INAUGURAL – DISSERTATION
zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der
Ruprecht - Karls - Universität
Heidelberg

Vorgelegt von
Diplom-Mathematiker Pavel Pavlov
aus Dobrich, Bulgarien
Tag der mündlichen Prüfung
25.09.2008

Analysis of motion in scale space

Gutachter:
Prof. Dr. Dr. h.c. mult. Willi Jäger
Prof. Dr. Dr. Bernd Jähne

Abstract

This work includes some new aspects of motion estimation by the optic flow method in scale spaces. The usual techniques for motion estimation are limited to the application of coarse to fine strategies. The coarse to fine strategies can be successful only if there is enough information in every scale. In this work we investigate the motion estimation in the scale space more basically.

The wavelet choice for scale space decomposition of image sequences is discussed in the first part of this work. We make use of the continuous wavelet transform with rotationally symmetric wavelets. Bandpass decomposed sequences allow the replacement of the structure tensor by the phase invariant energy operator. The structure tensor is computationally more expensive because of its spatial or spatio-temporal averaging. The energy operator needs in general no further averaging. The numerical accuracy of the motion estimation with the energy operator is compared to the results of usual techniques, based on the structure tensor. The comparison tests are performed on synthetic and real life sequences.

Another practical contribution is the accuracy measurement for motion estimation by adaptive smoothed tensor fields. The adaptive smoothing relies on nonlinear anisotropic diffusion with discontinuity and curvature preservation. We reached an accuracy gain under properly chosen parameters for the diffusion filter.

A theoretical contribution from mathematical point of view is a new discontinuity and curvature preserving regularization for motion estimation. The convergence of solutions for the isotropic case of the nonlocal partial differential equation is shown.

For large displacements between two consecutive frames the optic flow method is systematically corrupted because of the violence of the sampling theorem. We developed a new method for motion analysis by scale decomposition, which allows to circumvent the systematic corruption without using the coarse to fine strategy. The underlying assumption is, that in a certain neighborhood the grey value undergoes the same displacement. If this is fulfilled, then the same optic flow should be measured in all scales. If there arise inconsistencies in a pixel across the scale space, so they can be detected and the scales containing this inconsistencies are not taken into account.

Zusammenfassung

Diese Arbeit beinhaltet einige neue Aspekte zur Bewegungsbestimmung mittels der Methode des optischen Flusses in Skalenräumen. Übliche Ansätze für die Bewegungsanalyse beschränken sich auf die Anwendung von grob-zu-fein Strategien, die aber nur erfolgreich sein kann, wenn in allen Skalen genügend Information vorhanden ist. In dieser Arbeit wird die Bewegunganalyse im Skalenraum grundlegender untersucht.

Der erste Teil der Arbeit beschäftigt sich mit der Frage, welche Wavelets sich für die Skalenzerlegung bei Bildfolgen am besten eignen. Wir setzen die kontinuierliche Wavelettransformation mit rotationsinvariante Wavelets ein. Bandpasszerlegungen bieten die Möglichkeit

den Strukturtensor, der wegen der notwendigen räumlichen Mittelung recht rechenaufwändig ist, durch den phasenunabhängigen Energieoperator zu ersetzen, der keine weitere Mittelung benötigt. Die numerische Genauigkeit der Bewegungsschätzung im Vergleich zu den Standardverfahren, die auf dem Strukturtensor basieren, wird im Detail anhand von synthetischen und realen Bildfolgen untersucht.

Ein weiterer praktischer Beitrag ist die Messung der Genauigkeit von Bewegungsschätzung aus adaptiv geglätteten Tensorfeldern. Die adaptive Glättung beruht auf nichtlinearer anisotropier Diffusion mit Unstetigkeits- und Krümmungserhaltung. Dabei wurde eine Steigerung der Genauigkeit erreicht, wenn die Parameter für die Diffusionsfilter korrekt eingestellt werden.

Theoretisch mathematischer Natur ist eine neuartige Regularisierung für Bewegungsschätzung mit Unstetigkeits- und Krümmungserhaltung. Die Konvergenz von Lösungen für den isotropen Spezialfall der nichtlokalen partiellen Differentialgleichung wird gezeigt.

Die Methode des optischen Flusses wird bei hohen Verschiebungen von Bild zu Bild wegen der Verletzung des Abtasttheorems systematisch verfälscht. Die in dieser Arbeit entwickelte Skalenzerlegung der Bewegungsanalyse erlaubt einen neuen Ansatz der es gestattet ohne die übliche Grob-Fein-Strategie, diese systematischen Fehler zu vermeiden unter der Annahme, dass in einer Nachbarschaft die gleiche Bewegung stattfindet auf allen Skalenebenen stattfinden muss. Treten bei der Bewegungsanalyse Inkonsistenten über der Skala auf, so können detektiert und die entsprechenden Skalen aus der Synthese der Bewegungsschätzung über alle Skalen eliminiert werden.

Acknowledgements

Prof. Dr. Dr. h.c. mult. Willi Jäger was the supervisor of this thesis. The help of Prof. Jäger and discussions on PDE models and mathematical problems in image processing with him is gratefully acknowledged.

I would like to thank Prof. Dr. Dr. Bernd Jähne for his support and help and discussions on filter design, motion estimation and image processing.

I thank the whole team of the Digital Image Processing and Multidimensional Image Processing Group at the IWR, University of Heidelberg and especially to the following members.

Thanks go to Christoph Garbe for his help and cooperation within the LOCOMOTOR project and an external industry project and for discussions on optic flow, motion estimation, multigrid computation, inverse problems, image and signal processing, detection of temporal aliasing, the list can be arbitrary long...

I would like to thank Martin Schmidt for his support in all software and programming issues, cooperation within the LOCOMOTOR project and an external industry project, discussions on image processing, filter design, software development, programming languages, interpolation issues, time-frequency signal analysis, ...

I thank Marc Kirchner and Michael Kelm for the computer administration and ideas concerning grid models, camera calibration and signal detection.

My thanks go to Björn Menze and Markus Jehle for administrating the group server and discussions on optic flow, regularization, domain decomposition for variational fluid flow estimation, visualization and classification of image data.

My thanks go to Michael Klar for discussions and hints on camera calibration and recommendation of literature. Michael Klar did estimation of the cross positions in our test sequences with his grid locator software, which is gratefully acknowledged.

Special thanks to Achim Falkenroth for his support, corrections, suggestions and recommendations on this work.

I thank Prof. Fred Hamprecht for organizing seminars on signal processing and inviting me to participate in and for discussions on wavelets, time-frequency signal analysis, interpolation issues and optimal sampling grids.

I would like to express my gratitude to Ralf Schulz for conversations and some hints on visualization of tensor fields, superquadrics and reconstruction in laser medical imaging.

Special thanks are committed to Claudia and Daniel Kondermann for discussions and ideas on variational motion estimation, PDE solvers, multigrid methods and confidence measures.

I would like to express my gratitude to all the LOCOMOTOR cooperation partners from the University of Frankfurt, Forschungszentrum Jülich and University of Lübeck and in particular to the following collaborators.

I thank Prof. Rudolf Mester for his help and cooperation within the LOCOMOTOR project and hints concerning a probabilistic approach to large displacement estimation and algebraic classification of multiple orientation estimation.

I am grateful to Prof. Til Aach for his help, supervision and management of research within the LOCOMOTOR project.

I would like to thank Erhardt Barth for discussions on wavelets, signal processing, filter design, optic flow and multiple orientation estimation.

My thanks go to Kai Krajsek for fruitful discussions on optic flow, motion estimation, optimal and adaptive integration of the structure tensor, kriging, regularization issues, steerable filters, Lie groups, warping, ...

I thank Hanno Scharr for filter optimization with respect to wavelet decomposition and for computing the state of the art rotationally invariant filters for optical flow and motion estimation, for discussions on PDE, filter optimization, motion estimation, camera calibration, depth from multiple views,...

I would like to thank Ingo Stuke for providing the multiple motion algorithm, discussions on multiple motion and multiple orientation analysis.

I thank Cicero Mota for discussions on wavelet analysis, shereable and steerable filters.

The participants in the research initiative *Mathematical methods for time series analysis and digital image processing SPP 1114* of the *German Science Foundation* helped me too and I would like to express my gratitude to the following participants.

I thank Stephan Didas for discussions on optical flow, scale spaces and PDE methods in image processing.

I am grateful to Henning Thielemann for exchanging ideas on wavelet analysis and programming languages.

I would like to thank Mark Droske and Tobias Preusser for discussions on PDE methods and mathematical problems in image processing.

My thanks go to Per-Erik Forssen and Klas Nordberg for discussions on computer vision and image processing at the First International Workshop on Complex Motion 2004 in Reisensburg and the LOCOMOTOR Winter Workshop 2005 in Frankfurt.

I thank Hermann Lauer, Markus Riedinger and Stefan Riedel for administrating the compute servers of the IWR, University of Heidelberg.

I would like to thank Anette Braun, Barbara Werner, Karin Kubessa-Nasri, Gabriela Schocke and Elke Pietschmann for their help and assistance.

This research was funded by the *German Science Foundation*. The financial support by the *German Science Foundation* through the LOCOMOTOR project within the research initiative *Mathematical methods for time series analysis and digital image processing SPP 1114* is gratefully acknowledged.

Heidelberg, 5th October 2008

Pavel Pavlov

Danksagung

Herr Prof. Dr. Dr. h.c. mult. Willi Jäger war der Betreuer dieser Arbeit. Ich bedanke mich beim Herrn Prof. Jäger für seine Hilfe und Diskussionen über partielle Differentialgleichungen, mathematische Modelle und Probleme in der digitalen Bildverarbeitung.

Ich danke Herrn Prof. Dr. Dr. Bernd Jähne für seine Hilfe, Unterstützung und Diskussionen über Bildverarbeitung, Filterdesign und Bewegungsschätzung.

Ich danke allen Mitgliedern der Arbeitsgruppe Digitale Bildverarbeitung und Multidimensionale Bidverarbeitung am Interdisziplinären Zentrum für wissenschaftliches Rechnen der Ruprecht-Karls-Universität Heidelberg und insbesondere folgenden Mitarbeitern.

Ich bedanke mich bei Christoph Garbe for seine Hilfe und Zusammenarbeit am LOCOMOTOR-Projekt und an einem externen Industrieprojekt, sowie für Diskussionen und Ideen über optischen Fluss, Bewegungsschätzung, Mehrgitterverfahren, inverse Probleme, Bild- und Signalverarbeitung, Detektion von Aliasing in der Zeitrichtung, ...

Ich möchte mich ganz herzlich bei Martin Schmidt für seine Unterstützung in allen Fragen auf dem Gebiet der Softwareentwicklung und Programmierung bedanken, für seine Zusammenarbeit am LOCOMOTOR-Projekt und an einem externen Industrieprojekt, sowie für Besprechungen und Diskussionen über Bildverarbeitung, Filterdesign, Softwaredesign, Programmiersprachen, Interpolation, Zeit-Frequenz Signalanalyse, ...

Ich bedanke mich bei Marc Kirchner und Michael Kelm für das Administrieren der seven summits und Tipps bezüglich Gittermodellierung, Kamerkalibration and Signaldetektion.

Ganz besonderer Dank wird an Björn Menze und Markus Jehle für die Verwaltung des Gruppenservers und Diskussionen über optischen Fluss, Regularisierung, Visualisierung und Klassifikation von multidimensionalen Bilddaten ausgesprochen.

Ich danke Michael Klar für Diskussionen und Tipps zur Kamerakalibrierung and Empfehlung geiegneter Literaturquellen zum Thema. Michael Klar hat mit seiner grid-locator-Software Auswertungen zur Lokalisierung der Kreuzungspunkte in unseren Testsequenzen durchgeführt, wofür ich mich bei ihm bedanke.

Ein ganz besonderer Dank geht an Achim Falkenroth für seine Unterstützung, Korrekturen, Empfehlungen und Vorschläge beim Entstehen dieser Arbeit.

Ich möchte mich beim Herrn Prof. Fred Hamprecht für die Einladung zur Teilnahme am Seminar zur Signaldetektion und für Diskussionen über Wavelets, Zeit-Frequenz Signalanalyse, Interpolation und optimale Abtastgitter bedanken.

Bei Ralf Schulz bedanke ich mich für Gespräche und einige Vorschläge über Visualisierung von Tensorfeldern, Superquadrics und laserbasierte medizinische Bildgebung.

Ein besonderes Dankeschön wird an Claudia und Daniel Kondermann für Diskussionen und Ideen bezüglich Bewegungsschätzung mit Variationsmethoden, Bibliotheken zur Lösung von partiellen Differentialgleichungen, Mehrgitterverfahren und Konfidenzmaßen ausgesprochen.

Ich möchte meinen besonderen Dank allen Kooperationspartnern des LOCOMOTOR-Projekts an der Johann Wolfgang Goethe-Universität Frankfurt am Main, Forschungszentrum Jülich

und Universität zu Lübeck und insbesondere den folgenden Projektteilnehmern zum Ausdruck bringen.

Ich danke Herrn Prof. Rudolf Mester für seine Zusammenarbeit und Hilfe im Rahmen des LOCOMOTOR-Projekts und Tipps betreff wahrscheinlichkeitstheoretischen Zugangs zur Bewegungsschätzung and algebraischer Klassifikation von Mehrfachorientierungen.

Vielen Dank an Til Aach für seine Hilfe und für das Leiten und Betreuen der Forschung im Rahmen des LOCOMOTOR-Projekts.

Ich bedanke mich bei Erhardt Barth für Diskussionen über Wavelets, Signalverarbeitung, Filterdesign, optischen Fluss und Schätzung von Mehrfachorientierungen.

Mein besonderer Dank gilt Kai Krajsek für ergiebige und fruchtbare Diskussionen zum Thema optischer Fluss, Bewegungsschätzung, optimale und adaptive Glättung des Strukturtensors, Kriging, Regularisierung, steuerbare Filter, Lie-Gruppen, Warping, ...

Ich danke Hanno Scharr für die Optimierung von Ableitungsfiltern in Bezug auf die Wavlets-Skalen und für das freundliche zur Verfügung Stellen der aktuellen rotationsinvarianten Filter zur Bewegungsschätzung, sowie für Diskussionen über Filteroptimierung, partielle Differentialgleichungen, Bewegungsschätzung, Kamerakalibration, Entfernungsschätzung mit mereren Kameras,...

Ich bedanke mich bei Ingo Stuke dafür, dass er uns seinen Algorithmus zur Mehrfachbewegung zur Benutzung überlassen hat und für Diskussionen über Bewegungsschätzung und Analyse von Mehrfachorientierungen.

Ich danke Cicero Mota für Debatten über Waveletsanalysis, Scherungsfilter und steuerbare Filter.

Auch die Teilnehmer des Schwerpunkprogramms *Mathematische Methoden zur Zeitreihennalyse und digitaler Bildverarbeitung SPP 1114* der *Deutschen Forschungsgemeinschaft* haben mir geholfen und ich möchte meinen Dank den folgenden Teilnehmern aussprechen.

Ich danke Stephan Didas für Besprechungen zum Thema optischer Fluss, Skalenräume und Anwendungen von partiellen Differentialgleichungen in der Bildverarbeitung.

Ich bedanke mich bei Henning Thielemann für einige Ideen in Verbindung mit Waveletsanalysis and Programmiersprachen.

Besten Dank an Mark Droske und Tobias Preusser für Diskussionen über Anwendungen von partiellen Differentialgleichungen und mathematische Probleme in der digitalen Bildverarbeitung.

Vielen Dank an Per-Erik Forssen und Klas Nordberg für Gespräche und Erörterungen von Fragen des maschinellen Sehens und der digitalen Bildverarbeitung auf dem First International Workshop on Complex Motion 2004 in Reisensburg und auf dem LOCOMOTOR Winter Workshop 2005 in Frankfurt.

Ich danke Hermann Lauer, Markus Riedinger und Stefan Riedel für das Administrieren der compute servers am Interdisziplinären Zentrum für Wissenschaftliches Rechnen der Ruprecht-Karls-Universität Heidelberg.

Ich bedanke mich ganz herzlich bei Anette Braun, Barbara Werner, Karin Kubessa-Nasri, Gabriela Schocke und Elke Pietschmann für ihre Mitilfe und Unterstützung.

Diese Forschung wurde aus Mitteln der *Deutschen Forschungsgemeinschaft* im Rahmen des Schwerpunktprogramms *Mathematische Methoden zur Zeitreihenanalyse und digitaler Bildverarbeitung SPP 1114* über das LOCOMOTOR-Projekt finanziert, wofür ich mich bedanken möchte.

Heidelberg, 30.11.2007

Pavel Pavlov

Contents

Chapter 1

Introduction

1.1 PDE based image processing

Partial differential equations are a fundamental mathematical tool for description and modeling of processes in natural sciences, life sciences and engineering. In image processing it is possible to enhance, denoise, restore, inpaint, segment and even classify an image by means of partial differential equations, [Aubert and Kornprobst, 2002]. In the variational approach, the segmentation, registration, motion estimation or denoising of an image is expressed in terms of an energy functional according to the desired properties of the model. The energy functional consists of fidelity term, smoothness term and possibly additional constraints term, depending on the task to be solved. The minimization requirement for the energy functional leads to the Euler-Lagrange equations. The Euler-Lagrange equations together with the initial conditions can be solved numerically by grid discretization, finite element methods, level sets or some other integration scheme, which exploits the structure of the equations. Applications of PDE for motion estimation are described in subsection 1.3.2 and chapter 3.

Partial differential equations can also be used to generate a scale space for an image or an image sequence.

1.2 Scale spaces in computer vision

Structure and motion are important features for the human visual system and for digital image processing. The human beings are able to detect boundaries of objects by means of different structures without any other information, as shown in [Jähne, 2002, figure 13.1, Page 358] or [Jähne, Bernd, 2004, color plates 25–27]. Recognizing the ego motion and the motion of other objects is fundamental for our interaction with the rest of the world. Both, motion estimation and structure extraction from image data can be tackled by methods, which rely on computation of image derivatives. In mathematical analysis there is a nice definition for derivatives in the continuous formulation. In digital image processing the computation of derivatives is a question of scale, since we have discretized data on a rectangular grid and in the most practical applications the data is corrupted by noise. When solving an image processing task it is common to take into consideration the noise scale, the scale of the objects, the displacement scale, the structure scale, the texture scale etc.

Let's assume that we have to solve some image processing task for a given image u_0. The image processing task can be denoising, restoration, segmentation, inpainting, motion estimation, classification or whatever. One possible approach to handle the scale problem is to generate a stack of images at several scales and to solve our image processing task. Afterward

we can select the best result. The creation of the image family for the multi scale analysis should be made by taking into account some invariance properties for the generator. A good mathematical formulation for the natural invariance properties was stated by [Alvarez, L. et al., 1993]. They start from a set of axioms and prove that the filtered image must be the viscosity solution of a PDE. The multiscale analysis is defined in [Alvarez, L. et al., 1993] as a family of operators $T_{t t\geq 0}$, where t is the scale parameter. Applied to a given image $u_0(x)$, T generates a sequence of images $u(t, x) := T[u_0](x)$.

Let

$$T_t : C_b^\infty(\mathbb{R}^2) \to C_b^\infty(\mathbb{R}^2) \qquad \forall t \geq 0,$$

where C_b is the space of bounded continuous functions.

The axioms for the Alvarez-Guichard-Lions-Morel scale space theory are

- Recursivity

$$T_0(u) = u, \quad T_s \circ T_t(u) = T_{s+t}(u) \quad \forall u \in C_0^\infty .$$

- Regularity

$$\|T_t(u + hv) - T_t(u) + hv\|_{L^\infty} \leq cht$$
$$\forall h, t \in [0, 1], \qquad \forall u, v \in C_0^\infty .$$

- Locality

$$(T_t(u) - T_t(v))(x) = o(t), \quad t \to +0$$
$$\forall u, v \in C_0^\infty : \ D^\alpha u = D^\alpha v, \quad \forall \alpha .$$

- Comparison principle

$$T_t(u) \leq T_t(v) \text{ on } \mathbb{R}^2, \forall t \geq 0 \text{ and } u, v \in C_0^\infty : \ u \leq v \text{ on } \mathbb{R}^2 .$$

- Gray-level shift invariance

$$T_t(0) = 0, \ T_t(u + c) = T_t(u) + T_t(c) \quad \forall u \in C_0^\infty .$$

- Translation invariance

$$T_t(\tau_h \circ u) = \tau_h \circ T(u) \quad \forall h \in \mathbb{R}^2, t \geq 0 .$$

As examples for scale-spaces we can mention

- the linear diffusion,
- the continuous wavelet transform,
- the total variation flow and
- other curvature driven flows as discussed in chapter 3.

1.3 Optic flow and motion estimation

1.3.1 Fundamentals

Optic flow is the apparent motion of intensity patterns in an image sequence [Horn, 1986; Aubert and Kornprobst, 2002]. The optic flow is in general different from the projected motion field of the scene objects [Verri, A. and Poggio, T., 1987, 1989], [Aubert and Kornprobst, 2002, figure 5.3, page 185]. Determining the optic flow is a preprocessing step or a sub-task of motion estimation , segmentation from motion, reconstruction of 3D structure, separation of motion layers and other practical applications. Some image sequences and illustrations of magnitude and direction of the optic flow field are depicted below.

Figure 1.1: A real life test sequence and a synthetic sequence with ground truth.

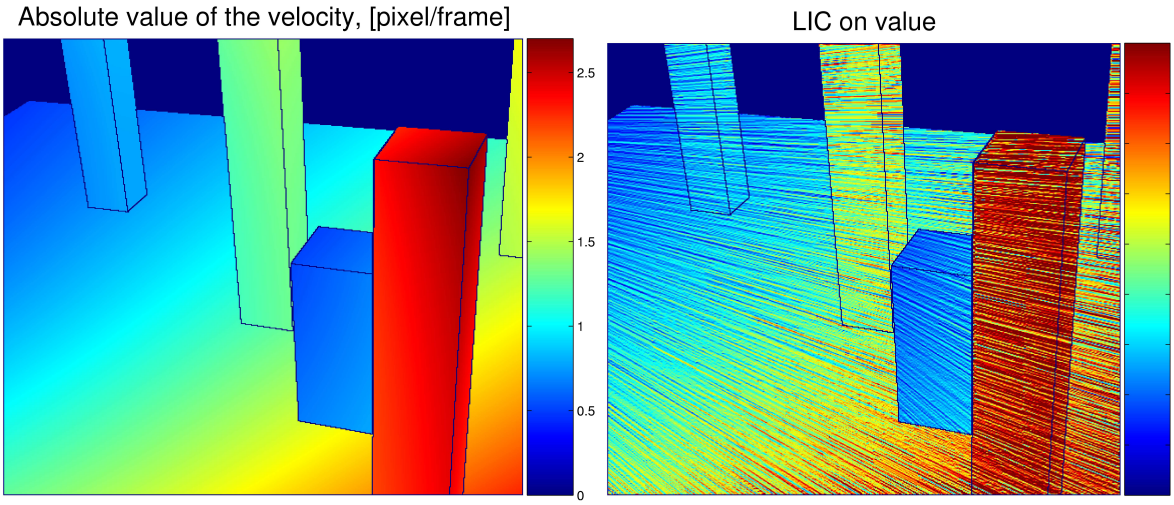


Figure 1.2: Absolute value of the velocity and visualization by Line Integral Convolution (LIC). LIC was introduced by [Cabral and Leedom, 1993].

In a number of natural scenes the problem of estimating motion is increasingly difficult, as commonly not only one motion of objects can be perceived. Many phenomena can cause

multiple motions . Among them occlusion and transparent motion are the most important in terms of their occurrence and significance in realistic imagery.

In this introduction first the different techniques for optic flow computation will be presented. Algorithms for estimating multiple motions will generally depend on these techniques by varying degrees. In the following section the general framework for computing multiple motions will be presented.

Techniques for the estimation of optical flow

Interesting reviews about different methods for recovering optic flow and motion estimation are [Beauchemin, S. S. and Barron, J. L., 1995]. Comparisons of these with error analysis can be found in [Barron, J. L. et al., 1994], [Mitiche, A. and Bouthemy, P., 1996] and [Haußecker, H. et al., 1999]. A good classification of motion estimation techniques can be found in [Jähne, 2002, chapter 14, page 395] and [Aubert and Kornprobst, 2002, section 5.1, pages 181–195]. Basic concepts of image motion, affine motion by the structure tensor and the relationship between motion and direction in image sequences are described in [Bigun, 2006, Chapter 12, page 245]. The methods for optic flow computation can be categorized into three groups:

- The class of *gradient based* techniques rely on computing spatio-temporal derivatives of image intensity, which can either be first order [Fennema, C. and Thompson, W., 1979; Horn, B. K. P. and Schunk, B., 1981] or second order [Nagel, 1983; Tretiak, O. and Pastor, L., 1984]. The gradient based methods yield good results with respect to accuracy and density of the flow field and they are widely used in practical applications.

- *Region-based matching* may be employed when under certain circumstances, aliasing, small number of frames, etc. it is inappropriate to compute derivatives of grey values. In this approach the velocity is defined as a shift giving the best fit between image regions at different times [Anandan, 1989; Glazer, F. et al., 1983; Little, J. J. and Verri, A., 1989]. The region based methods can be applied to data in which the optic flow between two frames has to be estimated or the time derivative can not be computed by filter masks.

- Feature-based methods rely on computation of the optic flow on a feature extracted from the sequence. The features can be corners, edges, phase level contours, curvature of the edges etc. As a representative of the feature-based approaches we will discuss the phase-based method. The phase based method is insensitive to light variations and intensity changes in the image data.

Gradient Based

A number of different gradient based techniques have been conceived. These methods are also termed *differential approaches*, as motion is locally modelled by constraint equations based on differentials of image intensities. The underlying assumption is that local changes in image intensities are due to image translation, there can be no occlusion, unless this is modeled, cf. section 1.3.2 and all objects in the scene are rigid. The optic flow equation for an image sequence $g = g(\vec{x}, t), \vec{x} \in \mathbb{R}^n$ is

$$g(\vec{x}, t) = g(\vec{x} - \vec{v}t, 0), \tag{1.1}$$

where $g(\vec{x}, t)$ is the image intensity at image location $\vec{x}$ and time t. In *first-order methods* [Fennema, C. and Thompson, W., 1979; Horn, B. K. P. and Schunk, B., 1981] the right hand side of Equation (1.1) is expanded as a Taylor series about $(\vec{x}, t)$ and discard the terms higher

than first-order. This leads to the *brightness change constraint equation (BCCE)* also known as the *gradient constraint equation*:

$$\frac{d}{dt}g(\vec{x},t) = g_x(\vec{x},t)v_1 + g_y(\vec{x},t)v_2 + g_t(\vec{x},t) = 0. \tag{1.2}$$

Here v_1 and v_2 are the velocities in x and y direction and g_n the derivative of g in n-direction. This equation can of course be equally well extended to problems of higher dimensionality, such as volumetric data sets.

The BCCE of Equation (1.2) presents one constraint equation for two unknowns. As such only the normal velocity of contours of constant intensity $\vec{v}_n = v_n\vec{n}$, where $\vec{n}$ is the direction of the spatial gradient, can be computed. The normal speed and direction are given by

$$v_n = \frac{-g_t(\vec{x},t)}{\|g_t(\vec{x},t), g_t(\vec{x},t)\|}, \tag{1.3}$$

$$\vec{n} = \frac{(g_x(\vec{x},t), g_y(\vec{x},t))^\top}{\|(g_x(\vec{x},t), g_y(\vec{x},t))\|}. \tag{1.4}$$

In order to attain the full velocity field further constraints or regularity assumptions are needed.

