11}$$

The angular variation is preserved and the radial function $h_n(r; s)$ is the Hankel transform of $H_{DOP}(\rho; s)$. Combined with the DOP bandpass filters, spherical harmonics of order 1 to 3 in the spatial domain are illustrated in Fig. 5.1, where the fine scale is one and the coarse scale takes two.

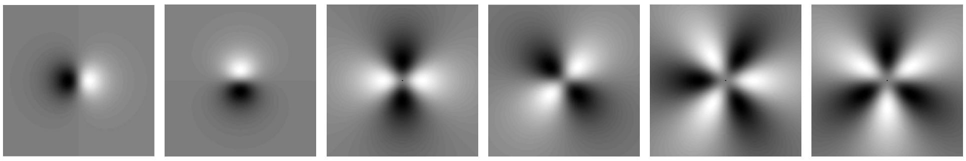

Fig. 5.1: From left to right are 2D spherical harmonic bandpasses of order 1 to 3 in the spatial domain, every spherical harmonic consists of two orthogonal components. White indicates positive one and black represents negative one.

It is obvious that the amplitude of the spherical harmonic equals always one, that is

$$|H_n^2(\alpha)| = \sqrt{(\cos(n\alpha))^2 + (\sin(n\alpha))^2} = 1 \ . \tag{5.12}$$

If the coordinate system rotates with the angle θ_0, spherical harmonics will rotate accordingly as

$$\begin{aligned} H_n'(\alpha) = {} & (\cos(n\theta_0)\cos(n\alpha) + \sin(n\theta_0)\sin(n\alpha)) + \\ & (-\sin(n\theta_0)\cos(n\alpha) + \cos(n\theta_0)\sin(n\alpha))\mathbf{e}_{12} \ . \end{aligned} \tag{5.13}$$

The amplitude of the rotated spherical harmonic is

$$\begin{aligned} |H_n'^2(\alpha)| = {} & [(\cos(n\theta_0)\cos(n\alpha) + \sin(n\theta_0)\sin(n\alpha))^2 + \\ & (-\sin(n\theta_0)\cos(n\alpha) + \cos(n\theta_0)\sin(n\alpha))^2]^{\frac{1}{2}} = 1 \ . \end{aligned} \tag{5.14}$$

This results from the fact that no matter which angle the spherical harmonic rotates, its amplitude is always one. Hence, the amplitude of any spherical harmonic is independent of the angular argument. Therefore, using spherical harmonics as basis functions gives access to a rotationally invariant signal representation.

5.2 Signal Modeling for Two-dimensional Image Structures

So far, we understood that the monogenic signal is derived from the monogenic extension of a scalar field. However, it is restricted to model only the i1D signals because only a minimum of information, i.e. the scalar value $f(x)$, is taken into consideration. If 2D images are interpreted as surfaces in $\mathbb{R}^3$, the first and second order fundamental theorems of differential geometry would deliver the most general local signal model in the classic framework. We will associate a curvature tensor instead of a scalar value to a location of interest. This results in a useful signal model for 2D image structures. For the moment we are neglecting the metric tensor. Hence, our operators for local signal analysis will be rotation invariant.

The proposed model can be regarded as the monogenic extension of the curvature tensor. Motivated from the differential geometry, this curvature tensor can be constructed. Therefore, a brief introduction to the differential geometry is given.

5.2.1 Basic Concepts of Differential Geometry

Differential geometry [22] is a well known methodology in the disciplines like physics, mechanical engineering and topography. Classical differential geometry deals with the mathematical description of curves and surfaces. In image processing field, Koenderink and van Doorn [63, 64] have introduced methods from differential geometry to analyze the local properties of signals. In such case, two dimensional intensity data can be represented as surfaces in 3D Euclidean space. Such surfaces in geometrical terms can be written as Monge patches. In the classic vector algebra, a Monge patch is a patch $\mathbf{f} : U \to \mathbb{R}^3$ of the form

$$\mathbf{f}(x, y) = (x, y, f(x, y)) \ , \tag{5.15}$$

where U is an open set in $\mathbb{R}^2$ and $f : U \to \mathbb{R}$ is a differentiable function. This notation of a surface is the most simple one for considerations in the framework of differential geometry because it enables to express the mentioned tensors with entries built from first and second order differentials of the image functions.

The second fundamental form of a Monge patch is given by

$$C = \frac{\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix}}{\sqrt{1 + f_x^2 + f_y^2}} = \frac{H}{\sqrt{1 + f_x^2 + f_y^2}} \ , \tag{5.16}$$

where H represents the Hessian matrix. For a Monge patch, the Gaussian curvature K and mean curvature M are obtained as the following

$$K = \frac{f_{xx} f_{yy} - f_{xy}^2}{(1 + f_x^2 + f_y^2)^2} = k_1 k_2 \tag{5.17}$$

$$M = \frac{(1 + f_y^2) f_{xx} - 2 f_x f_y f_{xy} + (1 + f_x^2) f_{yy}}{2(1 + f_x^2 + f_y^2)^{3/2}} = \frac{k_1 + k_2}{2} \ , \tag{5.18}$$

where k_1 and k_2 refer to the principal curvatures.

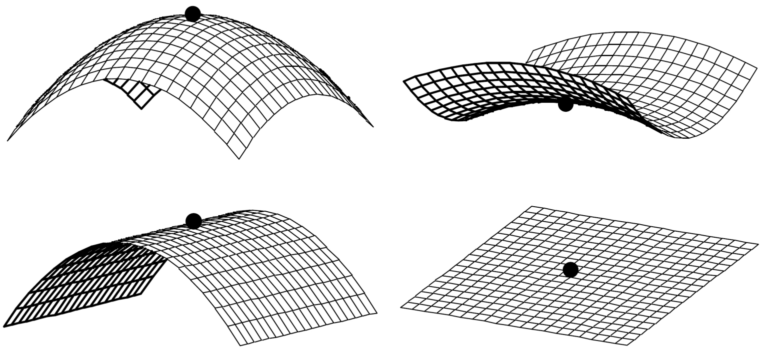

Fig. 5.2: Top row: from left to right are the elliptic surface ($K > 0$) and the hyperbolic surface ($K < 0$). Bottom row: from left to right are the parabolic surface ($K = 0$, $M \neq 0$) and the plane surface ($K = 0$, $M = 0$).

Because the determinant of the Hessian matrix equals the multiplication of its two eigenvalues, and the sum of these two eigenvalues is twice of the trace of the Hessian matrix, the Gaussian curvature and mean curvature can also be approximated as

$$\begin{aligned} K &\sim \det(H) && (5.19) \\ M &\sim \frac{\mathrm{trace}(H)}{2} . && (5.20) \end{aligned}$$

According to the Gaussian and mean curvatures, surfaces can be classified as four types (elliptic, hyperbolic, parabolic, plane) as shown in Fig. 5.2.

In the following, we will introduce basic concepts of differential geometry and the general 2D signal model in an algebraic framework with more powerful geometric meanings than $\mathbb{R}_3$. Because we are interested in a tensor representation of the image signal, our model will thus be represented in the matrix geometric algebra $M(2, \mathbb{R}_3)$ which results from the tensor product $\mathbb{R}_3 \times \mathbb{R}_3$.

The matrix geometric algebra $M(2, \mathbb{R}_3)$, see [103], is the geometric algebra of 2×2 matrices with entities in $\mathbb{R}_3$. For example, a general element P in this matrix geometric algebra can be written as

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix} , \qquad (5.21)$$

where $P \in M(2, \mathbb{R}_3)$ and a, b, c, $d \in \mathbb{R}_3$. Addition and multiplication of matrices in $M(2, \mathbb{R}_3)$ is the usual matrix addition and multiplication. The trace of this representation is

$$\mathrm{trace}_M(P) = \mathrm{trace}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a + d \tag{5.22}$$

and the corresponding determinant reads

$$\det{}_M(P) = \det{}_M \begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\overline{a}d\overline{d} + b\overline{b}c\overline{c} - (\overline{a}b\overline{d}c + \overline{c}d\overline{b}a)\ , \tag{5.23}$$

where $\overline{a}$, $\overline{b}$, $\overline{c}$ and $\overline{d}$ are the conjugations of a, b, c and d, respectively.

In this new algebraic framework, the Monge patch is thus formulated as

$$\mathbf{x} + \mathbf{f}(\mathbf{x}) = x\mathbf{e}_1 + y\mathbf{e}_2 + f(x,y)\mathbf{e}_3\ . \tag{5.24}$$

The primary first order differential quantity for an image, represented by the vector field $\mathbf{f}$, is the gradient defined as

$$\nabla\mathbf{f} = \mathbf{e}_1\frac{\partial}{\partial x}f(x,y)\mathbf{e}_3 + \mathbf{e}_2\frac{\partial}{\partial y}f(x,y)\mathbf{e}_3 = f_x\mathbf{e}_{13} + f_y\mathbf{e}_{23}\ . \tag{5.25}$$

For the second order geometry, the matrix of second order derivatives or Hessian H is given by

$$H_M = \begin{bmatrix} \frac{\partial^2}{\partial x^2}\mathbf{f} & \frac{\partial^2}{\partial x \partial y}\mathbf{f} \\ \frac{\partial^2}{\partial y \partial x}\mathbf{f} & \frac{\partial^2}{\partial y^2}\mathbf{f} \end{bmatrix} = \begin{bmatrix} \mathbf{e}_1\frac{\partial}{\partial x}f_x\mathbf{e}_{13} & \mathbf{e}_2\frac{\partial}{\partial y}f_x\mathbf{e}_{13} \\ \mathbf{e}_1\frac{\partial}{\partial x}f_y\mathbf{e}_{23} & \mathbf{e}_2\frac{\partial}{\partial y}f_y\mathbf{e}_{23} \end{bmatrix} = \begin{bmatrix} f_{xx}\mathbf{e}_3 & -f_{xy}\mathbf{e}_{123} \\ f_{xy}\mathbf{e}_{123} & f_{yy}\mathbf{e}_3 \end{bmatrix}\ . \tag{5.26}$$

The Hessian matrix is related to the curvature tensor, which describes the local deviation of the signal $\mathbf{f}$ from the tangent plane of the surface.

According to the derivative theorem of Fourier theory [87, 11], in the spectral domain, the second order derivative of $\mathbf{f}$ with respect to the x axis is given by

$$\mathcal{F}\{f_{xx}\mathbf{e}_3\} = -4\pi^2\rho^2\cos^2(\alpha)\mathbf{F} = -4\pi^2\rho^2\frac{1+\cos(2\alpha)}{2}\mathbf{F}\ , \tag{5.27}$$

where $\mathbf{F}$ denotes the Fourier transform of the original signal $\mathbf{f} = f(x,y)\mathbf{e}_3$. Analogously, the other second order derivatives are obtained as

$$\mathcal{F}\{f_{xy}\mathbf{e}_{123}\} = -4\pi^2\rho^2\cos(\alpha)\sin(\alpha)\mathbf{e}_{12}\mathbf{F} = -4\pi^2\rho^2\frac{\sin(2\alpha)}{2}\mathbf{e}_{12}\mathbf{F} \tag{5.28}$$

$$\mathcal{F}\{f_{yy}\mathbf{e}_3\} = -4\pi^2\rho^2\sin^2(\alpha)\mathbf{F} = -4\pi^2\rho^2\frac{1-\cos(2\alpha)}{2}\mathbf{F}\ . \tag{5.29}$$

Hence, in the spectral domain, the Hessian matrix reads

$$\mathcal{F}\{H_M\} = \begin{bmatrix} -4\pi^2\rho^2\frac{1+\cos(2\alpha)}{2}\mathbf{F} & (4\pi^2\rho^2\frac{\sin(2\alpha)}{2}\mathbf{F})\mathbf{e}_{12} \\ (-4\pi^2\rho^2\frac{\sin(2\alpha)}{2}\mathbf{F})\mathbf{e}_{12} & -4\pi^2\rho^2\frac{1-\cos(2\alpha)}{2}\mathbf{F} \end{bmatrix} . \tag{5.30}$$

According to Eq. (5.23), the determinant of the Hessian in this algebraic framework is obtained as

$$\det_M(H_M) = [(f_{xx}\mathbf{e}_3)(f_{yy}\mathbf{e}_3) - (-f_{xy}\mathbf{e}_{123})(f_{xy}\mathbf{e}_{123})]^2 = \left[f_{xx}f_{yy} - f_{xy}^2\right]^2 . \tag{5.31}$$

This determinant is identical to the square of the classic Hessian determinant in the vector algebra. In general, the determinant computation of the algebra $M(2, \mathbb{R}_3)$ can not be reduced to the original definition in the vector algebra. However, in the current case, due to the particular basis elements, the original definition of the determinant can still be used, see Appendix for details. Hence, the determinant of the Hessian is reformulated as

$$\det_R(H_M) = (f_{xx}\mathbf{e}_3)(f_{yy}\mathbf{e}_3) - (-f_{xy}\mathbf{e}_{123})(f_{xy}\mathbf{e}_{123}) = f_{xx}f_{yy} - f_{xy}^2 = \lambda_1\lambda_2 , \tag{5.32}$$

where λ_1 and λ_2 are two eigenvalues of the real valued Hessian matrix, which represent the principal curvatures. Thus, the Gaussian curvature K can be approximated as

$$K \sim \det_R(H_M) \tag{5.33}$$

or

$$K^2 \sim \det_M(H_M) . \tag{5.34}$$

The mean curvature M, obtained from the trace of the algebraically embedded Hessian matrix, takes the following form

$$M \sim \frac{1}{2}\mathrm{trace}_M(H_M) = \frac{1}{2}(f_{xx}\mathbf{e}_3 + f_{yy}\mathbf{e}_3) = \frac{1}{2}(\lambda_1 + \lambda_2)\mathbf{e}_3 . \tag{5.35}$$

Hence, both the Gaussian curvature and the mean curvature give rise to a rotation invariant local analysis of second order features. Combining the Gaussian curvature and mean curvature, a complete classification of the local structure into the types i2D (elliptic and hyperbolic regions), i1D (parabolic region) and i0D (planar region) in principle can be done, see Tab. 5.1.

5.2.2 Monogenic Extension of the Curvature Tensor

In order to build a general model for 2D structures with phase information contained, we follow the ideas of deriving the analytic or monogenic signal

Tab. 5.1: Surface type classification based on Gaussian curvature K and mean curvature M

Surface type	M	K
Elliptic (i2D)		$K > 0$
Hyperbolic (i2D)		$K < 0$
Parabolic (i1D)	$\|M\| \neq 0$	$K = 0$
Planar (i0D)	$\|M\| = 0$	$K = 0$

from a real valued 1D or 2D image signal. The holomorphic (1D) or monogenic (2D) completion of the signal results in an additional component which is in quadrature phase relation to the original signal. For a 2D image, every image point is now associated with a curvature tensor which is related to the Hessian matrix. It is necessary to find a conjugate matrix with quadrature phase relationship to the curvature tensor. In the following, we will introduce the general signal model based on 2D spherical harmonics.

As analyzed before, the Hessian matrix contains curvature information. Based on it, i0D, i1D and i2D structures can be easily separated. It is obvious that angular parts of the derivatives are related to spherical harmonics of even orders 0 and 2, see equations (5.27) to (5.30). These harmonics represent the even information of 2D structures. Therefore, we are motivated to construct a tensor T_e, which is related to the Hessian matrix, for the signal modeling. We will call T_e as the curvature tensor, although it is different to the curvature tensor of the second fundamental form of the differential geometry. This curvature tensor indicates the even information of 2D structures and is obtained from a tensor-valued filter H_e in the frequency domain, i.e. $T_e = \mathcal{F}^{-1}\{H_e \times_\tau \mathbf{F}\}$, where $\mathcal{F}^{-1}$ means the inverse Fourier transform and $\times_\tau$ indicates the geometric product between all elements of H_e and $\mathbf{F}$. Since the original 2D signal $f(x, y)$ is embedded as an $\mathbf{e}_3$-valued signal, the tensor-valued filter H_e, called the even Hessian operator, thus takes the following form

$$\begin{aligned} H_e &= \frac{1}{2}\begin{bmatrix} H_0 + \langle H_2\rangle_0 & -\langle H_2\rangle_2 \\ \langle H_2\rangle_2 & H_0 - \langle H_2\rangle_0 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} 1+\cos(2\alpha) & -\sin(2\alpha)\mathbf{e}_{12} \\ \sin(2\alpha)\mathbf{e}_{12} & 1-\cos(2\alpha) \end{bmatrix} = \begin{bmatrix} \cos^2(\alpha) & -\frac{1}{2}\sin(2\alpha)\mathbf{e}_{12} \\ \frac{1}{2}\sin(2\alpha)\mathbf{e}_{12} & \sin^2(\alpha) \end{bmatrix}. \end{aligned} \tag{5.36}$$

The entities of H_e are obtained from Eq. (5.30). For the convenience of analysis, the radial factors are ignored.

In this filter, the two elements $\cos^2(\alpha)$ and $\sin^2(\alpha)$ can be considered as two angular windowing functions which are the same as those of the orientation tensor in [47]. From them, two perpendicular i1D components of the 2D

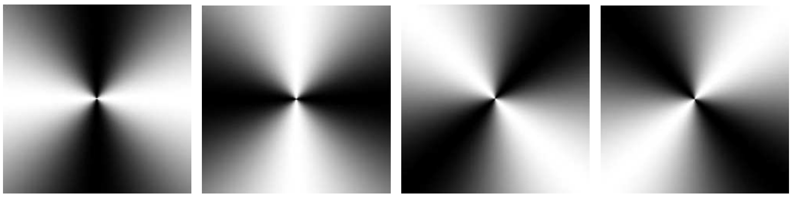

Fig. 5.3: From left to right are the angular windowing functions of $\cos^2(\alpha)$, $\sin^2(\alpha)$, $\cos^2(\alpha - \frac{\pi}{4})$ and $\sin^2(\alpha - \frac{\pi}{4})$. White indicates positive one and black represents zero.

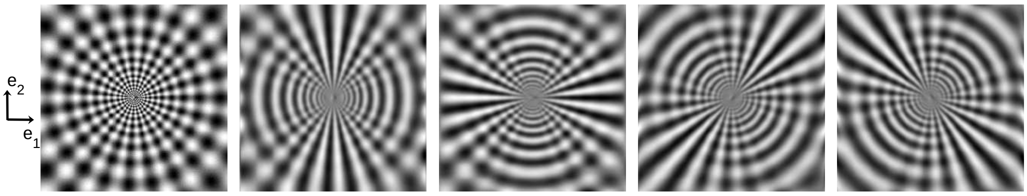

Fig. 5.4: From left to right are the test image, i1D signals result from the angular windowing functions of $\cos^2(\alpha)$, $\sin^2(\alpha)$, $\cos^2(\alpha - \frac{\pi}{4})$ and $\sin^2(\alpha - \frac{\pi}{4})$, respectively.

image, oriented along the $\mathbf{e}_1$ and $\mathbf{e}_2$ coordinates, can be obtained. The other component of the filter is also the combination of two angular windowing functions, i.e. $\frac{1}{2}\sin(2\alpha) = \frac{1}{2}(\cos^2(\alpha - \frac{\pi}{4}) - \sin^2(\alpha - \frac{\pi}{4}))$. These two angular windowing functions yield again two i1D components of the 2D image, which are oriented along the diagonals of the plane spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$. These four angular windowing functions, shown in Fig. 5.3, can also be considered as four differently oriented filters, which are basis functions to steer a detector for i1D structures [44]. They make sure that i1D components along different orientations are extracted, see Fig. 5.4. Consequently, the even Hessian operator H_e enables the extraction of four differently oriented i1D components of the 2D image. Hence, the superimpose of these four basis i1D signals will result in a rotation-invariant extraction of any arbitrary but even symmetric i1D signal.

The 2D quadrature filter, which was proposed in Chapter 4, employed also 2D spherical harmonics as basis functions. However, in that framework, the angular windowing functions constructed from these spherical harmonics are tightly related with the orientations of the two i1D signals which superimpose the i2D pattern. This means that orientations of these two i1D signals should first be estimated and then be used to steer two angular windowing functions for extracting the corresponding two i1D signals. In contrast to this, the curvature tensor consists of four entities, which can be regarded as angular windowing functions. They are tightly attached with the coordinate system. In this way, rotationally invariant extraction of the i2D pattern is easy to be implemented without steering.

The Riesz transform [36] is able to evaluate the corresponding conjugate information of the i1D signal, which is in quadrature phase relation with the i1D signal. Therefore, the odd representation of the curvature tensor, called the conjugate curvature tensor T_o, is obtained by employing the first order spherical harmonic h_1 to elements of T_e, which equals the Riesz transform of the curvature tensor T_e. Besides, the conjugate curvature tensor T_o results also from a tensor-valued filter H_o, called the odd Hessian operator.

$$T_o = h_1 *_\tau T_e = \mathcal{F}^{-1}\left\{H_o \times_\tau \mathbf{F}\right\} \ , \tag{5.37}$$

where $*_\tau$ represents the convolution of all elements of T_e with h_1. The odd Hessian operator H_o equals the Riesz transform of the even Hessian operator, i.e. $H_o = H_1 \times_\tau H_e$. In the spectral domain, the odd Hessian operator thus takes the following form

$$H_o \equiv \begin{bmatrix} H_{o11} & H_{o12} \\ H_{o21} & H_{o22} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} H_1(H_0 + \langle H_2\rangle_0) & H_1(-\langle H_2\rangle_2) \\ H_1(\langle H_2\rangle_2) & H_1(H_0 - \langle H_2\rangle_0) \end{bmatrix} \tag{5.38}$$

with

$$\begin{aligned} H_{o11} &= (\cos(\alpha) + \sin(\alpha)\mathbf{e}_{12})(\cos^2(\alpha)) && (5.39)\\ &= \frac{1}{4}\left[(3\cos(\alpha) + \cos(3\alpha)) + (\sin(\alpha) + \sin(3\alpha))\mathbf{e}_{12}\right] \\ H_{o21} &= -H_{o12} = (\cos(\alpha) + \sin(\alpha)\mathbf{e}_{12})(\frac{1}{2}\sin(2\alpha)\mathbf{e}_{12}) && (5.40)\\ &= \frac{1}{4}\left[(\cos(\alpha) - \cos(3\alpha)) + (\sin(\alpha) + \sin(3\alpha))\mathbf{e}_{12}\right] \\ H_{o22} &= (\cos(\alpha) + \sin(\alpha)\mathbf{e}_{12})(\sin^2(\alpha)) && (5.41)\\ &= \frac{1}{4}\left[(\cos(\alpha) - \cos(3\alpha)) + (3\sin(\alpha) - \sin(3\alpha))\mathbf{e}_{12}\right] \ . \end{aligned}$$

It is obvious that this tensor-valued filter consists of odd order spherical harmonics. Hence, the Riesz transform of the curvature tensor T_e gives its corresponding odd representation T_o. Combing the curvature tensor and its conjugate representation forms the general signal model of local 2D image structures,

$$T(\mathbf{x}) = T_e(\mathbf{x}) + T_o(\mathbf{x}) \ . \tag{5.42}$$

This signal model can also be regarded as the monogenic extension of the curvature tensor. Hence, it is called the monogenic curvature tensor.

In fact, the monogenic curvature tensor is an element of $M(2,\mathbb{R}_3)$ with monogenic entries. This representation is much more powerful than the monogenic signal. There for each pixel a real signal is extended to a Clifford valued signal. In our case, a Clifford valued tensor representation is gained with quadrature relations in each element of the tensor. There are different ways of evaluating the monogenic curvature tensor. One way is to evaluate the quadrature relationship of the tensor pair (T_e, T_o) separately, the other way is to evaluate the quadrature relation of the elements of $T = T_e + T_o$. We will present in this thesis the first way.

5.2.3 Local Representations for I1D and I2D Image Structures

Analogous with the real valued differential geometry approach, in our Clifford valued approach, 2D image structures can be classified by computing the Gaussian curvature and the mean curvature of the tensor pair T_e and T_o, see also Tab. 5.1. Since the non-zero Gaussian curvature indicates the existence of i2D structure, the even and odd parts of i2D structures are correspondingly obtained according to Eq. (5.33). The even part of i2D structures reads

$$\begin{aligned}
\mathbf{d}_e &= \det_R(T_e)\mathbf{e}_3 = (T_{e11}T_{e22} - T_{e12}T_{e21})\mathbf{e}_3 \qquad (5.43)\\
&= \mathcal{F}^{-1}\left\{\left[\left(\frac{H_0 + \langle H_2\rangle_0}{2}\mathbf{F}\right) * \left(\frac{H_0 - \langle H_2\rangle_0}{2}\mathbf{F}\right)\right.\right.\\
&\quad \left.\left. - \left(\frac{\langle H_2\rangle_2}{2}\mathbf{F}\right) * \left(\frac{-\langle H_2\rangle_2}{2}\mathbf{F}\right)\right]\mathbf{e}_3\right\}\\
&= \mathcal{F}^{-1}\left\{\left[\left(\frac{1 + \cos(2\alpha)}{2}\mathbf{F}\right) * \left(\frac{1 - \cos(2\alpha)}{2}\mathbf{F}\right)\right.\right.\\
&\quad \left.\left. - \left(\frac{\sin(2\alpha)}{2}\mathbf{F}\right) * \left(\frac{-\sin(2\alpha)}{2}\mathbf{F}\right)\right]\mathbf{e}_3\right\}\\
&= A\mathbf{e}_3\ ,
\end{aligned}$$

where T_{eij}, $i,j = 1,2$ are the corresponding components of the curvature tensor T_e. Because $\det_R(T_e)$ is scalar valued, similar as the monogenic signal, the even part of i2D structures is embedded as the $\mathbf{e}_3$ component in the 3D

Euclidean space. The odd part of i2D structures is given by

$$\begin{aligned}
\mathbf{d}_o &= \mathbf{e}_1 \mathrm{det}_R(T_o) = \mathbf{e}_1(T_{o11}T_{o22} - T_{o12}T_{o21}) \qquad (5.44)\\
&= \mathcal{F}^{-1}\left\{\mathbf{e}_1\left[\left(\frac{H_1(H_0 + \langle H_2\rangle_0)}{2}\mathbf{F}\right) * \left(\frac{H_1(H_0 - \langle H_2\rangle_0)}{2}\mathbf{F}\right)\right.\right.\\
&\quad \left.\left. - \left(\frac{H_1(\langle H_2\rangle_2)}{2}\mathbf{F}\right) * \left(\frac{-H_1(\langle H_2\rangle_2)}{2}\mathbf{F}\right)\right]\right\}\\
&= B\mathbf{e}_1 + C\mathbf{e}_2 \ ,
\end{aligned}$$

where T_{oij}, $i, j = 1, 2$ are the corresponding components of the conjugate curvature tensor T_o. Because $\mathrm{det}_R(T_o)$ is spinor valued, i.e. $\mathrm{det}_R(T_o) \in span\{1, \mathbf{e}_{12}\}$, by multiplying the $\mathbf{e}_1$ basis vector from the left, $\mathbf{d}_o$ takes a vector valued representation. A local representation for i2D structures is obtained by combining the even and odd parts of i2D structures. This local representation for i2D structures is called the generalized monogenic curvature signal and it takes the following form

$$\mathbf{f}_{i2D} = \mathbf{d}_e + \mathbf{d}_o = A\mathbf{e}_3 + B\mathbf{e}_1 + C\mathbf{e}_2 \ . \qquad (5.45)$$

The original signal $\mathbf{f}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^2$ is thus mapped to $\mathbf{f}_{i2D}(\mathbf{x}), \mathbf{x} \in \mathbb{R}^3$ as a local representation of i2D signals. The generalized monogenic curvature signal can be considered as a monogenic representation of the Gaussian curvature in real valued differential geometry.

The parabolic and planar surface patches, corresponding to i1D and i0D structures, have zero Gaussian curvatures. In order to separate them with each other, the trace of the tensor pair T_e and T_o is computed. Non-zero trace illustrates the existence of the i1D structure. Therefore, the combination of traces of T_e and T_o can be considered as the local representation of i1D structures. According to the combination and certain embedding, this representation is obtained as

$$\begin{aligned}
\mathbf{f}_{i1D} &= \mathrm{trace}_M(T_e) + \mathrm{trace}_M(T_o)\mathbf{e}_2 \qquad (5.46)\\
&= \mathcal{F}^{-1}\left\{\mathrm{trace}_M(H_e \times_\tau \mathbf{F}) + \mathrm{trace}_M(H_o \times_\tau \mathbf{F})\mathbf{e}_2\right\}\\
&= \mathcal{F}^{-1}\left\{[\mathrm{trace}_M(H_e) + (-\mathbf{e}_2)\mathrm{trace}_M(H_o)]\mathbf{F}\right\}
\end{aligned}$$

with

$$\begin{aligned}
\mathrm{trace}_M(H_e) + (-\mathbf{e}_2)\mathrm{trace}_M(H_o) &= 1 + (-\mathbf{e}_2)H_1 \qquad (5.47)\\
&= 1 + (-\mathbf{e}_2)(\cos(\alpha) + \sin(\alpha)\mathbf{e}_{12})\\
&= 1 + \cos(\alpha)(-\mathbf{e}_2) + \sin(\alpha)\mathbf{e}_1 = 1 + H_R \ ,
\end{aligned}$$

where H_R refers to the Riesz kernel in the spectral domain, see Eq. (5.2). Plugging Eq. (5.47) into Eq. (5.46), we will get the i1D structure representation as

$$\begin{aligned}\mathbf{f}_{i1D} &= \mathcal{F}^{-1}\left\{[\mathrm{trace}_M(H_e) + (-\mathbf{e}_2)\mathrm{trace}_M(H_o)]\mathbf{F}\right\} \\ &= \mathcal{F}^{-1}\left\{(1 + H_R)\mathbf{F}\right\} = \mathbf{f} + h_R * \mathbf{f} = f_3(\mathbf{x})\mathbf{e}_3 + f_1(\mathbf{x})\mathbf{e}_1 + f_2(\mathbf{x})\mathbf{e}_2\ ,\end{aligned} \tag{5.48}$$

where h_R is the spatial representation of the Riesz kernel and $\mathbf{f}$ refers to the $\mathbf{e}_3$ valued original signal. The obtained Eq. (5.48) is identical to Eq. (3.78) and Eq. (3.81), which indicates that the local representation for i1D structures, obtained from the general signal model, is the combination of the original signal and its Riesz transform. This means that the derived i1D structure representation is just the monogenic signal as proposed in [36]. Hence, the proposed signal model includes the monogenic signal as a special case, which is a monogenic measure of the mean curvature. In addition to this, it constitutes also the local representation for i2D structure. Therefore, the signal model of the monogenic curvature tensor is a general representation for 2D structures of any intrinsic dimension.

The current analysis of this novel model is based on only the combined traces and determinants of two separate tensors. It is very promising that more local features can be extracted via the component-wise eigensystem analysis of the monogenic curvature tensor.

5.2.4 Comparison with Some Related Work

The monogenic curvature tensor is a novel model for 2D image structures. However, there exist also some tensor-based approaches in the literature for image analysis. In this section, we will compare the proposed new method with some related work. For better understanding, all the formulations are presented in the classic vector algebra framework.

The structure tensor [43] is a well-known approach for local image analysis. It is constructed as the averaged outer product of the spatial gradient with itself. From it, local energy and orientation of the image structure can be easily obtained. Unfortunately, the structure tensor is not phase invariant since it reacts differently to edges and lines. Besides, it contains no phase information of local structures. A different approach, called the orientation tensor, was proposed in [47], see also Chapter 3 for a brief introduction. The authors are interested in i1D image structures, which are called simple signals. In the case of i2D structures, it is still unclear exactly how this tensor will behave. In contrast to the structure tensor, this method has the property of phase invariance. But it still lacks the capability of phase evaluation.

In [30], an extension of the orientation tensor was presented for non-simple signal orientation estimation, which takes the following form

$$T = AA^T + \gamma bb^T \ , \tag{5.49}$$

where A is a symmetric matrix which captures information about the even part of the signal (excluding the DC component), b is a vector which extracts information about the odd part of the signal. Details of the construction can be found in [30]. Possibilities for the construction include local polynomial fits, facet models, moment filters and Gaussian derivative filters. The choice of γ is problematic, which means it is impossible to find a single value that works well on the entire image. Consequently, no matter how A and b are estimated, this tensor is only phase invariant for a single frequency determined by γ. Same as the structure tensor, it also delivers no phase information.

The boundary tensor introduced in [65] integrates i1D and i2D structures detection into one framework. It is built based on a new generalization of quadrature filters to 2D using the Riesz transform. This tensor turns out to be structurally equivalent to the polynomial-based orientation tensor definition, but with a uniquely determined parameter $\gamma = 1$. It has the property of phase invariance for all frequencies in the same way as that of the quadrature filter.

Similar to the monogenic curvature tensor, 2D spherical harmonics are also employed as basis functions to construct the boundary tensor. They are combined with bandpass filters to form polar separable filters. However, the difference of Gaussian (DOG) is used instead of the difference of Poisson (DOP), which results in the loss of harmonic conjugate information. Besides, only up to second order 2D spherical harmonics are considered.

The polar separable filters used in this approach in the Fourier domain read

$$K(\rho, \varphi) = K_\varphi(\varphi)K_\rho(\rho) \ , \tag{5.50}$$

where ρ, φ denote the polar coordinates in the frequency domain, $K_\varphi(\varphi)$ are the angular parts which have the forms of $\cos(n\varphi)$ or $\sin(n\varphi)$ with n being the order of the 2D spherical harmonic, $K_\rho(\rho)$ indicates the DOG. In the spatial domain, the polar separable filters are given by

$$\begin{aligned} h_1^{(n)} &= \mathcal{F}^{-1}\left\{\cos(n\varphi)K_\rho(\rho)\right\} = \frac{i^n}{4\pi^2}\cos(n\theta)k_r^{(n)}(r) && (5.51) \\ h_2^{(n)} &= \mathcal{F}^{-1}\left\{\sin(n\varphi)K_\rho(\rho)\right\} = \frac{i^n}{4\pi^2}\sin(n\theta)k_r^{(n)}(r) \ , && (5.52) \end{aligned}$$

where r and θ are polar coordinates in the spatial domain. The radial functions $k_r^{(n)}$ are obtained by the nth order Hankel transform of $K_\rho(\rho)$. The filter response from these polar separable filters thus reads

$$C^{(n)} = \begin{bmatrix} C_1^{(n)} \\ C_2^{(n)} \end{bmatrix} = \begin{bmatrix} f * h_1^{(n)} \\ f * h_2^{(n)} \end{bmatrix} , \tag{5.53}$$

where f refers to the image signal. Based on this, the boundary tensor is constructed as

$$T = T_{\text{even}} + T_{\text{odd}} = AA^T + bb^T \tag{5.54}$$

with

$$A = \begin{bmatrix} C^{(0)} + C_1^{(2)} & C_2^{(2)} \\ C_2^{(2)} & C^{(0)} - C_1^{(2)} \end{bmatrix} \tag{5.55}$$

$$b = \begin{bmatrix} C_1^{(1)} \\ C_2^{(1)} \end{bmatrix} . \tag{5.56}$$

The structure of the boundary tensor looks very similar to the monogenic curvature tensor. They all consist of even and odd tensors. Looking into details, A can also be obtained from a tensor-valued filter H_A. Ignoring the radial parts results in a representation of H_A as the following

$$H_A = \begin{bmatrix} 1 + \cos(2\varphi) & \sin(2\varphi) \\ \sin(2\varphi) & 1 - \cos(2\varphi) \end{bmatrix} . \tag{5.57}$$

This is exactly two times the outer product of the Riesz kernel. It is obvious that H_A is the even Hessian operator (up to a factor), which is used in the monogenic curvature tensor. In the boundary tensor, A is called the generalized Hessian. But in our framework, this is named as the curvature tensor since it contains rich curvature information. In contrast to the monogenic curvature tensor, the even part of the boundary tensor is given by AA^T. Elements of this even tensor thus represent signal energies instead of filter responses, which makes it difficult to explore the corresponding harmonic conjugate information.

The construction of the odd part has big difference between the boundary tensor and the monogenic curvature tensor. Only the first order spherical harmonic is used in the boundary tensor, but the odd part of the monogenic curvature tensor is built based on the combination of the first and third order spherical harmonics. This delivers the possibility of detecting local structures with higher angular resolution. In addition to this, the odd part bb^T of the boundary tensor contains only the energy information. In such case, how the phase information can be extracted is not explicit at all.

By computing the trace of the boundary tensor, the boundary energy is obtained, which contains the energy information of i1D and i2D structures. According to the eigenvalue decomposition, energies for these two kinds of structures are able to be distinguished. Hence, they are directly used for detecting i1D and i2D structures. Unfortunately, due to the lack of phase information, the detection can not be implemented by the means of phase congruency. As a comparison, computing the trace of the monogenic curvature tensor results in a totally different entity, that is the monogenic signal. Besides, the interpretation of the determinant of the boundary tensor is still not well investigated. On the contrary, the determinant of the monogenic curvature tensor results in the generalized monogenic curvature signal as a novel model for i2D image structures.

It was reported in [66] that the boundary tensor exhibits very similar behavior as the derivative-based gradient energy tensor [32, 35]. The energy tensor is able to extract the phase information of i1D structures, however, it still delivers no information about the i2D parts. Compared with the boundary tensor and the energy tensor, the monogenic curvature tensor contains much rich information. In addition to the local energy and orientation information, local phases of i1D and i2D structures can also be estimated, which is a remarkable advantage compared with these previously mentioned tensor-based approaches.

5.3 Interpretation of the Generalized Monogenic Curvature Signal

Because the monogenic signal has been discussed in [36] in detail, we are more interested in the interpretation of the i2D structure representation, i.e. the generalized monogenic curvature signal.

5.3.1 Geometric Model and Feature Extraction

In the light of the introduction in Section 5.1.1, local features of the generalized monogenic curvature signal can be defined using the logarithm of $\mathbb{R}_3^+$. The spinor field which maps the $\mathbf{e}_3$ basis vector to the generalized monogenic curvature signal $\mathbf{f}_{i2D}$ is given by $\mathbf{e}_3\mathbf{f}_{i2D}$. According to Eq. (5.4) and Eq. (5.5), the local amplitude $A(\mathbf{x})$ and local phase representation $\Phi(\mathbf{x})$ are obtained as

$$A(\mathbf{x}) = |\mathbf{f}_{i2D}(\mathbf{x})| = \exp(\log(|\mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x})|)) = \exp(\langle\log(\mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x}))\rangle_0) \quad (5.58)$$

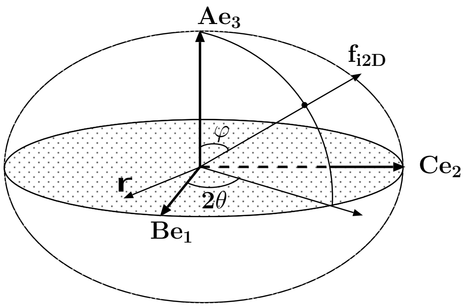

Fig. 5.5: The geometric model for the generalized monogenic curvature signal. Here, $A\mathbf{e}_3$ indicates the even information of the i2D structure, $B\mathbf{e}_1$ and $C\mathbf{e}_2$ are the two components of the odd part. The phase is represented by φ, 2θ denotes the main orientation in terms of double angle representation, $\mathbf{r}$ indicates the rotation vector.

$$\begin{aligned}\Phi(\mathbf{x}) &= \arg(\mathbf{f}_{i2D}(\mathbf{x})) = \langle \log(\mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x}))\rangle_2 \\ &= \frac{\langle \mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x})\rangle_2}{|\langle \mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x})\rangle_2|} \operatorname{atan}\left(\frac{|\langle \mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x})\rangle_2|}{\langle \mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x})\rangle_0}\right) ,\end{aligned} \tag{5.59}$$

where $\arg(\cdot)$ denotes the argument of the expression and $\operatorname{atan}(\cdot) \in [0, \pi)$. As the bivector part of the logarithm of the spinor field $\mathbf{e}_3\mathbf{f}_{i2D}$, this local phase representation describes a rotation from the $\mathbf{e}_3$ axis by a phase angle φ in the oriented complex plane spanned by $\mathbf{f}_{i2D}$ and $\mathbf{e}_3$, i.e. $\mathbf{f}_{i2D} \wedge \mathbf{e}_3$. The orientation of this phase plane indicates the local main orientation. Therefore, the local phase representation of the generalized monogenic curvature signal combines local phase and local orientation of i2D structures, just as in the case of the monogenic signal for i1D structures. Since the local phase representation $\Phi(\mathbf{x})$ is a bivector, its dual in $\mathbb{R}_3$ is a rotation vector that can be defined as

$$\mathbf{r}(\mathbf{x}) = (\Phi(\mathbf{x}))^* = \langle \log(\mathbf{e}_3\mathbf{f}_{i2D}(\mathbf{x}))\rangle_2^* \ . \tag{5.60}$$

The rotation vector $\mathbf{r}$ is orthogonal to the local orientation. The length of the rotation vector $|\mathbf{r}|$ indicates the phase angle φ of the i2D structure and its direction illustrates the rotation axis. According to the proposed algebraic embedding, a geometric model for the generalized monogenic curvature signal can be visualized as shown in Fig. 5.5. The geometric model is an ellipsoid, which looks very similar to that of the monogenic signal. However, each axis encodes totally different meaning. The even part of the i2D structure is encoded within the $\mathbf{e}_3$ axis, and the odd part is encoded within the $\mathbf{e}_1$ and $\mathbf{e}_2$ axes. The angle φ represents the phase and 2θ is the main orientation in a double angle representation form. The rotation vector $\mathbf{r}$ lies in the plane orthogonal to $\mathbf{e}_3$. Combining the local amplitude and local phase representation, the generalized monogenic curvature signal for i2D structures can be reconstructed as

$$\mathbf{f}_{i2D} = |\mathbf{f}_{i2D}| \exp\left(\arg\left(\mathbf{f}_{i2D}\right)\right) \ . \tag{5.61}$$

Having a definition for the i2D local phase, we recognize that the local phase representation contains additional geometric information, i.e. local orientation. Since local amplitude, local phase and local orientation are orthogonal to each other, the generalized monogenic curvature signal performs a split of identity.

From an alternative point of view, local amplitude, phase and orientation can also be obtained according to the relationship of the even and odd components in spherical coordinates. The local amplitude is computed as

$$|\mathbf{f}_{i2D}| = \sqrt{A^2 + B^2 + C^2} \ . \tag{5.62}$$

The local main orientation is given by

$$\theta = \frac{1}{2}\text{atan2}(B, C) \qquad \theta \in (-\frac{\pi}{2}, \frac{\pi}{2}] \ , \tag{5.63}$$

where $\text{atan2}(\cdot) \in (-\pi, \pi]$. And the local phase reads

$$\varphi = \text{atan2}\left(\text{sign}(B\mathbf{e}_1 + C\mathbf{e}_2)|B\mathbf{e}_1 + C\mathbf{e}_2|, A\right) \qquad \varphi \in (-\pi, \pi] \ . \tag{5.64}$$

Fig. 5.6 illustrates bandpassed local features extracted from the generalized monogenic curvature signal. A synthetic image which consists of a superposition of an angular and a radial modulation is used as the test image. The blobs in this test image are considered as i2D structures. The energy output of the generalized monogenic curvature signal, i.e. $\mathbf{d}_e^2 + \mathbf{d}_o^2$, can be regarded as i2D structure strength to detect points of interest. Besides, it also illustrates the rotation invariance property of the generalized monogenic curvature signal. The even and odd outputs also indicate the existence of i2D structures. Local main orientation of the generalized monogenic curvature signal denotes the main orientation of the i2D structure. Its minor orientation is simply perpendicular to the main orientation. Local phase contains the structure information. At the positions where the main orientation and the minor orientation wrap from zero to π, the estimated phase is inverted. This is called the orientation-phase wrapping. For another synthetic image, the corresponding local features are demonstrated in Fig. 5.7.

To show the difference between the monogenic signal and the generalized monogenic curvature signal, some experimental results are given in Fig. 5.8 and Fig. 5.9. Two patterns as the line and edge like intersections are

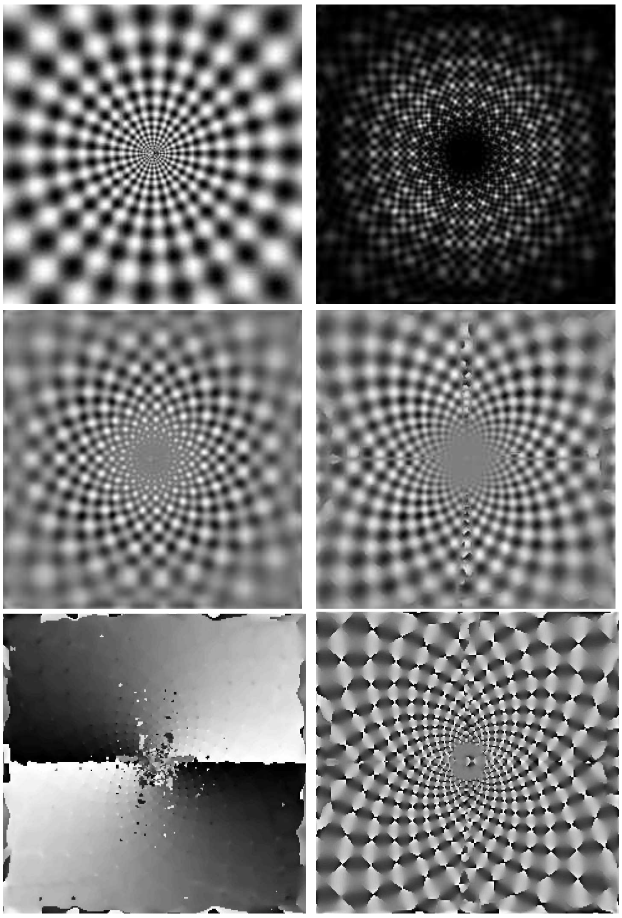

Fig. 5.6: Bandpassed local features. Top row: from left to right are the original image and the energy output of the generalized monogenic curvature signal. Middle row: even and odd parts of the generalized monogenic curvature signal. Bottom row: the estimated main orientation and phase.

employed as test images. Local energy output and phase information from the monogenic signal and the generalized monogenic curvature signal are extracted, respectively. These results indicate that the monogenic signal enables feature estimation of i1D signals, however, it delivers no information for the i2D part of the original signal. The generalized monogenic curvature signal in contrast gives access to local feature evaluation of i2D structures.

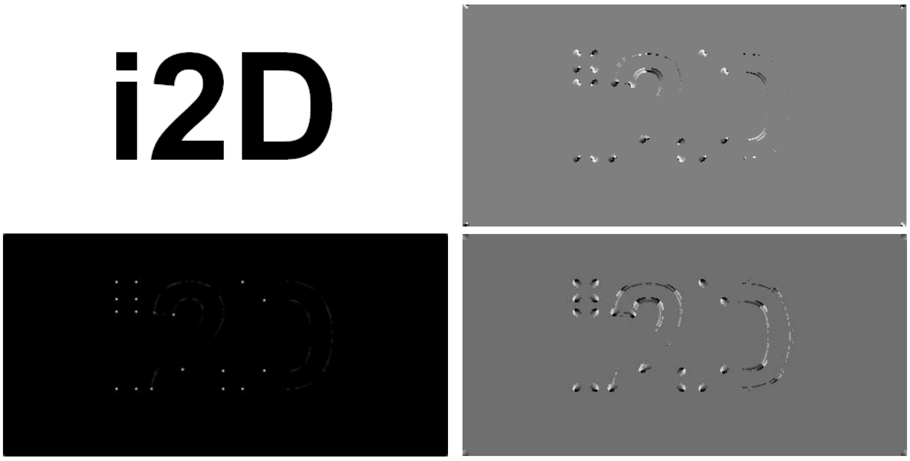

Fig. 5.7: Top row: the test image and its local main orientation estimation from the generalized monogenic curvature signal. Bottom row: the estimated energy and phase information.

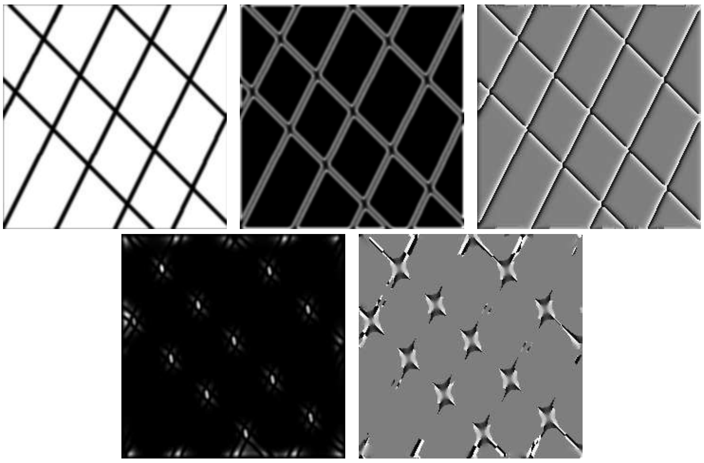

Fig. 5.8: Top row: test image, the energy of the monogenic signal and its phase. Bottom row: energy and phase of the generalized monogenic curvature signal.

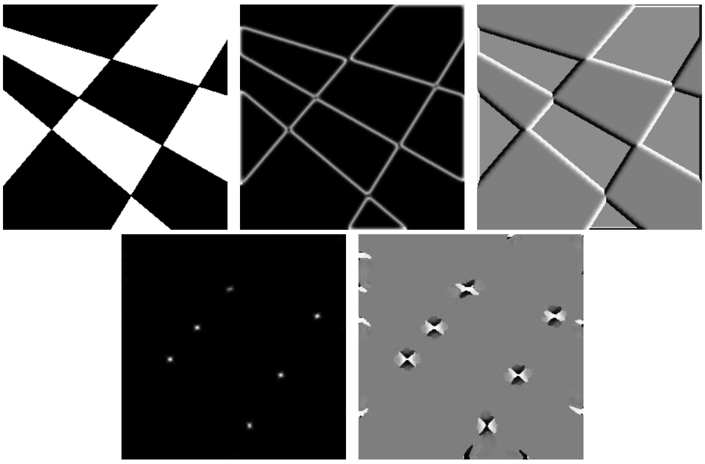

Fig. 5.9: Top row: test image, the energy of the monogenic signal and its phase. Bottom row: energy and phase of the generalized monogenic curvature signal.

5.3.2 Properties of the Generalized Monogenic Curvature Signal

According to Eq. (5.43), in the spectral domain, the even part of the generalized monogenic curvature signal $\mathbf{d}_e$ can be interpreted as

$$\begin{aligned}\mathcal{F}\{\mathbf{d}_e\} &= \mathcal{F}\{\det_R(T_e)\mathbf{e}_3\} \\ &= ((H_{e11}(\alpha)H_{DOP}(\rho;s)\mathbf{F}) * (H_{e22}(\alpha)H_{DOP}(\rho;s)\mathbf{F}) \\ &\quad -(H_{e12}(\alpha)H_{DOP}(\rho;s)\mathbf{F}) * (H_{e21}(\alpha)H_{DOP}(\rho;s)\mathbf{F}))\mathbf{e}_3 \ ,\end{aligned} \tag{5.65}$$

where $H_{eij}(\alpha)$, $i,j = 1,2$ are the corresponding components in the tensor-valued even Hessian operator H_e. Only if Eq. (5.65) delivers a non-zero output, the existence of i2D structures is indicated. For i0D and i1D structures, the result of Eq. (5.65) is zero. The i1D signals in the frequency domain are straight lines through the origin, therefore, the two dimensional convolutions in Eq. (5.65) can be reduced to 1D convolutions along the properly oriented axis, indicated by the fixed angle α_0. Then, Eq. (5.65) can be rewritten as

$$\begin{aligned}\mathcal{F}\{\mathbf{d}_e\} = ([H_{e11}(\alpha_0)H_{e22}(\alpha_0)]\,[(H_{DOP}(\rho;s)\mathbf{F}) * (H_{DOP}(\rho;s)\mathbf{F})] - \\ [H_{e12}(\alpha_0)H_{e21}(\alpha_0)]\,[(H_{DOP}(\rho;s)\mathbf{F}) * (H_{DOP}(\rho;s)\mathbf{F})])\mathbf{e}_3 \ .\end{aligned} \tag{5.66}$$

To ensure $\mathcal{F}\{\mathbf{d}_e\}$ to be zero for an i1D signal, the condition should be

$$H_{e11}(\alpha_0)H_{e22}(\alpha_0) = H_{e12}(\alpha_0)H_{e21}(\alpha_0) \ . \tag{5.67}$$

Hence, Eq. (5.67) makes sure that the even part is selective to the i2D structures and it was called the compensation equation in [122]. In the case of the generalized monogenic curvature signal, for its even part, we have

$$H_{e11}(\alpha_0)H_{e22}(\alpha_0) = \cos^2(\alpha_0)\sin^2(\alpha_0) = \frac{1}{4}\sin(2\alpha_0) \tag{5.68}$$

$$H_{e12}(\alpha_0)H_{e21}(\alpha_0) = \frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0) = \frac{1}{4}\sin(2\alpha_0)\ . \tag{5.69}$$

Therefore, the compensation Eq. (5.67) is satisfied and can be rewritten as

$$\cos^2(\alpha_0)\sin^2(\alpha_0) = \frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0)\ . \tag{5.70}$$

This equation decides that the even part of the generalized monogenic curvature signal is selective to the i2D structure and the i1D structure can be suppressed.