Instead of using the brightness function $g(\vec{x},t)$, in *multi-constraint methods* other functions invariant to motion can be used [Mitiche, A. et al., 1987]. From n such functions a system of n equations can be written and solved.

$$\nabla g^i(\vec{x},t)\vec{v} + g_t^i(\vec{x},t) = 0, \quad i = 1,\ldots,n, \tag{1.5}$$

where $\vec{v} = (v_1, v_2)^\top$ is the image velocity. Example of such functions:

- Multi-spectral images [Markandey, V. and Flinchbaugh, B. E., 1990]
- Operators (spatial operators, differential operators, linear spatio-temporal filters, wavelet transform) [Srinivasan, 1990; Sobey, P. and Srinivasan, M. V., 1991; Weber, J. and Malik, J., 1995]
- Multiple illumination sources [Woodham, 1990]

Second-order methods exploit second-order derivatives of the image intensities $g(\vec{x},t)$ to introduce further constraints [Nagel, 1983, 1987; Tretiak, O. and Pastor, L., 1984; Uras, S. et al., 1988; Girosi, F. et al., 1989; Simoncelli, 1993; Bainbridge-Smith, A. and Lane, R. G., 1997]. One possible constraint is given by [Uras, S. et al., 1988]

$$\begin{bmatrix} g_{xx}(\vec{x},t) & g_{yx}(\vec{x},t) \\ g_{xy}(\vec{x},t) & g_{yy}(\vec{x},t) \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} g_{tx}(\vec{x},t) \\ g_{ty}(\vec{x},t) \end{pmatrix} = 0. \tag{1.6}$$

This constraint can be deducted from the assumption that $\nabla g(\vec{x},t)$ is conserved, that is

$$d\nabla g(\vec{x},t)/dt = 0\ .$$

This type of constrain is a special case of the multi-constraint technique were the n functions are just the second derivatives of the grey values.

Another way of solving the aperture problem is assuming a single motion pattern locally in a spatio-temporal neighborhood. It is assumed that all parameters of the constraint equation are constant in the neighborhood. This leads to an over-determined system of equations

(one equation for each pixel in the neighborhood) which can be solved numerically. For example, [Lucas, B. and Kanade, T., 1981; Lucas, 1984] solve this system of equations as a weighted least squares problem, where weights are dependent on the pixel location in the spatio-temporal neighborhood. [Chu, C. H. and Delp, E. J., 1989] also use a local least squares approach but solve the system of equations in a total least squares framework [Van Huffel, S. and Vandewalle, J., 1991], that accounts for the errors in all the image intensity gradients g_x, g_y and g_t. The drawback of these local model is their poor performance in the presence of multiple motions within the neighborhood. This problem can be solved by employing a robust framework such as M-estimators [Huber, 1981; Hampel, F. R. et al., 1986] or the Least Median Squared of orthogonal distances [Rousseeuw, P. J. and Leroy, A., 1987](LMSOD) Estimator. These estimators have been successfully applied by a number of authors (eg. [Rousseeuw, P. J. and Van Aelst, S., 1999; Bab-Hadiashar, A. and Suter, D., 1997, 1998; Black, 1992; Black, M. J. and Anandan, P., 1996]).

Correlation Based Matching

Gradient based methods presuppose differentiability of time-varying image grey values. However, due to strong noise, severe aliasing or availability of only a small number of frames, accurate numerical differentiation may not always be possible. In this cases region-based matching techniques are often employed. These methods aim to find the best match between image regions in one frame with neighboring regions in subsequent frames. The degree of fit is formulated on some sort of correlation measure. Essentially, finding the best match is an optimization problem. One might maximize the normalized cross-correlation

$$S\left(g_0(\vec{x}), g_1(\vec{x};\vec{s})\right) = \frac{\langle g_0(\vec{x}), g_1(\vec{x};\vec{s})\rangle}{\|g_0(\vec{x})\|\,\|g_1(\vec{x};\vec{s})\|}, \tag{1.7}$$

or minimize the sum of squared difference (SSD)

$$SSD(g_0(\vec{x}), g_1(\vec{x};\vec{s})) = \left(g_0(\vec{x}) - g_1(\vec{x};\vec{s})\right)^2. \tag{1.8}$$

Some disadvantages of the technique include

- Exhaustive two-dimensional search
- Sensitive to illumination change
- Geometric distortions are problematic

It should be noted that the SSD in Equation (1.8) can be viewed as a first-order approximation to $\Delta t W(\vec{x} - \vec{x}_0) dg(\vec{x}, t)/dt$. Therefore, minimizing Equation (1.8) yields an average solution to the BCCE over the windowing function $W(\vec{x} - \vec{x}_0)$ [Anandan, 1989].

Edges can be extracted by one of several edge detection methods [Marr, D. and Hildreth, E. C., 1980; Canny, 1986]. From this relatively sparse collection of features the image velocity can be deducted from one image to the next. This type of approach is often referred to as *feature-based*. From the correspondence of features in one image to the ones in the next, the velocity is commonly defined as the perpendicular distance from corresponding contours in consecutive frames. The accuracy can be refined by iterative approaches [Wu, J. et al., 1989]. Due to the simplicity of the approach, considerable attention has been given to the theory of the motion of curves [Waxman, A. M. and Wohn, K., 1985; Bergholm, 1988; Faugeras, 1990].

There are several problems with contour-based approaches as well as specific problems with the approach suggested by [Waxman, A. M. et al., 1988]. The approaches assume that

features including edge locations can be well localized and are stable in subsequent images over time. Also, it is assumed that the edges are well enough isolated that they facilitate the correspondence process. Due to this isolation the velocity fields estimated from this technique are very sparse.

There are explicit methods such as least-squares fit to a parametric model [Waxman, A. M. and Wohn, K., 1985] and iterative methods such as those based on smoothness assumption [Hildreth, 1984; Gong, 1989] and [Jähne, 2002, p. 430].

A technique of this class is the motion estimation by using the Census transform [Stein, 2004]. By this method the images are transformed into binary images according to the sign of its first derivative. Then block matching is performed on the binary images. In this manner the computational speed is enhanced for real time applications, although the method yields sparse flow fields.

Feature-Based Techniques

Feature based techniques presume, that certain image features such as edges, corners, level phase contours are conserved in consecutive frames of the image sequence. Some features are relatively insensitive to illumination changes, for example the phase. As a representative of the feature-based methods, we can discuss the phase-based techniques.

The method developed by [Fleet, D. J. and Jepson, A. D., 1990] defines component velocity in terms of the instantaneous motion of level phase contours in the output of band-pass velocity tuned Gabor filters. These filters are used to decompose the input signal according to scale, speed and orientation. Each filter output is complex-valued and can be expressed as

$$R(x,t) = \rho(x,t)e^{i\phi(x,t)}, \tag{1.9}$$

where $\rho(x,t)$ and $\phi(x,t)$ are the amplitude and phase part of the output signal. The component 2d velocity in the direction normal to level phase contours is given by

$$\vec{v}_\perp = \frac{-\phi_t(x,t)\nabla\phi(x,t)}{\|\nabla\phi(x,t)\|_2^2}. \tag{1.10}$$

Here $\phi_t(x,t)$ is the temporal derivative of the phase and $\nabla\phi(x,t)$ is the spatial gradient. Phase derivatives are computed using the identity

$$\nabla\phi(x,t) = \frac{\mathrm{Im}[R^*(x,t)\nabla R(x,t)]}{|R(x,t)|^2}, \tag{1.11}$$

where $R^*(x,t)$ is the complex conjugate of $R(x,t)$ and Im denotes the imaginary part of a complex number. [Fleet, D. J. and Jepson, A. D., 1990] relate velocity to local phase information because of the relative insensitivity of the phase signal to amplitude variations due to changes in scene illumination.

Examples for complex motion

Examples for complex motion are listed in [Jähne, 1993, section 4.7.3, page 103]. The most important types of complex motion in image processing are

- occlusion,
- transparent motion,

- motion with illumination changes and
- disappearing or arising of moving objects.

Occlusion is a classical example for multiple motion. Even when we handle models for single motion estimation, it's possible to preserve the discontinuities on occlusion points to some extent.

1.3.2 Discontinuity and curvature preserving optical flow

In this section we give an overview of the computation of optic flow using discontinuity-preserving and curvature-preserving PDE's. Under the assumption of gray level constancy of the image features over time, the optical flow problem is ill-posed. To overcome this ambiguity called the *aperture problem* , we can add a piecewise smoothness constraint which respects the flow discontinuities. Combining the gray level constancy and the piecewise smoothness into a single energy functional leads to a classical minimization problem in the sense of calculus of variation. We investigate the existence and uniqueness of the solution under proper assumptions for image processing tasks. We propose a curvature-preserving PDE for optic flow computation to avoid over-smoothing on trajectory intersections or points of interest.

Related work:

[Horn, B. K. P. and Schunk, B., 1981] introduces a regularization scheme for variational optic flow computation with a quadratic smoothness term. This is an important work in flow field regularization. A lot of discontinuity-preserving methods for motion estimation are extensions of the method, described in this article.

[Snyder, 1989] derives a general form of the smoothness constraint under (i) Cartesian coordinate system invariance, (ii) positive definiteness and (iii) non-coupling different components of the optical flow. For quadratic terms depending on the first or second-order derivatives of the gray values and quadratic terms in the first-order derivative of the flow field, there are only four generic smoothness constraints. All other other possible constraints for quadratic regularizers are linear combinations of the four basic constraints. The smoothness term of [Horn, B. K. P. and Schunk, B., 1981], mentioned above, is just a special case.

[Schnörr, 1991] investigates the minimization of quadratic functionals, described in [Horn, B. K. P. and Schunk, B., 1981]. A proof is given for the existence, uniqueness and well posedness of the solution in Sobolev spaces. The author proposes the Ritz method for the discretization of the Euler-Lagrange equation.

[Proesmans, M. et al., 1994] introduces a nonlinear flow-driven modification of the energy functional. The spatial smoothness of the flow field is taken into account. A bidirectional flow field with inconsistency measures for the flow discontinuities lead to a system of six coupled reaction-diffusion equations.

[Schnörr, 1994] stated a general nonlinear formulation of the minimization problem for optic flow estimation.

The goal of [Weickert, J. et al., 1998] is the investigation of stable schemes for arbitrary time steps. The proposed linear-implicit updating rule is based on additive operator splitting. It ensures stability for arbitrary time step sizes.

[Weickert, 1998] introduces nonlinear rotationally invariant coupled reaction-diffusion equations for the optical flow problem.

[Aubert et al., 1999] proves the existence and uniqueness of the optical flow problem in the space of the functions of bounded variation $BV(\Omega)$ for Lipschitz data and in [Aubert and Kornprobst, 1999] they give a proof of existence and uniqueness for non-Lipschitz data. [Weickert,

J. and Schnörr, C., 1999] introduces a spatio-temporal regularization exploiting the coherency of the flow field in time.

[Weickert, J. and Schnörr, C., 2001a] presents a systematic classification of regularizers for optic flow with respect to their rotation invariance and their dependence on the image or the optic flow. A proof of the existence and uniqueness of the solution in $H^{1,2}(\Omega)$ is given.

[Weickert, J. and Schnörr, C., 2001b] spatio-temporal regularization exploiting the coherency of the flow field in time. The authors suggest a finite-difference approximation.

[Aubert and Kornprobst, 2002, pages 190-193] is a very good introduction to discontinuity preserving regularization. The book discusses the foundations of the PDE-based image processing.

[Bruhn, A. et al., 2005b] describe a combined local-global method for optic flow estimation. To circumvent the ill posedness of the problem, they introduce a mixed regularizer. The weight between the local and global term is set by the user.

[Bruhn, A. et al., 2005a] reveals insights in real-time implementation issues of isotropic discontinuity-preserving optic flow problem by using multi-grid techniques.

[Tschumperle, 2002, 2005; Tschumperle, D. and Deriche, R., 2005] developed a nice unified expression for a generic curvature-preserving regularization PDE for vector valued images.

Flow-Driven Smoothness Term

Let $\Omega \in \mathbb{R}^3$. We will denote the image data by f, $f \in L^2(\Omega)$ and the displacement field by $u : \Omega \to \mathbb{R}^2$.

The gray level constancy implies the optical flow constraint

$$f_x u_1 + f_y u_2 + f_t = 0 \tag{1.12}$$

The optical flow problem is ill-posed, because of the aperture problem. One can use regularization techniques for the displacement field u, in order to obtain a solution. We can assume, that the velocity u is piecewise smooth. The discontinuities of the optic flow should represent the object edges or occlusion boundaries.

An appropriate regularization model for the optic flow can be described by the minimization of the energy functional $E(u_1, u_2)$

$$E(u_1, u_2) := \int_\Omega (f_x u_1 + f_y u_2 + f_\theta)^2 \; + \; \lambda V(\nabla f, \nabla u_1, \nabla u_2)\, dxdy \to \min. \tag{1.13}$$

The term $V(\nabla f, \nabla u_1, \nabla u_2)$ is referred to as the *regularizing term* and λ denotes the regularization parameter. By choosing different values for λ we can steer the smoothness of the displacement field in our model for different applications.

The regularization part determines the behavior of the motion estimator. Let's consider the construction of the regularization term. There are several possibilities to design the regularization part. We can choose between image-driven or flow driven regularization on the one side and isotropic or anisotropic regularization on the other side. The classification of regularizers for optic flow models has been done in [Weickert, J. and Schnörr, C., 2001a, page 252] and is depicted in table 1.2. The key idea is to prevent smoothing across discontinuities.

Let's start with the simple case of isotropic smoothing. The corresponding terms in the

Euler-Lagrange equations to the functional in 1.13 are

$$\begin{aligned} \|\nabla u_i\|^2 &\sim 2 \triangle u_i \qquad i = 1, 2 \\ \Psi(\|\nabla u_i\|) &\sim \operatorname{div}\left(\frac{\Psi'(\|\nabla u_i\|)}{\|\nabla u_i\|}\nabla u_i\right) \qquad i = 1, 2 \ . \end{aligned}$$

We introduce local coordinates η and ξ according to the gradient direction, $\eta := \frac{\nabla u}{\|\nabla u\|}, \xi \perp \eta$. Locally the divergence term can be decomposed as a sum of two weighted directional derivatives in order to see the action of the diffusion operator along the directions ξ and η

$$\operatorname{div}\left(\frac{\Psi'(\|\nabla u\|)}{\|\nabla u\|}\nabla u\right) = \frac{\Psi'(\|\nabla u\|)}{\|\nabla u\|}u_{\xi\xi} + \Psi''(\|\nabla u\|)u_{\eta\eta} \ .$$

Isotropic smoothing condition inside homogeneous regions requires to encourage the desired diffusion behavior in gradient direction η and to prevent the smoothing across flow discontinuities. This leads to the following conditions on the weights.

- In regions of low gradients we have

$$\Psi'(0) = 0 \qquad \text{and} \qquad \lim_{\|\nabla u\|\to 0} \frac{\Psi'(\|\nabla u)}{\|\nabla u\|} = \lim_{\nabla\|u\|\to 0} \Psi''(\|\nabla u\|) = \Psi''(0) > 0.$$

- Near discontinuities the weights in both orthogonal directions should be set as:

$$\lim_{\|\nabla u\|\to\infty} \Psi''(\|\nabla u\|) = 0 \qquad \text{and} \qquad \lim_{\|\nabla u\|\to\infty} \frac{\Psi'(\|\nabla u\|)}{\|\nabla u\|} = \beta > 0.$$

The last two conditions near discontinuities are incompatible. For practical implementations, we can make a compromise on Ψ. For instance, $\Psi''(s)$ and $\Psi'(s)/s$ can both converge to zero as $s \to \infty$ but at different rates. Many functions Ψ can be found, which satisfy this compromise. Some of the functions, which have been used in the literature for regularization and computation of discontinuous flow fields are listed in the table below.

Name	$\Psi(s)$
Geman & Reynolds	$\frac{s^2}{1+s^2}$
Perona & Malik	$\log(1 + s^2)$
Green	$2 \log \cosh s$
Aubert	$2\sqrt{1 + s^2} - 2$
Schnörr & Weickert	$\varepsilon s^2 + (1 - \varepsilon)\kappa\sqrt{1 - \frac{s^2}{\kappa^2}}$

Table 1.1: Functions Ψ preserving discontinuities.

Let $\Psi(s^2)$ be an increasing smooth convex function, for instance

$$\Psi(s^2) := \varepsilon s^2 + (1 - \varepsilon)\kappa\sqrt{1 - s^2/\kappa^2} \ .$$

Let $A \in \mathbb{R}^{n\times n}$ be a matrix with orthonormal eigenvectors $w_1, ..., w_n$ and corresponding eigenvalues $\sigma_1, ..., \sigma_n$. We define a function of a matrix by

$$\Psi(A) := \sum_i \Psi(\sigma_i) w_i w_i^T.$$

Remark: $\Psi' = \sum \Psi'(\sigma_i) w_i w_i^T$.

$$V(\nabla f, \nabla u_1, \nabla u_2) := \text{tr } \Psi(\nabla u_1 \nabla u_1^T + \nabla u_2 \nabla u_2^T)$$

$$J := \nabla u_1 \nabla u_1^T + \nabla u_2 \nabla u_2^T$$

The reaction-diffusion system for $E(u)$ in 1.13 is given in the next proposition.

Proposition 1.1. *The reaction-diffusion system to*

$$E(u) = \int_\Omega (f_x u_1 + f_y u_2 + f_\theta)^2 \ + \ \lambda \text{ tr } \Psi(J) \, dxdy$$

is given by

$$\partial_t u_1 = \text{div } (\Psi'(J) \nabla u_1) - \frac{1}{\lambda} f_x (f_x u_1 + f_y u_2 + f_\theta)$$
$$\partial_t u_2 = \text{div } (\Psi'(J) \nabla u_2) - \frac{1}{\lambda} f_y (f_x u_1 + f_y u_2 + f_\theta)$$

Proof. By calculating the Euler-Lagrange equations. The computation steps can be seen in detail in [Weickert, J. and Schnörr, C., 2001a, page 251, Proposition 1]. □

Remarks:

1. If J is a symmetric and positive semi-definite 2x2 matrix then there exist orthonormal eigenvectors $v_1, v_2 \in \mathbb{R}^2$ of J and corresponding eigenvalues $\mu_1, \mu_2 \in \mathbb{R}$. The eigenvalues μ_1, μ_2 specify the contrast of the vector valued image (u_1, u_2) in the directions v_1 and v_2.

2. In [Zeidler, 1990b] there is a nice discussion about the interpretation of the time axis and the unknown function $u(x, y)$ in partial differential equations. It is just a question of interpretation, whether we treat the time axis equally with the spatial axis. We can compute the optic flow by spatio-temporal regularizers by introducing homogeneous coordinates and setting $u := (u_1, u_2, 1)$. The energy functional $E = E(u)$ for the entire space-time image volume is defined by

$$E(u_1, u_2) := \int_{\Omega \times [0,T]} ((f_x + f_y + f_\theta) \cdot u)^2 \ + \ \lambda V(\nabla f, \nabla u_1, \nabla u_2) \, dxdyd\theta.$$

The table below gives a classification of regularizers for optic flow models. The image-driven smoothing of the optic flow field leads to over-segmentation which is dependent on the object texture. If we compare the methods listed in the table, the best results are obtained with the anisotropic flow-driven regularization. This kind of smoothing depends on the solution u itself and is more elaborate and produces less over-segmentation on strongly textured image data.

Existence and Uniqueness of Solutions in $H^{1,2}(\Omega)$ and $\mathbf{BV}(\Omega)$

First we discuss the well posedness in L^2.

Let $\mathcal{H} := H^{1,2}(\Omega) \times H^{1,2}(\Omega)$ $\quad (u, v)_{\mathcal{H}} := \int_\Omega u^T v + tr \nabla u \nabla v^T \, dx.$

Theorem 1.2. *Assumptions:*

	Isotropic	Anisotropic
Image-Driven	$g(\|\nabla f\|^2) \sum \|\nabla u_i\|^2$	$\sum \nabla u_i D(\nabla f) \nabla u_i$
Flow-Driven	$\Psi(\sum \|\nabla u_i\|^2)$	$\text{tr}\, \Psi(\sum \nabla u_i \nabla u_i^T)$

Table 1.2: Classification of image and flow-driven regularizers.

1. Ψ *is differentiable and increasing.*
2. $\Psi(s^2)$ *is strictly convex in* s.
3. $\exists c_1, c_2 > 0 : c_1 s^2 \leq \Psi(s^2) \leq c_2 s^2, \quad \forall s \in \mathbb{R}$

Under these assumptions the functional $E(u)$ *admits unique minimizer in* $\mathcal{H}$. *This minimizer depends continuously on the image sequence* f.

Proof. The complete proof of the theorem can be found in [Weickert, J. and Schnörr, C., 2001a, page 253, Theorem 1]. The proof of the theorem can be split into two basic steps.

1. If u is solution of the minimum problem, then u is solution of the operator equation

   $E'(u) = 0.$ (Abstract Euler Equation)

   Proof of step 1:[Zeidler, 1990a, Proposition 25.11, Page 510]

2. If E' is strongly monotone and Lipschitz continuous on $\mathcal{H}$, then for each $b \in \mathcal{H}^*$, the operator equation

   $$E'u = b, \qquad u \in \mathcal{H}$$

   has a unique solution. The solution depends continuously on b.

   Proof of step 2:[Zeidler, 1990a, Theorem25.B ,Page504]

□

Let's now consider the choice of the regularization parameter λ. There is an optimal value for the regularization parameter λ, depending on the discontinuity preserving function V, the noise level in the data set and some technical details about derivative computation, e.g. filter size, amount of pre- and post-smoothing. The optimality means, that an error measure is minimal. In practice, the optimality of λ can be shown on image sequences with ground truth. Usually, the value of the regularization parameter is chosen manually and affects the optic flow measurements. The optimal choice of the regularization paraneter is discussed in citekrajsek2006. Under isotropic and homogeneous conditions one can compute the optimal value for the regularization parameter λ based on the maximum entropy principle.

$$\lambda \sim \frac{\sigma_t^2}{3\sigma_u^2},$$

where σ_t is the standard deviation of the time component of the image gradient and σ_u is the standard deviation of the optical flow u. Under anisotropic and non-homogeneous assumptions the value of λ can be determined numerically.

BV(Ω) Regularization for Non-Lipschitz data. The Algorithm for half-quadratic minimization

There is a variational formulation of the optic flow problem in $BV(\Omega)$

$$E(u) := \int_\Omega (\nabla f \cdot u + f_\theta + \lambda^r[\Psi(Du_1) + \Psi(Du_2)] + \lambda^h c(x)\|u\|^2)\, dx$$

The effect of the term

$$\int_\Omega c(x)\|u\|^2\, dx$$

is, that no visible motion should be detected in homogeneous regions.

Working with images requires discontinuous functions along curves, where the object boundaries are. A suitable choice can be the space of functions of bounded variation $BV(\Omega)$.

$$BV(\Omega) := \{f \in L^1(\Omega) | sup \int_\Omega f\ \mathrm{div}\varphi\, dx < \infty : \varphi \in C_0^1(\Omega), |\varphi| \leq 1\}$$

The next theorem states an existence result for the optic flow regularization problem in $BV(\Omega)$.

Theorem 1.3. *Let*

- *the data* $f \in H^{1,\infty}(\Omega \times [0,T])$ *be Lipschitz,*
- $\Psi : \mathbb{R} \mapsto \mathbb{R}^+$ *odd, convex and non decreasing,*
- $\exists\ b_1, b_2 > 0 : b_1 x + b_2 \leq \Psi(x) \leq b_1 x + b_2, \forall x \in \mathbb{R}^+,$
- $c \in C^\infty(\Omega), \exists\ m_c : \quad c(x) \in [m_c, 1], \forall x \in \Omega$

Then the minimization problem has a solution in $BV(\Omega)$.

Proof.
- $E(u)$ is coercive
- $E(u)$ is lower semi continuous for the BV-weak* topology
- $\Rightarrow \quad \exists\ u \in BV(\Omega) : \quad u$ minimizes E.

□

For non-Lipschitz data there arise two main difficulties:

- $$\int_\Omega u \cdot Df\, dx$$

  where the derivative Df to the set of the Radon measures $\mathcal{M}(\Omega)$ belongs.

  What is the product of an integrable function and a measure ?
  Aubert and Kornprobst give an integral representation in [Aubert and Kornprobst, 1999] .

- The global Energy $E(u)$ is no longer low semi continuous for the weak* topology of $BV(\Omega)$. One has to search for the relaxed problem.

[Aubert and Kornprobst, 1999] developed a convergent algorithm for half-quadratic minimization.

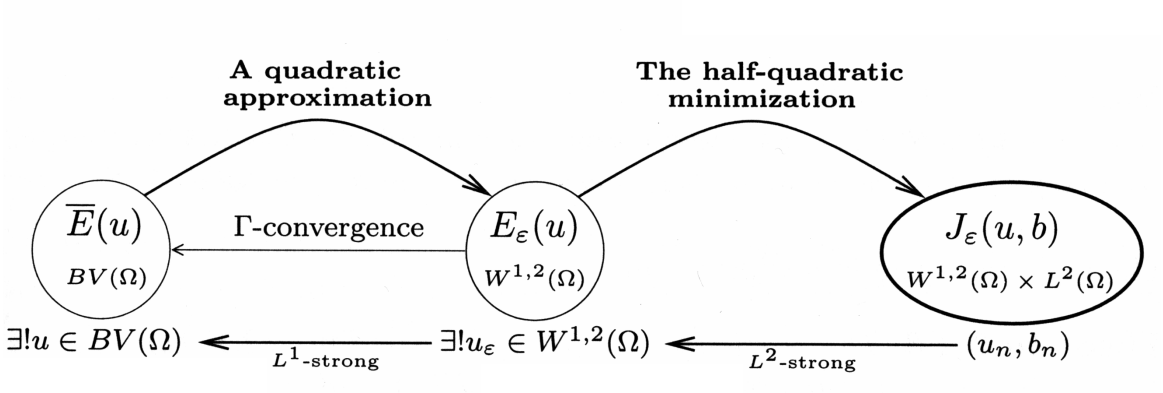


Figure 1.3: Half-quadratic minimization according to [Aubert and Kornprobst, 2002, Section 3.2.4], [Aubert et al., 1999, Section 4.2]

- New functional E_ε: unique solution in $H^{1,2}(\Omega)$.

$$u_\varepsilon \to u \quad in \quad L^2 \quad (\Gamma - convergence)$$

- For a fixed ε [Aubert and Kornprobst, 1999, 2002] propose a suitable numerical scheme: " half-quadratic minimization".
  Convergence in L^2 to the minimizer of E_ε.

Design principle and examples for anisotropic regularizers

[Weickert, J. and Schnörr, C., 2001a] proposed a design principle for anisotropic regularizers.

- The motivation arises from the observation, that the regularizer

  $$\Psi(\mu_1) + \Psi(\mu_2)$$

  anisotropic is, while the discontinuity preserving function

  $$\Psi(\mu_1 + \mu_2)$$

  is an isotropic one. Here μ_1, μ_2 are the eigenvalues of the diffusion tensor.

- The design principle states:
  Let $\Psi(\sum \|\nabla u_i\|^2)$ be an isotropic regularizer with a non quadratic function Ψ and let

  $$\sum \|\nabla u_i\|^2 = \sum \rho_j$$

  be a decomposition of its argument where ρ_j is rotationally invariant for all j. Then the regularizer $\sum(\Psi(\rho_j))$ is rotationally invariant and anisotropic.