Analogously, the odd part of the generalized monogenic curvature signal $\mathbf{d}_o$ in the Fourier domain reads

$$\begin{aligned}\mathcal{F}\{\mathbf{d}_o\} &= \mathcal{F}\{\mathbf{e}_1\mathrm{det}_R(T_o)\} \qquad (5.71)\\ &= \mathbf{e}_1((H_{o11}(\alpha)H_{DOP}(\rho;s)\mathbf{F}) * (H_{o22}(\alpha)H_{DOP}(\rho;s)\mathbf{F})\\ &\quad -(H_{o12}(\alpha)H_{DOP}(\rho;s)\mathbf{F}) * (H_{o21}(\alpha)H_{DOP}(\rho;s)\mathbf{F}))\ .\end{aligned}$$

In order to make sure $\mathcal{F}\{\mathbf{d}_o\}$ is zero for an i1D signal, the compensation equation for the odd part of the generalized monogenic curvature signal with the following form should be satisfied

$$H_{o11}(\alpha_0)H_{o22}(\alpha_0) = H_{o12}(\alpha_0)H_{o21}(\alpha_0)\ , \tag{5.72}$$

where $H_{oij}, i,j = 1,2$ indicates the corresponding components in the tensor-valued odd Hessian operator H_o. Hence, we are able to obtain

$$\begin{aligned}H_{o11}(\alpha_0)H_{o22}(\alpha_0) = {}& \cos(2\alpha_0)\left[\cos^2(\alpha_0)\sin^2(\alpha_0)\right] + \qquad (5.73)\\ & \sin(2\alpha_0)\left[\cos^2(\alpha_0)\sin^2(\alpha_0)\right]\mathbf{e}_{12}\end{aligned}$$

$$\begin{aligned}H_{o12}(\alpha_0)H_{o21}(\alpha_0) = {}& \cos(2\alpha_0)\left[\frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0)\right] + \qquad (5.74)\\ & \sin(2\alpha_0)\left[\frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0)\right]\mathbf{e}_{12}\ .\end{aligned}$$

It is obvious that the compensation Eq. (5.72) is satisfied and the odd part of the generalized monogenic curvature signal does only respond to i2D structures. Consequently, the generalized monogenic curvature signal is regarded as the local representation of the i2D structure.

Since the odd part of the generalized monogenic curvature signal has two components, the compensation of the odd part can in accordance be split into two parts, that is

$$\cos(2\alpha_0)[\cos^2(\alpha_0)\sin^2(\alpha_0)] = \cos(2\alpha_0)[\frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0)] \tag{5.75}$$

$$\sin(2\alpha_0)[\cos^2(\alpha_0)\sin^2(\alpha_0)]\mathbf{e}_{12} = \sin(2\alpha_0)[\frac{1}{2}\sin(2\alpha_0)\frac{1}{2}\sin(2\alpha_0)]\mathbf{e}_{12}\ . \tag{5.76}$$

These two parts, which are derived from $\det_R(T_o) = T_{o11}T_{o22} - T_{o12}T_{o21}$, determine the two components of the odd part $B\mathbf{e}_1$ and $C\mathbf{e}_2$, respectively. Comparing the compensation equations (5.70),(5.75) and (5.76), it is shown that the determinant of the conjugate curvature tensor T_o is obtained from the curvature tensor T_e by convolving it with the second order spherical harmonic, that is

$$\det_R(T_o) = h_2 * \det_R(T_e)\ . \tag{5.77}$$

Therefore, the odd part of the generalized monogenic curvature signal $\mathbf{d}_o$ can be derived from the even part $\mathbf{d}_e$ by employing the second order spherical harmonic operator, which means

$$\begin{aligned} \mathbf{f}_{i2D} &= \det(T_e)\mathbf{e}_3 + \mathbf{e}_1\det(T_o) = \mathbf{d}_e + \mathbf{d}_o \qquad (5.78) \\ &= \mathbf{d}_e + (\mathbf{e}_1 h_2 \mathbf{e}_3) * \mathbf{d}_e = A\mathbf{e}_3 + B\mathbf{e}_1 + C\mathbf{e}_2\ . \end{aligned}$$

The angle between $B\mathbf{e}_1$ and $C\mathbf{e}_2$ indicates the local main orientation in a double angle representation just as in the case of the structure tensor [43, 47]. It is introduced in Section 3.9 that the odd part of the monogenic signal is obtained from the even part by employing the Riesz transform, which is basically equivalent to the first order spherical harmonic. In case of the generalized monogenic curvature signal, we have comparable relations between the even and odd parts. Only the second order spherical harmonic occurs as a new operator which is another generalization of the Hilbert transform. This enables us to state that the Riesz transform is able to generalize the Hilbert transform with respect to i1D signals and the derived generalized Hilbert transform (the second order spherical harmonic) realizes the same with respect to i2D signals. This is valid for any dimension of the signal. Furthermore, this gives rise to the conjecture that a third order spherical harmonic will be responsible for generalizing the Hilbert transform in the case of i3D signals in the 3D case. Recently, several types of generalized Hilbert transforms have been derived in [13, 14]. Indeed, the second order spherical harmonic belongs to one of the considered types.

According to the Parseval theorem, the energy of the odd and even parts of the generalized monogenic curvature signal are related by

$$\begin{aligned} E_o &= ((\mathbf{e}_1 h_2 \mathbf{e}_3) * \mathbf{d}_e)^2 = \int\int |\mathbf{e}_1 H_2 \mathbf{e}_3 \mathbf{D}_e|^2 d\rho d\alpha \\ &= \int\int |\cos(2\alpha)\mathbf{e}_{13} + \sin(2\alpha)\mathbf{e}_{23}|^2 |\mathbf{D}_e|^2 d\rho d\alpha \\ &= \int\int (\cos^2(2\alpha) + \sin^2(2\alpha))|\mathbf{D}_e|^2 d\rho d\alpha = E_e \ , \end{aligned} \tag{5.79}$$

where $\mathbf{D}_e$ represents the Fourier transform of $\mathbf{d}_e$. It can be shown that the energy of the odd part equals that of the even part. Hence, the amplitudes of even and odd part are equivalent, i.e. $|\mathbf{d}_e| = |\mathbf{d}_o|$.

As a good local representation for the i2D signal, it requires that the generalized monogenic curvature signal has the property of rotation invariance. To analyze this property, we should start from Eq. (5.33). Hence, the determinant of the Hessian matrix can be rewritten as

$$\begin{aligned} K &= f_{xx}\mathbf{e}_3 f_{yy}\mathbf{e}_3 - (-f_{xy}\mathbf{e}_{123})(f_{xy}\mathbf{e}_{123}) \\ &= \frac{1}{4}[(f_{xx}\mathbf{e}_3 + f_{yy}\mathbf{e}_3)^2 - ((f_{xx}\mathbf{e}_3 - f_{yy}\mathbf{e}_3)^2 + (f_{xy})^2)] \\ &= \frac{1}{4}[(\Delta\mathbf{f})^2 - \varepsilon^2] \ , \end{aligned} \tag{5.80}$$

where $\Delta\mathbf{f}$ is the Laplacian of the signal and ε indicates the eccentricity. Because Eq. (5.80) must be zero for the i0D and i1D signals, the compensation equation of it now changes as

$$(\cos^2(\alpha) + \sin^2(\alpha))^2 = \cos^2(2\alpha) + \sin^2(2\alpha) \ . \tag{5.81}$$

Since i0D and i1D structures are suppressed, what left is the i2D structure. Hence, spherical harmonics which serve as the angular portions, decide the rotation invariance of the i2D structure output. As stated in Section 5.1.2, the amplitude of the nth order spherical harmonic is rotationally invariant. Thus, the filter response from an nth order spherical harmonic has the property of rotation invariance. Therefore, the left and right sides of Eq. (5.81) determine two isotropic operations. The determinant of the Hessian matrix thus can be considered as the difference of two isotropic operation outputs, which proves that the property of rotation invariance is fulfilled.

A similar analysis applies to the even part of the generalized monogenic curvature signal, that is

$$\begin{aligned} \mathbf{d}_e &= (T_{e11}T_{e22} - T_{e12}T_{e21})\mathbf{e}_3 \\ &= \frac{1}{4}\left[(T_{e11} + T_{e22})^2 - ((T_{e11} - T_{e22})^2 + T_{e12}T_{e21}))\right]\mathbf{e}_3 \ . \end{aligned} \tag{5.82}$$

The compensation for $\mathbf{d}_e$ now takes the same form as the compensation Eq. (5.81). Therefore, $\mathbf{d}_e$ is rotationally invariant. According to equations (5.78) and (5.79), the odd part is obtained by convolving the even part with the second order spherical harmonic and the energy of them are equivalent. Therefore, the amplitude of the odd part $\mathbf{d}_o$ is also rotationally invariant.

Hence, we can conclude the properties of the generalized monogenic curvature signal as follows

- It enables the simultaneous estimation of the local amplitude, local main orientation and local phase of i2D structures.
- It is rotationally invariant and no steering is needed.
- The odd part is transformed from the even part by the second order spherical harmonic, i.e. $\mathbf{d}_o = (\mathbf{e}_1 h_2 \mathbf{e}_3) * \mathbf{d}_e$.
- The energy of the even part is equal to that of the odd part, i.e. $\mathbf{d}_e^2 = \mathbf{d}_o^2$.
- Since the local amplitude and local phase representation are independent of each other, the generalized monogenic curvature signal fulfills the split of identity.

5.4 Parity Symmetry Analysis

5.4.1 Preliminaries

Parity symmetry, as an important local feature of qualitative signal analysis, is strongly related to the local phase of the signal [106]. In image processing, parity symmetry is a cue for the line-like or edge-like quality of a local image structure. Parity refers to the invariance of a process with respect to a reflection operation. It is well known that any real signal $f : \mathbb{R}^n \rightarrow \mathbb{R}$ at any location $\mathbf{x} \in \mathbb{R}^n$ can be decomposed into an even and odd part, i.e. $f(\mathbf{x}) = f_e(\mathbf{x}) + f_o(\mathbf{x})$ [47]. A real signal has even symmetry if $f(-\mathbf{x}) = f(\mathbf{x})$ and odd symmetry if $f(-\mathbf{x}) = -f(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$.

In 1D, the local phase of a signal is defined as the angular part of its analytic signal. If the local energy is zero, no phase analysis is available. Once the local energy exceeds a certain threshold, the parity symmetry would then enable a local structure analysis. The relation between the local structure and local phase is illustrated in Fig. 5.10. At a signal position with locally even symmetry, only the real valued even part of the quadrature filter matches. Thus, the phase is 0 for a peak like signal and π for a dip like signal. A similar reflection reveals the odd case for edge like signals. Only the odd,

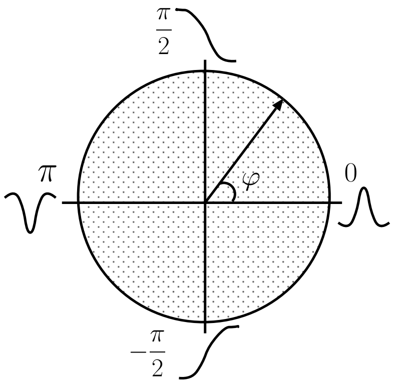

Fig. 5.10: The relation between the local signal structure and local phase in 1D case.

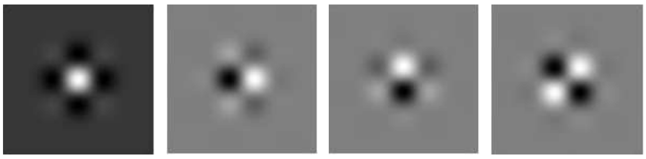

Fig. 5.11: Four components of the quaternionic Gabor filter.

and thus the imaginary, filter component matches the signal. Therefore, the signal structure has a phase of $\frac{\pi}{2}$ for decreasing slope and $-\frac{\pi}{2}$ for increasing slope. For i1D signals, the monogenic phase also indicates the even and odd symmetry as line and edge like structures in a rotation-invariant manner. The line-like structure is even symmetric with respect to its orientation vector, hence, the odd output of the monogenic signal is zero. The edge-like structure is pure odd symmetric with respect to its orientation vector. Therefore, no even output from the monogenic signal exists.

A phase concept of i2D signals has been investigated by Bülow and Sommer [20, 18]. In this approach, a two-dimensional signal is split into even and odd parts along the x-axis and along the y-axis, i.e. $\mathbf{f} = \mathbf{f}_{ee} + \mathbf{f}_{oe} + \mathbf{f}_{eo} + \mathbf{f}_{oo}$. Here, $\mathbf{f}_{ee}$ denotes the part of $\mathbf{f}$ that is even with respect to x and y, $\mathbf{f}_{oe}$ represents the part which is odd with respect to x and even with respect to y and so on. These four symmetries are obtained by convolving the signal with four components of a quaternionic Gabor filter which are shown in Fig. 5.11. Those four components have the even-even, odd-even, even-odd and odd-odd symmetries with respect to the coordinates. The estimated phase information can, to some extent, illustrate the intrinsic dimensionality of local structures. Unfortunately, this approach has the drawback of being not rotation invariant.

5.4.2 Parity Symmetry of the Generalized Monogenic Curvature Signal

The generalized monogenic curvature signal also enables the rotation invariant extraction of a phase for i2D signals. The presented phase interpretation strongly depends on the way the monogenic curvature tensor is analyzed. Here, we study only the relations between its even and odd components. The parity symmetry involved in our case can be analyzed from the even and odd parts of the generalized monogenic curvature signal, respectively. As indicated in Section 5.2, the even part is computed as,

$$\mathbf{d}_e = (T_{e11}T_{e22} - T_{e12}T_{e21})\mathbf{e}_3 = \lambda_1\lambda_2\mathbf{e}_3 \; . \tag{5.83}$$

According to [31], the eigenvalues λ_1 and λ_2 can be considered as a $\cos^2$-decomposition of the local amplitude. Hence, these two eigenvalues can be obtained from two angular windowing functions $\cos^2(\beta - \theta_0)$ and $\cos^2(\beta - (\theta_0 + \frac{\pi}{2})) = \sin^2(\beta - \theta_0)$, which are oriented along the principal axes. Here, θ_0 refers to the local main orientation. These two angular windowing functions are even with respect to the principal axes oriented along θ_0 and $\theta_0 + \frac{\pi}{2}$, respectively. Thus, $\mathbf{d}_e$ has the even-even symmetry with respect to the principal axes.

The odd part of the generalized monogenic curvature signal is computed as

$$\mathbf{d}_o = \mathbf{e}_1(T_{o11}T_{o22} - T_{o12}T_{o21}) = (\mathbf{e}_1 h_2 \mathbf{e}_3) * \mathbf{d}_e = B\mathbf{e}_1 + C\mathbf{e}_2 \; . \tag{5.84}$$

Hence, the two components of $\mathbf{d}_o$ are obtained by convolving the second order spherical harmonic with the even part. Therefore, the symmetry of the odd part can be determined by the second order spherical harmonic. As visualized in Fig. 5.12, the two components of the second order circular harmonic, $\cos(2\beta)$ and $\sin(2\beta)$, are odd-odd symmetric with respect to $\mathbf{e}_1$, $\mathbf{e}_2$ axes and the diagonals of the plane $\mathbf{e}_1 \wedge \mathbf{e}_2$, respectively. These two components together decide the local main orientation. Thus, the odd part of the generalized monogenic curvature signal, i.e. the vector sum of $B\mathbf{e}_1$ and $C\mathbf{e}_2$, has an odd-odd symmetry with respect to the principal axes.

Unlike the i1D signals, i2D signals have more degrees of freedom. There is no general way to completely specify the parity symmetry of an arbitrary i2D structure by applying a pre-defined phase model. Because the odd part of the generalized monogenic curvature signal has only odd-odd symmetry, the derived phase information can only be used to classify some specific i2D structures, even though the generalized monogenic curvature signal is a general local representation for all i2D structures. It can be assumed that a

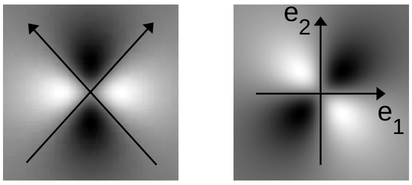

Fig. 5.12: Two orthogonal components of the bandpass bounded second order circular harmonic, $\cos(2\beta)$ and $\sin(2\beta)$ (white: +1, black: -1).

Fig. 5.13: Some i2D Structures: from left to right are the line-like intersection, edge-like intersection and a mixture of them.

component wise evaluation of the monogenic curvature tensor will result in a more detailed specification of the parity symmetry.

For any superimposed i2D signal, as shown in Fig. 5.13, there exist patterns such as the line-like intersection, the edge-like intersection and a mixture of them. The local neighborhood where two lines/edges intersect is considered as the i2D structure. It corresponds to elliptic or hyperbolic region with positive or negative Gaussian curvature, see also Tab. 5.1. Therefore, for the i2D neighborhood where two lines intersect, it belongs to the elliptic region which denotes local extreme. Hence, the even output of the generalized monogenic curvature signal is positive, i.e. $\mathbf{d}_e > 0$. Because the second order spherical harmonic cannot match the line-like intersection structure, in this case, the odd output would vanish, i.e. $\mathbf{d}_o = 0$. Therefore, the estimated local phase has a value of zero

$$\varphi = \operatorname{atan2}(\mathbf{d}_o, \mathbf{d}_e) = 0 \ . \tag{5.85}$$

Edge-like intersection structures are saddle points which correspond to hyperbolic regions. Therefore, the generalized monogenic curvature signals in such case have negative even outputs, i.e. $\mathbf{d}_e < 0$. Because the even parts are filtered results from original signals, in this case, edges in original signals will appear as peak or dip like structures. Thus, odd parts of the generalized monogenic curvature signals are non-zero since second order spherical harmonics could match the even outputs. Hence, due to the intersection of edges, the generalized monogenic curvature signal has not only even but also non-zero odd outputs. As mentioned before, the energy of the even part is

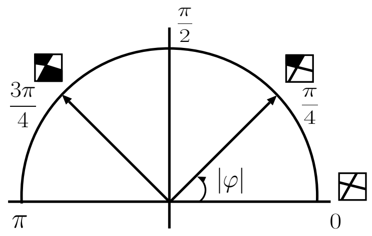

Fig. 5.14: Local phase and corresponding i2D structures. The horizontal axis is the real axis, the vertical axis indicates the imaginary axis.

identical to that of the odd part. Thereby, the amplitude of the even output equals that of the odd part, i.e. $|\mathbf{d}_e| = |\mathbf{d}_o|$. Hence, the absolute phase value $|\varphi|$ reads

$$|\varphi| = \text{atan2}(|\mathbf{d}_o|, \mathbf{d}_e) = \frac{3\pi}{4}\ . \tag{5.86}$$

The sign of the phase depends on the direction of $\mathbf{d}_o$.

The i2D structure which is the mixture of the line-like and edge-like intersection, has also a positive even output since it corresponds to the local extreme. Due to the edge-like intersection, its odd output is also non-zero. Therefore, the even part of the generalized monogenic curvature signal is positive, i.e. $\mathbf{d}_e > 0$. And the odd part of it is also equivalent to the even part, i.e. $|\mathbf{d}_e| = |\mathbf{d}_o|$. Thereby, for the mixture pattern, its local phase absolute value is obtained as

$$|\varphi| = \text{atan2}(|\mathbf{d}_o|, \mathbf{d}_e) = \frac{\pi}{4}\ . \tag{5.87}$$

Its sign also relies on the direction of $\mathbf{d}_o$. Consequently, some specific structures can be classified on the basis of the newly developed phase information. On the oriented complex plane, shown in Fig. 5.14, phase values clearly denote what kind of structure it is. The local phase is able to distinguish between line-like, edge-like intersection structures and the mixture pattern of them.

The two superimposed i1D signals are not necessarily to be perpendicular to each other. This is a meaningful extension of the structure multivector model as proposed in [31]. If the phase has a value of zero, the corresponding structure is indicating line-like intersection. A phase absolute value of $\frac{3\pi}{4}$ or $\frac{\pi}{4}$ implies that the corresponding local structure is edge-like intersection or a mixture pattern.

Fig. 5.15 illustrates the evaluated phase information for two superimposed patterns with flexible angle of intersections. In the case of very bright blobs, they can be regarded as line-like intersections, hence, their phases take zero values. For those blobs which are a bit darker than the bright blobs, they

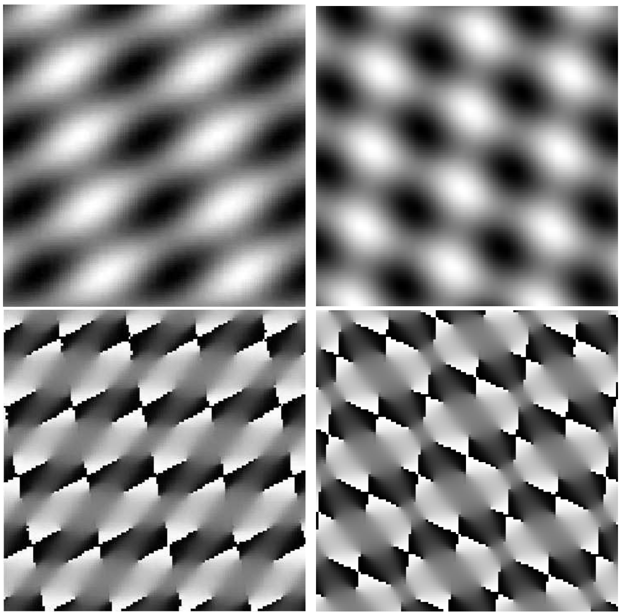

Fig. 5.15: Top row: two test images. Bottom row: corresponding phases of the two test images.

indicate edge-like intersections with phases of $\frac{3\pi}{4}$ or $-\frac{3\pi}{4}$. It can be assumed that an extensive analysis of the inner structure of the monogenic curvature tensor will yield more degrees of freedom. Thus, more phase angles will be obtained which can further help specifying i2D patterns.

5.5 Generalized Monogenic Curvature Scale-Space

5.5.1 Definition and Local Features

In Section 5.1.1, the way to embed a 2D signal into a 3D space has been introduced. So far, we discussed only the case with $z = 0$ for our general model. The half space $z > 0$ for the monogenic signal is the monogenic scale-space [38] which has already been investigated in detail. Now, we will sketch in this section the generalized monogenic curvature signal in the half space $z > 0$. Similar as in the case of the monogenic signal, z works as a scale parameter. Applying Poisson filtering to the monogenic curvature tensor results in a generalized monogenic curvature scale-space which enables the multi-scale processing of local structures. Since the generalized monogenic curvature signal consists of even and odd parts, the generalized monogenic curvature scale-space can also be formed by the Poisson curvature scale-space

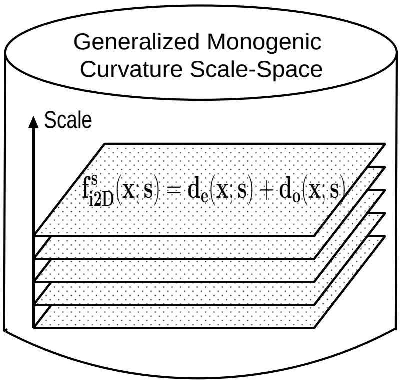

Fig. 5.16: The generalized monogenic curvature scale-space.

and the conjugate Poisson curvature scale-space as illustrated in Fig. 5.16.

Combining the Poisson kernel with 2D spherical harmonics forms damped spherical harmonics which are basis functions to build the generalized monogenic curvature scale-space. An nth order damped spherical harmonic in the frequency domain is given by

$$H_n(\rho, \alpha; s) = \exp(n\alpha\mathbf{e}_{12})\exp(-2\pi\rho s) \ , \tag{5.88}$$

where ρ, α indicates the spectral domain polar coordinates, s denotes the scale parameter. When s equals zero, the first order damped spherical harmonic is basically identical to the Riesz kernel. Let r, β be the polar coordinates in the spatial domain, the curvature tensor $T_e(r, \beta; s)$ results from a scale dependent tensor valued filter which in the Fourier domain reads

$$H_e(\rho, \alpha; s) = \begin{bmatrix} H_0(\rho, \alpha; s) + \langle H_2(\rho, \alpha; s)\rangle_0 & -\langle H_2(\rho, \alpha; s)\rangle_2 \\ \langle H_2(\rho, \alpha; s)\rangle_2 & H_0(\rho, \alpha; s) - \langle H_2(\rho, \alpha; s)\rangle_0 \end{bmatrix} . \tag{5.89}$$

The monogenic extension of the curvature tensor in the scale-space results in its conjugate part, i.e. $T_o(r, \beta; s)$. In this case, the conjugate part can be considered as the result of applying the conjugate Poisson filtering to the curvature tensor

$$T_o(r, \beta; s) = h_1(r, \beta; s) *_\tau T_e(r, \beta; s) \ , \tag{5.90}$$

where $h_1(r, \beta; s)$ indicates the conjugate Poisson kernel in the spatial domain. As an alternative, $T_o(r, \beta; s)$ can also be obtained from a scale dependent tensor valued filter $H_o(\rho, \alpha; s)$, that is

$$\begin{aligned} T_o(r, \beta; s) &= \mathcal{F}^{-1}\left\{H_o(\rho, \alpha; s) \times_\tau \mathbf{F}(\rho, \alpha)\right\} \qquad (5.91) \\ &= \mathcal{F}^{-1}\left\{H_1(\rho, \alpha; s) \times_\tau H_e(\rho, \alpha; s) \times_\tau \mathbf{F}(\rho, \alpha)\right\} \ . \end{aligned}$$

Hence, the filter $H_o(\rho, \alpha; s)$ results from the conjugate Poisson filtering of the filter $H_e(\rho, \alpha; s)$,

$$H_o(\rho, \alpha; s) = H_1(\rho, \alpha; s) \times_\tau H_e(\rho, \alpha; s) = \begin{bmatrix} H_{o11}(\rho, \alpha; s) & H_{o12}(\rho, \alpha; s) \\ H_{o21}(\rho, \alpha; s) & H_{o22}(\rho, \alpha; s) \end{bmatrix} \tag{5.92}$$

with

$$H_{o11}(\rho, \alpha; s) = H_1(\rho, \alpha; s)(H_0(\rho, \alpha; s) + \langle H_2(\rho, \alpha; s)\rangle_0) \tag{5.93}$$

$$H_{o12}(\rho, \alpha; s) = -H_{o12}(\rho, \alpha; s) = H_1(\rho, \alpha; s)(-\langle H_2(\rho, \alpha; s)\rangle_2) \tag{5.94}$$

$$H_{o22}(\rho, \alpha; s) = H_1(\rho, \alpha; s)(H_0(\rho, \alpha; s) - \langle H_2(\rho, \alpha; s)\rangle_0) \ . \tag{5.95}$$

Once the monogenic curvature tensor is unified within a scale-space framework, the computation of its trace delivers access to a scale-space which characterizes the i1D structures. This scale-space is given by

$$\mathbf{f}^s_{i1D}(r, \beta; s) = \mathrm{trace}(T_e(r, \beta; s)) + \mathrm{trace}(T_o(r, \beta; s))\mathbf{e}_2 \ . \tag{5.96}$$

In terms of the Cartesian coordinate, it takes the following form

$$\begin{aligned} \mathbf{f}^s_{i1D}(\mathbf{x}; s) &= \mathrm{trace}(T_e(\mathbf{x}; s)) + \mathrm{trace}(T_o(\mathbf{x}; s))\mathbf{e}_2 \\ &= p(\mathbf{x}; s) + \mathbf{q}(\mathbf{x}; s) \ . \end{aligned} \tag{5.97}$$

This means that the scale-space derived from the trace computation of the scale unified monogenic curvature tensor is exactly the monogenic scale-space [38]. From it, three orthogonal scale-spaces including the local amplitude, orientation and phase can be obtained for further image analysis.