Examples: Design of anisotropic regularizers

1. $J := \nabla u_1 \nabla u_1^T + \nabla u_2 \nabla u_2^T$.
   Let $\sigma(J) = \{\mu_1, \mu_2\}$. $\Rightarrow \Psi(trJ) = \Psi(\mu_1 + \mu_2)$ is an isotropic flow-driven regularization term.
   $\Rightarrow V_{AF} := \Psi(\mu_1) + \Psi(\mu_2)$ is an anisotropic one.

2. Using the Helmholtz decomposition of a flow: flow = laminar + rot + div or the identity

$$\|\nabla u_1\|^2 + \|\nabla u_2\|^2 = \frac{1}{2}(\text{div}^2 u + \text{rot}^2 u + \text{shear}^2 u)$$

   we can apply the design principle and derive the regularizer

$$V := \Psi(\text{div}^2 u) + \Psi(\text{rot}^2 u) + \Psi(\text{shear}^2 u) \; .$$

   This regularizer is suitable for motion estimation of fluid flow. A sophisticated approach with direct estimates for the divergence and vorticity rates is presented in [Corpetti, T. et al., 2002].

3. $V := \Psi(\|\nabla u_1\|^2) + \Psi(\|\nabla u_2\|^2)$

$$\partial_t u_1 = \text{div}(\Psi'(\|\nabla u_1\|^2)\nabla u_1) - \frac{1}{\lambda} f_x((f_x, f_y, f_\theta) \cdot u)$$
$$\partial_t u_2 = \text{div}(\Psi'(\|\nabla u_2\|^2)\nabla u_2) - \frac{1}{\lambda} f_y((f_x, f_y, f_\theta) \cdot u)$$

   This regularization term V leads to discontinuities at different locations for each channel u_1 and u_2.

To our best knowledge the most general formulation for nonlinear anisotropic diffusion PDE's with discussion of the coupling between the different channels is proposed in [Tschumperle, 2005].

Curvature-Preserving PDE's for optic flow estimation

The following discussion about curvature-preserving PDE's leans on [Tschumperle, 2002, 2005; Tschumperle, D. and Deriche, R., 2005].

Let D describe the local geometry, we want to smooth along. There are three possible formulations, which express the smoothing behavior in a model for image enhancement or optic flow regularization. The formulation can be expressed in terms of functionals, divergences and oriented Laplacians.

1. Functional minimization.
2. Diffusion equation.
3. Oriented Laplacians.

They are not equivalent. The relation is

$$(1) \Rightarrow (2) \Rightarrow (3).$$

Contras against the divergence formulation

$$\frac{\partial u_i}{\partial t} = \text{div}(D \nabla u_i), \qquad i = 1, ..., n \; :$$

- The PDE implicitly depends on the spatial variations of D.
- The PDE does not fully respect the geometry D.

There is a counterexample by [Tschumperle, 2005] , which shows, that the divergence PDE does not fully respect the underlying smoothing geometry D. Suppose we want to anisotropically smooth an image $u : \Omega \to \mathbb{R}$ along the gradient direction. We can define the local geometry tensor D as

$$D(x) := \frac{\nabla u}{\|\nabla u\|} \frac{\nabla u}{\|\nabla u\|}^T .$$

This leads to the diffusion equation

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{1}{\|\nabla u\|^2} \nabla u \nabla u^T \nabla u\right) = \text{div} \nabla u = \triangle u .$$

The same action can be caused by choosing $D := Id$, thus different tensor fields with different shapes determine the same regularization behavior. Anisotropic tensor D leads to an isotropic smoothing.

Let's now analyze the advantages and drawbacks of the trace-based PDE's. The trace-based PDE

$$\frac{\partial u_i}{\partial t} = \text{trace}(DH_i), \qquad H : \text{Hessian} \qquad i = 1, ..., n$$

respects better the local smoothing geometry D.

The trace is not a differential operator and it follows, that there is no a dependence on the spatial variations of D like in the divergence case. Two different tensors $D_1 \neq D_2$ lead to different smoothing. The interpretation is an oriented Gaussian smoothing whose strength and orientation is related to D.

There are some disadvantages of the trace-based PDE like its behavior on curved structures like corners. The effect is an over-smoothing on corners, which is typical for the mean curvature flow. Figure 1.4 illustrates this drawback. illustrates this disadvantage of the trace-based PDE's for image and flow regularization.

To avoid over-smoothing on corners we can use curvature-preserving PDE's for our image processing task. First we consider an image regularization along a vector field $w : \Omega \to \mathbb{R}^2$ and not along a tensor field. Curvature-preserving regularization can be constructed as follows. Let $C^x(a)$ be an integral curve of w starting from x parameterized by a. This is a one dimensional heat flow equation constrained on $C^x(a)$

$$\frac{\partial u_i(C^x(a))}{\partial t} = \frac{\partial u_i(C(a))}{\partial a^2}, \qquad i = 1, ..., n .$$

The constrained equation on the curve leads to

$$\frac{\partial u_i}{\partial t} = \text{trace}(ww^T H_i) + \nabla u_i \cdot J_w w, \qquad Jw : \text{Jacobian of the vector w} .$$

In the next step we will analyze the difference between trace and divergence formulation and we will formulate a curvature-preserving diffusion equation. Let $D := ww^T$ describes the local geometry of the smoothing process.

$$\text{div}(D \nabla u_i) = ... = \text{trace}(ww^T H_i) + \nabla u_i^T J_w w + \text{div}(w) \nabla u_i^T w$$

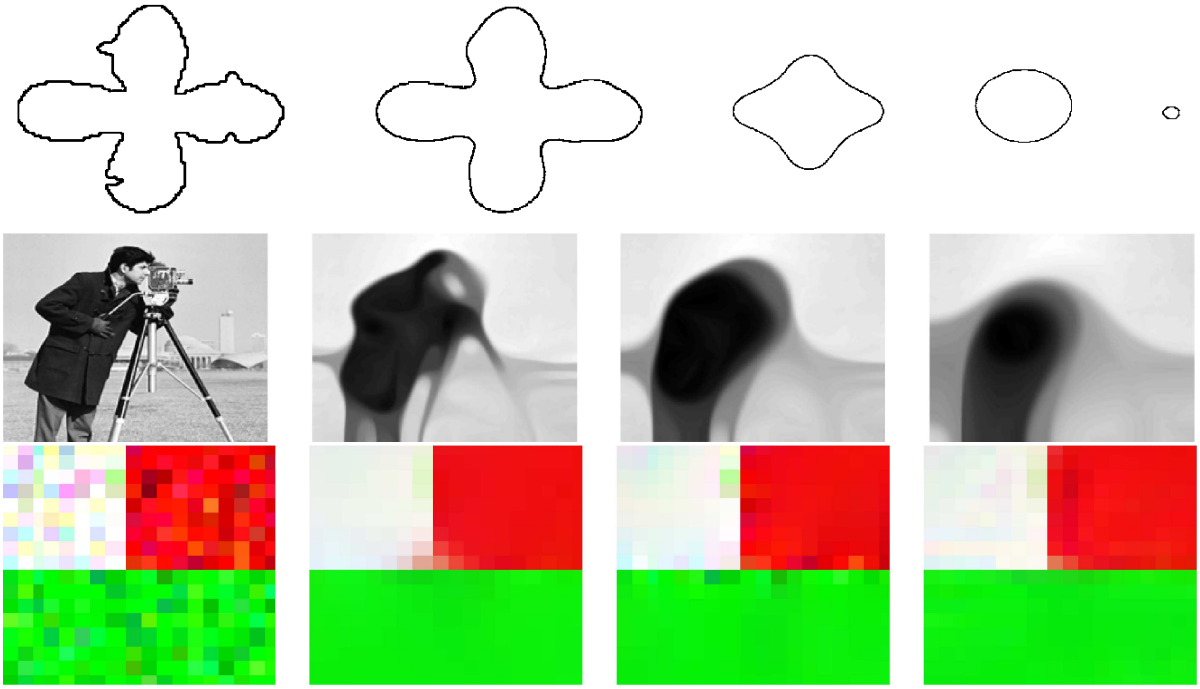


Figure 1.4: Action of trace-based PDE's: Mean Curvature Flow, figures from [Aubert and Kornprobst, 2002; Tschumperle, 2005; Petrovic, A. et al., 2004]

$$
\begin{aligned}
\operatorname{div}(D\nabla u_i) &= \operatorname{div}\begin{pmatrix} u^2\frac{\partial u_i}{\partial x} + uv\frac{\partial u_i}{\partial y} \\ uv\frac{\partial u_i}{\partial x} + v^2\frac{\partial u_i}{\partial y} \end{pmatrix} = \\
&= \left(u^2\frac{\partial^2 u_i}{\partial x^2} + 2uv\frac{\partial u_i}{\partial x \partial y} + v^2\frac{\partial^2 u_i}{\partial y^2}\right) + \nabla u_i^T \begin{pmatrix} 2u\frac{\partial u}{\partial x} + u\frac{\partial v}{\partial y} + v\frac{\partial u}{\partial y} \\ 2v\frac{\partial v}{\partial y} + u\frac{\partial v}{\partial x} + v\frac{\partial u}{\partial x} \end{pmatrix} = \\
&= \operatorname{trace}(ww^T H_i) + \nabla u_i^T \left[\begin{pmatrix} u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} \\ u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} \end{pmatrix} + \begin{pmatrix} u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} \\ u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} \end{pmatrix}\right] = \\
&= \operatorname{trace}(ww^T H_i) + \nabla u_i^T J_w w + \operatorname{div}(w)\nabla u_i^T w
\end{aligned}
$$

On the right hand side of the last equation there are three different terms.

1. The first term defines a trace PDE, that smoothes locally along w.

2. The two first terms can be considered as a curvature-constrained regularization PDE, that smoothes along w while taking the curvature of integral curves of w into account.

3. The three terms together represent a classical divergence PDE this is local diffusion along w.

   The tree terms together produce perturbations of the effective smoothing direction, cf. counterexample in this section.

   This behavior is not desirable for curvature-preserving image regularization.

Now we can formulate a curvature-preserving optic flow equation. Let D describes the local geometry of the desired smoothing behavior, for instance

$$D := \nabla u_1 \nabla u_1^T + \nabla u_2 \nabla u_2^T$$

or

$$D := \Psi(\lambda_-, \lambda_+, \theta_-, \theta_+),$$

where $\lambda_{+-}, \theta_{+-}$ are the spectral elements of the Gaussian smoothed version of the structure tensor of the flow field u.

The generic energy functional is given by

$$E(u) := \int_\Omega ((f_x, f_y, f_\theta) \cdot u)^2 \, dx + \lambda \int_\Omega \Psi(\lambda_+, \lambda_-) \, dx \, .$$

The Euler-Lagrange equations to $E(u)$ can be thought of as the steady state of the reaction-diffusion system

$$\frac{\partial u_i}{\partial t} = \operatorname{div}(D(u)\nabla u_i) - \frac{1}{\lambda} f_{x_i}((f_x, f_y, f_\theta) \cdot u), \qquad i = 1, 2$$

or equivalently

$$\frac{\partial u_i}{\partial t} = \operatorname{trace}(DH_i) + \nabla u_i^T J_u u + \operatorname{div}(u)\nabla u_i^T u - \frac{1}{\lambda} f_{x_i}((f_x, f_y, f_\theta) \cdot u), \; i = 1, 2 \, . \tag{1.14}$$

The term $\operatorname{div}(u)\nabla u_i^T u$ introduces perturbations and causes over-smoothing on points of interest. To avoid over-smoothing on trajectory intersections we propose the curvature-preserving PDE

$$\frac{\partial u_i}{\partial t} = \operatorname{trace}(DH_i) + \nabla u_i^T J_u u - \frac{1}{\lambda} f_{x_i}((f_x, f_y, f_\theta) \cdot u) \qquad i = 1, 2 \, . \tag{1.15}$$

The function Ψ can be chosen to respect the discontinuities of the optical flow u.

Nonlocal formulation in the mathematical sense:

Assuming the dependency of the tensor D on the neighborhood of the evaluated pixel in the space-time volume, the last two equations can be viewed as a nonlocal PDEs. The nonlocality is introduced by the smoothing of the structure tensor of u. A proof of the convergence of the asymptotic solution to a stationary solution in $H^{1,2}(\Omega)$ for (1.14) in the isotropic case $D = D(\|\nabla u\|)$ can be found in [Chipot, 2004, Theorem 1.1, Page 2]. In terms of image processing, this smoothing is a local or a neighborhood operation. Usually a nonlocal operator in image processing is meant to depend on the whole image or at least on a bigger subregion of the image. An example for a nonlocal operator in image processing is the Hilbert transform.

1.3.3 Multiple motion estimation

A good introduction to multiple motion and multiple orientations estimation can be found in [Stuke, 2006, pages 5–10].

Techniques that compute multiple motion can be classified into methods which separate motion by segmentation and those, that require no segmentation. Dependent on the space in which the analysis of multiple motion is fulfilled there are image domain methods, frequency domain methods and parametric space approaches. In the image domain one can use the gradient based method, described in section 1.3.1.

Occlusion and transparency are the most important and significant phenomena in realistic image processing. Their information content is useful for later steps of processing such as motion segmentation, depth ordering and 3D surface reconstruction. Occlusion causes motion discontinuities, so that the derivatives on occlusion border are not well defined. Correlation-based techniques are also sensitive to occlusion, since image structure may appear or disappear

in time. Therefore the issue of choosing a strategy for estimating discontinuous optical flow is crucial in motion analysis. Transparency is some kind of superposition of motion patterns. [Shizawa, M. and Mase, K., 1990] introduced a superposition principle to multiple motion and extended existing algorithms for optical flow to handle many motion distributions simultaneously.

In some sense the opposite approach to the superposition principle is to segment iteratively the image using the displacement information . One can use particular motion models for the different image regions. Each iteration step consists of motion parameter estimation and segmentation, where the motion models vary depending on the obtained segmentation. Pixels are aligned to coherent moving regions according to the Bayesian classifier, convergence speed of the prediction error or other probabilistic criteria. Segmentation can be fulfilled by considering each pixel together with its velocity as an element of an multidimensional space and grouping together the regions with similar features such like curvature, orientation and distance between the elements in the multidimensional space. One can segment an image sequence into regions by using level sets. The level sets are obtained according to partial differential equations. In this manner the motion boundaries are computed without relying on intensity discontinuities and this process is purely motion based.

Analysis of multiple motion without segmentation has been suggested using the dominant motion approach. The dominant motion can be estimated in a coarse-to-fine manner and then the dominant motion is suppressed in order to detect the next moving object. By repeating recursively this steps one can extract the motion information from the video data. To recover transparencies the detected objects can be tracked throughout long image sequences. Then one can segment the regions by classifying the pixels as moving or stationary.

Multiple motion in the spatial domain corresponds to multiple surfaces in the frequency domain, hence motion analysis converts to detecting orientation of structure in the Fourier space. There are several different methods based on directional filtering, outliers detection and orientation analysis for multiple motion estimation in the frequency domain. In case of transparency [Stuke,I. et al., 2003] separate the motion layers by solving a linear system in the Fourier space.

Layered representation of video data is another possible way to treat the problem of multiple motion. The image is decomposed into a set of layers, where possibly disconnected regions of similar motion are grouped together and represent one single layer. Regularization assumptions can be used to produce a smoothed layered motion representation.

Parametric space methods allow us to combine the advantages of probabilistic and analytic motion estimation techniques in order to develop computationally efficient and robust algorithms for motion feature description and detection of occlusion boundaries. Parametric models can describe discontinuous optical flow and the motions of large image regions may be described with a single set of parameters.

Error Metrics and Statistics

A good description of error metrics for optic flow measurements can be found in [McCane, B. et al., 2001, page 133]. There are a lot of error metrics defined in the literature. The most importan ones are

- The angular error e:

$$e := \arccos(u_{correct} \cdot u_{estimated})$$

  in homogeneous coordinates, i.e. $u = (u_x, u_y, 1)$.

- The normalized magnitude of the vector difference.

Also some statistics are used to benchmark the accuracy of motion estimation methods.

- Average or variance of the angular error or normalized magnitude.
- Cumulative histogram graphs, [McCane, B. et al., 2001, pages 135–140, figures 6,7,9–11].

Numerical experiments for the model in $BV(\Omega)$ are depicted in the table below.

$\Psi(s)$	$\mathbb{E}(ang.error)$	$\sigma^2(ang.error)$	SNR
$2\sqrt{1+s^2}-2$	3.87	9.20	12.27
$\log 1+s^2$	4.32	9.24	12.51

Table 1.3: Error estimates for the model proposed in [Aubert et al., 1999]

The next table represents state of the art results for different types of regularizers for motion estimation.

Sequence	2D Linear	2D Nonlin.	3D Linear	3D Nonlin
Marble	5.30	5.14	2.06	1.70
Office	4.33	4.13	3.60	3.24
Yosemite	7.14	6.03	6.18	5.18

Table 1.4: Average angular error for linear and nonlinear combined local-global methods according to [Bruhn, A. et al., 2005b]

Discontinuity-Preserving Real-Time Computations

In [Bruhn, A. et al., 2005a] the authors achieved real-time performance for an isotropic flow-driven regularization.

$$E(u) := \int_{\Omega} (\nabla f \cdot u + f_\theta)^2 + \lambda \sqrt{|\nabla u_1|^2 + |\nabla u_2| + \varepsilon^2}\, dx$$

The image size is 160×120 pixel. The obtained flow fields are dense with a computational speed of 12 frames per second.

1.4 Time-frequency signal analysis

1.4.1 Uncertainty principles for the continuous wavelet transform

Definition 1. *A function $\psi \in L^2(\mathbb{R})$, which fulfills the admissibility condition*

$$0 < c_\psi := 2\pi \int_{\mathbb{R}} \frac{|\hat{\psi}(\omega)|^2}{|\omega|}\, d\omega < \infty$$

is called a wavelet . The wavelet transform $L_\psi[f]$ of a function $f \in L^2(\mathbb{R})$ associated with a wavelet ψ is given by

$$L_\psi[f](a,b) := \frac{1}{\sqrt{c_\psi}} |a|^{-1/2} \int_{\mathbb{R}} f(t)\psi\left(\frac{t-b}{a}\right) dt, \qquad a \in \mathbb{R} \setminus \{0\}, b \in \mathbb{R} .$$

Theorem 1.4. *The Fourier Transform*

$$\hat{f}(\omega) = (2\pi)^{-\frac{1}{2}} \int_{\mathbb{R}} f(x) e^{-i\omega x}\, dx$$

for any $f \in L^1(\mathbb{R})$ exists as an ordinary Lebesgue integral with the following properties

1. $\|\hat{f}\|_\infty \leq \|f\|_1 (2\pi)^{-\frac{1}{2}}$,
2. $\hat{f} \in C(\mathbb{R})$,
3. $\|f|\_2 = \lim_{|\omega|} \to \infty \hat{f}(\omega) = 0$,
4. $\mathcal{F}(f * g) = \sqrt{2\pi}\hat{f}\hat{g}$,
5. $\hat{f} = 0 \Leftrightarrow f = 0$.

Proof. [Prassad, L. and Iyengar, S.S., 1997, Theorem 4.18, page 97] □

The Fourier transform $\hat{\psi}$ is continuous by the theorem 1.4. From the admissibility condition in the wavelet definition it follows, that $\hat{\psi}(0) = 0$, otherwise there will be a contradiction. This means, that a wavelet is zero mean, because

$$0 = \hat{\psi}(0) = (2\pi)^{-1/2} \int_{\mathbb{R}} \psi(t)\, dt.$$

In certain applications in physics, signal and image processing there is a need of information about the frequency distribution of a function over time. For a one dimensional signal f we would like to associate a function $D[f](t,\omega)$, which shows the frequency ω at time t of the signal f. For multidimensional signals $\mathbf{f}$ the distribution $D[\mathbf{f}](t,\omega)$ describes the wavenumber vector ω at time t. The set of all pairs $\{(t,\omega) \mid t,\omega \in \mathbb{R}\}$ is called the phase space and D is called the phase space representation of f. Sometimes, for one dimensional signals the phase space is referred to as the time-frequency plane.

Definition 2. *Let $g \in L^2(\mathbb{R})$. The quantities*

$$\int_{\mathbb{R}} (t - t_0)\, |g(t)|^2\, dt$$

and

$$\int_{\mathbb{R}} (\omega - \omega_0)\, |\hat{g}(\omega)|^2\, dt$$

are called the duration and the bandwidth of the signal $g(t)$.

THEOREM 1.5. *(Uncertainty principle) Let* $g \in L^2(\mathbb{R}), \|g\|_{L^2(\mathbb{R})} = 1$. *Then*

$$\int_{\mathbb{R}} (t - t_0)^2 \, |g(t)|^2 \, dt \int_{\mathbb{R}} (\omega - \omega_0)^2 \, |\hat{g}(\omega)|^2 \, d\omega \geq \frac{1}{4} \, .$$

Proof. [Neumann, 1968] □

DEFINITION 3. *Let* $g \in L^2(\mathbb{R}), \|g\|_{L^2(\mathbb{R})} = 1$ *of duration* t_0 *and bandwidth* ω_0. *Then we say, that* g *is localized around the point* (t_0, ω_0) *with the uncertainty*

$$\mu(g) := \int_{\mathbb{R}} (t - t_0)^2 \, |g(t)|^2 \, dt \int_{\mathbb{R}} (\omega - \omega_0)^2 \, |\hat{g}(\omega)|^2 \, d\omega.$$

The motivation for the uncertainty relation and localization comes from quantum mechanics. The interpretation of the uncertainty principle in signal processing is, that a signal can not be localized arbitrarily well in time and frequency. Some interesting remarks about the interpretation of the uncertainty principle can be found in [Cohen, 1989, pages 970–971]. The relation between operators in quantum mechanics and signal processing is listed in [Cohen, 1989, table 4, page 970].

Linear time-invariant techniques and the short time Fourier transform

The Fourier transform yields no localization in time. Usually, time invariant techniques for signal analysis like the Fourier transform change differential equations into algebraic equations. Then the linear algebra methods are used to solve the equations. The eigenfunctions of the linear time-invariant differential operators are of the form

$$c \, e^{-i \, \omega t}, \qquad c \in \mathbb{R} \, .$$

The next step in signal and image analysis is investigation of signals with frequency content, which vary with time. As the eigenfunctions $e^{-i \, \omega t}$ characterize only the frequency ω and not the time t or equivalently, the Fourier transform yields no localization in time, we need another analysis tool to study the phase space representation $D[f]$ of a signal f. Historically, the short time Fourier transform was first introduced to investigate the behavior of a signal in the phase space. The short time Fourier transform is defined by

$$\mathcal{F}_\psi[f](p, q) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \, e^{-i \, qx} \, h(x - p) \, dx,$$

where h is a windowing function.

The wavelet transform

The magnitude squared short time Fourier transform is called the spectrogram. The spectrogram often represents serious difficulties. If the analysis window h is made small enough to capture rapid changes in the signal, it becomes impossible to resolve signal components that are close in frequency within the analysis frequency duration. An alternative approach with variable resolution in the phase space is the wavelet transform. Let ψ be a wavelet with $\|\psi\|_{L^2(\mathbb{R})} = 1$ and

$$\int_{\mathbb{R}} t \, |\psi| \, dt = 0.$$

We introduce the central frequency ω^+ of the wavelet by

$$\omega^+ := \int_0^\infty \omega \, |\psi(\hat{\omega})| \, d\omega.$$

The dilated and shifted version

$$\psi_{ab} := a^{-1/2} \psi \left(\frac{t-b}{a} \right)$$

of a wavelet ψ is a localization around

$$t_0^{ab} = \frac{1}{a} \int_{\mathbb{R}} t \, (\psi \left(\frac{t-b}{a} \right))^2 \, dt$$

$$\omega_0^{ab} = a \int_0^\infty \omega \, |\hat{\psi}(a\omega)|^2 \, d\omega$$

With $(a, b) \in \mathbb{R}^2$ $a \neq 0$ we can interpret

$$L_\psi[f](a, b) = D[f] \left(b, \frac{\omega_0 t}{a} \right)$$

as a phase space localization of f.
The inversion formula for the wavelet transform is given by the following lemma.

Lemma 1.6. *The adjoint operator*

$$\begin{aligned} L_\psi^* : L^2(\mathbb{R}) &\to L^2(\mathbb{R}) \\ g &\mapsto c_\psi^{-1/2} \int_{\mathbb{R}^2} |a|^{-1/2} \psi \left(\frac{t-b}{a} \right) g(a, b) \frac{dadb}{a^2} \end{aligned}$$

gives the inversion of the wavelet transform onto its range.

Proof. [Alfred Karl Louis et al., 1994, page 22, theorem 1.19] □

Remark: The function with the minimal uncertainty, centered around (t_0, ω_0) is

$$g_{t_0\omega_0}(t) = \pi^{-1/4} \, e^{-i\omega_0 t} \, e^{-(t-t_0)^2/2}$$

and

$$\mu(g_{t_0\omega_0}) = 1/4 \, .$$

If a wavelet has a maximum in ω_0, then $L_\psi[f](a, b)$ and $F[L_\psi f](a, \omega)$ are determined by the contribution of the frequencies around ω_0/a to the function f.

The information about f is contained in $D[f](t, \omega)$. It is possible to reconstruct f from its phase space representation

$$\int_{\mathbb{R}^2} Df(s, \omega) f_{s\omega}(t) ds d\omega = \omega_0 \int_{\mathbb{R}^2} L_\psi f(a, b) a^{-1/2} \psi \left(\frac{t-b}{a} \right) \frac{dadb}{a^2}.$$

In the following, we introduce the notion of scale for the wavelet transform.

Definition 4. *The scale a is the inverse frequency with respect to a reference frequency* ω_0.

$$a := \frac{\omega_0}{\omega}$$

Here the ω_0 is the central frequency of the wavelet.

For $\omega \neq 0$ the phase space can be transformed into a time scale plane by using the definition for scale. The extension of wavelet transforms to time-scale energy distributions is discussed in [Rioul, Olivier and Flandrin, Patrick, 1992]. At the intersection of the time -frequency and time-scale distributions is the Wigner-Ville transform. It is known also as the Cohen's class or the quadratic class of transforms. The Wigner-Ville transform is a classical tool for time-frequency signal analysis.

Cohen's class of transforms

The Cohen's class of transforms or the quadratic class of time-frequency transforms is defined with respect to a kernel $\Phi(\theta, \tau)$. By varying the kernel, we can obtain different time-frequency transforms of the quadratic class such as Wigner-Ville, Richczek, Page and Choi-Williams transform. The kernels for some transforms are listed in [Cohen, 1989, table 1, page 952].

Definition 5. *For a signal* f *the Cohen's class of transforms* $C[f](t, \omega; \Phi)$ *is defined by*

$$C[f](t, \omega; \Phi) := \frac{1}{2\pi} \int_{\mathbb{R}^3} e^{i(\theta t - \tau\omega + \theta u)} \Phi(\theta, \tau) f(u + \tau/2) f^*(u - \tau/2) du d\tau d\theta \,,$$

where $\Phi(\theta, \tau)$ *is the kernel of the transform.*

The Wigner-Ville transform $W[f]$ can be obtained from the last definition by setting the kernel $\Phi(\theta, \tau) := 1$.

$$W[f](t, \omega) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t + \tau/2) f^*(t - \tau/2) e^{-i\,\omega\tau}\, d\tau$$

The quadratic class can be interpreted as the Fourier transform of the signal's autocorrelation function with respect to the delay variable. This interpretation is suitable for the analysis and design of reduced inference transforms of the Cohen's class. Usually, the Wigner-Ville transform of a signal contains phantom answers in regions between two real peaks. The reduced inference transforms suppress to some extent the phantom answers and localizes better the real peaks. To sketch out the idea of reduced inference time-frequency transforms we need the definition of the autocorrelation and ambiguity function, because the suppression of interferences is fulfilled in the ambiguity domain.

The instantaneous autocorrelation $R[f](t, \tau)$ of a complex signal f is defined as

$$R[f](t, \tau) := f(t + \tau/2) f^*(t - \tau/2) \,.$$

In terms of the autocorrelation function R, the Wigner-Ville transform $W[f](t, \omega)$ can be expressed as the Fourier transform of R with respect to the variable τ.

$$W[f](t, \omega) := \mathcal{F}[f(t + \tau/2) f^*(t - \tau/2)] = \mathcal{F}[R[f](t, \tau)]$$

The symmetrical ambiguity function $A[f](\theta, \tau)$ is introduced as the inverse Fourier transform of the instantaneous autocorrelation $R(t, \tau)$ with respect to the time variable t.

$$A[f](\theta, \tau) := \mathcal{F}^{-1}[f(t + \tau/2) f^*(t - \tau/2)] = \mathcal{F}^{-1}[R(t, \tau)]$$

In the ambiguity domain, the contributions to the phantom answers come from regions far away from the origin. To suppress these interferences, [Debnath, 2001, pages 381–418] proposes to convolve the ambiguity function by a two dimensional binomial or Gaussian kernel, centered at the origin. [R. Hamila et al., 1997, 1999] reveal a relation between the ambiguity function and the energy operator, cf. chapter 4.

Uncertainty principles for the continuous wavelet transform

In the previous sections we discussed some aspects of the Heisenberg uncertainty principle in time-frequency and time-scale signal analysis. As a consequence it follows, that a nonzero signal cannot be arbitrarily well localized simultaneously in time and frequency domain or in time and scale domain as well. Our aim of wavelet decomposition is optic flow and motion estimation. This means, it would be desirable to use wavelets, which minimize the uncertainty relation. Therefore, when working with wavelet decomposed image data, the question of optimally localized wavelets arise.

The minimizing function of an affine uncertainty principle in one and two dimensions was stated in [Dahlke, S. and Maass, P., 1995]. As a main result of the analysis therein, it was found, that for the wavelet transform the equivalent of the Gaussian function in the windowed Fourier transform is the Mexican hat function. We give the following statement for the wavelet choice with respect to minimization of the uncertainty relation.

Statement 1.7. *When there is no* a priori *information about the signal, it can be decomposed by using the continuous wavelet transform with the Mexican hat wavelet, because in two dimensions it is optimal with respect to minimization of the uncertainty principle, [Dahlke, S. and Maass, P., 1995].*

Further uncertainty inequalities of Heisenberg type are stated in [Singer, 1999] and [Wilczok, 2000].

On the existence of optimally localized wavelets

The optimal localization properties of wavelets can be formulated in terms of the reproducing kernel of the wavelet analysis. For example, the Heisenberg uncertainty principle means, that there is no wavelet such that the associated reproducing kernel is compactly supported. In [Holschneider and Teschke, 2005] the authors defined optimality criteria and proved the existence of optimally localized reproducing kernel. The cost functional to be minimized is weak lower semi-continuous and possesses a minimum over any weak * compact set. Since the problem to solve is nonlinear and there is no closed form analytical solution, the authors provided a numerical strategy to compute a critical point of the functional. If the algorithm is initialized with a function, which is not far away from the expected solution, the method yields at least a critical point. Under additional assumptions on the solution and on the functional itself, the critical point is a global minimizer.

1.5 Conclusion

The goal of this chapter was to introduce the common methods for optic flow computation and regularization by parabolic PDE's, which respect the discontinuity of the solution. We sketched out the construction of the discontinuity-preserving regularization term and showed the existence and uniqueness of the solution in $L^2(\Omega)$ and $BV(\Omega)$. Real-time discontinuity-preserving computation is possible.

The smoothness assumption models in optic flow and motion estimation with discontinuity preservation lead to parabolic PDEs.

Parabolic PDEs can be used to generate a scale-space for an image sequence. According to the scale-space axioms of [Alvarez, L. et al., 1993], the scale-space can be generated only by parabolic PDEs.

The Mexican hat wavelet yields in two dimensions the best localization in the phase space. If there is no a priori knowledge about the image data, the data can be decomposed by the Mexican hat wavelet.

At the end we mention two open questions. Are there other more useful rotation invariant convex regularizers? How can we prove the existence of solutions for the curvature-preserving PDE (1.15)?

Chapter 2

Motivation: Scale-spaces for motion estimation

There are some major limitations of optical flow analysis which can be overcome by processing a scale-space stack of the given image sequence.

One of this limitations is, that only small displacements can be estimated by local differential techniques, otherwise temporal aliasing can occur. The conditions for temporal aliasing appearance are discussed in section 5.1, page 67.

Another limitation is, that the standard coarse-to-fine strategy does not work for many types of complex motion. At points of occlusion, transparent motion, disappearing or arising objects it doesn't make much sense to warp the image by the estimated flow field because the flow field itself is inaccurate or can not be estimated at those points.

An open issue is the optimal degree of smoothing for the local structure tensor. Some other orientation estimation operators like the energy operator are defined only for a narrow band signals and have some nice properties with respect to phase invariance compared to the structure tensor. The scale-space decomposition of an image sequence allows us to use the energy operator instead of the simple local structure tensor.

An intermediate step in the spatio-temporal differential based optic flow and motion estimation is the computation of a tensor field for the entire image sequence. Since this tensor field is calculated from the partial derivatives of the image data it contains noise. By a proper denoising of the tensor field we can improve the accuracy of the estimated flow field. The denoising can be fulfilled on the manifold $S\mathbb{O}(n)$ of the special orthogonal group. If we use some kind of nonlinear diffusion to denoise the tensor field, this can be viewed as a scale-space of the feature extracted from the image data. By generating the scale-space on the Riemannian manifold $S\mathbb{O}(n)$ we introduce an additional constraint for the scale space computation and respect better the properties of the tensor field by using the natural Riemann metrics. Another manifold, which respects the properties of the linear model for motion estimation by local differential techniques is the manifold of matrices with given set of eigenvalues. The scale space, restricted on this manifold can be generated by the isospectral flow and it represents a trade-off between accuracy and computational complexity.

The scale-space analysis of the image data offers a possible method for detection of temporal aliasing. The main reason for temporal aliasing occurrence is leakage of high frequencies into the low frequency part of the Fourier spectrum of the spatio-temporal image data. In the scale-space stack there are only narrow band images. By estimating the optic flow in parallel we can search for inconsistencies in the flow field across the scales. In this manner we can detect regions of the image sequence, where temporal aliasing occurs.

2.1 Related Work

Here are some important developments in motion estimation in the scale-space and motion estimation with wavelet-based techniques from mathematical and image processing point of view.

- [Simoncelli, 1993] introduced a coarse-to-fine strategy with warping techniques
- [Martin Lefébure and Laurent D. Cohen, 2001]
  Theoretical results on warping with applications to
  - motion estimation
  - image registration
  - rigidity and non rigid deformations
- [Bernard, 2001]
  Discrete wavelet analysis for optic flow computation
- [Demonceaux and Kachi-Akkouche, 2004]
  Motion detection using wavelet analysis and robust techniques
- [Mujica et al., 2000]
  Continuous wavelet transform for motion parameter estimation

The coarse-to-fine strategy is well described in [Simoncelli, 1993, pages 60–65]. The technique is inspired by a biological visual system, where the sensors can track a given object in the imagery scene. By the tracking, the object velocity decreases or vanishes completely. This mechanism can be used only in an interactive setting. For already acquired sequences we can warp the image in a direction opposite to the motion. The motion field for the warping can be computed at a coarser scale. This step is called motion compensation . If the motion compensation is done by the velocity vector with the biggest magnitude, then this simple correction step is called linear motion compensation or motion compensation without warping .

The coarse-to-fine strategy suffers from a serious drawback. If the coarse-scale estimates are missing or incorrect we can neither warp the image nor correct the errors. This situation can occur at points of complex motion such as occlusion, transparency, appearing or vanishing objects.

[Bernard, 2001] developed a more elaborated coarse-to-fine strategy for motion estimation based on compactly supported wavelets. In the following there is a brief overview of the proposed method, which is inspired by the concept of the weak derivatives and the distribution theory.

Let $\Omega \subset \mathbb{R}^2, T \in \mathbb{R}$. Given $f \in L^2(\Omega, [0,T])$ we have to estimate the velocity $v \in C^2$.

Let $\{\psi^k\}_{0<k<K}$ be a family of complex mother wavelets with compact support.

$$\psi^k_{u,s}(x) := \frac{1}{s}\psi^k\left(\frac{x_1-u_1}{s}, \frac{x_2-u_2}{s}\right)$$

Assumption:

$$v(x,t) = v(u,t) + \varepsilon \quad \text{where } \varepsilon \text{ is an error term .}$$
$$\nabla f \cdot v = -\frac{\partial f}{\partial t} \quad \text{optical flow equation}$$

$$\left\langle f, \frac{\partial \psi^k_{u,s}}{\partial x_1} \right\rangle v_1 + \left\langle f, \frac{\partial \psi^k_{u,s}}{\partial x_2} \right\rangle v_2 = \frac{\partial}{\partial t} \langle f, \psi^k_{u,s} \rangle + \varepsilon_s(u,t) \qquad \forall k = 0..K$$

$\left\langle f, \frac{\partial \psi^k_{u,s}}{\partial x_i} \right\rangle$ are the wavelet coefficients of f in a basis, consisting of derivatives.

[Bernard, 2001] proved, that if $v \in C^2$ and if f is Lipschitz at u, then $\varepsilon_s \to 0$ with $s \to 0, \forall u, t$.

To overcome problems with temporal aliasing, the assumption is made, that the displacement between two consecutive frames is small compared to the wavelet support.

Discrete differential quotient

$$\frac{\partial}{\partial t} \langle f, \psi \rangle = \frac{1}{\Delta t} \langle f_{p+1} - f_p, \psi \rangle \, .$$

$$\left\langle \frac{f_p + f_{p+1}}{2}, \frac{\partial \psi^k_{u,s}}{\partial x_1} \right\rangle v_1 + \left\langle \frac{f_p + f_{p+1}}{2}, \frac{\partial \psi^k_{u,s}}{\partial x_2} \right\rangle v_2 = \frac{\partial}{\partial t} \langle f, \psi^k_{u,s} \rangle \quad \forall k$$

For the estimate of the velocity $v_{u,s}$ we obtain the K by 2 system

$$W_{u,s} v_{u,s} = D_{u,s} \, . \tag{2.1}$$

Least squares solution of (2.1) yields the estimated optic flow field. Eigenvalue analysis of

$$\mathrm{Real}(W^*_{u,s} W_{u,s})$$

is also performed.

Five separable analytic wavelets calculated with Daubechies conjugate mirror filters are used.

The method of [Bernard, 2001] is a multi-scale approach with motion compensation without warping between the different scales.

[Demonceaux and Kachi-Akkouche, 2004] extended this technique to an algorithm with warping and robust M-estimator.

[Mujica et al., 2000] proposed a frame-to-frame tracking technique. The motion estimation takes place in the Fourier domain. The algorithm works as described below.

A linear motion defines a velocity plane

$$(k^T \; \omega) \cdot (v \; 1)^T = 0 \, .$$

Let ψ be a mother wavelet used to derive the wavelet basis by a set of parameters.

Transformation of the wavelet basis to match the motion characteristics of the tracked object.

Transformations:	spatio-temporal translation	T	b, τ
	scaling	D	a
	speed adaptation	Λ	c
	rotation	R	θ

We can make a suitable choice of the parameters $p := (b, \tau, a, c, \theta)$ to have a norm preserving map

$$\psi \mapsto D^a \Lambda^c R^\theta T^{b,\tau} \, \psi \;\; =: \psi_p .$$

The continuous wavelet transform of input image f is defined by

$$W_p f := \frac{1}{\sqrt{c}} \langle \psi_p, f \rangle = \int_{\mathbb{R}^3} \psi_p(x,t) f(x,t) \, dx \, dt \, .$$

If $c = 1$ then

$$\|f\|^2 = \|\mathcal{F}(f)\|^2 = \int_{\mathbb{R}^5} W_p f \; dp \; .$$

Energy densities which can be used to derive local estimates of the motion parameters:

$$\begin{aligned} E^1_{a_0,\tau_0}(c,\theta): &\quad \text{speed orientation energy density} \\ E^2_{a_0,c_0,\theta_0,\tau_0}(b): &\quad \text{spatial energy density} \\ E^3_{c_0,\theta_0,\tau_0}(a): &\quad \text{scale energy density} \end{aligned}$$

$$\text{For instance} \quad E^1_{a_0,\tau_0}(c,\theta) := \int_{\mathcal{B}} |\; W_p f \;|^2 \; db \; .$$

Minimization of the energies is performed to update the state vector $L(t_i) := (v_{t_i}, x_{t_i}, a_{t_i})$.
The wavelets can be seen as filters controlled by the motion parameters.
Separable Morlet wavelets were used: $\psi(k,\omega) = K(k)\Omega(\omega)$.

2.2 Scale-Space Analysis with 2D Rotationally Invariant Wavelets

The Continuous Wavelet Transform

Definition: A wavelet is a function $\psi \in L^2(\mathbb{R})$, centered in the neighborhood of 0 with

$$\int_{\mathbb{R}} \psi(t) \; dt = 0$$

and

$$\|\psi\|_{L^2} = 1 \; .$$

The continuous wavelet transform is defined by

$$\psi_{u,s}(t) := \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

$$Wf(u,s) := \langle f, \psi_{u,s} \rangle := \int_{\mathbb{R}} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) \; dt \; .$$

The continuous wavelet transform offers us optimal spatial smoothing and a continuous scale space representation.

Rotationally Invariant Wavelets in the discrete case

A classical decomposition scheme, with separable 2D Wavelets is given by [Mallat, 1999] or [Heijmans and Goutsias, 2000]. The wavelet transform with separable wavelets is direction dependent.

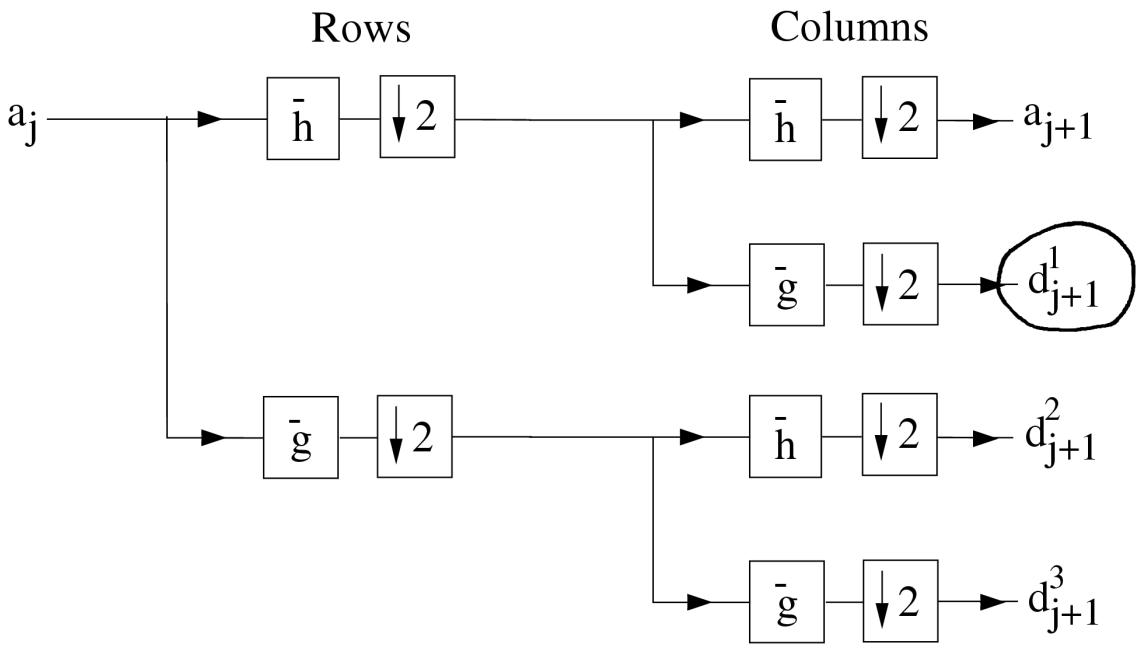


Figure 2.1: Separable 2D wavelets, classical decomposition scheme.

We can illustrate the direction dependence by decomposing the picture above by the separable wavelet transform and compare the results for the coefficient d^1 at different scales.

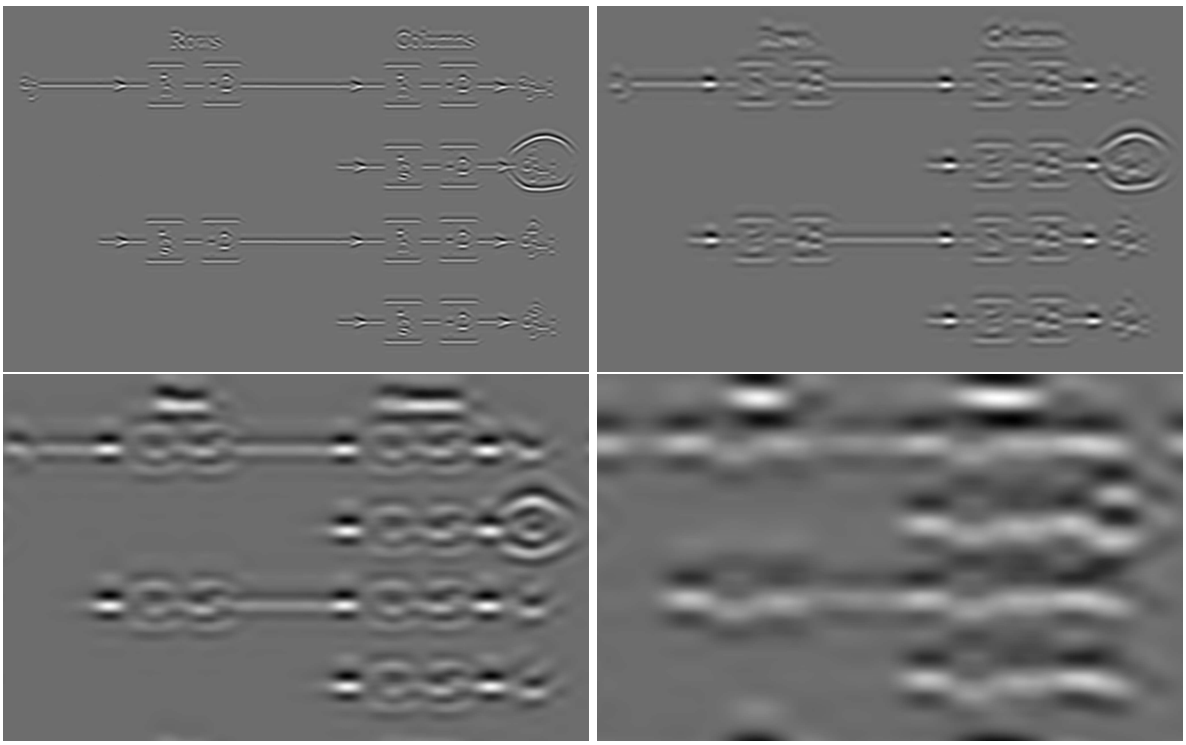
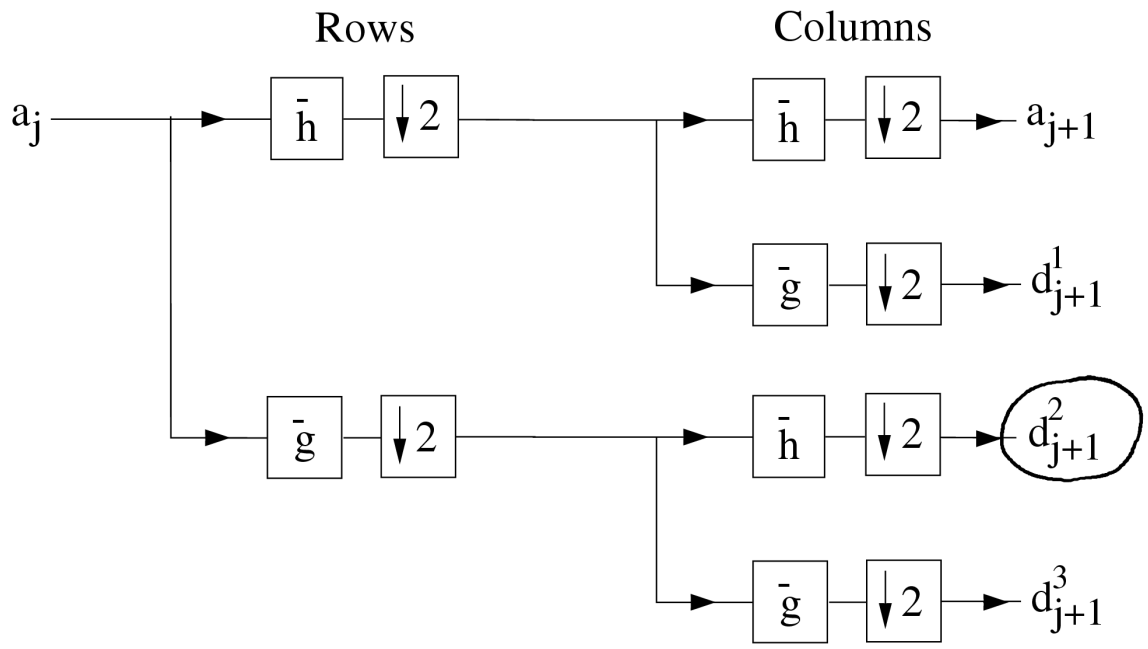


Figure 2.2: Direction dependence in the co-set d^1 at wavelet scales 4, 5, 6 and 7. The horizontal direction is preferred

If we pick up the coefficients d^2 for this scheme



we obtain results, which show high pixel values only at vertical edges.

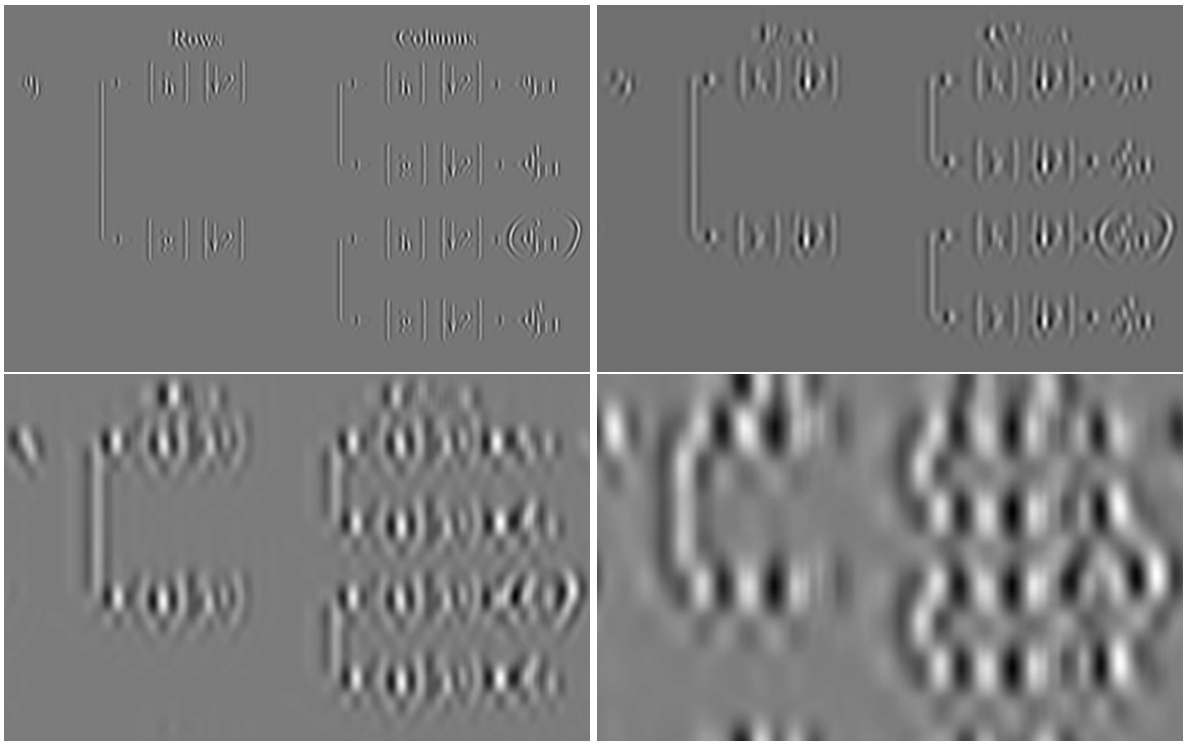


Figure 2.3: Direction dependence in the co-set d^2 at wavelet scales 4, 5, 6 and 7. The vertical direction is preferred

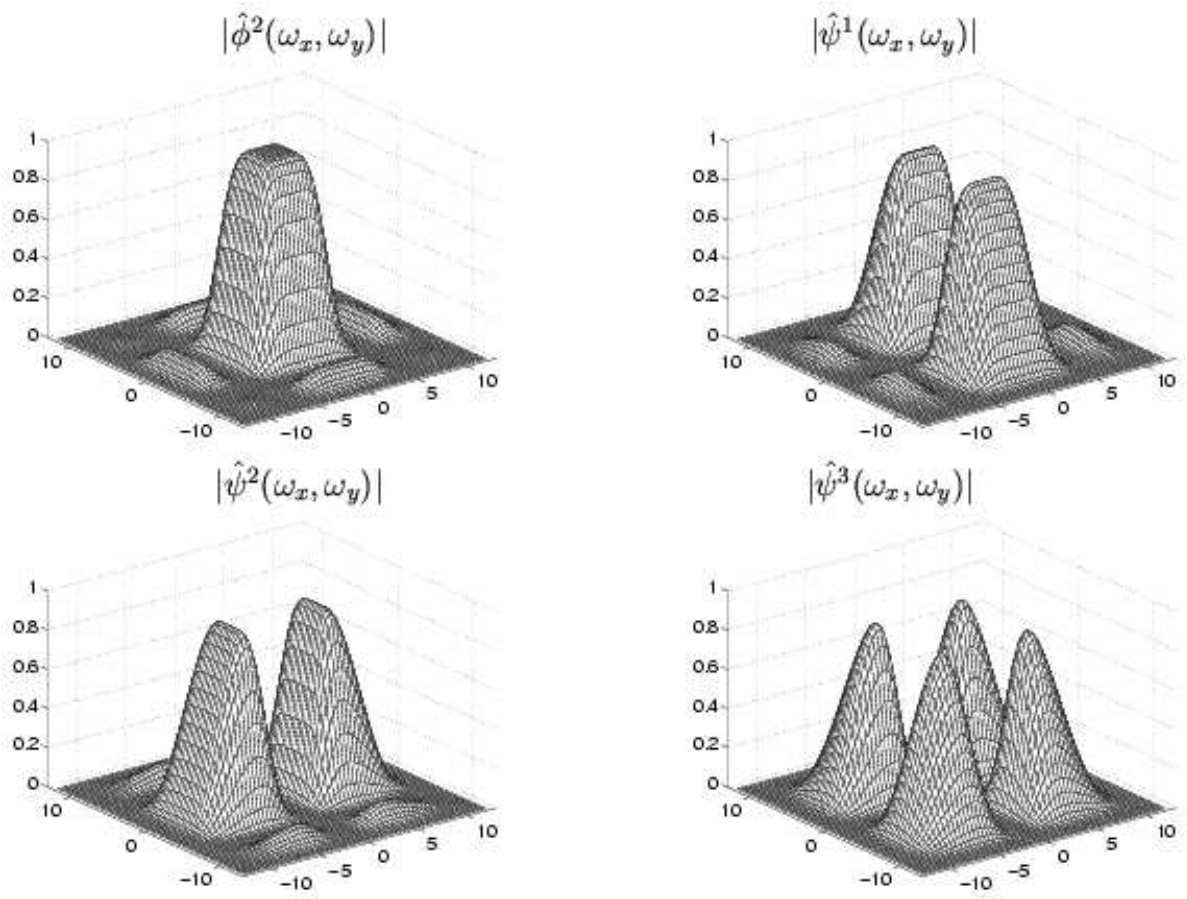


Figure 2.4: Fourier transform of separable 2D Wavelets

Separable wavelets lead to preferred directions.