By computing the Gaussian curvature of $T_e(r, \beta; s)$, the Poisson curvature scale-space can be obtained as

$$\begin{aligned} \mathbf{d}_e(r, \beta; s) &= \det_R(T_e(r, \beta; s))\mathbf{e}_3 \\ &= (T_{e11}(r, \beta; s)T_{e22}(r, \beta; s) - T_{e12}(r, \beta; s)T_{e21}(r, \beta; s))\mathbf{e}_3 \\ &= A(r, \beta; s)\mathbf{e}_3 \ . \end{aligned} \tag{5.98}$$

Hence, the conjugate Poisson curvature scale-space is given by

$$\begin{aligned} \mathbf{d}_o(r, \beta; s) &= \mathbf{e}_1\det_R(T_o(r, \beta; s)) \\ &= \mathbf{e}_1(T_{o11}(r, \beta; s)T_{o22}(r, \beta; s) - T_{o12}(r, \beta; s)T_{o21}(r, \beta; s)) \\ &= B(r, \beta; s)\mathbf{e}_1 + C(r, \beta; s)\mathbf{e}_2 \ . \end{aligned} \tag{5.99}$$

Combining the Poisson curvature scale-space and the conjugate Poisson curvature scale-space, the generalized monogenic curvature scale-space can be obtained, which in terms of Cartesian coordinates is given as

$$\mathbf{f}_{i2D}^{s}(\mathbf{x};s) = \mathbf{d}_e(\mathbf{x};s) + \mathbf{d}_o(\mathbf{x};s) \ . \tag{5.100}$$

According to the analysis in Section 5.1.1, local amplitude and local phase representation of the generalized monogenic curvature scale-space are extracted as

$$A(\mathbf{x};s) = |\mathbf{f}_{i2D}^{s}(\mathbf{x};s)| = \exp(\log(|\mathbf{f}_{i2D}^{s}(\mathbf{x};s)|)) = \exp(\langle \log(\mathbf{e}_3\mathbf{f}_{i2D}^{s}(\mathbf{x};s))\rangle_0) \tag{5.101}$$

$$\Phi(\mathbf{x};s) = \arg\left(\mathbf{f}_{i2D}^{s}(\mathbf{x};s)\right) = \langle \log\left(\mathbf{e}_3\mathbf{f}_{i2D}^{s}(\mathbf{x};s)\right)\rangle_2 \ . \tag{5.102}$$

As the dual of the local phase representation, the local rotation vector is defined as

$$\mathbf{r}(\mathbf{x};s) = \arg\left(\mathbf{f}_{i2D}^{s}(\mathbf{x};s)\right)^{*} = \langle \log\left(\mathbf{e}_3\mathbf{f}_{i2D}^{s}(\mathbf{x};s)\right)\rangle_2^{*} \ . \tag{5.103}$$

The length of the rotation vector $\mathbf{r}(\mathbf{x};s)$ indicates the angle of phase at a certain scale, and it is also perpendicular to the local main orientation. Instead of the generalized monogenic curvature scale-space, local amplitude, main orientation and phase form three orthogonal scale-spaces which should be applied for real applications. Fig. 5.17 illustrates local energies and local phases extracted from a synthetic image at three scales in the generalized monogenic curvature scale-space. For the feature extraction of the bandpassed version in the generalized monogenic curvature scale-space, one may refer to Fig. 5.6. All the extracted features are scale dependent, however, local amplitude, phase and orientation are still independent of each other at each scale. Hence, the split of identity is still preserved in the generalized monogenic curvature scale-space.

To investigate these local features in detail, two real images, shown in Fig. 5.18, are employed for feature extraction in the generalized monogenic curvature scale-space. In the real applications, only the local feature scale-spaces are considered for the analysis. Figs. 5.19 and 5.20 illustrate the corresponding three orthogonal scale-spaces. At the finer scale, structures are much more clear and easy to be detected. Hence, local features represent rich information of the image but sensitive to the noise. With the increase of the scale, less structures can be identified, and the curvature information is correspondingly changed. Thus local features are blurred and locations varied with respect to the scale. Using multi-scale processing strategy, many applications such as shape description, image segmentation, matching, pattern

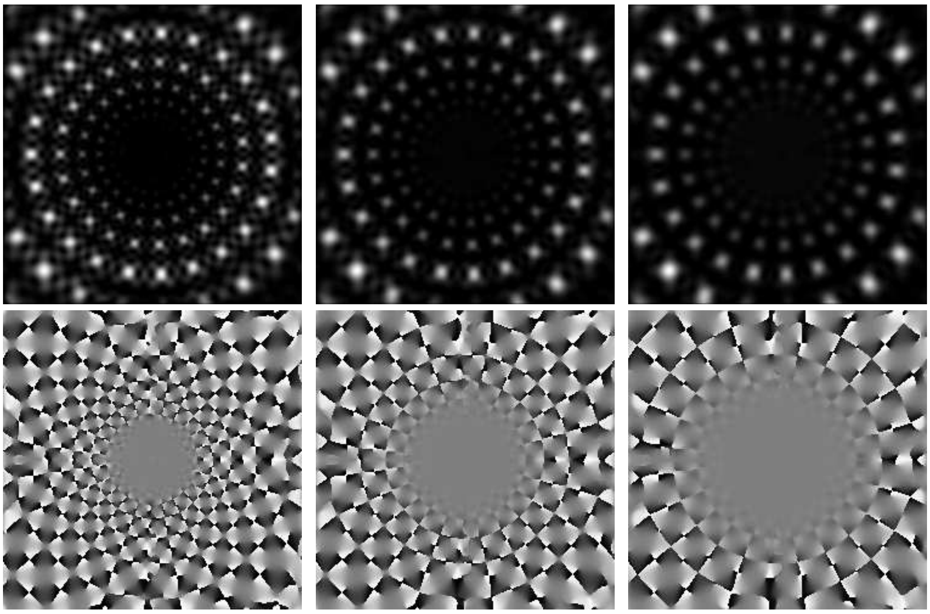

Fig. 5.17: Top row: local energies extracted at three scales in the generalized monogenic curvature scale-space. Bottom row: local phases evaluated at three scales.

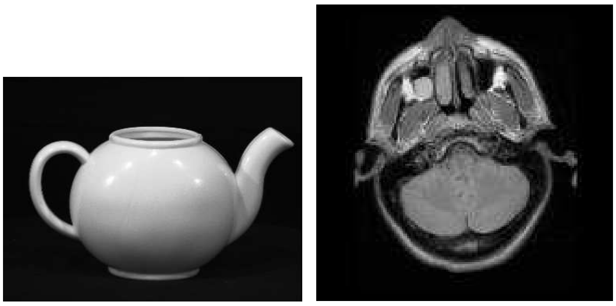

Fig. 5.18: Two test images.

recognition and so on can be well implemented in the generalized monogenic curvature scale-space.

Up to now, we understand how the local features vary with respect to the scale. But when the scale changes, what would happen to the properties of local structures is still unclear. To this end, a synthetic image representing line-crossing is used for the investigation. Besides, the relation between the angle of intersection variation and the type of local structures is also investigated. Fig. 5.21 demonstrates the test image and the confidence measure with respect to the scale and angle of intersection variation. The confidence

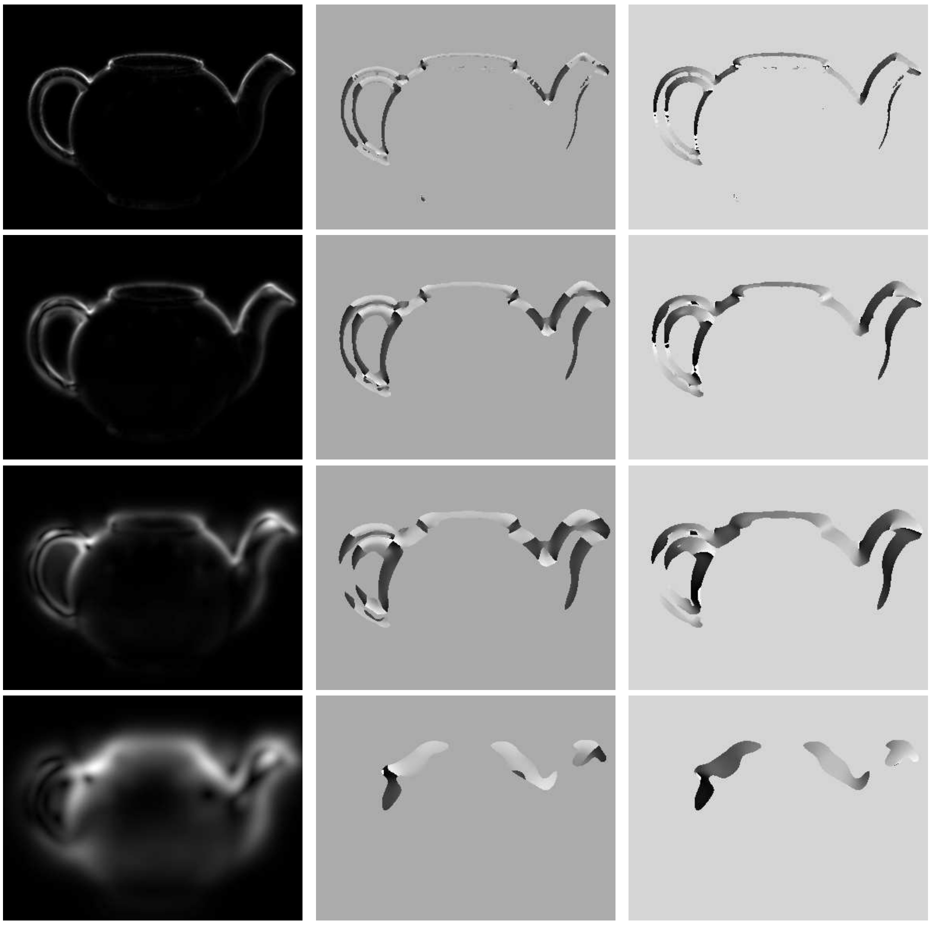

Fig. 5.19: From top row to the bottom row are the local features extracted at four different scales. Column one represents the local amplitude, local phase and main orientation are shown in columns two and three. Note that the local phase and orientation are thresholded according to the energy information.

measure is computed according to the local energy. Zero indicates that the local structure is i1D, where for the i2D structure, the confidence measure takes one. It is shown that the original i2D local structure is gradually changed as the i1D structure with the increase of the scale. This is because of the blurring effect resulting from the scale increase. When the angle of intersection is 90^0, the two lines are perpendicular to each other. Hence, the local energy at the crossing point is very strong which denotes a confidence measure of one. With the decreasing of this angle of intersection, two lines are becoming much close to each other. Thus the local energy at the crossing

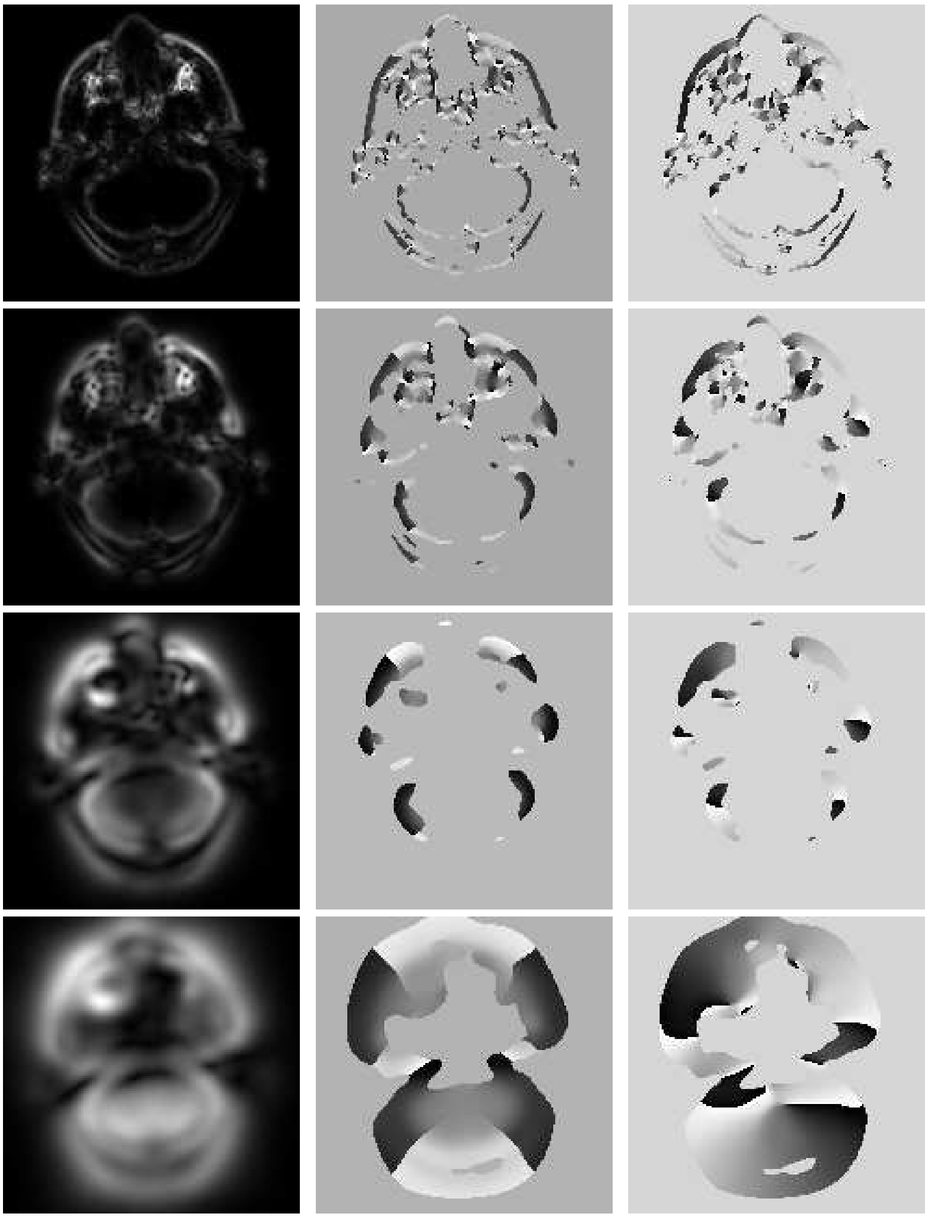

Fig. 5.20: From top row to the bottom row are the local features extracted at four different scales. Column one represents the local amplitude, local phase and main orientation are shown in columns two and three. Note that the local phase and orientation are thresholded according to the energy information.

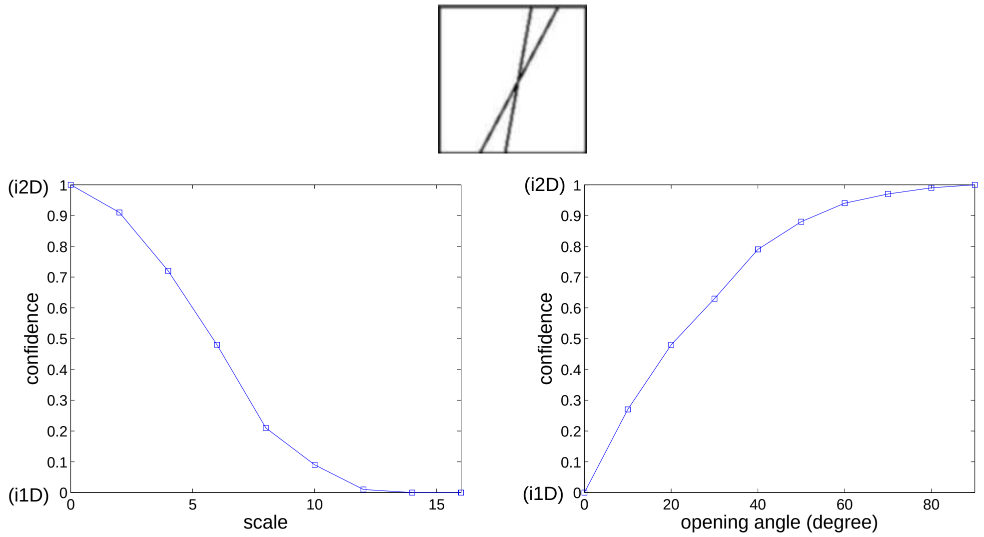

Fig. 5.21: Top row: the test image. Bottom row: confidence measure with respect to the scale and angle of intersection variation.

point decreases correspondingly and finally the confidence measure reaches a value of zero.

5.5.2 Comparison with the Gaussian Curvature Scale-Space

Since the generalized monogenic curvature signal is derived from the Gaussian curvature, the scale-space representation of the generalized monogenic curvature signal is sensitive to the curvature of local structures at the chosen scale. A similar concept is the Gaussian curvature scale-space [79], which is also suitable for recovering invariant geometric features of a signal at multiple scales.

Instead of considering the surface curvature, the Gaussian curvature scale-space focuses on the planar curve curvature. A planar curve is a set of points whose position vectors are the values of a continuous vector-valued function, which can be represented by a parametric vector equation

$$\mathbf{r}(u) = (x(u), y(u)) \; . \tag{5.104}$$

From the planar curve, the isophote curvature is defined as

$$k(u) = \frac{x'(u)y''(u) - y'(u)x''(u)}{(x'(u)^2 + y'(u)^2)^{3/2}} \; . \tag{5.105}$$

By convolving with a Gaussian kernel $g(u,\sigma)$, an evolved version of the planar curve can be obtained as

$$\mathbf{r}(u,\sigma) = (x(u,\sigma),\, y(u,\sigma))\ , \tag{5.106}$$

where $x(u,\sigma) = x(u) * g(u,\sigma)$ and $y(u,\sigma) = y(u) * g(u,\sigma)$.

Hence, the evolved isophote curvature of the planar curve reads

$$k(u,\sigma) = \frac{x_u(u,\sigma)y_{uu}(u,\sigma) - y_u(u,\sigma)x_{uu}(u,\sigma)}{(x_u(u,\sigma)^2 + y_u(u,\sigma)^2)^{3/2}} \tag{5.107}$$

with

$$\begin{aligned} x_u(u,\sigma) &= x(u) * g_u(u,\sigma) \quad y_u(u,\sigma) = y(u) * g_u(u,\sigma) \\ x_{uu}(u,\sigma) &= x(u) * g_{uu}(u,\sigma) \quad y_{uu}(u,\sigma) = y(u) * g_{uu}(u,\sigma)\ , \end{aligned} \tag{5.108}$$

where $g_u(u,\sigma)$ and $g_{uu}(u,\sigma)$ denote the first and second derivatives of the Gaussian kernel, respectively. The solution of the equation

$$k(u,\sigma) = 0 \tag{5.109}$$

thus forms the Gaussian curvature scale-space image of the planar curve.

Compared with the generalized monogenic curvature scale-space, the Gaussian curvature scale-space has a totally different curvature definition. In addition to this, it applies a Gaussian smoothing kernel instead of a Poisson kernel which results in the loss of harmonic conjugate information. Hence, the unique advantage of the generalized monogenic curvature scale-space is the odd part being in quadrature phase relation to the even part at each scale. Besides, local amplitude, phase and orientation, which are unified with the scale concept, can be simultaneously estimated. These advantages give access to a multi-scale phase-based processing in many computer vision tasks.

5.6 Summary

A general signal model for 2D image structures was proposed in this chapter. In order to obtain more degrees of freedoms for modeling, a 2D signal is embedded into a certain geometric algebra of the Euclidean 3D space. Coupling methods of tensor algebra, differential geometry, monogenic signal and quadrature filter, we are able to design a curvature tensor and its monogenic extension. The monogenic extension of the curvature tensor contains rich information for 2D structures as the generalization of the analytic signal to 2D case. Therefore, it is regarded as the general model for 2D image structures. Based on it, local representations for i1D and i2D structures are obtained as the monogenic signal and the generalized monogenic curvature signal by computing the mean and Gaussian curvatures of the general model. Hence, the monogenic signal for i1D structures can be considered as a special case of the proposed general model.

From the generalized monogenic curvature signal, three features can be extracted. They are the local amplitude, local phase and local orientation. These features are independent of each other, hence, the generalized monogenic curvature signal performs a split of identity. The local amplitude derived from that model represents the energetic information and it indicates the existence of i2D structures. Structure information of i2D structures is contained in the local phase, and some specific types of i2D structures can be classified by the local phase. Local orientation denotes the main orientation and it illustrates the geometric information of i2D structures. Compared with other approaches for the i2D structure representation, the generalized monogenic curvature signal has the remarkable advantage of simultaneous estimation of local amplitude, local phase and local orientation in a rotation-invariant manner.

A generalized monogenic curvature scale-space is built by applying the Poisson kernel to the monogenic curvature tensor. The local amplitude, local phase and local orientation are scale dependent, but they are still independent of each other at every scale. In contrast to the Gaussian curvature scale-space, not only the energetic but also the structure and geometric information can be extracted at the same time in a multi-scale way. This advantage delivers access to various applications of the generalized monogenic curvature signal in computer vision tasks.

6. SCALE-SPACE BASED IMAGE ANALYSIS

To understand more about the presented theoretic framework, some scale-space based image analysis applications are demonstrated in this chapter. The first one is about image reconstruction using the phase partial information. The second one describes detecting i2D image structures using local phase information. The last one illustrates phase based optical flow estimation.

6.1 Phase Based Image Reconstruction in the Monogenic Scale-Space

In this section, image reconstruction based on the partial information of phase in a scale-space framework will be introduced. The presented approach is simple and robust, which outperforms other related approaches by showing the comparisons.

6.1.1 Motivation and Background

In the past decades, signal reconstruction from partial information has been an active area of research. Partial information such as zero crossing, Fourier magnitude and localized phase are considered to represent important features of the original signal. Therefore, we are able to reconstruct the original signal based on only the partial information. The variety of results on signal reconstruction has a major impact on the research fields like image processing, communication and geophysics.

Reconstruction from zero crossings in the scale-space is investigated by Hummel [57]. He has demonstrated that reconstruction based on zero crossings is possible but can be unstable, unless gradient values along the zero crossings are added. In [85], it is proved that many features of the original image are clearly identifiable in the phase only image but not in the magnitude only image, and reconstruction from Fourier phase is visually satisfying.

However, the application of this approach is rather limited in practice due to the computational complexity. Behar et al. have stated in [8] that image reconstruction from localized phase only information is more efficient and faster than that from the global phase. The reconstruction errors produced by this method can be very small. However, compared with this approach, the way of image reconstruction presented in this paper is more easier and faster.

In this chapter, we present an approach of image reconstruction from the local phase vector in the monogenic scale-space. Image reconstruction is easy, fast, accurate and stable when compared with the above mentioned approaches. In [36], Felsberg and Sommer proposed the first rotationally invariant 2D analytical signal. As one of its features, the monogenic phase vector preserves most of the important information of the original signal. The local phase vector contains not only the local phase but also the orientation information of the original signal, which enables the evaluation of structure and geometric information at the same time. The embedding of local phase and local orientation into monogenic scale-space improves the stability and robustness. However, in the Gaussian scale-space, there is no common filter set which could evaluate the local orientation and local phase simultaneously. To show the advantage of our approach, we replace the Gaussian kernel with the Gabor filter [45] for phase evaluation, the reconstruction results of these two approaches are also compared in this thesis.

6.1.2 Relationship Between the Local Attenuation and the Local Phase Vector

The relationship between 1D attenuation and phase was reported in [87], which indicates that for a minimum-phase system, the attenuation and the phase response are related by the Hilbert transform. In this context, a minimum-phase system is defined by having no zeros and poles in the positive half-plane of the Laplace domain. Under certain conditions, this relation could also be generalized to 2D.

For a 2D signal with an intrinsic dimension of one, if the scale-space representation has no zeros in the half space with $s \geq 0$, then the local attenuation a and the local phase vector $\mathbf{r}$ form a Riesz triplet [38]

$$\mathbf{r}(\mathbf{x}; s) \approx (h_R * a)(\mathbf{x}; s) \ . \tag{6.1}$$

In practice, images are in general not globally i1D signals. However, they commonly have lots of i1D neighborhoods which makes the reconstruction from the local phase vector available. In most practical applications zeros occur in the positive half-space, but as we can see from [38], the influence of the zeros can mostly be neglected.

6.1.3 Image Reconstruction Based on the Monogenic Phase

To recover the amplitude information from only the phase vector information, we take the inverse Riesz transform of the local phase vector. By definition, the Riesz transform of the local phase vector is DC free. This means that the transformed output has no DC component. Consequently, the DC-free local attenuation in the scale-space is approximated by the following form

$$a(\mathbf{x};s) - \overline{a}(\mathbf{x};s) \approx -(h_R * \mathbf{r})(\mathbf{x};s) \ , \tag{6.2}$$

where $\overline{a}(\mathbf{x};s)$ indicates the DC component of the local attenuation that should be calculated beforehand. Hence, the original image reconstruction based on the local phase vector reads

$$f(\mathbf{x}) = \exp(\overline{a}(\mathbf{x};0))\exp(-(h_R * \mathbf{r})(\mathbf{x};0))\cos(|\mathbf{r}(\mathbf{x};0)|) + C_{DC} \ , \tag{6.3}$$

where C_{DC} denotes a further DC correction term corresponding to a gray value shift. To reconstruct a real image, we use only the real part of the local phase vector $\cos(|\mathbf{r}(\mathbf{x};0)|)$. The above introduction indicates that image reconstruction from the local phase vector can be easily and quickly implemented, no iterative procedure is needed.