Motion estimation needs isotropic image decomposition. Otherwise there will be a bias in the computed motion field because of the anisotropic wavelet transform.

There are different ways to obtain isotropic or nearly isotropic wavelets for the wavelet transform.

- new non-separable almost symmetric wavelets
  [Belogay and Wang, 1999] describe an algorithm for construction of arbitrarily smooth wavelets in $\mathbb{R}^2$. Some of the computed wavelets are rather symmetric, see figure 2.5.

- optimization of separable wavelets
  [Lemaur, 2003]: optimality criteria and optimization algorithms

- by generalization of the 1D case: rotation around the z-axis
  Pet-hat wavelet, Wheel wavelet, 2D Mexican-hat wavelet

[Belogay and Wang, 1999] computed a smooth scaling function $\in C^7$.

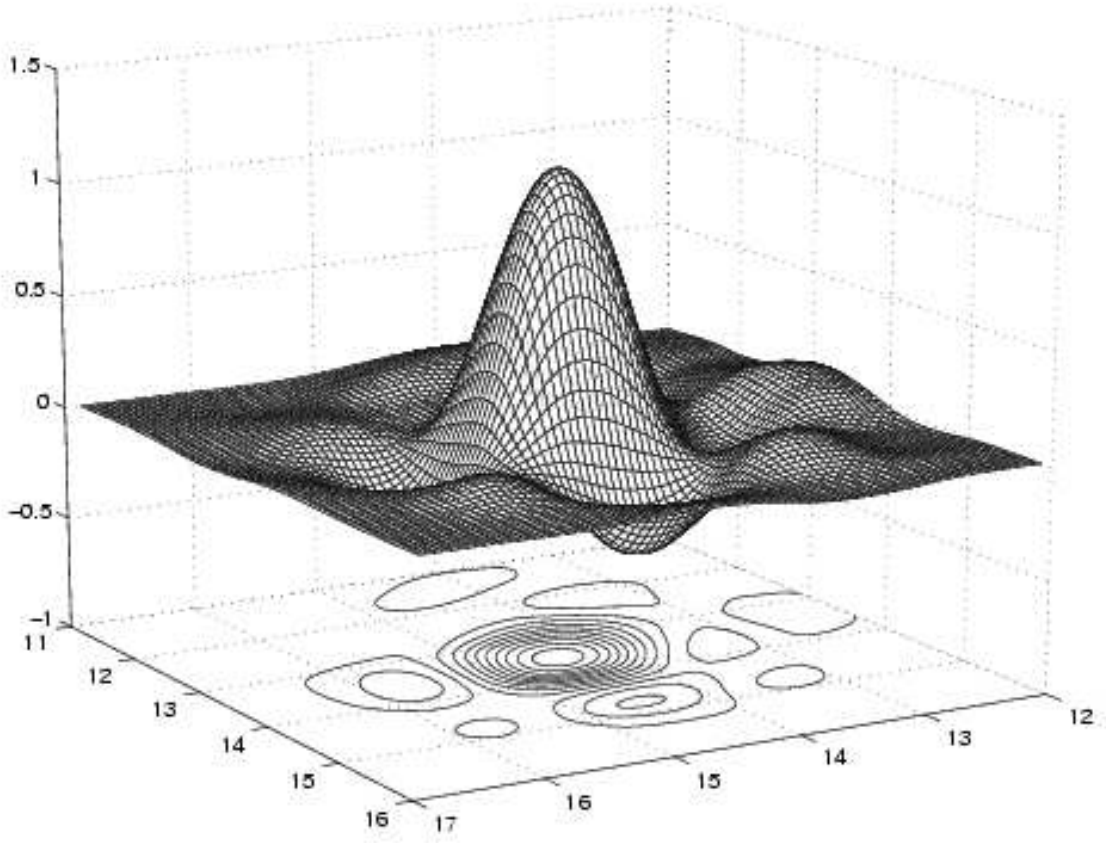


Figure 2.5: [Belogay and Wang, 1999]: Scaling function $\in C^7$

The scaling function shows better rotation invariance properties than the separable 2D wavelet transform.

[Lemaur, 2003] optimizes separable wavelets with respect to isotropy. This is a sketch of the idea in the thesis of [Lemaur, 2003].

We are only interested in separable functions of the form

$$F(x,y) = F_x(x)F_y(y) \qquad \text{with} \qquad F_x = F_y \,,$$

otherwise the function F cannot be isotropic. The function should represent a FIR-filter.

THEOREM 2.1. *All solutions of our problem are of the form* $F_x(x) := ae^{-\frac{x^2}{\sigma}}$.

We are searching for a filter *as close as possible* to a Gaussian, where the closeness is the usual scalar product in $L^2(\mathbb{R})(\mathbb{R}^n)$.

$$\textit{closeness measure} : \frac{\langle F_x, G_x \rangle}{\|F_x\| \cdot \|G_x\|}$$

The optimization is fulfilled by genetic algorithms, random walk algorithms or hill climbing.

2.2.1 Non separable wavelet bases

When we deal with optic flow and motion estimation from image sequences, we should encounter the rotation invariance of the scale decomposition method. Isotropic decomposition can be achieved by the non separable discrete wavelet transform or by the continuous wavelet transform, both with rotationally invariant wavelets.

THEOREM 2.2. *Let $\{V\}_{m\in\mathbb{Z}}$ be a multi resolution analysis of $Ł^2(\mathbb{R}^2)$ with dilation matrix A. Then there exist |det A|-1 wavelets, which build an orthonormal basis of the orthogonal compliment of V_0 in V_1.*

Proof. [Meyer, 1992] □

Image analysis requires isotropy and small number of wavelets. It will be good, if |det A|=2. In this case for the scaling function ψ and the mother wavelet function ψ the following relation holds.

$$\psi(x) = |\det A|^{\frac{1}{2}} \sum_{k\in\mathbb{Z}^2} h_k \phi(Ax-k)$$
$$\phi(x) = |\det A|^{\frac{1}{2}} \sum_{k\in\mathbb{Z}^2} g_k \phi(Ax-k)$$

If we choose

$$g_k := (-1)^{\varepsilon(k)} h_{z-k}, \qquad \text{where } \varepsilon(k) := \begin{cases} 0, & for\ k \in A\mathbb{Z}^2 \\ 1, & for\ k \notin A\mathbb{Z}^2 \end{cases}$$

it can be shown, that this will lead to an orthogonal function

$$\psi(x) = |\det A|^{\frac{1}{2}} \sum_{k} g_k \phi(Ax-k),$$

for which

$$\text{span}\{\psi(\cdot - k)\}_{k\in\mathbb{Z}^2} = W_0\ .$$

For |det A|=2 there are only 3 possible grid patterns in $\mathbb{R}^2$.

- Row grid $A\mathbb{Z}^2 = \{(z_1, z_2) \in \mathbb{Z}^2 \mid z_2 \text{ is even}\}$
- Column grid $A\mathbb{Z}^2 = \{(z_1, z_2) \in \mathbb{Z}^2 \mid z_1 \text{ is even}\}$
- Quincunx $A\mathbb{Z}^2 = \{(z_1, z_2) \in \mathbb{Z}^2 \mid z_1 + z_2 \text{ is even}\}$

By mirroring along the diagonal of the grid one can prove that the row and column grids are equivalent.

Simple not equivalent dilation matrices with |det A|=2 are

$$R := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \qquad S := \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The eigenvalues of R are different from the eigenvalues of S, consequently R is not equivalent to S.

S is equivalent to $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ and leads to separable wavelets.

Thus $R := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ is interesting as dilation matrix. R causes a rotation by an angle of $-\frac{\pi}{4}$ and a dilation by a factor of $\sqrt{2}$.

Norm equivalence between discrete Besov norms and weighted sequences of discrete wavelet coefficients in the non separable case is established in [Lindemann, 2005].

Isotropic Wavelets

After extensive study of different approaches for nearly isotropic wavelet transforms we decided to use the continuous wavelet transform with rotationally symmetric wavelets. To ensure optimal localization in space and frequency domain, the two dimensional Mexican hat wavelet can be used. Depending on the desired bandpass form alternatives can be the two dimensional Pethat-Wavelet or the two dimensional Wheel-Wavelet. The Fourier transforms of the chosen wavelets are depicted on figures 2.6 and 2.7.

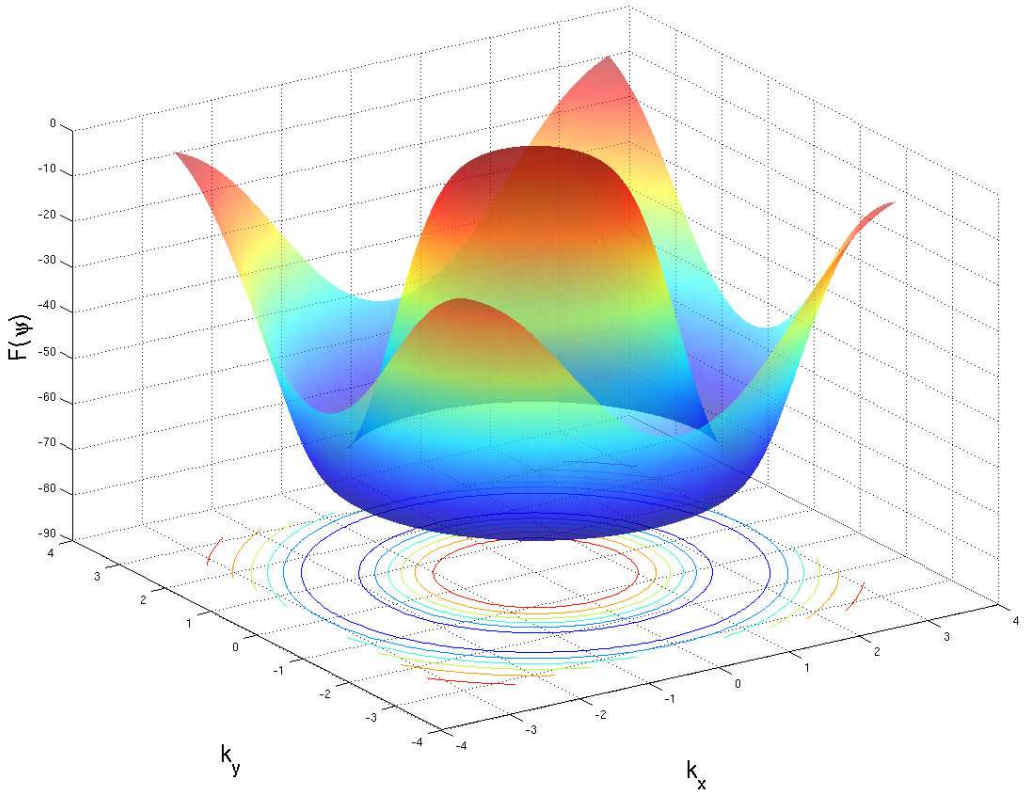


Figure 2.6: Mexican Hat-Wavelet, $\hat{\psi}(k) := -|k|^2 e^{-\frac{1}{2}|Ak|^2}$

The Mexican hat wavelet can be defined by its transfer function

$$\hat{\psi}(k) := -|k|^2 e^{-\frac{1}{2}|Ak|^2}.$$

When there is no a priori knowledge about the images, the Mexican hat wavelet can be used to decompose the image data. [Dahlke, S. and Maass, P., 1995] show that it is optimally localized in space and frequency.

The Fourier transform of the pethat wavelet is given by

$$\hat{\psi}(k) := -\cos^2\left(\frac{\pi}{2}\log_2\left(\frac{|k|}{\sqrt{2}}\right)\right).$$

When a narrow bandpass with sharp cut-off border is needed, one can use the pethat wavelet.

The wheel wavelet has a smooth transition towards the high frequency and a sharp cut off border towards the low frequency. The wheel wavelet is defined by its transfer function

$$\hat{\psi}(k) := -\cos^2\frac{\pi}{2}\frac{\log_2(|k|)}{\log_2(\sigma)}.$$

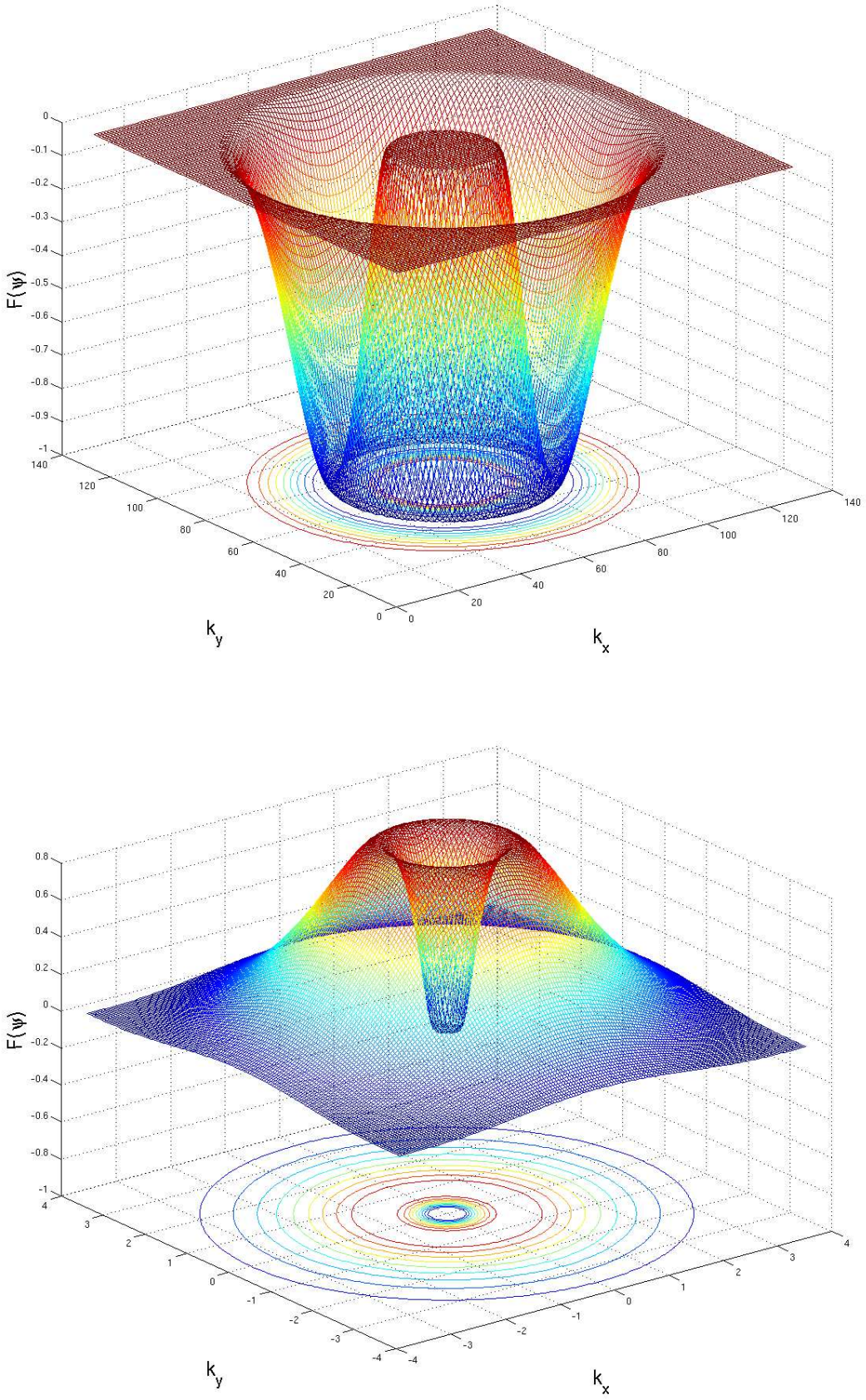


Figure 2.7: Pethat-Wavelet and Wheel-Wavelet

2.3 Conclusion

To our best knowledge there is no method for detection of temporal aliasing in the literature. We studied approaches for optic flow and motion estimation in the scale space and particularly wavelet-based optic flow computation. We decided to make use of the continuous wavelet transform with rotationally invariant wavelets for motion estimation in the scale space in our experiments in chapters 4 and 5. The rotation symmetry of the wavelets is important for direction independent optic flow computation.

The discrete wavelet transform is suitable for compression of multidimensional signal and image data. For analysis purposes, the continuous wavelet transform is a better tool than the discrete one, because it generates a continuous scale representation of the image sequence. It is easy to construct a fully isotropic multidimensional wavelet just by rotation of the a one dimensional one. Additionally, we have the freedom to choose the cut-off behavior of the wavelet filter according to the desirable properties for the image processing task. In the discrete case of the multidimensional wavelet transform based on tensor products there is a trade-off between the frequency response of the wavelet and the isotropy of the scaling function, [Feilner et al., 2005]. The anisotropy is available even for the case of the quincunx decomposition grid, which ensures more isotropic scores than the classical dyadic scheme. The cut-off is coupled with the rotation invariance of the scaling function, [Feilner et al., 2005, fig. 3, fig. 4].

Chapter 3

Tensor scale space on a Riemannian manifold for complex motion analysis

The structure tensor reveals information about the orientation in image data. Optic flow can be estimated by means of the structure tensor. Under the assumption of the gray value conservation, the apparent motion takes place in the direction of the eigenvector to the smallest eigenvalue. This differential technique is widely used in practice and yields good results with respect to accuracy. Also some weaker assumptions for the intensity conservation over time can be made. In heat flux measurements it is physically funded to make the assumption, that the gray level decays over time according to a diffusion process. This model is discussed in [Garbe, 2001, pages 115–117]. The tensor method for optic flow estimation works well with this kind of weakened assumptions for the intensities and has wide application in heat flux measurement at the sea surface and water transport in plant leaves.

Because of the ill posedness of the optic flow problem for a single pixel, the tensor should be smoothed out or alternatively some regularization techniques should be applied. Combined local-global strategies are proposed in [Andrés Bruhn et al., 2002; Bruhn, A. et al., 2005b]. In this chapter we will discuss local strategies for integration of the structure tensor.

Although the motion estimation problem is quite mature, maybe even one of the best investigated in image processing, the amount of smoothing for the structure tensor is still an open issue. The smoothing region determines the scale, at which some feature is extracted from the tensor. Too small smoothing radius won't supply enough structure information, on the other hand too large smoothing regions will smear out edges, corners and motion boundaries.

Hence it is desirable to integrate the tensor field in an adaptive manner along spatio-temporal tubes representing the object trajectories.

We will investigate a method for adaptive integration of tensor fields with respect to motion estimation. The smoothing is fulfilled by a suitable nonlinear diffusion strategy, which is then applied on the manifold of matrices with given set of eigenvalues, cf. [Tschumperle, 2002].

3.1 Related work

As already discussed in 1.3.2 and [Tschumperle, 2002, 2005; Tschumperle, D. and Deriche, R., 2005] there is a general smoothing on corners, when we use the nonlinear diffusion, based on total variation flow. In order to circumvent this drawback, [Brox, Thomas et al., 2006] defines a coherence dependent map, which stops the diffusion near corners. For optic flow estimation, we propose the formalism of curvature preserving PDE's for the tensor field, [Tschumperle, D. and Deriche, R., 2005] to avoid over-smoothing on corners. This is just a result of the analysis

in [Tschumperle, 2002, 2005; Tschumperle, D. and Deriche, R., 2005]. It is a direct approach simply following the formalism for the heat flow equation, constrained on a curve. The adaptive, curvature dependent metrics drives the diffusion according to the desired behavior by itself.

The next theoretical part of this chapter is to choose a proper integration scheme, which constraints the diffusion flow on the manifold $S\mathbb{O}(3)$ of the orthogonal matrices. We represent different flows and conduct experiments with the isospectral flow . In optic flow and motion estimation not only the computation of the flow field is important, but also the confidence of the estimated flow. The most of the confidence measures rely on the eigenvalues The isospectral flow leaves the eigenvalues of second order tensors untouched. This is an intrinsic property of the flow, at least analytically. This means it preserves the confidence measure locally. That's why we decided to employ an integration scheme, based on the isospectral flow . Additionally it should be mentioned, that the isospectral flow represents a good trade-off between performance and computational costs, [Tschumperle, 2002, page 149].

The interested reader is encouraged to take a look in [Brox, Thomas et al., 2006] , where a good overview of adaptive and nonlinear smoothing techniques for the structure tensor is presented.

3.2 Tensor field regularization by diffusion

Diffusion tensor regularization can be used to general symmetric and semi-positive definite matrices such as structure tensors or covariance matrices.

Let $\Omega \in \mathbb{R}^n$. $T : \Omega \to P^{n\times n}$, P: positive semi-definite matrices. Multi-valued regularization process in variational, divergence and trace-based formulation.

$$\int_\Omega \phi(\|\nabla T\|)\, dx \to \min!$$

$$\frac{\partial T_i}{\partial t} = \operatorname{div}\left(\frac{\phi'\|\nabla T_i\|}{\|\nabla T_i\|}\nabla T_i\right), \qquad (i = 1, ..., n)$$

$$\frac{\partial T_i}{\partial t} = \operatorname{trace}(DH_i), \qquad (i = 1, ..., n)$$

Direct approach

Direct approach for tensor field regularization Analogy with cromaticity denoising in color images.

$$\frac{\partial T_i}{\partial t} = \operatorname{trace}(DH_i) \quad H_i : \text{ Hessian of } T_i$$

D is the smoothing geometry, for instance

$$D := \frac{1}{\sqrt{\sum \lambda_i}}, \qquad \lambda_i \in \operatorname{spec}(\nabla T_i \otimes \nabla T_i).$$

The PDE intrinsically preserves the matrix symmetry. Application only on the upper right triangular part of T. Numerical preservation of the semi-positive definiteness: reprojection into the semi-positive cone after each PDE iteration.

Preservation of the semi-positive definiteness

$$T = U\,\Gamma\,U^T, \quad U \in \mathrm{SO(n)}, \quad \Gamma := \operatorname{diag}(\lambda_1, ..., \lambda_n)$$

Projection of T into the semi-positive cone.

$$\mathcal{P}(T) := U \operatorname{diag}(\tilde{\lambda}_1, ..., \tilde{\lambda}_n)\, U_T, \quad \tilde{\lambda}_i := \lambda_i,\ \lambda_i \geq 0$$
$$\tilde{\lambda}_i := 0,\ \lambda_i < 0$$

PROPOSITION 3.1. *The projection P minimizes the distance between T and $\mathcal{P}(T)$ in the sense of the Frobenius norm.*

Proof. Cf. [Tschumperle, 2002, Page 139]. □

Some PDE flows preserve the semi-positive definiteness, cf [Weickert, 1996]
There are some drawbacks of the direct approach.

- The reprojection requires a time consuming spectral decomposition of T.
- There isn't any direct control on the spectral elements of T, which are the essential features of interest. Orientation U and diffusivity Γ are the relevant data and decompose the tensor information.

As a consequence of the direct approach we can mention, that the PDE regularizes the orientations and the diffusivities in a coupled way. Orthogonal neigbour tensors swell instead of aligning themselves. This leads to an eigenvalue swelling effect as a result of the regurarisation, which is not desirable. Orthogonal neigbour tensors converge to identity matrices and in this way essential orientation information may be lost.

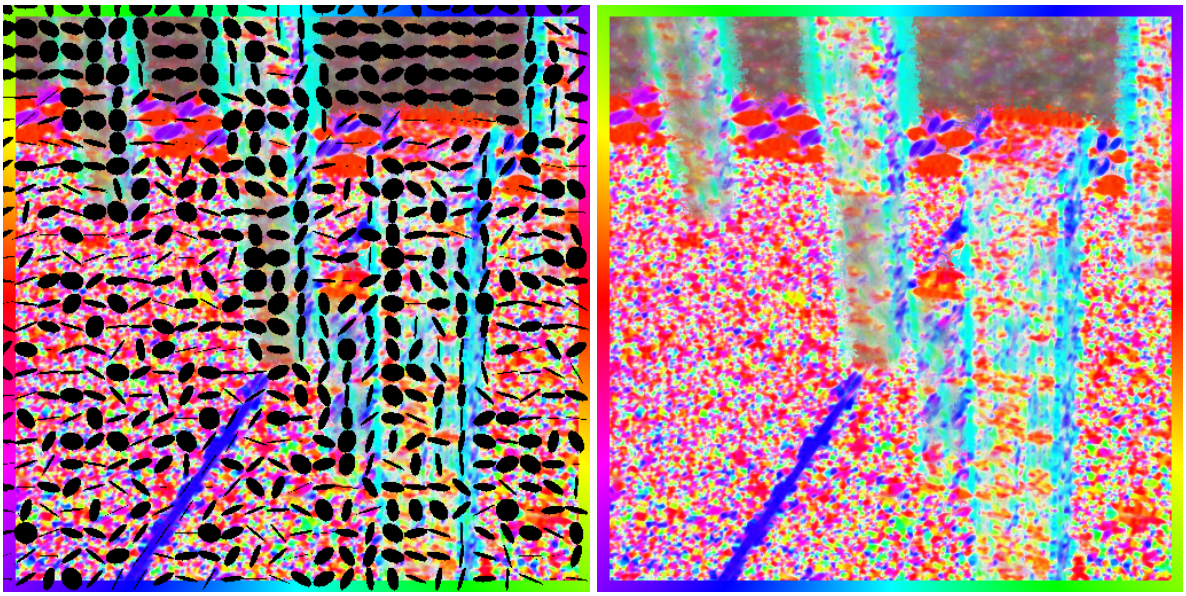


Figure 3.1: Positive definite flow in two dimensions: Orientation map, hue:=angle, saturation:=anisotropy, luminance:=diffusivity.