In order to improve the reconstruction robustness and stability, a scale pyramid structure, see [21], is employed for investigating the image reconstruction in the monogenic scale-space. Fig. 6.1 shows the diagram of the reconstruction in the monogenic scale-space. Given an input image, the differences of monogenic signals at adjacent scales are first computed as the bandpass decomposition at different frequencies in the monogenic scale-space. The information of different bandpasses forms a Laplacian like pyramid. Local phase vectors of the corresponding bandpass information are considered as the partial information. Signals can thus be reconstructed in the scale-space by a coarse to fine way. Let $g^{(s)}$ denote the representation of the image in the pyramid at scale s, then the one scale higher representation reads $g^{(s+1)}$. By interpolation, $g^{(s+1)}$ is expanded as $\widehat{g}^{(s+1)} = T_I g^{(s+1)}$, where T_I refers to the operation of interpolation and $\widehat{g}^{(s+1)}$ has the same size of $g^{(s)}$. The difference of adjacent scales can then be computed as

$$l^{(s)} = g^{(s)} - \widehat{g}^{(s+1)} = g^{(s)} - T_I g^{(s+1)} \ , \tag{6.4}$$

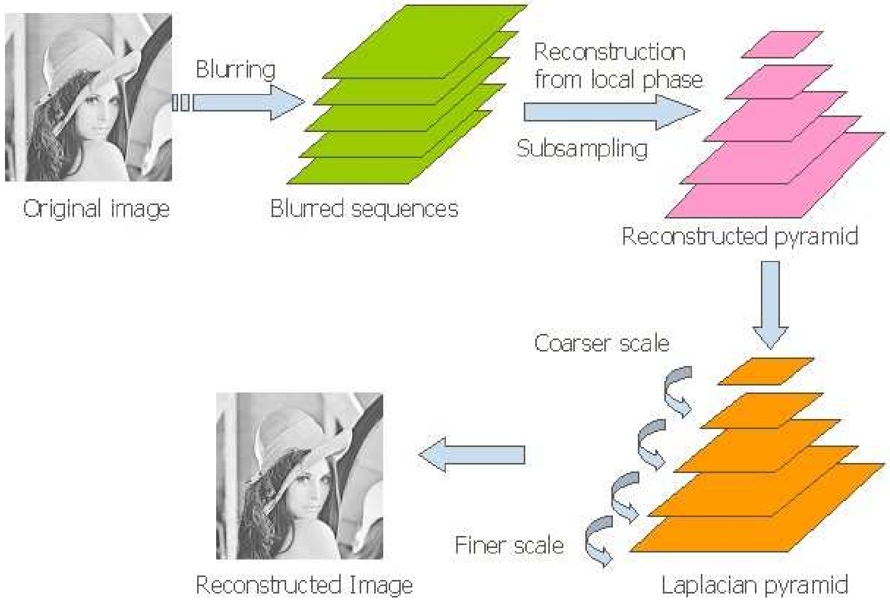

Fig. 6.1: The image reconstruction process based on the monogenic phase in the monogenic scale-space.

where $l^{(s)}$ can be regarded as a bandpass decomposition of the original image. Based on only the local phase vector of the intermediate representation, the reconstruction at different scales can be implemented as follows

$$\widetilde{l}^{(s)} = \exp(\overline{a}(\mathbf{x};s))\exp(-(h_R * \mathbf{r})(\mathbf{x};s))\cos(|\mathbf{r}(\mathbf{x};s)|) + C_{DC} \ , \tag{6.5}$$

where $\widetilde{l}^{(s)}$ describes the reconstructed result at a certain scale. By means of a coarse to fine approach, all the scale-space images can be combined together to make the final reconstruction of the original image. Starting from the most coarse level, the recovery of one scale lower image takes the following form

$$\widetilde{g}^{(s)} = \widetilde{l}^{(s)} + T_I \widetilde{g}^{(s+1)} \ . \tag{6.6}$$

This is an iterative procedure. It will end until s goes to zero, hence, $\widetilde{g}^{(0)}$ indicates the final reconstruction.

6.1.4 Image Reconstruction Based on the Gabor Phase

In contrast to the monogenic scale-space, there is not common filter set in the Gaussian scale-space which enables the simultaneous evaluation of the local phase and local orientation. However, phase information can be estimated when the Gaussian kernel is replaced by the Gabor filter [45]. As a comparison, image reconstruction based on the Gabor phase is also included.

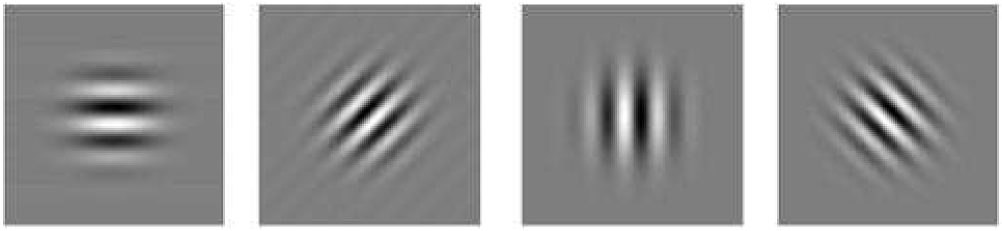

Fig. 6.2: Imaginary parts of 2D Gabor filters with orientations of 0^0, 45^0, 90^0 and 135^0.

Gabor filters have been shown to be a useful tool in different image processing and analysis tasks. They are quite useful wherever one is interested in local properties of a signal. The main advantage of the Gabor filter is its optimal localization with respect to the uncertainty principle in the spatial and the frequency domain simultaneously [45]. In this section, phases of the Gabor filtered images are extracted as partial information to reconstruct the original image.

Let $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2$ and $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ be the spatial and Fourier domain coordinates, respectively. A two dimensional Gabor filter is a linear shift invariant filter with the following impulse response

$$h(\mathbf{x}, \mathbf{u}, \sigma_x, \sigma_y) = g(x, y)\exp(i2\pi(xu + yv)) \tag{6.7}$$

with

$$g(x, y) = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{\sigma^2}\right)\right] , \tag{6.8}$$

where σ is the standard derivation. The orientation of the Gabor function is obtained as $\theta = \text{atan}(\frac{v}{u})$.

Let ω denote the angular frequency, $2\pi u$ and $2\pi v$ can thus be replaced by $\omega 1$ and $\omega 2$. If the research emphasis lies on Gabor wavelets, ω should be substituted by $\frac{c}{\sigma}$ with c being a constant [18]. The reason is that Gabor filter does not change its shape when σ varies while c is fixed. Thus, all Gabor filters with the same value c can be derived from one filter by scaling. To satisfy the wavelet property, DC response must be subtracted from the Gabor filter. Hence, the impulse response of such a 2D Gabor filter reads

$$h = \frac{1}{2\pi\sigma^2}\exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{\sigma^2}\right)\right]\left[\exp\left(i\left(\frac{c_1 x + c_2 y}{\sigma}\right)\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right] . \tag{6.9}$$

Fig. 6.2 illustrates the imaginary parts of 2D Gabor filters with orientations of 0^0, 45^0, 90^0 and 135^0. By convolving a 2D image with the Gabor filter, local phases are estimated. Based on the partial information of local phases,

Fig. 6.3: Three test images.

the original image is able to be constructed. The signal reconstruction is to estimate the original form of a signal in a functional space from partial information, i.e. the Gabor phase. It is based on an iterative procedure, which has a similar form with Gerchberg Saxton algorithm [58]. By alternatively imposing constrains in the spatial and frequency domains, the original image can be reconstructed according to the local Gabor phase in an iterative way. As for the reconstruction in a multi-scale way, a similar reconstruction process as in the case of monogenic scale-space is used.

6.1.5 Experimental Results

In this section, some experiments are presented to check the performance of image reconstruction based on the local phase vector in the monogenic scale-space. Three images (lena, bird and circles) used for the experiment are shown in Fig. 6.3.

The performance evaluation criteria is the normalized mean square error (NMSE) which is defined as

$$NMSE = \frac{1}{N}\sum_{i=0}^{N-1}\left[\frac{s_i}{\sqrt{\frac{1}{N}\sum_{i=0}^{N-1} s_i^2}} - \frac{\widehat{s}_i}{\sqrt{\frac{1}{N}\sum_{i=0}^{N-1} \widehat{s}_i^2}}\right]^2 , \qquad (6.10)$$

where N denotes the total number of pixels in the image, s_i and $\widehat{s}_i$ indicate the signal intensity of pixel i in the original and reconstructed images, respectively. Image reconstruction of the lena image in the monogenic scale-space is illustrated in Fig. 6.4. Although pyramid structures are used for scale-space reconstruction, the results shown at different scales are scaled to the same size as the original one. The top row shows the original image and its smoothed versions at three different scales. Bottom row demonstrates the reconstructed images at different scales. The corresponding absolute error images multiplied by a factor of 10 are shown in the middle row. The left image in the bottom row is the final result, which is reconstructed by a coarse

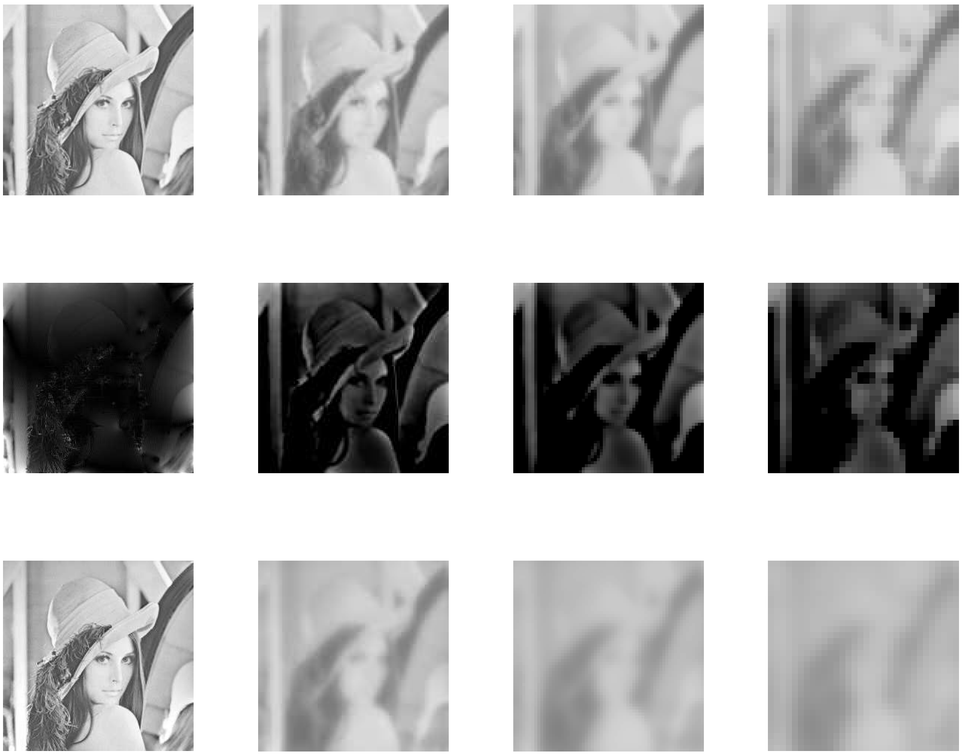

Fig. 6.4: Top row: the original image and its smoothed versions of three different scales. Middle row: absolute error images (multiplied by a factor of 10) at different scales. Bottom row: reconstructed images at different scales.

to fine way. The final reconstruction has a NMSE of 0.0018 when compared with the original one. This demonstrates that image reconstruction can be implemented accurately from the local phase vector. For the bird image, the reconstructed results are illustrated in Fig. 6.5, the corresponding NMSE is 0.005. A successful reconstruction from partial information requires a stable output. To investigate the performance of reconstruction from the local phase vector, another experiment is conducted by adding noise to contaminate the input images and checking the outputs. Fig. 6.6 demonstrates the reconstruction for noisy images. Additive Gaussian noise with the signal noise ratio (SNR) of 3 and 15 is added to the lena and bird images, respectively. The corresponding reconstructed results indicate that the reconstruction based on the monogenic phase in the scale-space is still stable to some extent. For a quantitative measurement, the reconstruction under the noise environment is investigated in depth. Additive Gaussian noise with SNR from zero to 20 is added to both the lena and bird images, the corresponding NMSEs are

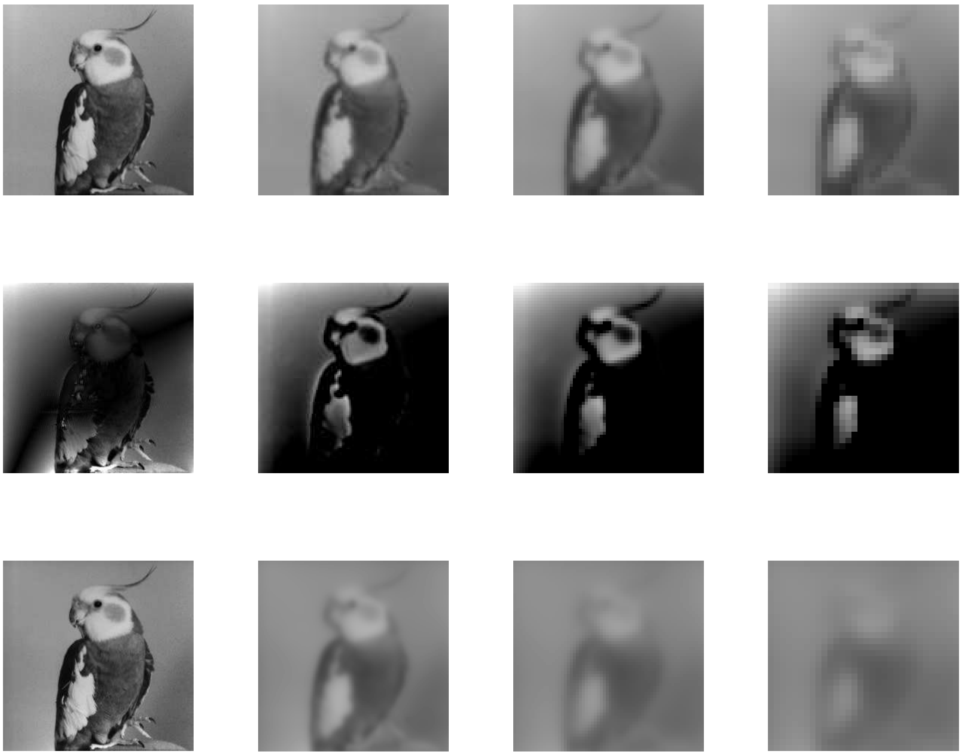

Fig. 6.5: Top row: the original image and its smoothed versions of three different scales. Middle row: absolute error images (multiplied by a factor of 10) at different scales. Bottom row: reconstructed images at different scales.

measured to plot the relationship between the reconstructed results and the added noise, see Fig. 6.7. The NMSEs increase when the signal noise ratio is reduced. However, for both cases, the proposed approach results in limited reconstruction errors even the SNR is set to zero. The results indicate that reconstruction based on the local phase vector is a stable process, hence, the local phase vector can be regarded as a stable representation of the original signal. In contrast to this, reconstruction from only zero crossings is proved to produce unstable results [57], unless the gradient data along the zero crossings are combined for reconstruction.

There is no common filter set in the Gaussian framework to evaluate the phase and orientation simultaneously. However, phase information can be estimated when the Gaussian kernel is replaced by the Gabor filter. To show the advantage of the proposed approach, we compare the results of the proposed method with that of the Gabor phase based case. A certain orientation must be assigned to the Gabor filter beforehand. In this case,

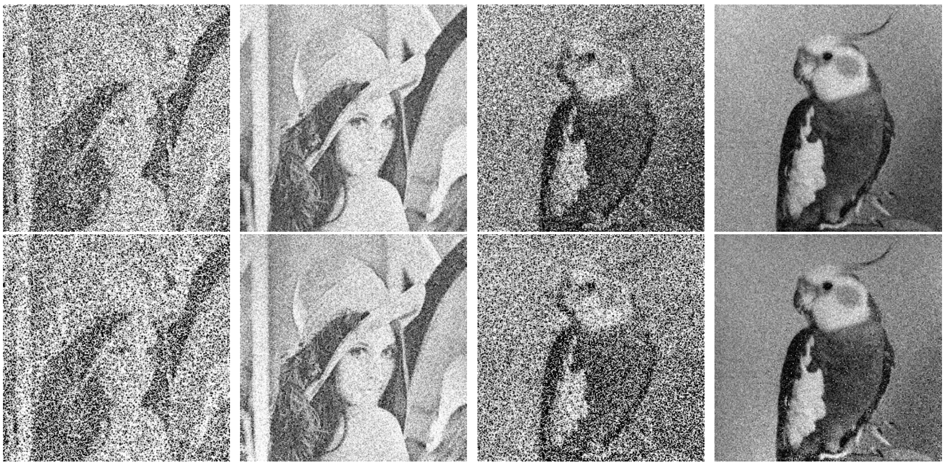

Fig. 6.6: Top row: noise contaminated original images. The first and third columns are images added with Gaussian additive noise (SNR=3). The second and fourth columns are images added with Gaussian additive noise (SNR=15). Bottom row: reconstructed results for the noise contaminated images.

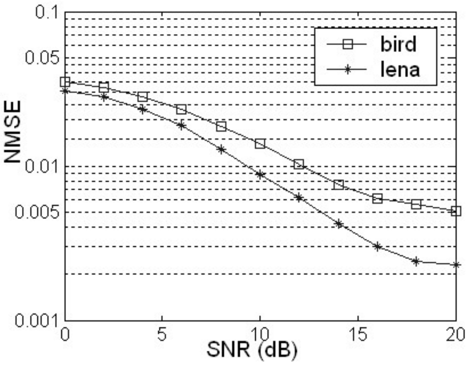

Fig. 6.7: Normalized mean square error with respect to the signal noise ratio.

the orientation is independent with the scale-space, local orientation estimation does not change when the scale is changed. Superior to the Gabor phase, the monogenic phase vector enables the estimation of structural and geometric information simultaneously at each scale-space. In the monogenic scale-space, local phase vector and local attenuation form a Riesz triplet, which means that the amplitude can be easily recovered from the local phase vector simply by using the inverse Riesz transform. Unfortunately, the Gabor phase and the local amplitude do not have such relationship. Hereby, we have to employ an iterative algorithm to reconstruct the image based on

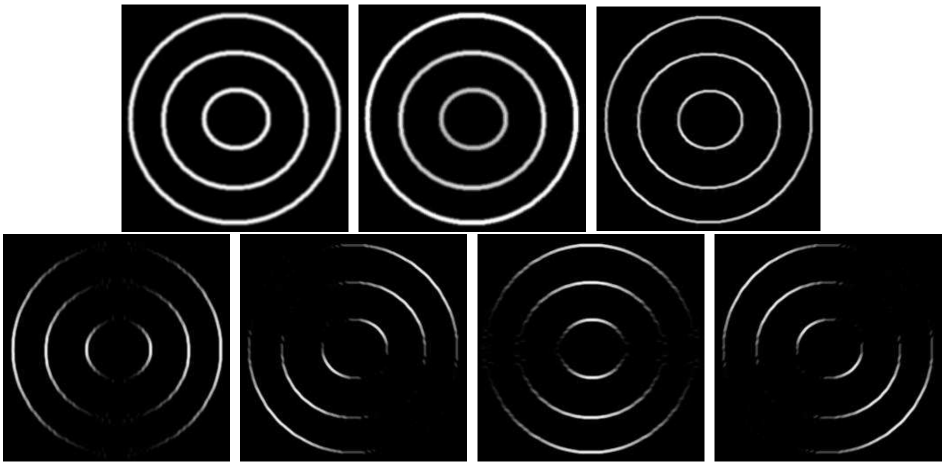

Fig. 6.8: Top row: from left to right are the original image, the reconstructed result based on the monogenic phase and the sum of the reconstructions from Gabor phases with orientations of 0, 45^0, 90^0 and 135^0. The corresponding NMSEs are 0.0014 and 0.0132. Bottom row: from left to right are reconstructed results based on Gabor phases with orientations of 0, 45^0, 90^0 and 135^0, respectively. The corresponding NMSEs are 0.0812, 0.0833, 0.0815 and 0.0836.

local Gabor phases. The iterative reconstruction procedure is similar to the Gerchberg Saxton algorithm [58]. By alternatively imposing constrains in the spatial and frequency domains, an image could be reconstructed in an iterative way. The comparison results are illustrated in Fig. 6.8, four channels with orientations of 0^0, 45^0, 90^0 and 135^0 are considered, the corresponding normalized mean square errors are 0.0812, 0.0833, 0.0815 and 0.0836, respectively. The sum of these four reconstructions shows a much better performance with a NMSE of 0.0132. It is obvious that Gabor phase only preserves the information at the given orientation, however, the monogenic phase results in an accurate and isotropic outcome with an NMSE of 0.0014. Due to the rotation invariant property of the monogenic signal, signals can be well reconstructed in the isotropic way.

6.1.6 Conclusions

A novel approach to reconstruct an image in the monogenic scale-space based on the local phase vector was presented. According to the estimated local structural and geometric information, an image can be easily and quickly reconstructed in the monogenic scale-space by a coarse to fine way. Experimental results show that accurate reconstruction is available. In contrast to the reconstruction from zero crossings, a stable reconstruction can be achieved based on the local phase vector. Furthermore, the very nice property of local orientation adaptivity can result in a much better reconstruction when compared with that of the orientation selective Gabor phase.

6.2 Detecting Intrinsically Two-dimensional Image Structures Using Local Phase

This section mainly depicts a novel approach towards detecting i2D image structures based on only local phase information. The presented method shows better performance when compared with the classic approaches.

6.2.1 Motivation and Background

Local image structures play important roles in many computer vision tasks. They can be associated with the term intrinsic dimensionality [121], which, as a local property of multidimensional signal, expresses the number of degrees of freedom necessary to describe local structures. For 2D images, there exist three type of structures. The intrinsically zero dimensional (i0D) structures are constant signals. Intrinsically one dimensional (i1D) structures represent lines and edges. Corners, junctions, line ends, etc. are all intrinsically two dimensional (i2D) structures which all have certain degree of curvature. It is well known that these i2D structures are of high significance in object recognition, motion estimation, image retrieval, etc. Consequently, correct detection of i2D structures under image deformations is very important.

There exists a lot of work concerning the detection of i2D structures based on intensity information, see [7, 62, 43, 49, 102]. These intensity based approaches are sensitive to variations in image illumination. Hence, it is necessary to find some features of local structures which are invariant with respect to image brightness change for a robust and reliable detection. Phase is such a good candidate, which carries most essential structure information of the original signal and has the advantage of being invariant to illumination variation [85]. Detecting local structures can be realized by means of the phase congruency. Using phase congruency to detect edges has been reported in [67, 38]. However, i2D structure detection based on its local phase has not yet been well investigated, although Kovesi proposed to use i1D local phase to detect i2D points by constructing the phase moments [68].

In this section, we will present a novel approach to detect i2D image structures using local phase information. The local phase of the i2D structure is derived from a curvature tensor and its conjugate part in a rotationally

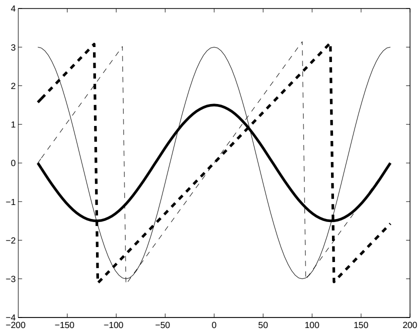

Fig. 6.9: Two 1D signals and their corresponding phases. The solid line indicates the signal, its phase is represented by the dashed line.

invariant way. By employing 2D damped spherical harmonics as basis functions, the local phase is unified with a scale concept. The i2D structures can be detected as points of stationary phases in this scale-space by means of the so called phase congruency. Experimental results illustrate that the proposed approach outperforms Harris and Susan detectors under illumination change and noise contamination.

6.2.2 Phase Congruency

Since the local phase is independent of the local amplitude, it thus has the advantage of being not sensitive to illumination change. Hence, detecting i2D image structures can be done by looking for points of stationary phases in the scale-space. This approach is commonly called phase congruency and is based on comparisons of the local phase at certain distinct scales [93, 67]. Phase congruency is a very appealing concept for general feature detection because it permits detection independent of the actual feature type. The reason is that phase is constant or congruent over some scales at the location of what the human visual system would perceive. This means those feature points are all in phase as shown in Fig. 6.9. There exist two 1D signals (indicated by the solid lines) at different scales and their corresponding phases (represented with the dashed lines). It is obvious that at those places where the signal peak, dip, increasing slop and decreasing slop exist, phases are all the same for each type of feature.

For the phase congruency, we take a similar idea as those reported in [67, 68]. However, there are some differences. First, our local phase information can be evaluated in a rotation-invariant manner. Therefore, no orientation sampling is required. Second, the local phase directly indicates the phase information of the i2D structure. Thus, there is no need to construct principal moments of the phase congruency to determine i2D structures.

Morrone and Owens [80] define the phase congruency function in terms of the Fourier series expansion of a signal at a local position $\mathbf{x}$ as

$$PC = \max_{\overline{\phi}\in(0,2\pi]} \frac{\sum_n A_n \cos(\phi_n - \overline{\phi})}{\sum_n A_n} , \tag{6.11}$$

where A_n represents the amplitude of the nth Fourier component, ϕ_n denotes the local phase of the Fourier component at position $\mathbf{x}$ and $\overline{\phi}$ is the amplitude weighted mean local phase angle of all the Fourier terms at the position being considered. The measure has a value between zero and one. A phase congruency of value one means that there is an edge or a line, zero phase congruency indicates there is no structure. However, this measure results in poor localization and is also sensitive to noise. Hence, Kovesi [67] developed a modified version of the phase congruency. In this measure, the local phase is obtained from the logarithmic Gabor wavelet. Due to its lack of rotation invariance, orientation sampling must be employed to make sure that features at all possible orientations are treated equally. Hence, the new measure of phase congruency reads

$$PC = \frac{\sum_o \sum_n W_o \lfloor A_{no}(\cos(\phi_{no} - \overline{\phi}_o) - |\sin(\phi_{no} - \overline{\phi}_o)|) - T_o \rfloor}{\sum_o \sum_n A_{no} + \varepsilon} , \tag{6.12}$$

where n and o refer to the scale parameter and the index over orientations, respectively. And W_o denotes a factor that weights for frequency spread along certain orientation and ε is added to avoid division by zero. The terms A_{no} and ϕ_{no} are the local amplitude and local phase at a certain scale and orientation, respectively. The mean local phase at a certain orientation is represented as $\overline{\phi}_o$. Only energy values that exceed the estimated noise influence T_o can be taken into consideration. The symbols $\lfloor$ and $\rfloor$ indicate that the enclosed entity equals itself when its value is positive and zero otherwise. This new phase congruency measure produces a more localized response and it also incorporates noise compensation. However, the estimated local phase is only valid for the i1D signal. Hence, using phase congruency to detect i2D structures requires the construction of principal moments of the phase congruency [68].

According to the classical moment analysis, the following entities are computed based on the phase congruency of i1D structures:

$$a = \sum (PC(\theta)\cos(\theta))^2 \tag{6.13}$$

$$b = 2\sum (PC(\theta)\cos(\theta))(PC(\theta)\sin(\theta)) \tag{6.14}$$

$$c = \sum (PC(\theta)\sin(\theta))^2 , \tag{6.15}$$

where $PC(\theta)$ refers to the phase congruency value determined at the orientation θ, and the sum is performed over the discrete set of all the orientations used. The maximum and minimum moments are then obtained as

$$\begin{aligned} M &= \frac{1}{2}\left(c+a+\sqrt{b^2+(a-c)^2}\right) && (6.16)\\ m &= \frac{1}{2}\left(c+a-\sqrt{b^2+(a-c)^2}\right)\ . && (6.17) \end{aligned}$$

If the minimum moment is bigger than a given threshold, then the corresponding point is indicated as the i2D structure.

In contrast to this, we have now a rotationally invariant evaluation of the local phase for the i2D structure, no orientation sampling is needed. Hence, the computation of phase congruency can be simplified as the following

$$PC = \frac{\sum_s W\lfloor A_s(\cos(\phi_s - \overline{\phi}) - |\sin(\phi_s - \overline{\phi})| - T)\rfloor}{\sum_s A_s + \varepsilon}\ , \qquad (6.18)$$

where s denotes the scale parameter, W is also a factor weighting for frequency spread, A_s and ϕ_s represent the local amplitude and local phase of the i2D structure point, respectively. This new measure can be directly applied to detect i2D image structures. Any point with a phase congruency value higher than a certain threshold can be considered as an i2D point.

The above mentioned phase congruency approaches are based on comparisons of the local phase at certain distinct scales. Nevertheless, there exist some drawbacks. Since local features are scale relative, an algorithm using distinct scales has to contain heuristics to judge whether the structure is present or not if the phase is only congruent in some of the considered scales. Besides, it is not straightforward, how to map at different scales estimated phases to a certainty measure. Hence, a new method called the differential phase congruency was proposed [38] to detect i1D structures in a more simple and efficient way. Those points in the monogenic scale-space, where the differentials of their phase vectors are zeros, are called points of differential phase congruency and then identified as i1D structures.

Let $\mathbf{r}(\mathbf{x}; s)$ be the local phase vector of a monogenic scale-space representation, see also Chapter 3, its scale derivative reads

$$\partial_s \mathbf{r}(\mathbf{x}; s) = \frac{p(\mathbf{x}; s)\partial_s \mathbf{q}(\mathbf{x}; s) - \mathbf{q}(\mathbf{x}; s)\partial_s p(\mathbf{x}; s)}{p(\mathbf{x}; s)^2 + |\mathbf{q}(\mathbf{x}; s)|^2}\ . \qquad (6.19)$$

Points where $\partial_s \mathbf{r}(\mathbf{x}; s) = 0$ are of differential phase congruency and hence considered as i1D structures. To find those points, one has to find the zeros of the two components of the numerator in Eq. (6.19). By a linear regression, these zeros are easily obtained with subpixel accuracy. The differential phase congruency is quite useful since it yields a higher accuracy and a significant speedup of the derivative computation compared to a finite difference approximation.

Following this approach, it is very natural to extend our approach of detecting i2D structures by means of differential phase congruency of i2D structures. Let $\mathbf{\Phi}(\mathbf{x}; s)$ be the i2D phase representation in the generalized monogenic curvature scale-space. Points where $\partial_s \mathbf{\Phi}(\mathbf{x}; s) = 0$ are of differential phase congruency and regarded as i2D structures. In comparison to Eq. (6.19), the definition now has some differences

$$\partial_s \mathbf{\Phi}(\mathbf{x}; s) = \frac{\mathbf{d}_e(\mathbf{x}; s)\partial_s \mathbf{d}_o(\mathbf{x}; s) - \mathbf{d}_o(\mathbf{x}; s)\partial_s \mathbf{d}_e(\mathbf{x}; s)}{\mathbf{d}_e(\mathbf{x}; s)^2 + |\mathbf{d}_o(\mathbf{x}; s)|^2} = 0 \quad \text{if} \quad \frac{\partial \theta}{\partial s} = 0 \, , \tag{6.20}$$

where θ indicates the local main orientation. The corresponding components are replaced with the i2D entities. In this way, i2D structure points can be detected by finding the zeros of the two components of the numerator in the above equation.

6.2.3 Performance Evaluation Criteria

In the literature, many detectors are designed for detecting i2D image structures. However, most of them show only qualitative experimental results. Because computer vision tasks require more robust and reliable detection results, there has been an increasing emphasis on quantitative performance evaluation. There also exists a number of research for assessing the detector performance. The measure suggested by Schmid et al. [100] is based on the idea of repeatability. Rockett [95] and Martinez-Fonte et al. [75] proposed a more empirical method for accessing. In their research, examples of true corners and non-corners are provided. For each threshold level, the corner detection probability and the false alarm rate are estimated to plot an ROC curve. In [23], Carneiro et al. assessed the detector performance by two measures, namely, the precision and recall rates.

The repeatability evaluation delivers the number of points repeated between two images with respect to the total number of detected points. However, this measure does not consider those correctly or wrongly detected points which do not repeat at all. The ROC curve plots the relation between the detection rate and false alarm rate with respect to the threshold variation, but it is not easy to show the detection performance with respect to image deformations like illumination change, rotation change and so on. In this thesis, we follow the measures in [23].

The recall rate measures the probability of finding an i2D point in a de-

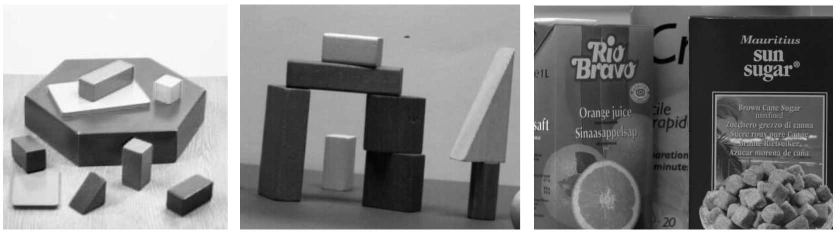

Fig. 6.10: Two test images (blox and blocks) and one frame of a boxes image sequence.

formed image given that it is detected in the reference image. The definition of the recall rate is given by

$$R = \frac{TP}{TP + FN} , \tag{6.21}$$

where TP denotes the true positive and FN is the false negative. Since it is not easy to identify the ground truth, in this case, the true positive means the number of correctly matched points. Given a point $\mathbf{x}_i$ in the reference image and a point $\mathbf{x}_j$ in the deformed image, let $M(\cdot)$ represent the deformation transform, if the Euclidean norm condition is satisfied, i.e. $\|M(\mathbf{x}_i) - \mathbf{x}_j\| < 1.5$, then these two points are correctly matched. False negative is the number of points in the reference image which cannot be matched with any points in the deformed image.

The precision rate indicates the probability that an i2D point detected in a deformed image is actually an i2D point in the reference image. Its definition reads

$$P = \frac{TP}{TP + FP} , \tag{6.22}$$

where FP is false positive, it means the number of points in the deformed image which cannot be matched with any points in the reference image. Both the recall and precision rates have values between zero and one. If the rate is higher, the detection performance is better.

6.2.4 Experimental Results

In this section, some experimental results are presented. As shown in Fig. 6.10, two test images and one image sequence are employed for the experiments. The first experiment aims to illustrate some qualitative comparison results between our approach and the well-known Harris detector. The blox image is used for detection under the rotation change and the additive Gaussian noise contamination (standard derivation is 10). For the illumination change, we use the blocks image to show the detection difference. Fig. 6.11 demonstrates the detection results of our approach and the Harris detector under the rotation change, the noise contamination and the illumination change. According to the false positives and false negatives, it can be shown that our approach performs better than the Harris detector when the illumination changes and the noise is added to some degree.

The second experiment is to show some quantitative comparison results. We follow the evaluation criteria of recall and precision rates to compare the performances of the i2D differential phase congruency method, the improved i2D phase congruency, the phase moment approach [68], the Harris detector and also the well-known Susan detector. Ten frames of the boxes image sequence are employed for this experiment. Image deformations of rotation change, additive Gaussian noise contamination and the illumination variation are considered. For each deformation, the averaged values of ten frames are recorded to plot the recall and precision rates. Fig. 6.12 demonstrates comparison results between these five approaches according to the performance assessment criteria of recall and precision rates. Here, **Normal PC** means the improved i2D phase congruency, **Differential PC** represents the i2D differential phase congruency and **Phase Moment** is the approach proposed in [68]. Note that recall and precision rates have different scales for different image deformations.

The top row shows detection results under the rotation change. **Normal PC** and **Differential PC** have comparable results with the Harris detector, **Phase Moment** performs worse than these three approaches and the Susan detector gives the worst result. It is obvious that all approaches produce lower recall and precision rates at the rotation angles of 45^0 and 135^0. This is because of the bigger discretization errors at these two angles. The second row are recall and precision rates for the illumination change. The phase congruency is a dimensionless quantity which is in theory invariant to the illumination change, although it is not absolutely invariant to brightness variation in practice, it is still less sensitive to the illumination variation than those intensity based approaches. Results indicate that phase based approaches perform much better than the Harris and Susan detectors especially in the case of higher illumination change. Due to the advantage of differential phase congruency, **Differential PC** is more robust with respect to brightness variation. Bottom row shows the additive Gaussian noise contaminated results. Since the phase congruency takes several scales into consideration and it also incorporates noise compensation, phase based approach demonstrates a better performance than that of the Harris detector. **Differential PC** is more sensitive to the noise compared with **Normal PC** because of the computation of scale derivative. And the Harris detector is less sensitive to the noise when compared with that of the Susan detector due to the Gaussian smoothing in the local neighborhood.

6.2.5 Conclusions

A novel approach towards detecting i2D image structures using local phase information was presented. The local phase of the i2D structure can be derived from a curvature tensor and its conjugate part in a rotation invariant manner. The i2D image structures are detected as those points with stationary phases in the scale-space by means of the improved phase congruency and differential phase congruency. The recall and precision rates are employed as detection performance assessment criteria. Experimental results illustrate that our approach outperforms the phase moment approach, the Harris and Susan detectors when the illumination changes and the images are contaminated by the additive Gaussian noise. For the deformation of rotation change, our approach shows a comparable result with the Harris detector.

6.3 Optical Flow Estimation from the Monogenic Curvature Tensor

This section describes a novel optical flow estimation approach based on the local phase information derived from the monogenic curvature tensor. By replacing the intensity information with the phase in the constancy assumption, the estimated optical flow fields demonstrate very good performance of this new method. Compared with the intensity based approach, it performs outstandingly under the illumination change situation.

6.3.1 Motivation and Background

Optical flow estimation is one of the key problems gathering the interest of researchers for decades in the computer vision community. It has a wide application in motion estimation, object recognition, tracking, surveillance and so on.

Various approaches have been proposed to estimate the optical flow. Significant improvements [16, 88] have been obtained since the pioneering work of Horn and Schunck [56] and Lucas and Kanade [74]. In [6], Barron et al. made the performance evaluation of optical flow techniques. The local phase-based method [40] was proven to be the best performed due to its subpixel accuracy and its robustness with respect to smooth contrast changes and

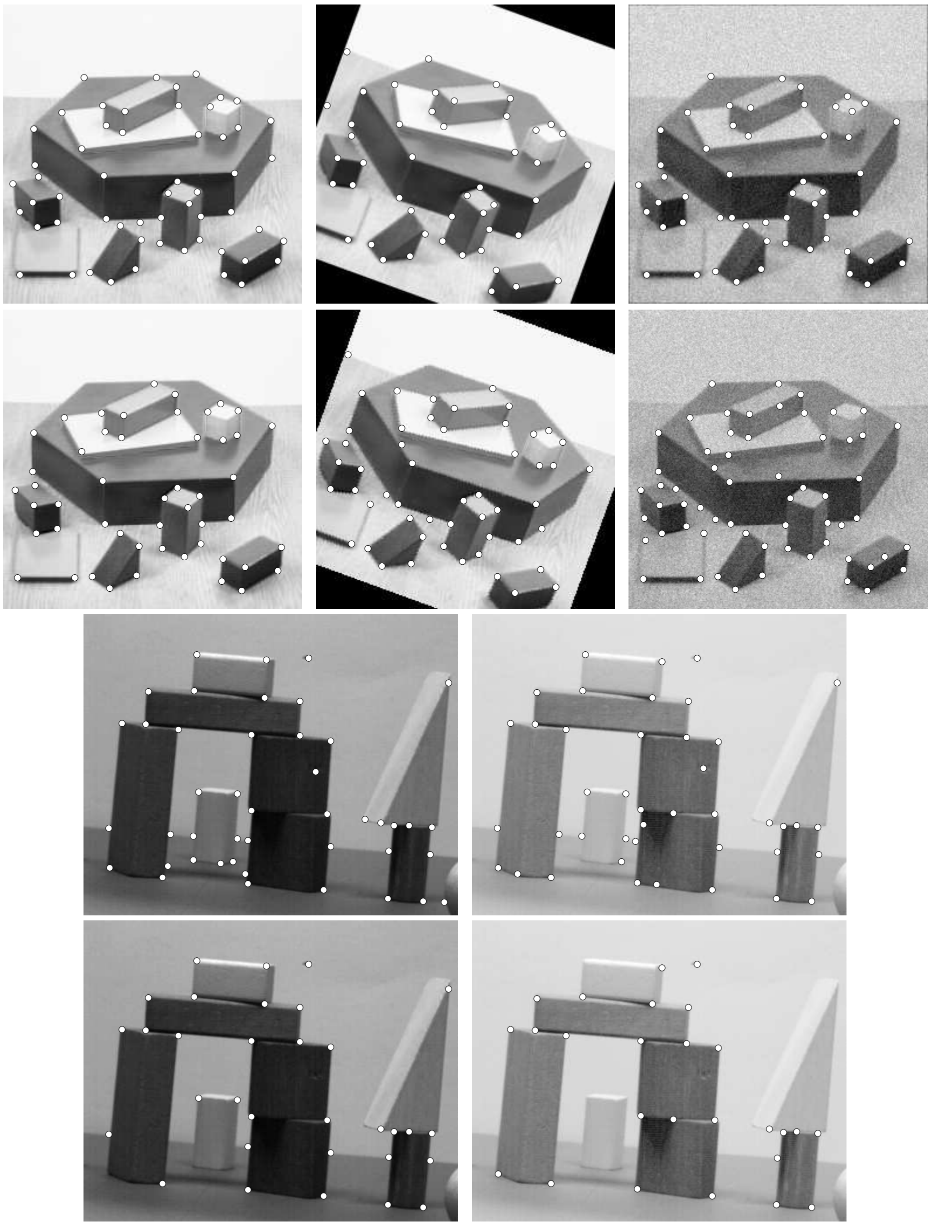

Fig. 6.11: Top row shows the detection results using our approach for the original image, the rotated image and the Gaussian noise contaminated image. The second row demonstrates the results from the Harris detector for the original image, the rotated one and the noise contaminated one. Results shown in the third row are detections for the original image and the illumination varied one by using our approach. Bottom row illustrates results from the Harris detector for the original image and the illumination changed one.

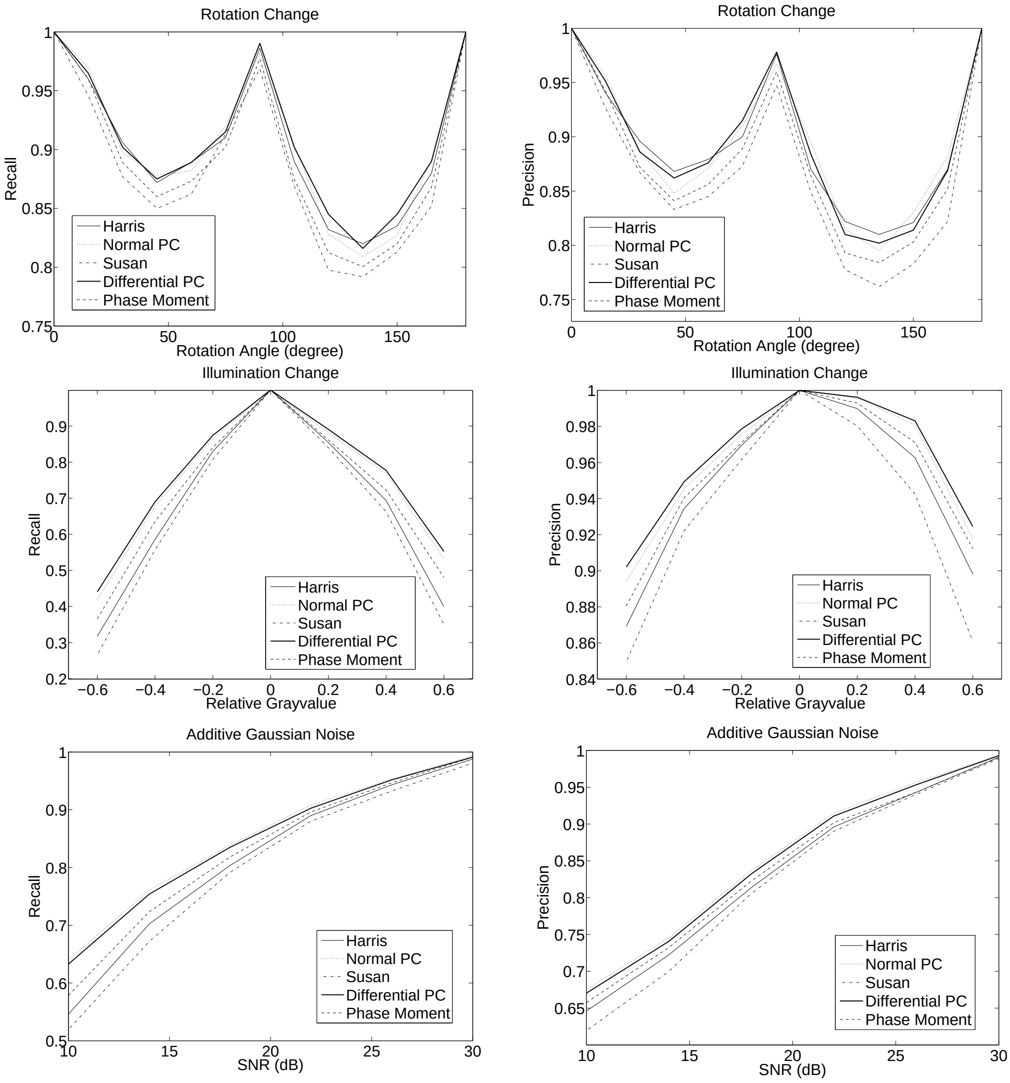

Fig. 6.12: First column: from top to bottom are recall rates under the rotation change, illumination variation and the additive Gaussian noise contamination. Second column: from top to bottom are precision rates under the rotation change, illumination variation and the additive Gaussian noise contamination.

affine deformations. Differential methods, on the other hand, have become the most frequently used techniques for optical flow estimation because of the simplicity and good performance. Among the differential methods, there exist two classes. They are local methods such as that of the Lucas and Kanada and global methods such as that of the Horn and Schunk. Local methods are known to be more robust under noise, while global approaches yield 100% dense flow fields. Hence, Bruhn et al. [17] proposed the combined local-global (CLG) approach to yield dense optical flow fields which is robust against noise.

In order to have accurate and robust estimation of dense optical flow fields against noise and brightness variation, we propose a novel approach based on the monogenic curvature tensor, a new image model. In contrast to the classical phase computation in [40], the monogenic curvature tensor can generate multi-scale local phases of image structures in a rotation invariant way. Thus, the proposed approach combines the advantages of the phase-based method and the CLG method. Experiments with synthetic and real image sequences demonstrate the favorable performance of the proposed method when compared with the related work.

6.3.2 Local Amplitudes and Phases of Image Structures

As mentioned in Chapter 5, the monogenic curvature tensor is a novel 2D image model, from which multi-scale local phases of image structures can be obtained in a rotation invariant way. It is well known that the phase has the advantage of being invariant to the illumination change [85]. In this application, we will adapt the CLG method to this framework for the dense optical flow estimation.

Similar as the Hessian matrix, we are able to compute the trace and determinant of the curvature tensor T_e and its harmonic conjugate part T_o for knowing the existence of the intrinsically one dimensional (i1D) and intrinsically two dimensional (i2D) structures.

Consequently, a novel model for the i1D structures is obtained by combing the traces of T_e and T_o, this is exactly the monogenic scale-space, as proposed in [38]

$$\mathbf{f}_{i1D}(\mathbf{x}; s) = \mathrm{trace}(T_e(\mathbf{x}; s)) + \mathrm{trace}(T_o(\mathbf{x}; s))\mathbf{e}_2 \ . \tag{6.23}$$

Hence, the multi-scale local amplitude and local phase vector for i1D structures are given by

$$a(\mathbf{x}; s) = \sqrt{\mathrm{trace}^2(T_e(\mathbf{x}; s)) + \mathrm{trace}^2(T_o(\mathbf{x}; s))\mathbf{e}_2} \tag{6.24}$$

$$\boldsymbol{\varphi}(\mathbf{x};s) = \frac{\mathrm{trace}(T_o(\mathbf{x};s))\mathbf{e}_2}{|\mathrm{trace}(T_o(\mathbf{x};s))\mathbf{e}_2|}\mathrm{atan}\left(\frac{|\mathrm{trace}(T_o(\mathbf{x};s))\mathbf{e}_2|}{\mathrm{trace}(T_e(\mathbf{x};s))}\right) , \qquad (6.25)$$

where $\frac{\mathrm{trace}(T_o(\mathbf{x};s))\mathbf{e}_2}{|\mathrm{trace}(T_o(\mathbf{x};s))\mathbf{e}_2|}$ denotes the local orientation of the i1D structure.

Correspondingly, combing the determinants of T_e and T_o results in a novel model for the i2D structure, which is called the generalized monogenic curvature scale-space $\mathbf{f}_{i2D}(\mathbf{x};s)$,

$$\mathbf{f}_{i2D}(\mathbf{x};s) = \det(T_e(\mathbf{x};s))\mathbf{e}_3 + \mathbf{e}_1\det(T_o(\mathbf{x};s)) . \qquad (6.26)$$

From it, the local amplitude for the i2D structure is obtained as

$$A(\mathbf{x};s) = \sqrt{\det{}^2(T_e(\mathbf{x};s))\mathbf{e}_3 + \mathbf{e}_1\det{}^2(T_o(\mathbf{x};s))} \qquad (6.27)$$

and the local phase vector takes the following form

$$\Phi(\mathbf{x};s) = \frac{\det(\mathbf{e}_1 T_o(\mathbf{x};s))}{|\det(\mathbf{e}_1 T_o(\mathbf{x};s))|}\mathrm{atan}\left(\frac{|\mathbf{e}_1\det(T_o(\mathbf{x};s))|}{\det(T_e(\mathbf{x};s))\mathbf{e}_3}\right) , \qquad (6.28)$$

where $\frac{\det(\mathbf{e}_1 T_o(\mathbf{x};s))}{|\det(\mathbf{e}_1 T_o(\mathbf{x};s))|}$ decides the local main orientation of the i2D structure.

Since the local phase information of the i1D and i2D structures contains not only phase information but also the local orientation, the evaluation can be done in a rotation-invariant way.

6.3.3 Dense Optical Flow Estimation

Differential methods have become the most widely used techniques for optical flow computation. By combining the advantages of local methods and global methods, Bruhn et al. [17] proposed a new method (CLG), which could yield flow fields with 100% density and have the robustness against noise. Since the phase-based approach was shown to perform very good with the advantage of being robust against brightness change [40, 6], it is very natural to combine the advantages of the phase-based approach and the CLG method. In this section, we will adapt the CLG method into our model framework to estimate two-frame optical flow fields.

2D Combined Local-Global (CLG) Method

Many differential methods are based on the assumption that the grey values of image sequences $\mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2 + t\mathbf{e}_3) = f_1\mathbf{e}_1 + f_2\mathbf{e}_2 + f_3\mathbf{e}_3$ in subsequent frames do not change over time

$$\mathbf{f}((x+u)\mathbf{e}_1 + (y+v)\mathbf{e}_2 + (t+1)\mathbf{e}_3) = \mathbf{f}(x\mathbf{e}_1 + y\mathbf{e}_2 + t\mathbf{e}_3) , \qquad (6.29)$$

where the displacement field $\mathbf{u} = u\mathbf{e}_1 + v\mathbf{e}_2$ denotes the optical flow. Following this, the spatial CLG method aims to minimize an energy function for estimating the flow field

$$E(\mathbf{w}) = \int_\Omega (\psi_1(\mathbf{w}^T J_\rho(\nabla_3 \mathbf{f})\mathbf{w}) + \alpha\psi_2(|\nabla \mathbf{w}|^2))dxdy \tag{6.30}$$

with

$$\begin{aligned}
\nabla &= \mathbf{e}_1\partial_x + \mathbf{e}_2\partial_y && (6.31)\\
\nabla_3 &= \mathbf{e}_1\partial_x + \mathbf{e}_2\partial_y + \mathbf{e}_3\partial_t && (6.32)\\
\mathbf{w} :&= u\mathbf{e}_1 + v\mathbf{e}_2 + \mathbf{e}_3 && (6.33)\\
|\nabla \mathbf{w}|^2 :&= |\nabla u\mathbf{e}_1|^2 + |\nabla v\mathbf{e}_2|^2 && (6.34)\\
J_\rho(\nabla_3 \mathbf{f}) :&= K_\rho * (\nabla_3 \mathbf{f} \nabla_3 \mathbf{f}^T) \ , && (6.35)
\end{aligned}$$

where Ω denotes the image domain, α serves as regularization parameter, K_ρ means a Gaussian kernel with standard deviation ρ, $\psi_1(\cdot)$ and $\psi_2(\cdot)$ indicate two nonquadratic penalisers with the following form

$$\psi_i(s^2) = 2\beta_i^2 \sqrt{1 + \frac{s^2}{\beta_i^2}} \quad i \in \{1,\, 2\} \tag{6.36}$$

with β_1 and β_2 as scaling parameters to handle outliers.