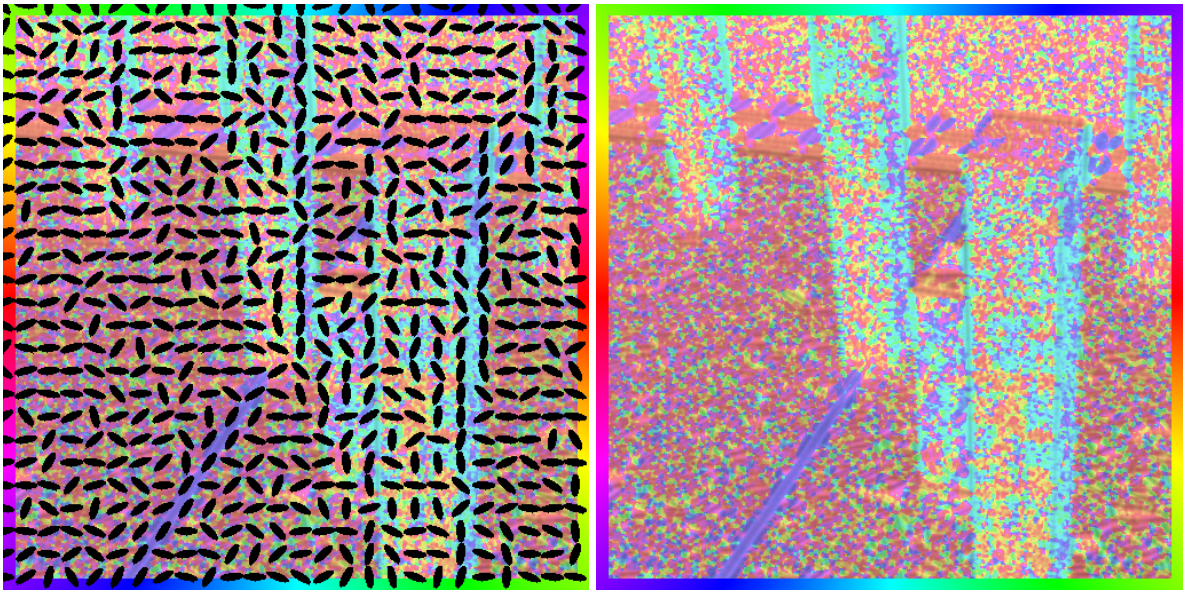


Figure 3.2: Positive definite flow in 2 dimensions: Diffusivity map with an anisotropy=0.8.

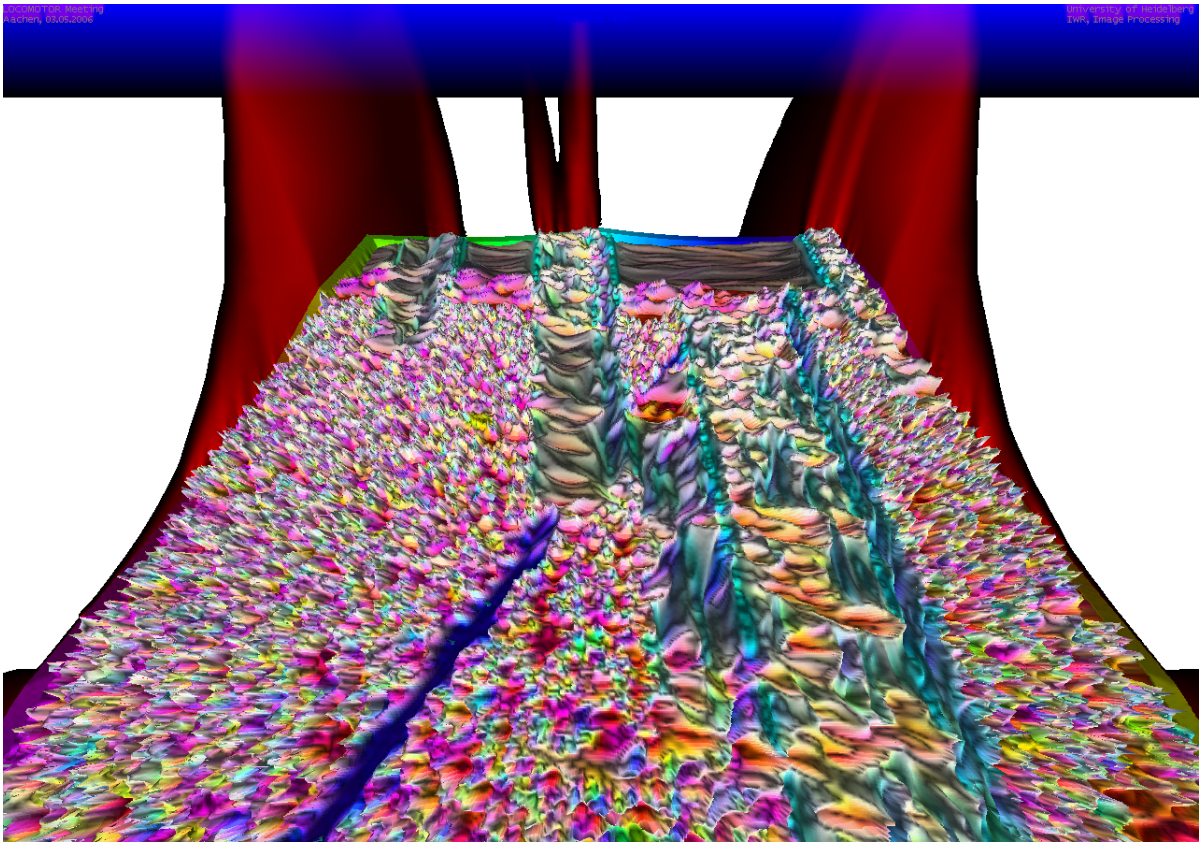


Figure 3.3: Two dimensional orientation representation as a height map. The same height represents the same orientation.

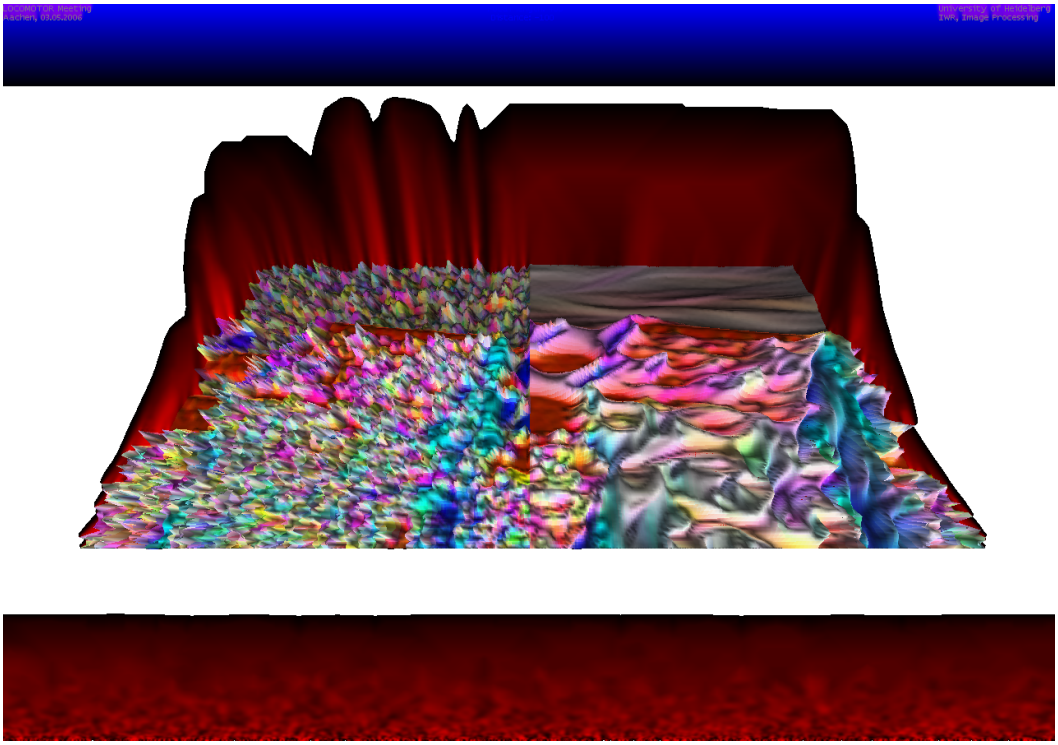


Figure 3.4: Two dimensional orientation representation after an anisotropic integration of the structure tensor for the column top of the Marble sequence. Left half: initial structure tensor orientation, right half: after an anisotropic integration. Discontinuities are well preserved, while the background noise is suppressed.

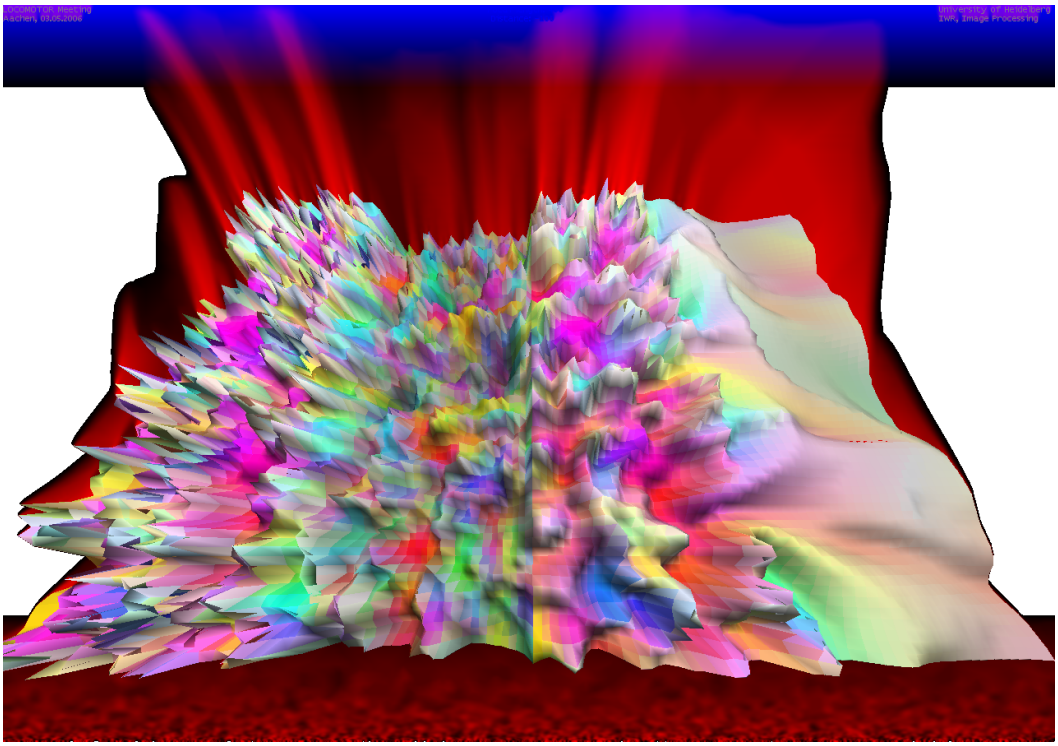


Figure 3.5: Two dimensional orientation representation after an anisotropic integration of the structure tensor for the column border of the Marble sequence. Left: initial value for the structure tensor orientation, right: after anisotropic integration.

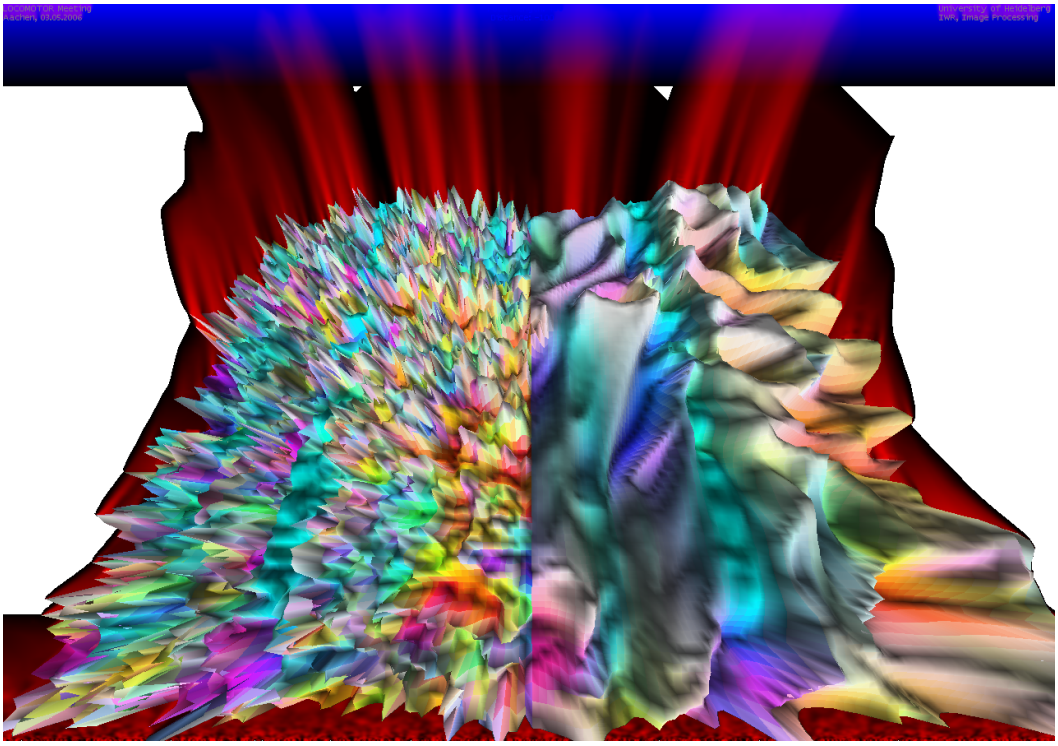


Figure 3.6: Two dimensional orientation representation after an anisotropic integration of the structure tensor for the column border of the Marble sequence. Left: initial value for the structure tensor orientation, right: after anisotropic integration.

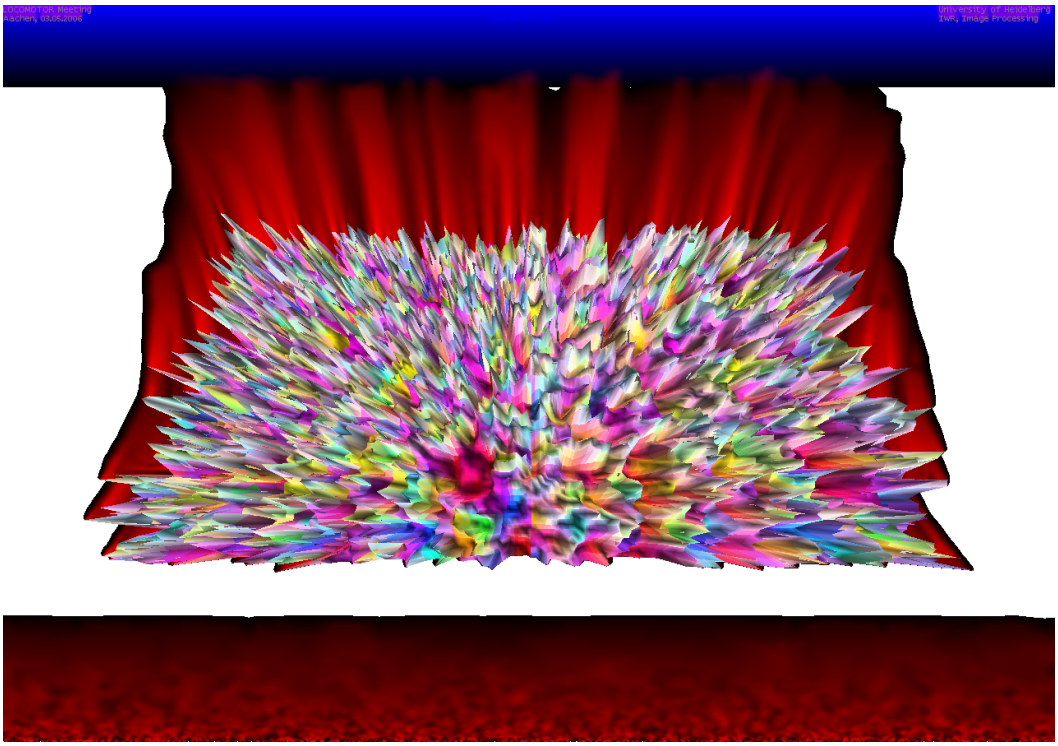


Figure 3.7: Two dimensional orientation representation after an anisotropic integration of the structure tensor for the bottom of the Marble sequence. Left: initial value for the structure tensor orientation, right: after anisotropic integration. The structure of the bottom is well preserved.

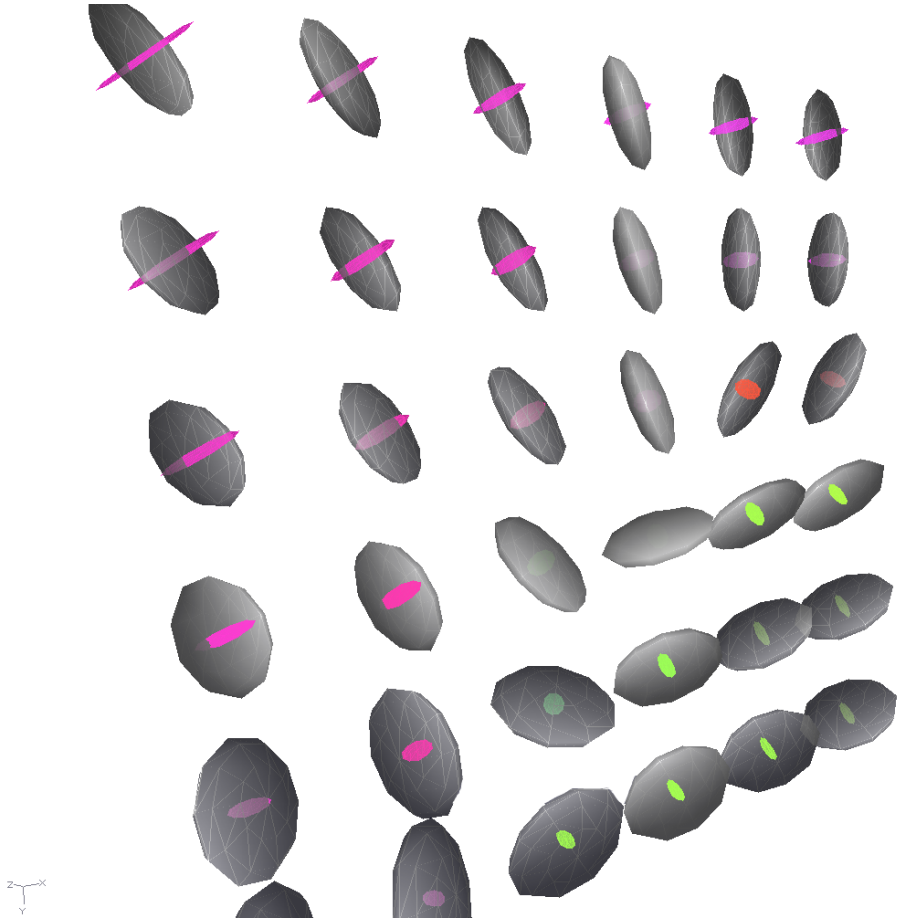


Figure 3.8: Three dimensional orientation representation and the resulting diffusion tensors, computed from the original tensor field. The duffusion tensors are represented in grey. The orientation is color coded. Red: x-axis, green: y-axis, blue: z-axis.

Our application of interest is motion estimation by adaptive smoothing of the underlying tensor field. Retrieving the trajectories of the objects by following the main directions of a tensor field defined on $\Omega \subset \mathbb{R}^3$ is important. Using direct regularization would loose some trajectory directions, which is an undesired property. There are two other approaches to adaptive integration of tensor fields. These are the spectral approach and the orthonormal preserving flow.

Spectral approach

The spectral approach relies on direct processing on the spectral decomposition.

$$T = U\,\Gamma\,U^T, \quad U \in \mathrm{SO(n)}, \quad \Gamma := \mathrm{diag}(\lambda_1, ..., \lambda_n)$$

For structure tensors of multichannel images, Γ and U measure the channel variations and their corresponding directions.

The method consists of two constrained and coupled regularizations acting on Γ and U, [Tschumperle, 2002].

For the smoothing of the tensor diffusivities there are several strategies. Tensor diffusivity integration schemes can be described as follows.

- Don't change the eigenvalues! Leave them untouched!
- Each eigenvalue is processed separately.
- Vector-valued diffusion PDE's for the vector $(\lambda_i)_i$.
- A-priori spectral information inside the diffusion equation for specific applications.

Orthonormal preserving flow

The regularization of orthonormal vector sets can be derived from the unconstrained regularization scheme by adding constraints for the orthogonal group SO(n). Minimization of

$$E(T^k) := \int_\Omega \alpha \|T^k - T_0^k\| + \psi(\lambda_i^k)\, dx, \qquad \lambda_i^k \in \sigma(G^k)$$

leads by computaion of the Euler-Lagrange equations to the system of PDE's

$$T^k = T_0^k\,|_{t=0} \qquad (k = 1...n)$$
$$\frac{\partial T^k}{\partial t} = \alpha(T_i^k - T_0^k) + div\left(\sum \frac{\partial \psi}{\partial \lambda_i^k} \theta_i^k \otimes \theta_i^k \nabla T_i^k\right) \qquad (k = 1...n).$$

This PDE system is decoupled and doesn't preserve the orthonormal properties.

The regularization of orthonormal vector sets can be fulfilled by constrained minimization:

$$\sum E(T^k) \to \min!$$

subject to the orthonormal constraint

$$T^p \cdot T^q = \delta_{pq} \qquad \text{(Kronecker delta)}$$

We introduce Lagrange multipliers λ_{pq} and perform unconstrained minimization of

$$E^*(\mathcal{T}_0, \lambda) := E(\mathcal{T},) + \int_\Omega \sum \lambda_{pq}(I_p \cdot I_q - \delta_{pq})\, dx.$$

This leads to the system of PDE's for the tensor channels T^k, cf [Tschumperle, 2002]

$$\frac{\partial T^k}{\partial t} = \mathcal{L}^k + \sum_l (\mathcal{L}^l \cdot T^k) T^l, \qquad (k = 1, ..., n)$$

where $\mathcal{L}^k$ is the unconstrained diffusion force, acting on T^k.

Remarks:

1. The last PDE system is a set of n coupled vector PDE's with the coupling term
$$\sum_l \mathcal{L}^l \cdot T^k T^l.$$
2. For unit norm vector sets, the PDE velocity is orthogonal to the vector and tangent to the unit sphere.
3. A possible application is the regularization of structure tensors or covariance matrices.

4. Equivalent matrix PDE notation for 3x3 matrices with bases $\mathcal{B} := I, J, K$ could be

$$\frac{\partial R}{\partial t} = -L \; + \; R\, L^T\, R \quad \text{with } R := (I \mid J \mid K) L := (\mathcal{L}^I \mid \mathcal{L}^J \mid \mathcal{L}^K).$$

When perfoming the smoothing of the tensor orientations

$$T = U \; \mathrm{diag}(\lambda_1, ..., \lambda_n) \; U^T$$

the important part is the preservation of the orthogonality of U during the diffusion flow. This yields in matrix notation the PDE

$$\frac{\partial U}{\partial t} = -L + U L^T U \tag{3.1}$$

L is the matrix, corresponding to an unconstrained Lagrangian, describing the regularization process.

Some suitable examples for L, which represent different diffusion straegies are, [Tschumperle, 2005]

- $$L_{ij} := \mathrm{trace}(D H_{ij})$$
- $$L_{ij} := \mathrm{trace}(D H_{ij}) + \frac{2}{\pi} \nabla T^T_{ij} \int_0^\pi J^{ij}_{\sqrt{D_\alpha}} \sqrt{D_\alpha}\, d\alpha,$$

where H denotes the Hessian and J denotes the Jacobian of the tensor T.

The local alignment method for the orthonormal preserving flow

When implementing a straightforward iterative solution for (3.1), one have to consider, that the PDE velocity is orthogonal to the current vector in each iteration step. The iteration step is a rotation of the vector, cf [Tschumperle, 2002]. Some of the eigenvectors of the tensor can change sign due to numerical computations. This may lead to artifitial discontinuities, althogh neigbouring vectors are well aligned. This fact requires an additional local alignment step in a vicinity of the processed tensor.

An alternative solution is to represent the rotation corresponding to the evolution equation by the rotation vector

$$\omega := (I \times \mathcal{L}^I) + (J \times \mathcal{L}^J) + (K \times \mathcal{L}^K).$$

For given ω the rotation matrix Γ is computed by the Rodriguez' formula. This ensures

$$R_t \in SO(3) \Rightarrow R_{t+dt} \in SO(3)$$

.

A suitable numerical scheme for the isospectral flow

The isospectral flow is an evolution flow that preserves the tensor diffusivities, while regularizing the orientations. Let there is a given manifold $\mathcal{N}$ of the matrices with a set of eigenvalues. Then the isospectral flow PDE on $\mathcal{N}$ is given by

$$\frac{\partial T}{\partial t} = [T, [T, -\mathcal{L} + \mathcal{L}^T]] \, .$$

A suitable integration scheme is

$$T_{t+dt} := A_t(x)^T \; T_t(x) \; A_t(x) \tag{3.2}$$

$$A_t(x) := e^{-dt[\mathcal{L}(x)\mathcal{L},T]} \, . \tag{3.3}$$

The advantage is the direct application of the iterative step on the matrix coefficients. There is no need of the eigenvalue and eigenvector computation of the tensor field and no local alignment step is required. In the next section we present results of motion estimation by adaptive integration af the tensor field by the isospectral flow.

3.3 Numerical Experiments

The motion estimation was fulfilled by total least squares estimation of the velocity from the tensor filed. We compare the results of an isotropically smoothed against adaptive integrated tensor field. The nonlinear diffusion is constrained on the manifold of matrices with given set of eigenvalues by using the isospectral flow. The integration was computed by matrix exponentials according to the scheme 3.2.

Intermediate visualization of the tensor fields is used to study the effect of the nonlinear diffusion flow on the tensor fields. The tensor fields can be visualized by ellipsoids and by color encoding, where the colors may represent either orientation or shape or both. For tensor field visualization, the reader is referred to [Christopher R. Johnson and Charles D. Hansen, 2004].

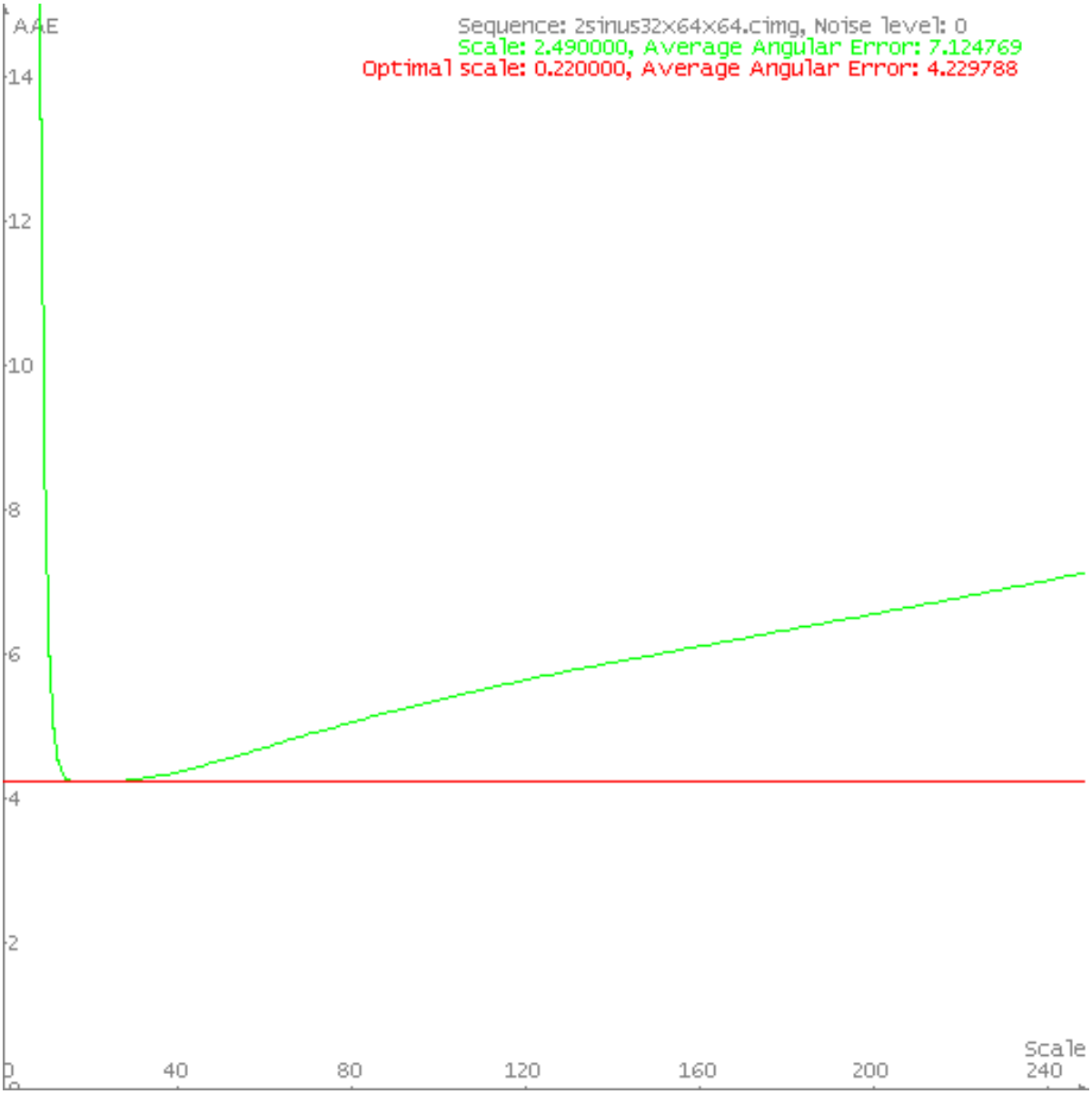


Figure 3.9: Optimal average angular error measurement. Dependence of the error on the smoothing amount of the structure tensor. Red: best result for the isotropic smoothing, green: Average angular error against smoothing scale.

First we measured the performance of both techniques on a synthetic sequence without noise, containing a moving sinus pattern with discontinuity. The diffusion time or the scale parameter is represented on the x-axis, the three dimensional average angular error on the y-axis. Both methods reach the same accuracy. The dependence of the error on the scale parameter for the nonlinear diffusion is colored in green.

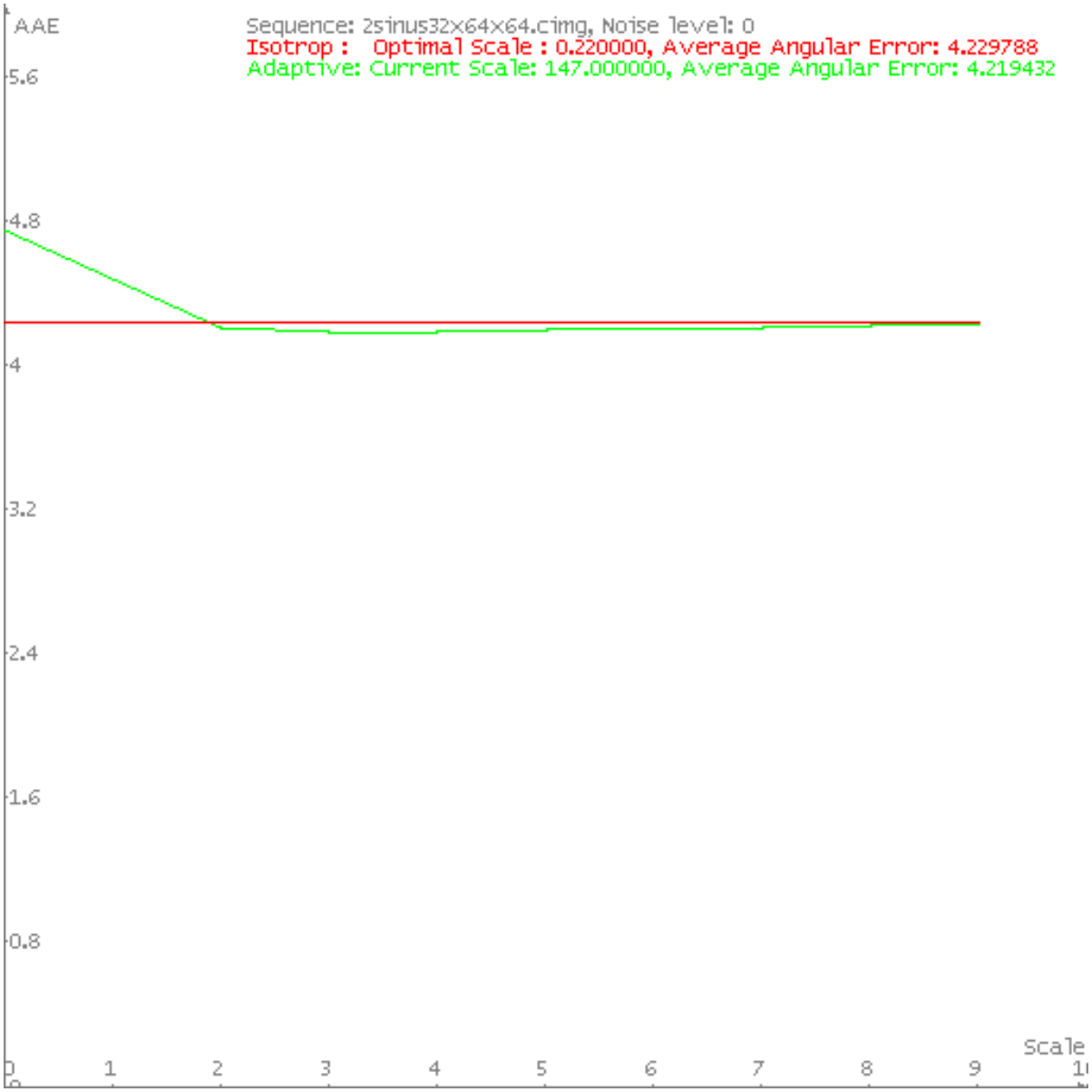


Figure 3.10: Average angular error against smoothing amount of the structure tensor. Red: best result for the isotropic smoothing, green: nonlinear regularization of the structure tensor

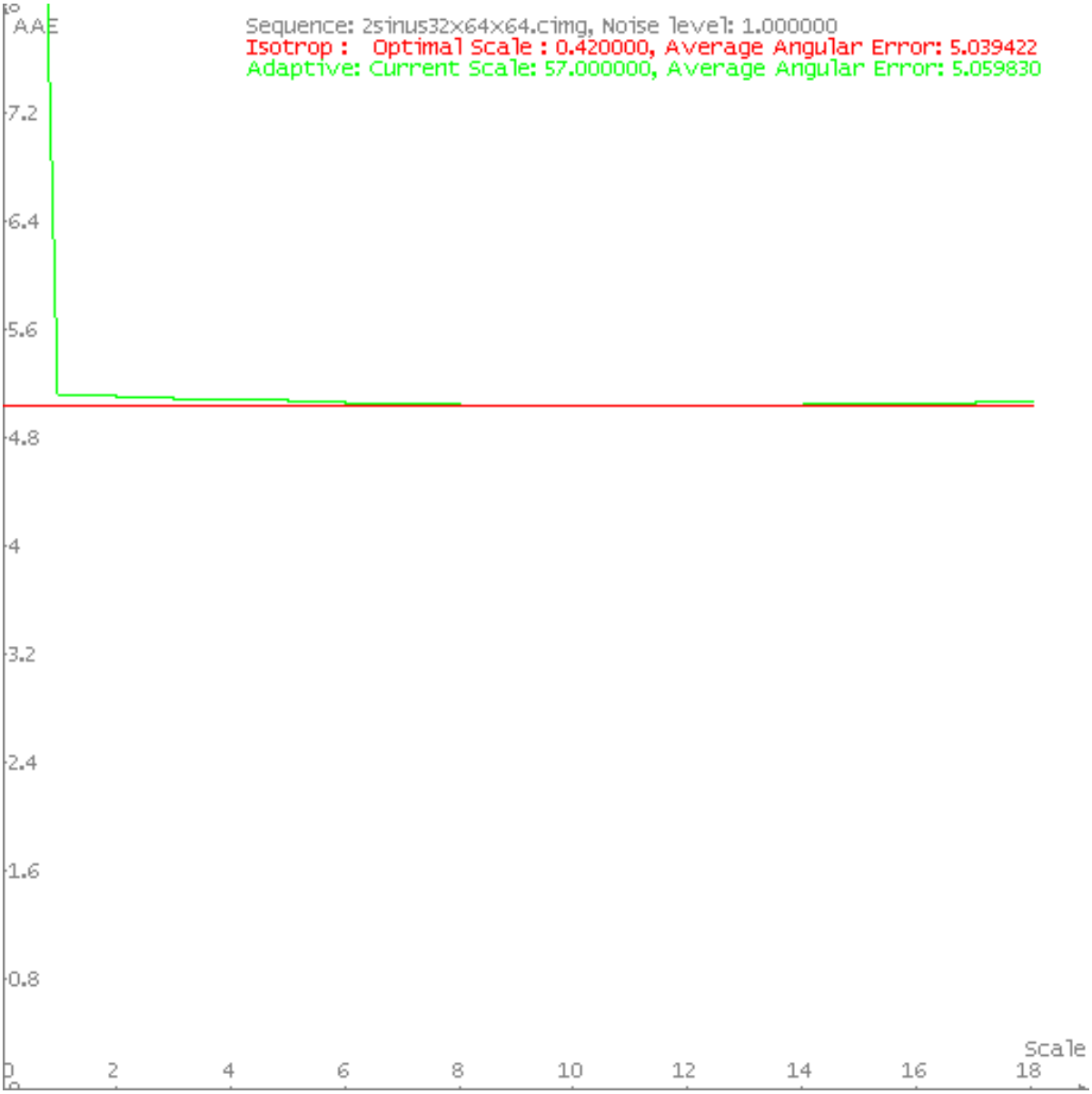


Figure 3.11: Average angular error against smoothing amount of the structure tensor after extensive search and tuning for the diffusion parameters. The accuracy improvement is marginal. Red: best result for the isotropic smoothing, green: nonlinear regularization of the structure tensor.

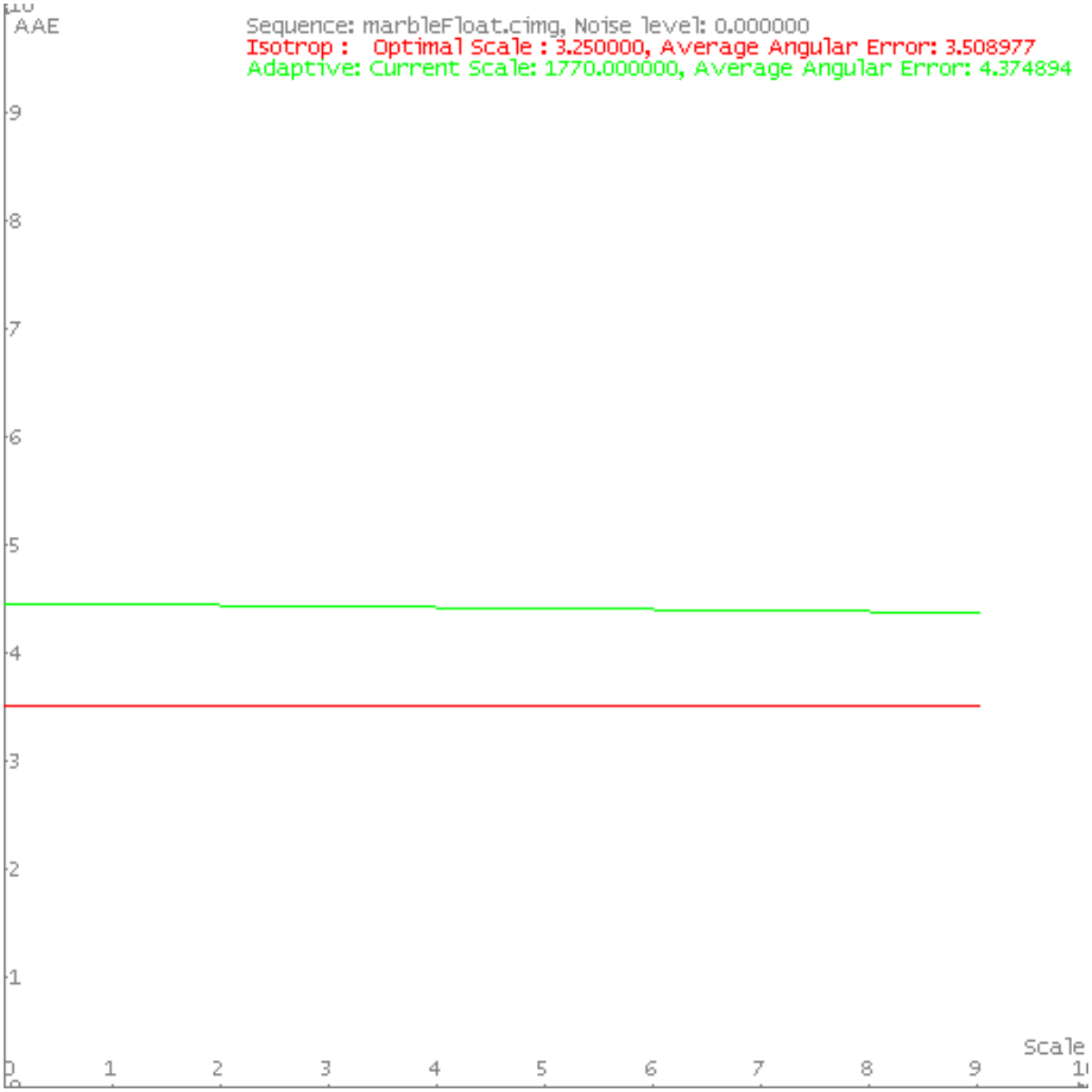


Figure 3.12: Average angular error against smoothing amount of the structure tensor for the Marbled Block sequence. Red: best result for the isotropic smoothing, green: nonlinear regularization of the structure tensor

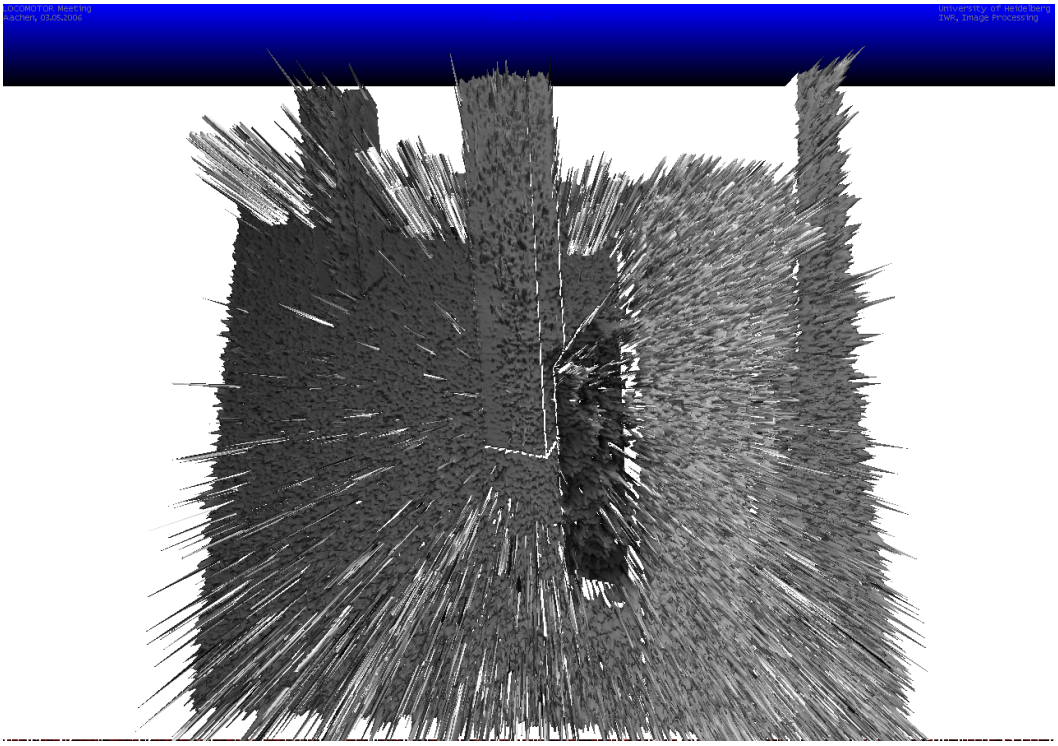


Figure 3.13: x-component of the estimated velocity from the Marbled Block sequence after nonlinear tensor regularization as height map. The spikes show, that the noise is enhanced.