For the spatial temporal CLG method, more than two image frames are considered. In such case, the 3D energy function to estimate the flow field replaces $\nabla \mathbf{w}$ with $\nabla_3 \mathbf{w}$, other components remain the same as the 2D case. Hence, the 3D CLG approach aims to minimize the following energy function

$$E(\mathbf{w}) = \int (\psi_1(\mathbf{w}^T J_\rho(\nabla_3 \mathbf{f})\mathbf{w}) + \alpha\psi_2(|\nabla_3 \mathbf{w}|^2))dxdydt \ . \tag{6.37}$$

New Energy function with Phase Constraints

In order to combine the phase-based approach with the 2D CLG method, the classical brightness constancy assumption will be replaced by new phase constraints. Two local phase vectors of i1D and i2D structures can be derived from the monogenic curvature tensor, one can assume that local phases of image sequences in subsequent frames do not change over time. This results in the following new constancy assumptions

$$\begin{aligned}
\boldsymbol{\varphi}((x+u)\mathbf{e}_1 + (y+v)\mathbf{e}_2 + (t+1)\mathbf{e}_3) &= \boldsymbol{\varphi}(x\mathbf{e}_1 + y\mathbf{e}_2 + t\mathbf{e}_3) && (6.38)\\
\Phi((x+u)\mathbf{e}_1 + (y+v)\mathbf{e}_2 + (t+1)\mathbf{e}_3) &= \Phi(x\mathbf{e}_1 + y\mathbf{e}_2 + t\mathbf{e}_3) \ . && (6.39)
\end{aligned}$$

For small displacements, we may perform a first order Taylor expansion yielding the optical flow constraints:

$$\begin{aligned} \boldsymbol{\varphi}_x u\mathbf{e}_1 + \boldsymbol{\varphi}_y v\mathbf{e}_2 + \boldsymbol{\varphi}_t\mathbf{e}_3 &= 0 && (6.40)\\ \Phi_x u\mathbf{e}_1 + \Phi_y v\mathbf{e}_2 + \Phi_t\mathbf{e}_3 &= 0\,. && (6.41) \end{aligned}$$

Hence, we propose to minimize the following energy function

$$E(\mathbf{w}) = \int_\Omega (\psi_1(\mathbf{w}^T J_\rho(\nabla_3\boldsymbol{\varphi} + \gamma\nabla_3\Phi)\mathbf{w}) + \alpha\psi_2(|\nabla\mathbf{w}|^2))dxdy \tag{6.42}$$

with

$$J_\rho(\nabla_3\boldsymbol{\varphi} + \gamma\nabla_3\Phi) = K_\rho * \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \tag{6.43}$$

$$\begin{aligned} M_{11} &= |\boldsymbol{\varphi}_x|^2 + \gamma|\Phi_x|^2 && (6.44)\\ M_{12} &= M_{21} = |\boldsymbol{\varphi}_x \cdot \boldsymbol{\varphi}_y| + \gamma|\Phi_x \cdot \Phi_y| && (6.45)\\ M_{13} &= M_{31} = |\boldsymbol{\varphi}_x \cdot \boldsymbol{\varphi}_t| + \gamma|\Phi_x \cdot \Phi_t| && (6.46)\\ M_{22} &= |\boldsymbol{\varphi}_y| + \gamma|\Phi_y|^2 && (6.47)\\ M_{23} &= |\boldsymbol{\varphi}_y \cdot \boldsymbol{\varphi}_t| + \gamma|\Phi_y \cdot \Phi_t| && (6.48)\\ M_{33} &= |\boldsymbol{\varphi}_t|^2 + \gamma|\Phi_t|^2\,. && (6.49) \end{aligned}$$

In this energy function, γ is employed to adjust the trade-off between the i1D and i2D structures. According to the new energy function, the minimizing flow field $\mathbf{u}$ will satisfy the following Euler-Lagrange equations

$$\begin{aligned} &\mathrm{div}(\psi_2'(|\nabla\mathbf{w}|^2)\nabla u\mathbf{e}_1) - && (6.50)\\ &\frac{1}{\alpha}\psi_1'(\mathbf{w}^T J_\rho(\nabla_3\boldsymbol{\varphi} + \gamma\nabla_3\Phi)\mathbf{w})(J_{11}u\mathbf{e}_1 + J_{12}v\mathbf{e}_2 + J_{13}\mathbf{e}_3) = 0 \\ &\mathrm{div}(\psi_2'(|\nabla\mathbf{w}|^2)\nabla v\mathbf{e}_2) - && (6.51)\\ &\frac{1}{\alpha}\psi_1'(\mathbf{w}^T J_\rho(\nabla_3\boldsymbol{\varphi} + \gamma\nabla_3\Phi)\mathbf{w})(J_{21}v\mathbf{e}_2 + J_{22}u\mathbf{e}_1 + J_{23}\mathbf{e}_3) = 0 \end{aligned}$$

with

$$\psi_i'(s^2) = \frac{1}{\sqrt{1 + \frac{s^2}{\beta_i^2}}} \quad i \in \{1,\, 2\}\,. \tag{6.52}$$

The estimation of optical flow field can thus be obtained iteratively by using an SOR [111] scheme.

Computation of Phase Derivatives

In order to avoid phase wrapping, phase derivatives are computed from the filter responses in the monogenic scale-space and generalized monogenic curvature scale-space. Let $t_e = \text{trace}(T_e(\mathbf{x};s))$, $\mathbf{t}_o = \text{trace}(T_o(\mathbf{x};s))\mathbf{e}_2$, $d_e = \det(T_e(\mathbf{x};s))$ and $\mathbf{d}_o = \mathbf{e}_1\det(T_o(\mathbf{x};s))$, the spatial derivatives of i1D and i2D local phase vectors are thus given by

$$\nabla\boldsymbol{\varphi} = \frac{t_e\nabla\mathbf{t}_o - \mathbf{t}_o^T\nabla^T t_e}{t_e^2 + |\mathbf{t}_o|^2} \tag{6.53}$$

$$\nabla\Phi = \frac{d_e\nabla\mathbf{d}_o - \mathbf{d}_o^T\nabla^T d_e}{d_e^2 + |\mathbf{d}_o|^2}\ . \tag{6.54}$$

The temporal derivatives of these local phase vectors read

$$\boldsymbol{\varphi}_t = \frac{t_e^t\mathbf{t}_o^{t+1} - t_e^{t+1}\mathbf{t}_o^t}{|t_e^t\mathbf{t}_o^{t+1} - t_e^{t+1}\mathbf{t}_o^t|}\text{atan}\left(\frac{|t_e^t\mathbf{t}_o^{t+1} - t_e^{t+1}\mathbf{t}_o^t|}{t_e^t t_e^{t+1} + \mathbf{t}_o^t\cdot\mathbf{t}_o^{t+1}}\right) \tag{6.55}$$

$$\Phi_t = \frac{d_e^t\mathbf{d}_o^{t+1} - d_e^{t+1}\mathbf{d}_o^t}{|d_e^t\mathbf{d}_o^{t+1} - d_e^{t+1}\mathbf{d}_o^t|}\text{atan}\left(\frac{|d_e^t\mathbf{d}_o^{t+1} - d_e^{t+1}\mathbf{d}_o^t|}{d_e^t d_e^{t+1} + \mathbf{d}_o^t\cdot\mathbf{d}_o^{t+1}}\right)\ , \tag{6.56}$$

where t_e^t, $\mathbf{t}_o^t$, d_e^t, $\mathbf{d}_o^t$ denote the filter responses of the image frame at time t and t_e^{t+1}, $\mathbf{t}_o^{t+1}$, d_e^{t+1}, $\mathbf{d}_o^{t+1}$ are the filter responses of the next image frame.

Multi-scale Optical Flow Estimation

The linearized optical flow constraint, stated in Section 6.3.3, is based on the phase constancy assumption. As a consequence, it requires that u and v are relatively small so that the linearization holds. However, this is not always the case for an arbitrary sequence. Hence, multi-scale optical flow estimation technique should be employed to deal with large displacements.

In this thesis, we use an incremental coarse to fine strategy. In contrast to the classical multi-scale approach, the estimated flow field at a coarse level is used to warp the image sequence instead of serving as initialization for the next finer scale. This compensation results in a hierarchical modification which requires to compute only small displacement. Once this is done from coarse to the finest scale, much more accurate estimation will be obtained.

Let $d\mathbf{w}^s$ denote a displacement increment at scale s. For the coarsest scale ($s = S$), the initial data of the optical flow field is assigned to be zero. Hence, $d\mathbf{w}^s$ is given by minimizing the following energy function

$$\begin{aligned} E(d\mathbf{w}^s) &= \int_\Omega (\psi_1((d\mathbf{w}^s)^T J_\rho(\nabla_3\boldsymbol{\varphi}(\mathbf{x}+\mathbf{w}^s) + \gamma\nabla_3\Phi(\mathbf{x}+\mathbf{w}^s))d\mathbf{w}^s) \\ &+ \alpha\psi_2(|\nabla(\mathbf{w}+d\mathbf{w}^s)|^2))dxdy\ , \end{aligned} \tag{6.57}$$

where $\mathbf{x} = x\mathbf{e}_1 + y\mathbf{e}_2 + t\mathbf{e}_3$ and $\mathbf{w}^{s+1} = \mathbf{w}^s + d\mathbf{w}^s$. Note that local phase vectors will be warped as $\boldsymbol{\varphi}(\mathbf{x}+\mathbf{w}^s)$ and $\Phi(\mathbf{x}+\mathbf{w}^s)$ via bilinear interpolation. The final result will be obtained when the minimization is done to the finest scale.

6.3.4 Experimental Results

In order to evaluate the performance of the proposed approach, optical flow estimations on both synthetic and real-world image data are given in this section. We use the so-called average angular error (AAE) [6] as the quantitative quality measure. Given the estimated flow field $\mathbf{u}_e = u_e\mathbf{e}_1 + v_e\mathbf{e}_2$ and the ground truth $\mathbf{u}_c = u_c\mathbf{e}_1 + v_c\mathbf{e}_2$, the AAE is defined as

$$AAE = \frac{1}{N}\sum_{i=1}^{N} \arccos\left(\frac{u_{ci}u_{ei} + v_{ci}v_{ei} + 1}{\sqrt{(u_{ci}^2 + v_{ci}^2 + 1)(u_{ei}^2 + v_{ei}^2 + 1)}}\right) , \tag{6.58}$$

where N denotes the total number of pixels.

The *Yosemite* sequence with clouds created by *Lynn Quam* [92] is employed as the synthetic data for the experiment. This sequence combines divergent and translational motion under varying illumination and hence is usually regarded as a benchmark for the optical flow estimation. Fig. 6.13 demonstrates the ground truth, the estimated magnitudes and optical flow fields from our approach and the 2D CLG model. It is obvious that our approach produces more accurate result than that of the 2D CLG method. Especially in the clouds region, where the illumination varies, the proposed approach shows more stable estimation. Even if we compare with 2D CLG where the intensity is replaced by the gradient, our approach also performs better.

Detail comparisons with other approaches according to the measurement AAE are given in Tab. 6.1, where STD indicates the standard deviation. Our approach demonstrates much better performance with lower AAE and STD when compared with the related methods. When $\gamma = 0$, only i1D phase information is included for the constraint, the AAE now takes 3.37^0, which is 1.49^0 lower than that of the 2D CLG method. Interestingly, this result is even lower than that of the 3D CLG method. When γ is set to non-zero, the i2D phase is also contained to strengthen the constraint. One optimal value of γ is 0.1, which is obtained from several experiments. Hence, estimation with even lower error can be obtained. For this experiment, we also extend the two-frame estimation to multi-frame by adding the temporal information. Results also indicate the good performance of the proposed approach.

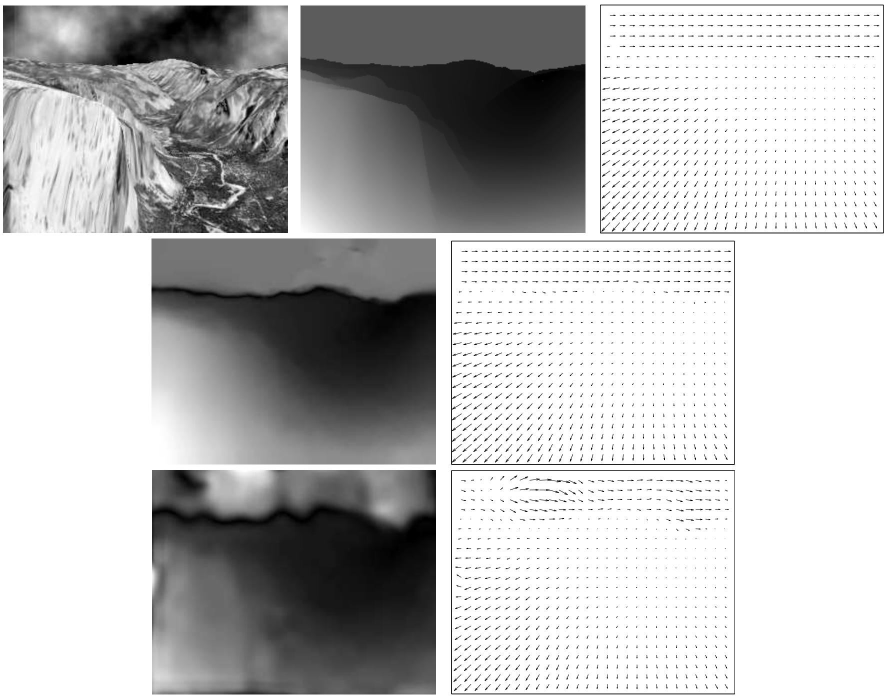

Fig. 6.13: Top row: from left to right are one frame of the Yosemite image sequence, the magnitude and flow field of the ground truth. Middle row: from left to right are the magnitude and flow field estimated from our approach. Bottom row: from left to right are the magnitude and flow field estimated from the 2D CLG method.

Even much better results have been reported in [16, 88]. However, they do not perform a first order Taylor expansion of the intensity assumption to yield the optical flow constraint. Thus, it is very promising that our approach can also yield comparably good results by using the non-linearized constancy assumption.

For the following experiments, we simply focus on two-frame flow field estimation. To investigate the robustness of our approach against noise, the 8th frame of the *Yosemite* sequence is degraded with additive Gaussian noise.

The noise contaminated image (signal noise ratio: SNR=10dB) and the estimated flow field are shown in Fig. 6.14. It is obvious that the original image is seriously degraded, nevertheless, the estimation also shows good performance with AAE=14.16^0. More detail information can be found in Fig. 6.15. When the SNR decreases from 40dB to 10dB, much more noise is added to the original image. However, the estimated result is still not very sensitive to noise. This indicates that employing the local method and

Tab. 6.1: Optical flow estimation comparisons between different approaches (100% density). AAE (average angular error), STD (standard deviation).

Approach	AAE	STD
Horn/Schunck (Barron et al., 1994) [6]	31.69^0	31.18^0
Nagel (Barron et al., 1994) [6]	10.22^0	16.51^0
Uras et al. (Barron et al., 1994) [6]	8.94^0	15.61^0
2D CLG (2005) [17]	4.86^0	8.48^0
Mémin and Pérez (1998) [77]	4.69^0	6.89^0
3D CLG (2005) [17]	4.17^0	7.72^0
Our 2D approach ($\gamma = 0$)	3.37^0	8.27^0
Our 2D approach ($\gamma = 0.1$)	3.25^0	8.22^0
Our 3D approach ($\gamma = 0$)	2.74^0	7.17^0
Our 3D approach ($\gamma = 0.1$)	2.67^0	7.12^0

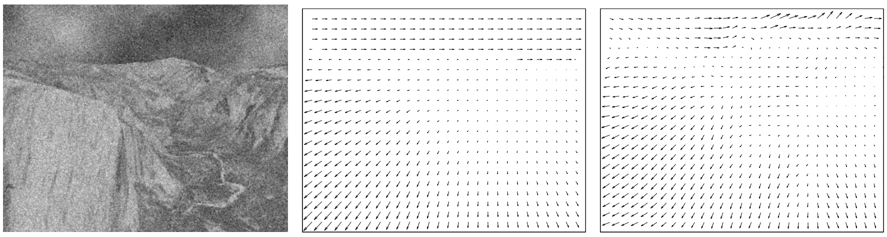

Fig. 6.14: From left to right are the noise degraded image (SNR=10dB), ground truth and the estimated flow field (AAE=14.16^0, STD=12.76^0).

multi-scale technique into our approach does result in a robust estimation against noise.

As mentioned in [40], the phase-based approach has the advantage of being not sensitive to the illumination variation. And the proposed approach adapts the CLG method into the framework of the monogenic curvature tensor. As a consequence, this new method combines the advantages of phase-based approach and the CLG method. In this way, our approach should also be robust under illumination change within some limits. To this end, another experiment is conducted to test the performance of our approach for the brightness variation. Fig. 6.16 shows the performance comparison between our approach and the 2D CLG method under brightness change. The 8th frame of the synthetic sequence is degraded with brighter and darker illumination changes of 50%, respectively. Experimental results denote that

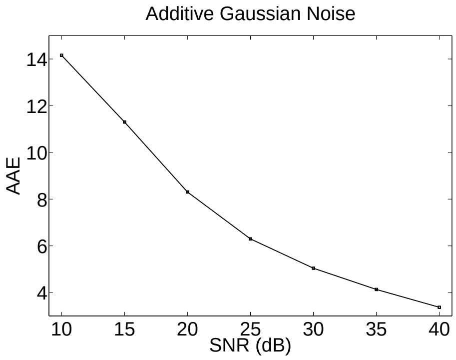

Fig. 6.15: The estimated results with respect to additive Gaussian noise change.

our approach is much more robust against illumination variation when compared with that of the 2D CLG method. To evaluate the performance of

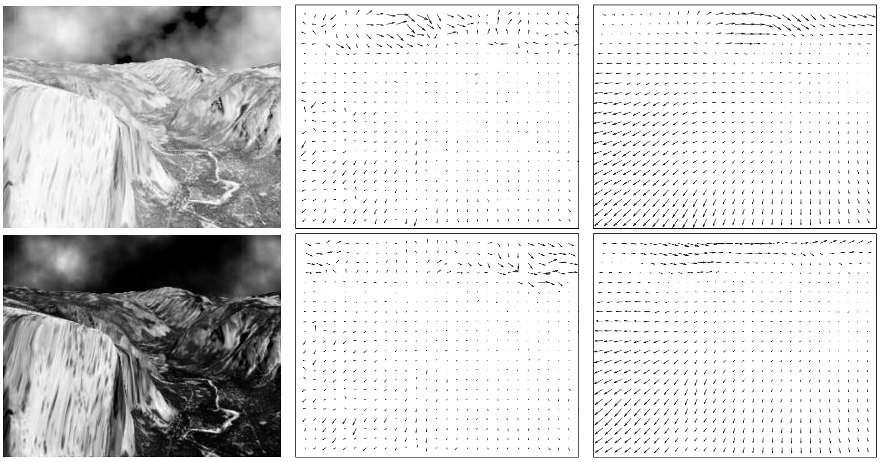

Fig. 6.16: Top row: from left to right are frame 8 degraded with brighter illumination change of 50%, estimated flow fields from the 2D CLG (AAE=46.94^0, STD=39.97^0) and our approach (AAE=13.50^0, STD=17.42^0). Bottom row: from left to right are frame 8 degraded with darker illumination variation of 50%, optical flow estimations from the 2D CLG (AAE=52.14^0, STD=46.63^0) and our method (AAE=15.83^0, STD=19.71^0).

the proposed approach under illumination change in detail, the 8th frame are degraded with different brighter and darker brightness variations. The estimated AAEs with respect to the relative grayvalue changes are shown in Fig. 6.17.

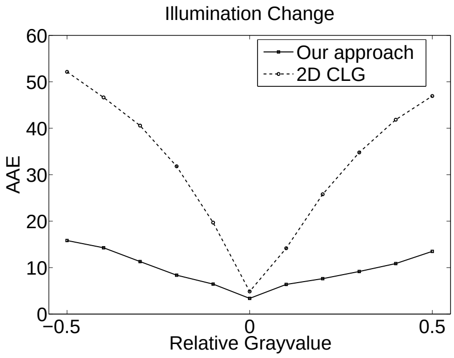

Fig. 6.17: The estimated results with respect to illumination change.

Results indicate that our approach is very robust against illumination change. However, the 2D CLG method is very sensitive to it.

The last experiment aims to examine the performance of the proposed approach for real image sequences. In this experiment, two sequences are used. They are the well-known Hamburg taxi sequence and the Ettlinger Tor traffic sequence [84]. Estimated flow fields are illustrated in Fig. 6.18, it is clear that the proposed approach also yields realistic optical flow for real-world data.

6.3.5 Conclusions

We have presented a novel approach for estimating two-frame dense optical flow field in this section. This new approach adapts the CLG approach to the monogenic curvature tensor, a new framework which enables multi-scale local phase evaluation of i1D and i2D image structures in a rotation invariant way. Hence, our approach takes both the advantages of phase-based approach and the CLG approach. In this way, the proposed method produces accurate estimations with 100% density and is robust against noise. Compared with the intensity based approach, our method performs much better under illumination variation.

6.4 Summary

In this chapter, some scale-space based image analysis were presented to demonstrate the applications of the newly developed theoretic framework. They include the phase-based image reconstruction, detecting i2D image structures using local phase and optical flow estimation from the monogenic curvature tensor. Since the remarkable advantage of the proposed image

Fig. 6.18: Top row: one frame of the Hamburg taxi sequence and the estimated optical flow from our approach. Bottom row: one frame of the Ettlinger Tor traffic sequence and the flow field from our approach.

model lies in the phase evaluation of local image structures. All of these phase based applications show much better performance when compared with the related work.

7. CONCLUSIONS

In this chapter, a summary of the proposed theoretic framework and some applications will be given. Open problems and future possible work will also be discussed.

7.1 Summary

This thesis aims to introduce novel signal representations for image analysis in the scale-space. In the past literatures, there exists a bunch of work for representing image information based on intensity information. Unfortunately, most of these intensity based approaches suffer seriously from the illumination change and hence perform usually unstable. Phase information, on the other hand, has the advantage of being invariant to brightness variation. As a consequence, it is desirable to take both the intensity and phase information into consideration to develop novel image representation with robust performance. In the past decades, there exists also some research work related to this purpose. For example, the analytic signal, quadrature filter, partial and total Hilbert transform, quaternionic analytic signal, monogenic signal and so on. However, these previously presented approaches either lack the property of rotation invariance or have the drawback of being able to handle all 2D image structures.

Combing the rotation invariance of the monogenic signal and the symmetric decomposition of the quaternionic analytic signal, a rotationally invariant 2D quadrature filter was first proposed as a novel image representation to extract the local amplitudes and phases of the signal. The assumed signal model is valid for the i2D signal superimposed by two differently oriented i1D signals. This means that the two i1D signals are not necessarily to be perpendicular to each other. Hence, it is regarded as an extension of the structure multivector, since the i2D pattern with a flexible angle of intersection can be correctly handled. However, this approach requires a previously estimated orientation to steer the corresponding quadrature filter.

Consequently, a more powerful image representation framework was proposed. This novel image model is rotationally invariant and is able to handle all 2D structures within a multi-scale framework. It is called the monogenic curvature tensor, consisting of the curvature tensor and its harmonic conjugate part. Coupling methods of differential geometry, tensor algebra, monogenic signal and quadrature filter, the monogenic curvature tensor can be derived. Motivated from the second fundamental theorem of the differential geometry, the curvature tensor is constructed from even order spherical harmonics using a so-called hybrid matrix geometric algebra. The corresponding harmonic conjugate part is obtained by employing the Riesz transform. From the monogenic curvature tensor, the monogenic scale-space can be obtained by computing the trace as the i1D signal model. Combining determinants of the curvature tensor and its conjugate part results in a generalized monogenic curvature scale-space. Interestingly, the odd part of it can be obtained from the even part via the convolution with the second order spherical harmonic. Hence, the second order spherical harmonic can be interpreted as a generalization of the Hilbert transform for the i2D case. From the generalized monogenic curvature scale-space, local amplitude, orientation and phase are extracted in a multi-scale way. Hence, the property of split of identity is completely fulfilled. Compared with the related work, this novel image representation has remarkable advantage of rotationally invariant estimation of local phase information for all 2D image structures. This delivers access to many phase-based computer vision applications.

Several computer vision applications were presented in this thesis. Experiential results demonstrate the favorable performance of the presented framework when compared with the related work.

7.2 Outlook

The newly developed monogenic curvature tensor is a powerful representation of the 2D image. However, the current investigation ignores the metric tensor, which should also be taken into consideration for further analysis. Besides, the monogenic scale-space and the generalized monogenic curvature scale-space are derived by combing the traces and determinants of the curvature tensor and its conjugate part. The component-wise analysis of the monogenic curvature tensor should be explored since it may contain much rich information of local image structures. Furthermore, elements of the monogenic curvature tensor may be fed into an artificial neural network to find out more information via different connections of those components. The presented framework has the capability to handle all 2D image structures, however, to model 3D images, further extension is needed. To this end, the construction of the curvature tensor should be based on the Riemann tensor, its conjugate part is also gained from the Riesz transform. The above mentioned aspects are all open problems and can be further developed in the future for more advanced image analysis.

APPENDIX

.1 Eigenvalues of Commutative Hypercomplex Matrices

Commutative hypercomplex algebras (**HCA**) are the hypercomplex algebras which have similar construction rules as the Clifford algebra but with commutative multiplication rules.

In [25], Davenport proposed the four dimensional commutative hypercomplex algebras ($\mathbf{HCA}_4$). It was shown that $\mathbf{HCA}_4$ are isomorphic to the two-fold tensor product and the Cartesian product of the complex algebras $\mathbb{C}^2$. Later on, Felsberg et al. [33] extended the **HCA** to any 2^n dimensions and proved that the 2^n-dimensional **HCA** are isomorphic to $\mathbb{C}^{2^{n-1}}$. The four dimensional commutative hypercomplex algebras are also isomorphic to the reduced biquaternions which are used for the signal processing [101, 89].

A four dimensional commutative hypercomplex number is defined as

$$z = z_1 + z_2 i + z_3 j + z_4 k = (z_1 + z_2 i) + (z_3 + z_4 i) j \ , \tag{0.1}$$

where i, j, k are unit basis which obey the following rules,

$$ij = ji = k,\ jk = kj = i,\ ik = ki = -j,\ i^2 = k^2 = -1,\ j^2 = 1 \ . \tag{0.2}$$

In contrast to the definition in [25], only the square of j is chosen to be positive one, which results in more simple computations.

According to the above rules, z can be reformulated as

$$\begin{aligned} z &= \{(z_1 + z_3) + i(z_2 + z_4)\} \left[\frac{1+j}{2}\right] + \{(z_1 - z_3) + i(z_2 - z_4)\} \left[\frac{1-j}{2}\right] \\ &= \xi r_1 + \eta r_2 \end{aligned} \tag{0.3}$$

with

$$\xi = (z_1 + z_3) + i(z_2 + z_4) \tag{0.4}$$

$$\eta = (z_1 - z_3) + i(z_2 - z_4) \tag{0.5}$$

$$r_1 = \frac{1+j}{2} \tag{0.6}$$

$$r_2 = \frac{1-j}{2} \ . \tag{0.7}$$

Correspondingly, a four dimensional commutative hypercomplex matrix $M_{\mathbf{HCA}_4}$ can be represented as the linear composition of two complex matrices with the following form

$$M_{\mathbf{HCA}_4} = M_\xi r_1 + M_\eta r_2 \ . \tag{0.8}$$

For an $n \times n$ $M_{\mathbf{HCA}_4}$ matrix, there are n eigenvalues of each of the corresponding complex matrix. Hence, there are n^2 eigenvalues for this $M_{\mathbf{HCA}_4}$ matrix.

In the following, we restrict ourselves to the eigenvalues computation of a 2×2 matrix. Let $\lambda_{\xi 1}$, $\lambda_{\xi 2}$, $\lambda_{\eta 1}$ and $\lambda_{\eta 2}$ be the eigenvalues of complex matrices M_ξ and M_η respectively, then the eigenvalues of the commutative hypercomplex matrix are given by

$$\begin{aligned} \lambda_1 &= \lambda_{\xi 1} r_1 + \lambda_{\eta 1} r_2 \\ \lambda_2 &= \lambda_{\xi 1} r_1 + \lambda_{\eta 2} r_2 \\ \lambda_3 &= \lambda_{\xi 2} r_1 + \lambda_{\eta 1} r_2 \\ \lambda_4 &= \lambda_{\xi 2} r_1 + \lambda_{\eta 2} r_2 \ . \end{aligned} \tag{0.9}$$

For example, given an 2×2 $M_{\mathbf{HCA}_4}$ matrix

$$\begin{aligned} M_{\mathbf{HCA}_4} &= \begin{bmatrix} 1 & i \\ j & k \end{bmatrix} = \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & i \end{bmatrix} j = \begin{bmatrix} 1 & i \\ 1 & i \end{bmatrix} r_1 + \begin{bmatrix} 1 & i \\ -1 & -i \end{bmatrix} r_2 \\ &= M_\xi r_1 + M_\eta r_2 \ . \end{aligned} \tag{0.10}$$

The eigenvalues of matrices M_ξ and M_η are respectively obtained as follows

$$\lambda_{\xi 1} = 0,\ \lambda_{\xi 2} = 1 + i,\ \lambda_{\eta 1} = 0,\ \lambda_{\eta 2} = 1 - i \ . \tag{0.11}$$

Hence, according to Eq. (0.9), the four eigenvalues of the commutative hypercomplex matrix read

$$\lambda_1 = 0 \quad \lambda_2 = \frac{1-i}{2} - \frac{1-i}{2} j \quad \lambda_3 = \frac{1+i}{2} + \frac{1+i}{2} j \quad \lambda_4 = 1 + k \ . \tag{0.12}$$

.2 Eigenvalue Computation of the Hessian Matrix

According to the definition of the 4D commutative hypercomplex number, we can identify its unit basis i, j and k with the following basis elements in $\mathbb{R}_3$

$$i \rightarrow \mathbf{e}_{12} \quad j \rightarrow \mathbf{e}_3 \quad k \rightarrow \mathbf{e}_{123} \ . \tag{0.13}$$

The property of the unit basis based on the chosen elements can be checked as follows

$$\begin{aligned} i^2 = (\mathbf{e}_{12})^2 = k^2 = (\mathbf{e}_{123})^2 = -1 \quad j^2 = (\mathbf{e}_3)^2 = 1 \\ ij = \mathbf{e}_{12}\mathbf{e}_3 = ji = \mathbf{e}_3\mathbf{e}_{12} = \mathbf{e}_{123} = k \\ jk = \mathbf{e}_3\mathbf{e}_{123} = kj = \mathbf{e}_{123}\mathbf{e}_3 = \mathbf{e}_{12} = i \\ ik = \mathbf{e}_{12}\mathbf{e}_{123} = ki = \mathbf{e}_{123}\mathbf{e}_{12} = -\mathbf{e}_3 = -j \ . \end{aligned} \tag{0.14}$$

Hence, the Hessian matrix in the algebraically extended framework can be considered as a commutative hypercomplex matrix, that is

$$H_M = \begin{bmatrix} f_{xx}\mathbf{e}_3 & -f_{xy}\mathbf{e}_{123} \\ f_{xy}\mathbf{e}_{123} & f_{yy}\mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} f_{xx}j & -f_{xy}k \\ f_{xy}k & f_{yy}j \end{bmatrix} . \tag{0.15}$$

The above equation can further be written as

$$\begin{bmatrix} f_{xx}j & -f_{xy}k \\ f_{xy}k & f_{yy}j \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} f_{xx} & -f_{xy}i \\ f_{xy}i & f_{yy} \end{bmatrix} j . \tag{0.16}$$

Two corresponding complex matrices M_ξ and M_η are obtained as follows

$$M_\xi = \begin{bmatrix} f_{xx} & -f_{xy}i \\ f_{xy}i & f_{yy} \end{bmatrix} \quad M_\eta = \begin{bmatrix} -f_{xx} & f_{xy}i \\ -f_{xy}i & -f_{yy} \end{bmatrix} . \tag{0.17}$$

The eigenvalues of M_ξ read

$$\lambda_{\xi 1} = m \quad \lambda_{\xi 2} = n . \tag{0.18}$$

These two eigenvalues are equivalent to those of the real valued Hessian matrix H_R, that is

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \quad \text{with} \quad \det_R(H) = mn . \tag{0.19}$$

Accordingly, eigenvalues of M_η are given by

$$\lambda_{\eta 1} = -n \quad \lambda_{\eta 2} = -m . \tag{0.20}$$

Combining eigenvalues of the two matrices M_ξ and M_η, the four eigenvalues of the algebraically extended Hessian matrix are

$$\lambda_1 = \frac{m-n}{2} + \frac{m+n}{2}j \quad \lambda_2 = \frac{m-m}{2} + \frac{m+m}{2}j = mj \tag{0.21}$$

$$\lambda_3 = \frac{n-n}{2} + \frac{n+n}{2}j = nj \quad \lambda_4 = \frac{n-m}{2} + \frac{n+m}{2}j . \tag{0.22}$$

Hence, the determinant of the Hessian in the extended algebraic framework reads

$$\begin{aligned} \det_M(H_M) &= [(f_{xx}\mathbf{e}_3)(f_{yy}\mathbf{e}_3) - (-f_{xy}\mathbf{e}_{123})(f_{xy}\mathbf{e}_{123})]^2 = \left[f_{xx}f_{yy} - f_{xy}^2\right]^2 \\ &= \lambda_1\lambda_2\lambda_3\lambda_4 = (mn)^2 = (\det_R(H_M))^2 . \end{aligned} \tag{0.23}$$

It can be concluded that the determinant computation of the Hessian in this framework can be reduced to the determinant definition as the one in the vector algebra.

The corresponding trace of the Hessian equals half of the sum of eigenvalues

$$\operatorname{trace}_M(H_M) = f_{xx}\mathbf{e}_3 + f_{yy}\mathbf{e}_3 = \frac{1}{2}(\lambda_1+\lambda_2+\lambda_3+\lambda_4) = (m+n)j = (m+n)\mathbf{e}_3 . \tag{0.24}$$

BIBLIOGRAPHY

[1] T. Aach, I. Stuke, C. Mota, and E. Barth. Estimation of multiple local orientations in image signals. In *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 553–556, 2004.

[2] R. Ablamowicz. *Clifford Algebras with Numeric and Symbolic Computations*. Birkhäuser, Boston, 1996.

[3] R. Ablamowicz and B. Fauser. A maple package for Clifford algebra computations with 'bigebra', 'cliplus', 'define', 'gtp', 'octonion', Jan. 2006. `http://math.tntech.edu/rafal/cliff8/`.

[4] M. Ashdown and S. Furuta. Geometric algebra software. `http://www.mrao.cam.ac.uk/~clifford/pages/software.htm`.

[5] H. Bårman, G. H. Granlund, and H. Knutsson. A new approach to curvature estimation and description. In *Proceedings of 3rd International Conference on Image Processing and its Applications*, pages 54–58, Warwick, Great Britain, 1989. IEE.

[6] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[7] P. R. Beaudet. Rotationally invariant image operators. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 579–583, 1978.

[8] J. Behar, M. Porat, and Y. Y. Zeevi. Image reconstruction from localized phase. *IEEE Trans. on Signal Processing*, 40:736–743, 1992.

[9] D. Betten and E. Montanucci. An elementary introduction to lie groups. Technical Report 0312, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, 2003.

[10] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings of the IEEE First International Conference on Computer Vision*, pages 433–438, London, Great Britain, June 1987.

[11] R. Bracewell. *Fourier Analysis and Imaging.* Kluwer Academic / Plenum Publishers, New York, 2003.

[12] F. Brackx, R. Delanghe, and F. Sommen. *Clifford Analysis.* Pitman, Boston, 1982.

[13] F. Brackx, B. De Knock, and H. De Schepper. Generalized multi-dimensional Hilbert transforms involving spherical monogenics in the framework of Clifford analysis. In *In: Extended abstracts of the IC-NAAM 2005 Conference*, Rhodos, Greece, 2005.

[14] F. Brackx, B. De Knock, and H. De Schepper. Generalized multi-dimensional Hilbert transforms in Clifford analysis. *International Journal of Mathematics and Mathematical Sciences*, 2006:Article ID 98145, 19 pages, 2006.

[15] L. Brand. *Vector and Tensor Analysis.* John Wiley and Sons, New York, 1947.

[16] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *Proc. 8th European Conference on Computer vision - (ECCV'04)*, volume IV, pages 25–36, 2004.

[17] A. Bruhn, J. Weickert, and C. Schnörr. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *International Journal of Computer Vision*, 61(3):211–231, 2005.

[18] T. Bülow. Hypercomplex spectral signal representations for the processing and analysis of images. Technical Report No. 9903, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, 1999.

[19] T. Bülow and G. Sommer. Algebraically extended representation of multi-dimensional signals. In *Proc. of the 10th Scand. Conf. on Image Analysis*, pages 559–566, 1997.

[20] T. Bülow and G. Sommer. Hypercomplex signals - a novel extension of the analytic signal to the multidimensional case. *IEEE Transactions on Signal Processing*, 49(11):2844–2852, 2001.

[21] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. on Communications*, 31(4):532–540, April 1983.

[22] M. P. Do Carmo. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, 1976.

[23] G. Carneiro and A. D. Jepson. Multi-scale phase-based local features. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR03)*, pages 736–743, 2003.

[24] P. E. Danielsson, Q. Lin, and Q. Z. Ye. Efficient detection of second-degree variations in 2d and 3d images. *Journal of Visual Communication and Image Representation*, 12(3):255–305, 2001.

[25] C. M. Davenport. A commutative hypercomplex algebra with associated function theory. In *Clifford Algebras with Numeric and Symbolic Computations*, pages 213–227. Birkhauser Boston Inc., Cambridge, MA, USA, 1996.

[26] C. Doran, D. Hestenes, and F. Sommen. Lie groups as spin groups. *Math. Phys.*, 34(8):3642–3669, August 1993.

[27] L. Dorst, S. Mann, and T. Bouma. Gable: A matlab geometric algebra tutorial. `http://staff.science.uva.nl/~leo/clifford/gable.html`.

[28] T. Caelli E. Barth and C. Zetzsche. Image encoding, labeling and reconstruction from differiental geometry. *CVGIP: Graphical Models and Image Processing*, 55(6):428–446, 1993.

[29] R.R. Ernst et al. Two dimensional NMR spectroscopy. In *Proc. 4th Ampere Int. Summer School*, Pula, Yugoslavia, 1976.

[30] Gunnar Farnebäck. *Polynomial Expansion for Orientation and Motion Estimation.* PhD thesis, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 2002. Dissertation No 790, ISBN 91-7373-475-6.

[31] M. Felsberg. Low-level image processing with the structure multivector. Technical Report 2016, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, 2002.

[32] M. Felsberg. The GET operator. Technical Report LiTH-ISY-R-2633, Dept. EE, Linköping University, SE-581 83 Linköping, Sweden, October 2004.

[33] M. Felsberg, T. Bülow, and G. Sommer. Commutative hypercomplex Fourier transforms of multidimensional signals. In G. Sommer, editor, *Geometric Computing with Clifford Algebra*, pages 209–229. Springer-Verlag, Heidelberg, 2001.

[34] M. Felsberg and E. Jonsson. Energy tensors: Quadratic phase invariant image operators. In *DAGM 2005*, volume 3663 of *LNCS*, pages 493–500. Springer, 2005.

[35] M. Felsberg and U. Köthe. Get: The connection between monogenic scale-space and gaussian derivatives. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *LNCS*, pages 192–203. Springer, 2005.

[36] M. Felsberg and G. Sommer. The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136–3144, December 2001.

[37] M. Felsberg and G. Sommer. Image features based on a new approach to 2d rotation invariant quadrature filters. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Computer Vision - ECCV 2002*, volume 2350 of *Lecture Notes in Computer Science*, pages 369–383. Springer, 2002.

[38] M. Felsberg and G. Sommer. The monogenic scale-space: A unifying approach to phase-based image processing in scale-space. *Journal of Mathematical Imaging and Vision*, 21:5–26, 2004.

[39] D. J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Opt. Soc, Amer. A.*, A4:2379–2394, 1987.

[40] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, 1990.

[41] L. Florack. *Image Structure.* Kluwer Academic Publishers, 1997.

[42] L. M. J. Florack, B. M. ter Haar Romeny, J. J. Koenderink, and M. A. Viergever. Cartesian differential invariants in scale-space. *Math. Imaging Vision*, 3(4):327–348, 1993.

[43] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987.

[44] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

[45] D. Gabor. Theory of communication. *Journal of the IEE*, 93:429–457, 1946.

[46] G. H. Granlund. Hierarchical computer vision. In E. Masgrau L. Torres and M. A. Lagunas, editors, *Proceedings of EUSIPCO 90, Fifth European Signal Processing Conference*, pages 73–84, Barcelona, Spain, September 1990.

[47] G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision.* Kluwer Academic Publishers, Dordrecht, 1995.

[48] S. L. Hahn. *Hilbert Transforms in Signal Processing.* Artech House, Boston, London, 1996.

[49] C. Harris and M. Stephen. A combined corner and edge detector. In *Proceedings of 4th Alvey Vision Conference*, pages 147–151, Manchester, 1988.

[50] S. Haykin. *Communication Systems.* Wiley, New York, 2nd edition edition, 1987.

[51] J. H. Heinbockel. Introduction to tensor calculus and continuum mechanics, Feb. 2006. `http://www.math.odu.edu/~jhh/counter2.html`.

[52] D. Hestenes. *New Foundations for Classical Mechanics.* Reidel Publishing, Dordrecht, 1986.

[53] D. Hestenes, H. Li, and A. Rockwood. New algebraic tools for classical geometry. In G. Sommer, editor, *Geometric Computing with Clifford Algebras*, pages 3–23. Springer-Verlag, 2001.

[54] D. Hestense and G. Sobczyk. *Clifford Algebra to Geometric Calculus.* Reidel, Dordrecht, 1984.

[55] E. M. S. Hitzer and KamiWaAi. Interactive java application for three dimensional geometrical sketching, Jan. 2006. `http://sinai.mech.fukui-u.ac.jp/gcj/software/KamiWaAi/`.

[56] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[57] R. Hummel and R. Moniot. Reconstruction from zero crossing in scale space. *IEEE Trans. on Acoustic, Speech, and Signal Processing*, 37(12):2111–2130, 1989.

[58] N. E. Hurt. *Phase Retrieval and Zero Crossings Mathematical Methods in Image Reconstruction.* Kluwer Academic Publishers, 1989.

[59] B. Jähne. *Digital Image Processing.* Springer-Verlag, Berlin, 2001.

[60] R. E. Critchlow Jr. A geometric calculator, Jan. 2006. `http://www.elf.org/calculator/`.

[61] D. C. Kay. *Theory and Problems of Tensor Calculus.* Schaum's outline series. McGraw-Hill, 1988.

[62] L. Kitchen and A. Rosenfeld. Grey-level corner detection. *Pattern Recognition Letters*, pages 95–102, 1982.

[63] J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[64] J. J. Koenderink and A. J. van Doorn. Two-plus-one-dimensional differential geometry. *Pattern Recogn. Lett.*, 15(5):439–443, 1994.

[65] U. Köthe. Integrated edge and junction detection with the boundary tensor. In *Proceeding of 9th Intl. Conf. on Computer Vision*, volume 1, pages 424–431, 2003.

[66] U. Köthe and M. Felsberg. Riesz-transforms vs. derivatives: On the relationship between the boundary tensor and the energy tensor. In R. Kimmel, N. Sochen, and J. Weickert, editors, *Scale Space and PDE Mathods in Computer Vision*, volume 3459 of *LNCS*, pages 179–191. Springer, 2005.

[67] P. Kovesi. Videre: A image features from phase congruency. *Journal of Computer Vision Research*, 1(3), 1999.

[68] P. Kovesi. Phase congruency detects corners and edges. In *Proc. The Australian Pattern Recognition Society Conference*, pages 309–318, December 2003.

[69] W. Krabs. *Mathematical Foundations of Signal Theory.* Heldermann Verlag, 1995.

[70] M. Krause and G. Sommer. A 3d isotropic quadrature filter for motion estimation problems. In *Proc. Visual Communications and Image Processing, Beijing, China*, volume 5960, pages 1295–1306. The International Society for Optical Engineering, Bellingham, 2005.

[71] G. Krieger and C. Zetzsche. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Transactions on Image Processing*, 5(6):1026–1042, June 1996.

[72] N. Krüger and M. Felsberg. A continuous formulation of intrinsic dimension. In *Proc. of British Machine Vision Conference*, pages 260–270, 2003.

[73] P. Lounesto. *Clifford Algebras and Spinors.* London Mathematical Society Lecture Note Series. Cambridge University Press, 1997.

[74] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. 7th International Joint Conference on Artifical Intelligence*, pages 674–679, Vancouver, Canada, 1981.

[75] L. Martinez-Fonte, S. Gautama, and W. Philips. An empirical study on corner detection to extract buildings in very high resolution satellite images. In *Proc. of IEEE-ProRisc*, pages 288–293, November 2004.

[76] A. Mcintosh. Clifford algebras, Fourier theory, singular integrals and harmonic functions on Lipschitz domains. In J. Ryan, editor, *Clifford Algebras in Analysis and Related Topics*, Studies in Advanced Mathematics, pages 33–88. CRC Press, Boca Raton, 1996.

[77] E. Mémin and P. Pérez. A multigrid approach for hierarchical motion estimation. In *Proc. 6th International Conference on Computer Vision*, pages 933–938, Bombay, India, Jan. 1998. Narosa Publishing House.

[78] F. Mokhtarian and M. Bober. *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization.* Kluwer Academic Publishers, 2003.

[79] F. Mokhtarian and R. Suomela. Robust image corner detection through curvature scale space. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 20(12):1376–1381, 12 1998.

[80] M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313, 1987.

[81] C. Mota, T. Aach, I. Stuke, and E. Barth. Estimation of multiple orientations in multi-dimensional signals. In *Proc. of the IEEE International Conference on Image Processing*, pages 2665–2668, 2004.

[82] M. N. Nabighian. The analytic signal of two-dimensional magnetic bodies with polygonal cross-section: its properties and use for automated anomaly interpretation. *Geophysics*, 37(3):507–517, June 1972.

[83] M. N. Nabighian. Towards a three-dimensional automatic interpretation of potential field data via generalized hilbert transform: Fundamental relations. *Geophysics*, 49(6):780–786, June 1984.

[84] H.-H. Nagel. Ettlinger tor traffic sequence. Available at `http://i21www.ira.uka.de/image\_sequences/`.

[85] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *IEEE Proceedings*, 69:529–541, May 1981.

[86] R. Mikkola P. Lounesto and V. Vierros. Clical, Jan. 2006. `http://users.tkk.fi/~ppuska/mirror/Lounesto/CLICAL.htm`.

[87] A. Papaulis. *The Fourier Integral and its Application.* McGraw-Hill, New York, 1962.

[88] N. Papenberg, A. Bruhn, T. Brox, S. Didas, and J. Weickert. Highly accurate optic flow computation with theoretically justified warping. *International Journal of Computer Vision*, 67(2):141–158, 2006.

[89] S. C. Pei, J. H. Chang, and J. J. Ding. Commutative reduced biquaternions and their Fourier transform for signal and image processing applications. *IEEE Trans. on Signal Processing*, 52(7):2012–2031, 2004.

[90] C. Perwass. Clifford algebra library and utilities. `http://www.clucalc.info/`.

[91] C. Perwass and D. Hildenbrand. Aspects of geometric algebra in Euclidean, projective and conformal space. Technical Report Number 0310, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, September 2003.

[92] L. Quam. Yosemite image sequence. Available at `ftp://csd.uwo.ca/pub/vision`.

[93] D. Reisfeld. The constrained phase congruency feature detector: simultaneous localization, classification and scale determination. *Pattern Recogn. Lett.*, 17(11):1161–1169, 1996.

[94] J. Riel. Glyph-a maple package for Clifford algebra, Jan. 2006. `http://bargains.k-online.com/~joer/glyph/glyph.htm`.

[95] P. Rockett. Performance assessment of feature detection algorithms: A methodology and case study on corner detectors. *IEEE Transactions on Image Processing*, 12(12):1668–1676, 2003.

[96] C. Roran and A. Lasenby. Physical applications of geometric algebra. `http://www.mrao.cam.ac.uk/~clifford/ptIIIcourse/`.

[97] I. Rorteous. *Clifford Algebras and the Classical Groups*. Cambridge University Press, 1995.

[98] S. J. Sangwine. Fourier-transforms of color images using quaternion or hypercomplex numbers. *Electronic Letters*, 32(21):1979–1980, 1996.

[99] D. H. Sattinger and O. L. Weaver. *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*. Springer-Verlag, New York, 1993.

[100] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.

[101] H. D Schütte and J. Wenzel. Hypercomplex numbers in digital signal processing. In *Proc. IEEE Int. Symp. Circuits Syst.*, volume 2, pages 1557–1560, 1990.

[102] S. M. Smith and J. M. Brady. Susan – a new approach to low level image processing. *International Journal of Computer Vision*, 23(1):45–78, 1997.

[103] G. Sobczyk and G. Erlebacher. Hybrid matrix geometric algebra. In H. Li, P. J. Olver, and G. Sommer, editors, *Computer Algebra and Geometric Algebra with Applications*, volume 3519 of *LNCS*, pages 191–206. Springer-Verlag, Berlin Heidelberg, 2005.

[104] G. Sommer, editor. *Geometric Computing with Clifford Algebras*. Springer-Verlag, Heidelberg, 2001.

[105] G. Sommer and D. Zang. Local analysis of multi-dimensional signals. In T.E. Simos, G. Pshoyios, and Ch.Tsitouras, editors, *Internaltional Conference on Numerical Analysis and Applied Mathematics (ICNAAM)*, pages 628–629. WILEY-VCH Verlag, Weinheim, 2006.

[106] G. Sommer and D. Zang. Parity symmetry in multi-dimensional signals. *Journal of Communications on Pure and Applied Analysis*, 6(3):829–852, 2007.

[107] E. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton University Press, New Jersey, 1971.

[108] T. Tao. Lecture about harmonic analysis in the phase plane: Review of Fourier transform, phase space, uncertainty principle. `http://www.math.ucla.edu/~tao/254a.1.01w/notes1.dvi`.

[109] B. Wegman and C. Zetzsche. Statistical dependence between orientation filter outputs used in a human vision based image code. *Visual Communication and Image Processing*, SPIE 1360:909–923, 1990.

[110] B. Wegman and C. Zetzsche. Visual system based polar quantization of local amplitude and local phase of orientation filter outputs. *Human Vision and Electronic Imaging: Models, Methods and Applications*, SPIE 1249:306–317, 1990.

[111] D. M. Young. *Iterative Solution of Large Linear Systems.* Academic Press, 1971.

[112] D. Zang and G. Sommer. Phase based image reconstruction in the monogenic scale space. In C.E. Rasmussen, H.H. Bülthoff, M.A. Giese, and B. Schölkopf, editors, *26. Symposium für Mustererkennung, DAGM 2004, Tübingen*, volume 3175 of *LNCS*, pages 171–178. Springer-Verlag, Berlin Heidelberg, 2004.

[113] D. Zang and G. Sommer. An operator for the analysis of superimposed intrinsically two dimensional patterns. In *12th International Workshop on Systems, Signals and Image Processing, Chalkida, Greece, 22.-24.9.05*, pages 77–81, 2005.

[114] D. Zang and G. Sommer. Algebraically extended 2d image representation. In *Proceedings (digital) of the 17th International Conference on the Applications of Computer Science and Mathematics in Architecture and Civil Engineering, IKM, Weimar*, 2006.

[115] D. Zang and G. Sommer. The monogenic curvature scale-space. In R. Reulke, U. Eckardt, B. Flach, U. Knauer, and K. Polthier, editors, *11th International Workshop on Combinatorial Image Analysis, IWCIA'06, Berlin*, volume 4040 of *LNCS*, pages 320–332. Springer-Verlag, Berlin, Heidelberg, 2006.

[116] D. Zang and G. Sommer. A novel representation for two-dimensional image structures. Technical Report Number 0606, Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, June 2006.

[117] D. Zang and G. Sommer. Signal modeling for two-dimensional image structures. *Journal of Visual Communication and Image Representation*, 18(1):81–99, 2 2007.

[118] D. Zang, L. Wietzke, and G. Sommer. Dense optical flow estimation from the monogenic curvature tensor. In *Proc. of International Conference on Scale-space and Variational Methods*, 2007. accepted.

[119] Di Zang and Gerald Sommer. Detecting intrinsically two-dimensional image structures using local phase. In Katrin Franke, Klaus-Robert Müller, Bertram Nickolay, and Ralf Schäfer, editors, *Pattern Recognition, 28th DAGM Symposium*, volume 4174 of *Lecture Notes in Computer Science*, pages 222–231. Springer, 2006.

[120] C. Zetzsche. Statistical properties of the representation of natural images at different levels in the visual system. *Perception*, 17:359, 1988.

[121] C. Zetzsche and E. Barth. Foundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–1117, 1990.

[122] C. Zetzsche and E. Barth. Image surface predicates and the neural encoding of two-dimensional signal variation. In B. Rogowitz, editor, *Proc. SPIE Human Vision and Electronic Imaging: Models, Methods and Applications*, volume 1249, pages 160–177, 1990.

[123] C. Zetzsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. B. Watson, editor, *Digital Images and Human Vision*, pages 107–138. MA: MIT Press, Cambridge, 1993.