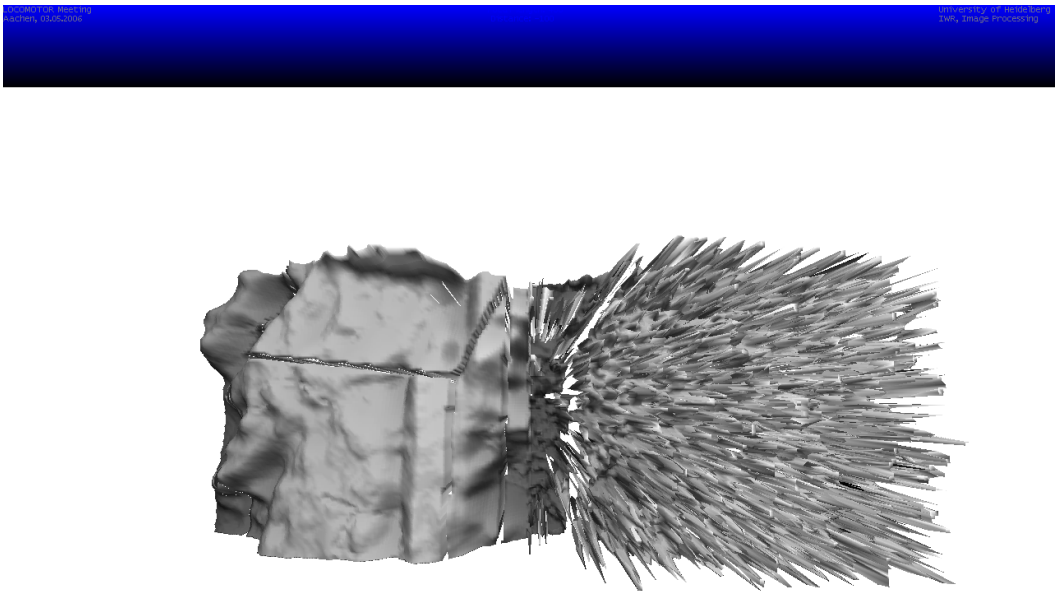


Figure 3.14: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

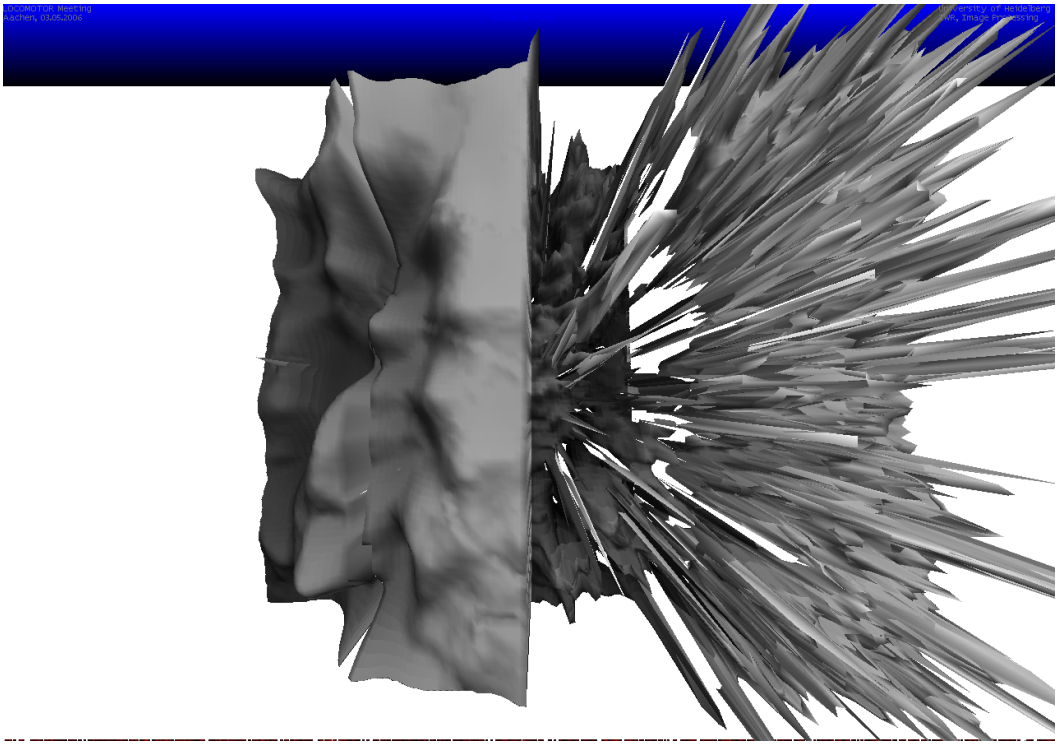


Figure 3.15: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

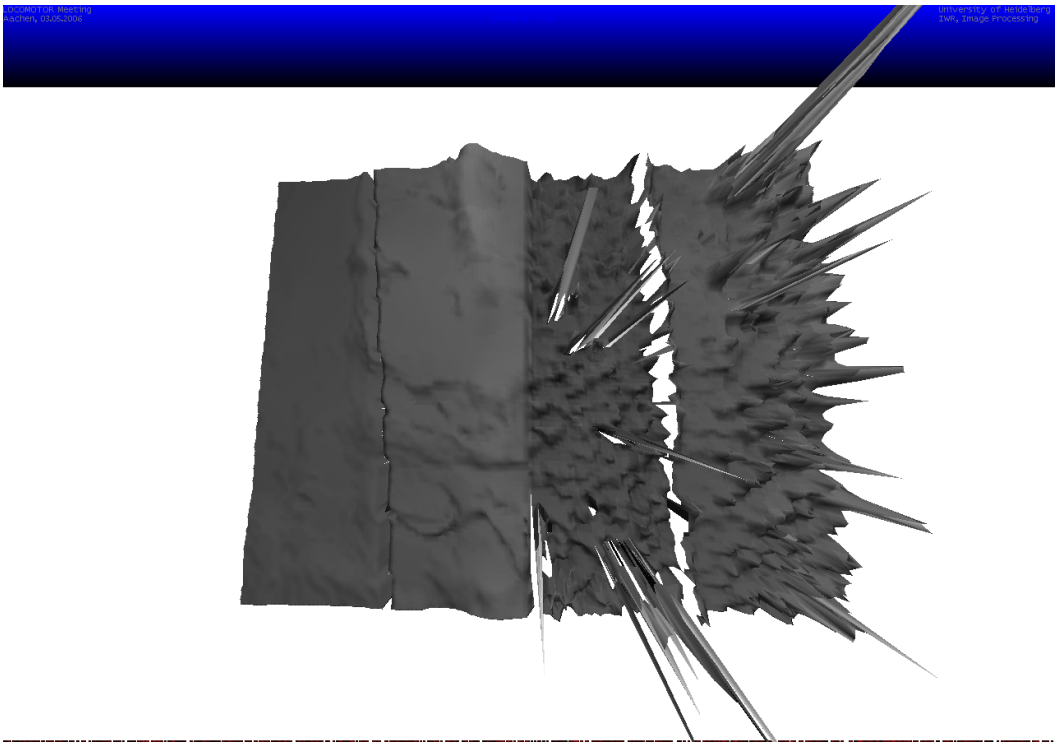


Figure 3.16: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

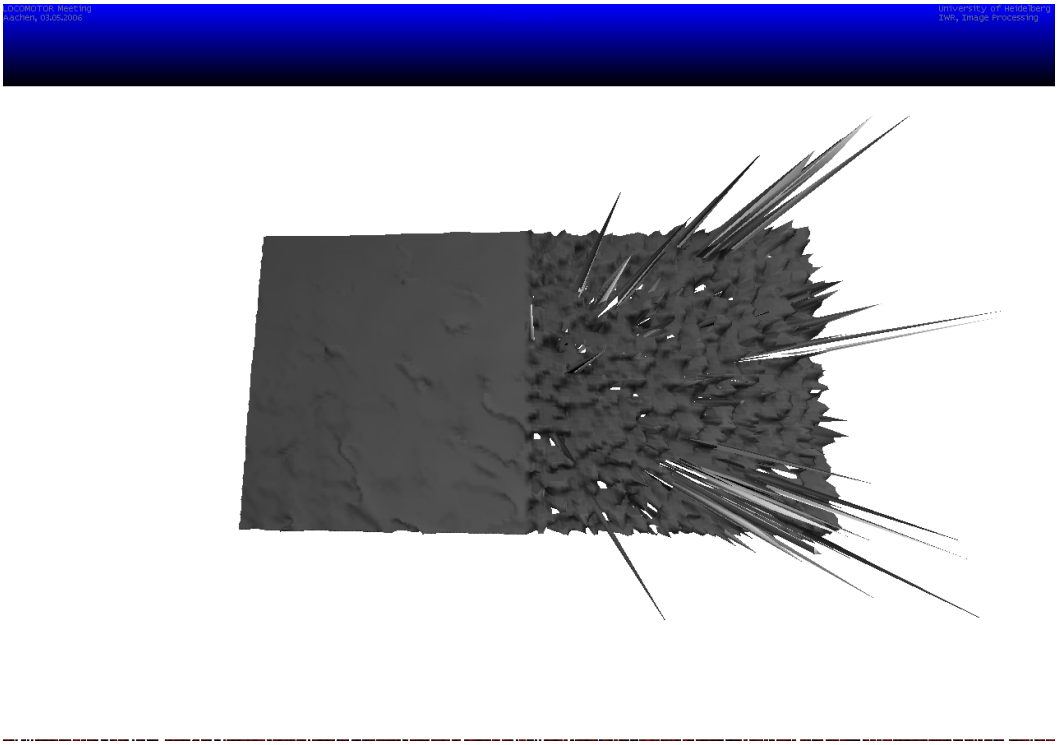


Figure 3.17: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

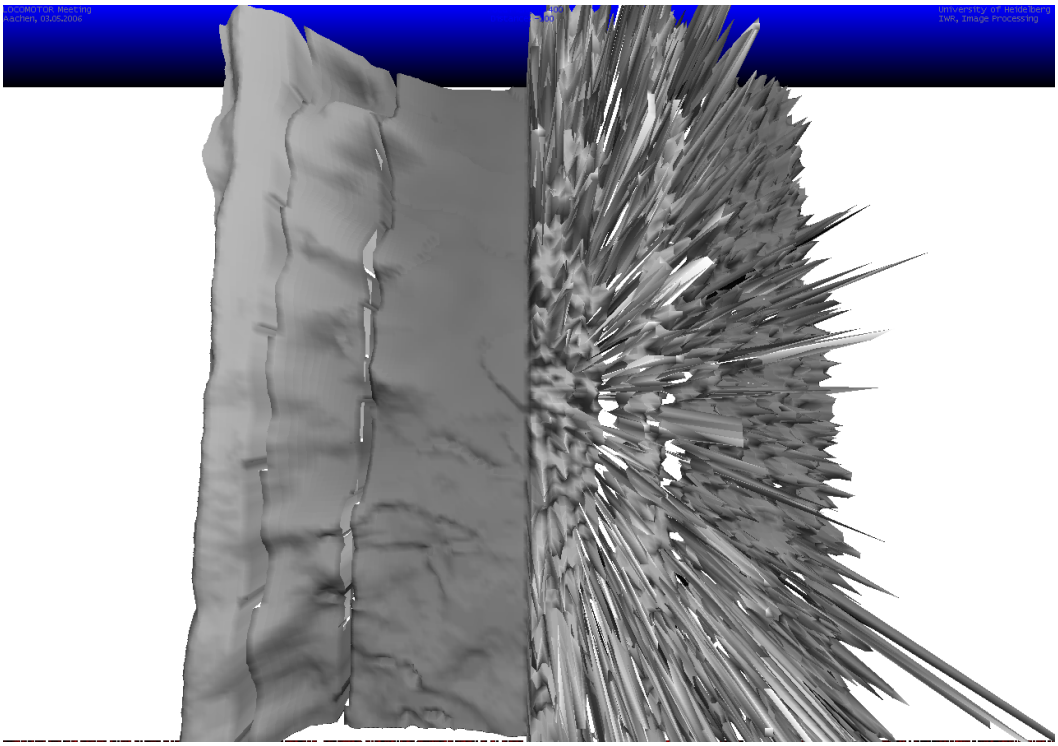


Figure 3.18: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

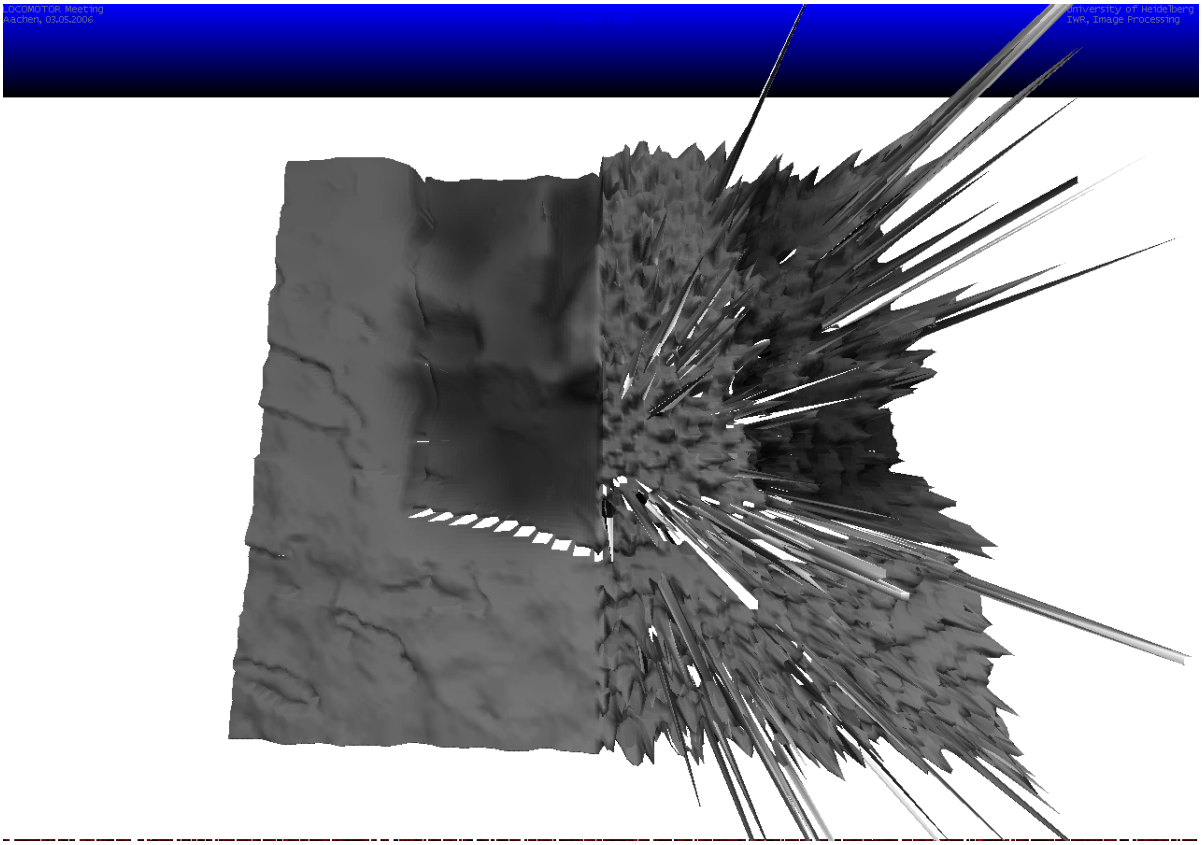


Figure 3.19: x-component of the estimated velocity from a segment of the the Marbled Block sequence after nonlinear tensor regularization as height map. Left half: isotropic smoothing, right half nonlinear smoothing. The spikes show, that the noise is enhanced.

3.4 Conclusion

The goal of this chapter was to study methods for tensor field integration by parabolic PDE's, which respect not only the discontinuity, but also the curvature in the image data. Currently known techniques for motion estimation stop the over-smoothing on corners by a coherence dependent function, which inhibits the diffusion on positions with high curvature. We discussed the issue of tensor field regularization for optic flow computation and motion estimation.

A practical contribution is the accuracy measurement for motion estimation by adaptive smoothed tensor fields. The adaptive smoothing relies on nonlinear anisotropic diffusion with discontinuity and curvature preservation. We reached an accuracy gain under properly chosen parameters for the diffusion filter.

Here are some open questions concerning the regularization of tensor valued images.

1. Fair comparison between different diffusion flows?

   Maybe a fair comparison of different smoothing methods will be possible in an invariant metrics, cf [Xavier Pennec and Nicholas Ayache, 1998; Xavier Pennec et al., 2006].

2. Spatial gradient of tensor fields?

   There could be an improvement of the method, if the derivatives of the tensor field are computed on the manifold and not in the flat Euclidean metrics as done in our computations.

3. ϕ-function formalism in Riemannian metrics? Discontinuity preserving regularization was introduced in section 1.3.2. The idea was to replace the regularization term

$$\int_{\Omega} \|\nabla u\|^2 \, dx$$

by a discontinuity preserving term

$$\int_{\Omega} {}'Psi(\|\nabla u\|) \, dx$$

with suitable Ψ. Under regularity assumptions for Ψ we end up with

$$\operatorname{div}\left(\frac{\Psi'(\|\nabla u\|)}{\|\nabla u\|}\nabla u\right)$$

for the regularization term. An interesting open question is, how can we adapt this formalism to a Riemannian or an invariant metrics.

Chapter 4

Optic flow with the energy operator

The energy operator was developed in 1983 by Herbert M. Teager and Shushan M. Teager and later by James F. Kaiser to express the energy of the *source*, which *generates* a simple sinusoidal signal or signal $f = f(t)$ of the form

$$f(t) = a(t)\, cos\, \varphi(t)$$

with $a(t)$ being the amplitude or envelope and $\phi(t) := w_0(t) + \Phi(t)$ being the phase . The frequency w_0 can be viewed as the carrier frequency and $w := \frac{d\varphi}{dt}$ is the instantaneous frequency.

4.1 Related work in the literature

The whole story of the energy operator in signal processing is described in the wonderful thesis [Kvedalen, 2003, Pages 3–7]. We give a brief overview of the historical development of the energy operator in signal and image processing below.

4.1.1 Signal processing literature

The energy operator was first applied for speech modeling in 1983 by Herbert M. Teager and Shushan M. Teager. That's why it is also known as the Teager energy operator. James F. Kaiser presented in [Kaiser, 1990a] an algorithm to compute the energy of a discrete signal. James F. Kaiser called this algorithm "Teager's energy algorithm" although he derived the algorithm alone. Because of this fact the energy operator is also called the Teager-Kaiser energy operator. The extension to continuous signals was published in [Kaiser, 1990b].

Further applications of the energy operator to speech and signal processing and important theoretical results were presented at the ICASSP 1993. Some of them are [Maragos, P. et al., 1993], [Alan C. Bovik et al., 1993] and [Kaiser, 1993]. [Maragos, P. et al., 1993] was the first paper where the energy operator was deeply investigated with respect to its AM-FM demodulation properties. [Alan C. Bovik et al., 1993] stated theorems that give approximations on AM-FM demodulation. Some researchers like [Larkin, 2005, Page 7] call the energy operator the Teager-Kaiser-Maragos-Quatieri-Bovik operator, in order to honer all of the inventors of this computationally efficient tool for signal analysis. From now on we will stick to the name *energy operator* when we mean an operator

$$E(f) := D(f)^2 - f \cdot D^2(f) \tag{4.1}$$

of Teager-Kaiser-Maragos-Quatieri-Bovik type in multidimensional signal or image processing for discrete or continuous signals. Here D denotes an abstract derivative or pseudo-derivative operator.

[Larkin, 1996] applied the energy operator to parameter estimation of bandpass filtered AM signals.

Another theoretical work is [A. Bovik et al., 1997], where error bounds for approximations of system responses to AM-FM signals are derived and particularly for the energy operator. One application is the selection of optimal bandpass filters for preprocessing of discrete AM-FM signals. The optimality is in the sense of the uncertainty principle.

[Michael Moore et al., 1997] adjusted the sensitivity of the energy operator in dependence of the local mean of the signal. In this manner one can develop an adaptive energy operator in dependence of the local average intensity of the signal.

[Santhanam, 2004] discusses demodulation of wide-band signals.

The interested reader is encouraged to take a look in [Kvedalen, 2003, Pages 3–7] for more detailed list of papers about the development of the energy operator in the early eighties and nineties of the past century.

4.1.2 Applications in image processing

In 1991 the Teager-Kaiser energy operator was generalized to two dimensions as an image processing operator in [Mitra, S.K. et al., 1991]. Thus the energy operator is known in the image processing community at the latest since 1991. A very good review of the 2-D extension considerations is given in [Larkin, 2005, Section 3.2, page 8].

[Maragos, P. et al., 1992] extended the energy operator to arbitrary dimensions with applications to texture analysis and multi-spectral images.

[Maragos and Bovik, 1995] estimated the parameters of texture images with the 2D energy operator. The texture was modeled as two dimensional AM-FM image with slowly varying amplitude and frequency.

[Larkin, 2001, Pages 121–131] discusses the relationship between the energy operator, the Riesz transform, Wigner-Ville transform and the spiral phase transform as tools for estimation of instantaneous frequency . For orientation estimation [Larkin, 2001, Page 131] recommends the energy operator and tensor methods.

[Felsberg and Granlund, 2004], [Felsberg and Jonsson, 2005] have proposed a 2-D energy tensor , where the differential operator D in (4.1) is replaced by the structure tensor, [Jähne, Bernd, 2005, Page 364] and the differential operator of second order D^2 is replaced by the Hessian. Therein the dual frames for a discrete implementation of the energy operator have been computed too and some basic performance comparison with respect to orientation estimation have been presented.

Ullrich Köthe developed in his thesis [Köthe, 2000] the idea of reusable software in image and signal processing. [Felsberg and Köthe, 2005] proposed the gradient energy tensor based on Gaussian derivatives. So one can reuse derivative filters based on Gaussians to compute the energy operator.

The relationship and differences between the gradient energy tensor and the boundary tensor , a well known corner and edge operator, have been investigated in [Köthe and Felsberg, 2005]. It have been shown, that both operators yield similar results for properly chosen scales, however, the gradient energy operator is computationally more efficient. The gradient energy operator is theoretically phase invariant , see page 61 or [Larkin, 2005, page 7], while the uniformity of the boundary tensor relies on the phase congruence over many frequencies, [Kovesi, 1999] on edges and corners of natural images.

[Larkin, 2005] used the demodulation properties of the energy operator to estimate orientation and phase in fringe patterns. A new phase invariant non-local energy operator for real and complex signals is proposed in the paper. Besides the very good introduction to the

generalization of the energy operator to two dimensions there is also a discussion about the extension of the analytic signal to images and the different approaches to the extension.

4.2 Definition and properties of the energy operator

A good systematic and self contained introduction to the energy operator for one dimensional signals is given in [Kvedalen, 2003, pages 15–30].

4.2.1 Definition

For continuous signals $f = f(x)$ the energy operator E is defined as

$$E(f) := \left(\frac{df}{dx}\right)^2 - f \cdot \frac{d^2}{dx^2} f \; .$$

For a simple sinusoidal signal

$$f(x) := a \cdot \cos(\omega x)$$

the output of the energy operator is

$$E(f) = (-a\omega \sin(\omega x))^2 - a\cos(\omega x)(-\omega^2 a \cos(\omega x)) = a^2\omega^2(\sin^2(\omega x) + \cos^2(\omega x)) = a^2\omega^2 \; .$$

The main property of the energy operator is, that the output is proportional to the square of the amplitude and the square of the frequency of an oscillating signal. As stated in [Larkin, 2005, page 7], the energy operator generates two oscillatory terms in phase quadrature. The oscillation exactly cancels out, resulting in a constant energy $\omega^2 a^2$. This is a phase invariant estimation or phase invariant demodulation.

Energies of well known signals with plots for comparison between the input signal and the response are presented in [Kvedalen, 2003, pages 18-22].

Signal type	Signal $f(x)$	Response $E[f](x)$
sine	$a \cdot \cos(\omega x)$	$a^2\omega^2$
exp	$e^{-\alpha x}$	0
AM	$a(x) \cdot \cos(\omega_c x)$	$a^2(x)\omega_c^2 + \cos^2(\omega_c x)E(a(x))$
exp sine	$e^{-\alpha x}a \cdot \cos(\omega x)$	$e^{-2\alpha x}\omega^2 a^2$
FM	$a \cdot \cos\varphi(x)$	$a^2(\varphi'(x) + \frac{1}{2}\varphi''(x)\sin 2\varphi(x))$
AM-FM	$a(x) \cdot \cos\varphi(x)$	$(a(x)\varphi(x))^2 + \frac{1}{2}a^x\varphi''(x)\sin 2\varphi(x) + \cos^2\varphi(x) \cdot E(a(x))$

4.2.2 Some analytic properties

The energy operator was designed to track the energy of a source generating given signal. As the energy is always non negative, the question arises weather the output of the energy operator is non negative and if not, for which signal this happens.

A counterexample by [Kvedalen, 2003, page 28 and figure 3.8] shows, that the output of the energy operator can become negative for certain signals. Hence the conditions for non-negativity should be investigated more carefully. This have been already done in [Maragos, P. et al., 1993, pages 1546–1547], [Bovik, Alan C. and Maragos, Petros, 1994] and [R. Hamila et al., 1999, page 261]. As a main result of simple analysis it comes out, that for real signals, the output of the energy operator is non negative, if any of the following conditions hold

- $f(x) = 0, \qquad \forall x \in \mathbb{R}$
- $f''(x) = 0, \qquad \forall x \in \mathbb{R}$
- $f(x) < 0$ and $f''(x) > 0$
- $f(x) > 0$ and $f''(x) < 0$.

Another statement for the no-negativity of the energy operator follows from the main result of [R. Hamila et al., 1999]. The output of the energy operator of a signal is proportional to the second order conditional moment in frequency of the Wigner-Ville transform, see also section 1.4 *Time-frequency signal analysis*. Sometimes it happens, that the second conditional moment of the Wigner-Ville transform is negative. As the proportionality factor is π it follows, that the output of the energy operator for such signal is negative too.

Some interesting remarks and open questions on the relation between the positivity of the energy operator and the negative convexity of the logarithm of the estimated signal are published in [Larkin, 2005, page 10].

4.3 Implementation for optic flow estimation

Historically, the discrete definition of the energy operator in signal and image processing was published before its continuous counterpart. Thus the well known Teager-Kaiser algorithm, [Kaiser, 1990a], computes the output from three adjacent samples.

In one dimension there are experiments reported in [Larkin, 1996, 2001], where the energy operator performed well with a huge rate of undersampling up to a factor of three. This is because of the compensation of aliasing in the *gradient squared* term alone and the *signal times the Hessian* term alone. Both terms alone are band enlarging operators, see also 5.1, [Bigun, 2006, pages 114-117] or [Köthe, 2003, pages 26–27].

According to [Larkin, 2005, page 8109] the discretization of the gradient form of the 2D energy operator with small kernels and autoconvolution for the Hessian leads to an anisotropic response. Therein larger kernels are recommended. We choose the 5x5x5 optimized Sobel filter by [Scharr, H. et al., 1997; Scharr, 2000; Scharr, H. and Weickert, J., 2000]. It is also possible to implement the Hessian by a filter mask of the same size as the gradient mask, anyway the both filters should be adjusted to each other, otherwise the phase invariance will no longer apply [Larkin, 2005, page 8110].

4.4 Numerical experiments

4.4.1 Error measurements

We implemented the energy operator in three dimensions for optic flow estimation based on three dimensional derivative filters in the spatio-temporal domain in a straightforward manner

using (4.1) by replacing D^2 with the Hessian of the sequence. We used the modified Sobel filters as described in [Scharr, H. et al., 1997] and [Scharr, 2000, page 155]. For comparison between different implementations, the Hessian of the spatio-temporal image data was computed either by twice applying first order derivative filters, [Jähne, B. et al., 2007, page 100] or second order derivative masks.

In our experiments we computed the energy operator on original resolution for every wavelet scale. We conducted measurements of the average angular error for a synthetic sequence without noise, for synthetic sequences with noise and a real world test sequence acquired by a camera. The local total least squares approach was used for the 3D structure tensor and for the 3D energy operator.

Since we estimated image sequences with ground truth, we compared the best results of a total least squares local approach for the energy operator and the structure tensor. First the optic flow estimation was fulfilled with the energy operator in the scale-space and we determined by brute force the optimal integration size and wavelet scale for the best average angular error. Then we performed motion estimation at the optimal wavelet scale by the well known structure tensor and measured the best average angular error in dependence on the integration scale for the tensor field. In this way we compared the best possible results for the both operators at given wavelet scale, because our task was to investigate, whether the energy operator is better or not. The results are listed in the table below.

Synthetic sequence: sinus pattern with discontinuity

Type of the derivative filter	Optimal integration scale for the tensor	Average angular error
structure tensor	0.24	4.589143
energy operator by ... first order derivative	3.54	10.2803

Marbled Block sequence

Type of the derivative filter	Optimal integration scale for the tensor	Average angular error
structure tensor	3.2	3.509298
energy operator by ... first order derivative	2.59	3.215193

Street sequence

Type of the derivative filter	Optimal integration scale for the tensor	Average angular error
structure tensor	1.57	4.589143
energy operator by ...		
first order derivative	6.1	10.251023

4.4.2 On the effect of the bandpass filtering: filter bandwidth versus wavelet scale

We investigated the dependency of the average angular error as a function of the bandwidth and wavelet scale simultaneously. Here we mean the bandwidth of the bandpass filter or the spread of the wavelet, used to filter the input image sequence. As a result it comes out, that there is an optimal point in the bandwidth-scale plane which minimizes the error measure, see also figure 4.4.2.

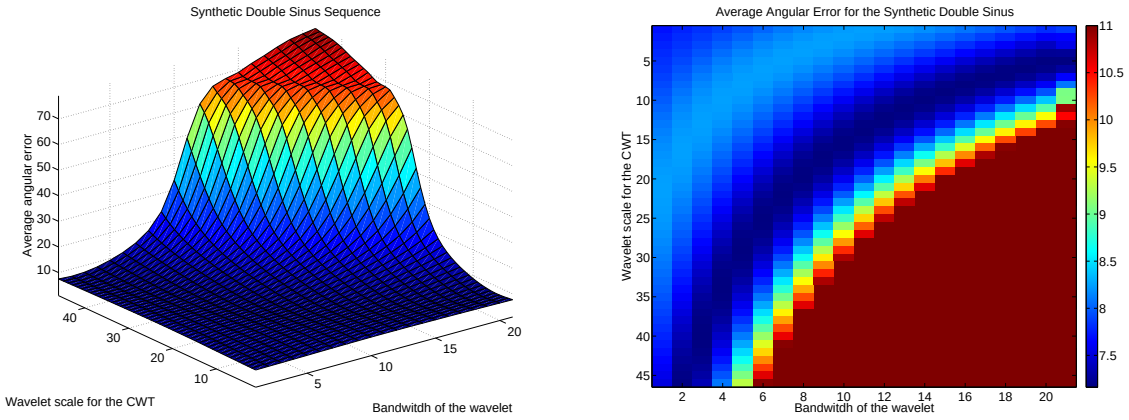


Figure 4.1: The optimal average angular error as a function of the bandwidth and scale of the Mexican hat wavelet for a synthetic sinus pattern sequence with discontinuity.

4.5 Conclusion

Despite the preprocessing of the image sequence, our experiments showed, that there is a need of post-integration for the energy operator to achieve optimal average angular error of the estimated flow fields. For orientation estimation with the energy operator [Felsberg and Jonsson, 2005, page 498] reported similar results and applied a Gaussian post-filtering with $\sigma = 1$ and a smoothing window size of 7x7. The accuracy gain for the real world Marbled Block sequence is in the first digit behind the decimal point or approximately 7.7 %. This improvement is achieved by optimal parameter setting for both operators, the structure tensor and the energy operator.

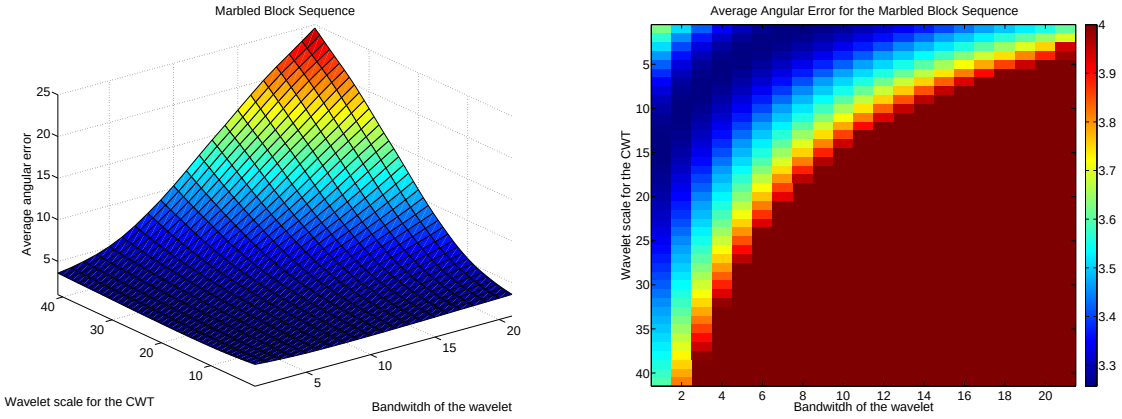


Figure 4.2: The optimal average angular error as a function of the bandwidth and scale of the Mexican hat wavelet for the Marbled Block sequence.

In general for motion estimation purposes the energy tensor is comparable to the structure tensor, but requires less computational steps. The energy tensor has three essential advantages:

- In comparison with the structure tensor it possesses higher spatial resolution and is auto adapted to the underlying spatial structures due to its phase invariance.
- Unlike in the case of the structure tensor, [Bigun, 2006, pages 114-117] or [Köthe, 2003, pages 26–27], there is no need of oversampling. The higher spatial resolution allows in one dimension an undersampling up to three times without significant performance reduction, [Larkin, 2005, page 8111].
- It needs less computational steps do determine the energy tensor, if there is no need of averaging. The averaging of six tensor components is more expensive than the computation of the Hessian. To obtain the Hessian we need only three additional computational steps for the spatio-temporal derivatives of second order. For an averaging by binomial filters with equivalent standard deviation of 1 and derivative masks of size 5x5x5, in three dimensions the speed acceleration is approximately threefold or 2.9069767 times if we consider the number of multiplications as the significant time consuming computational step.

Chapter 5

Detection of temporal aliasing

5.1 Sampling of band enlarging operators

Translational motion is the simplest motion model. In this section we will give a brief overview of the conditions for occurrence of temporal aliasing in image sequences under the assumption that the objects undergo a translational motion. Translational and affine motion parameter estimation can be derived in the continuous domain, as described in 1.3.1. Conditions for temporal aliasing appearance are investigated in [Bigun, 2006, Section 12.7, pages 264–267].

A simple illustration of the temporal aliasing effect is described below. Given a one dimensional sine signal $s(t) = \sin t$ we can determine the translation in an unambiguous way up to the period of 2π. Translations by a length l greater than 2π will be recognized as a shift

$$l - n\,2\pi \qquad n \in \mathbb{N}$$

with suitable n.

Given a image sequence $f = f(x, y, t)$

$$f : \mathbb{R}^3 \to \mathbb{R}$$

let's assume, that there is a translational motion $v = (v_x, v_y)$ in the image plane between two image frames $f(x, y, 0)$ and $f(x, y, t)$

$$f(x, y, t) = f(x - vt, y - vt, 0), \qquad t \in \mathbb{R}.$$

Let $g(x, y)$,

$$g : \mathbb{R}^2 \to \mathbb{R}$$

is a band-limited image. Band-limited means, that the Fourier transform $\mathcal{F}[g]$ has a compact support. Translating g(x,y) to produce a 3 dimensional spatio-temporal image sequence $f = f(x, y, t)$ is a band enlarging operation but $f(x, y, t)$ still remains band limited if the modulus of the translational speed v is bounded.

If we denote by K_s the upper bound for the spatial wavenumber k_s of $g(x, y)$

$$|k_s| > K_s \quad \Rightarrow \quad \mathcal{F}[g](k_x, k_y) = 0$$

then a upper bound K_t for the spatio-temporal wavenumbers k_t of $f(x, y, t)$ is given by

$$|k_t| \leq |v| K_s =: K_t \,.$$

Thus we have seen, that if the translational speed v is bounded, then the spatio-temporal image sequence $f(x, y, t)$ is band limited too. Similar result can be derived if we analyze 3

dimensional voxel images, depth maps or other multidimensional image data with respect to motion estimation and aliasing detection.

If the temporal axis is sampled with the sampling period $\frac{2\pi}{2K_t}$ or tighter, then the speeds not greater than v will be recoverable. In this case the motion planes generated by these shifts will have a smaller inclination angles with the (k_x, k_y)-plane. The maximal speed v_{max} of the translation is bounded by

$$v_{max} \leq \frac{K_t}{K_s} \,. \tag{5.1}$$

This result is summarized in [Bigun, 2006, Lemma 12.8, page 265].

A well known strategy for avoidance of temporal aliasing in spatio-temporal image sequences follows from the above analysis. The strategy consists of spatial low-pass filtering of the image frames and is described in [Jähne, Bernd, 2005, page 293], [Christmas, 2000] and [Bigun, 2006, pages 265–266]. In this manner the high-frequency spatial content is removed and higher speeds are recoverable in an unambiguous way. A drawback of spatial low-pass filtering is occurrence of errors at object boundaries in the following image processing steps such as gradient calculation, motion estimation, segmentation from motion etc. It should be mentioned that temporal aliasing will potentially affect any temporal filtering operation. The temporal derivative is an important value for the gradient based optic flow estimation, so it will be desirable to remove temporal aliasing. It will be even better to detect aliasing occurrence and to label the regions, it occurs in.

We have seen that the support of the spatio-temporal Fourier spectrum of an image sequence depends on the upper bound v_{max} for the modulus of the translational velocity $|v|$. In practice we don't know in advance the upper bound v_{max}. We propose a practical method to detect regions of temporal aliasing by optic flow estimation in the scale space.

5.2 Experimental Results

Experimental Results

An example with ground truth for a real life sequence

To prepare the scene for more complicated experiments and to study the behavior of aliasing occurrence in the scale-space, we investigated the dependence of estimated optic flow on the scale parameter. We decomposed a simple real life test sequence with ground truth. The test sequence shows a known horizontal translational motion of a scale grid paper, see also figure 5.1.

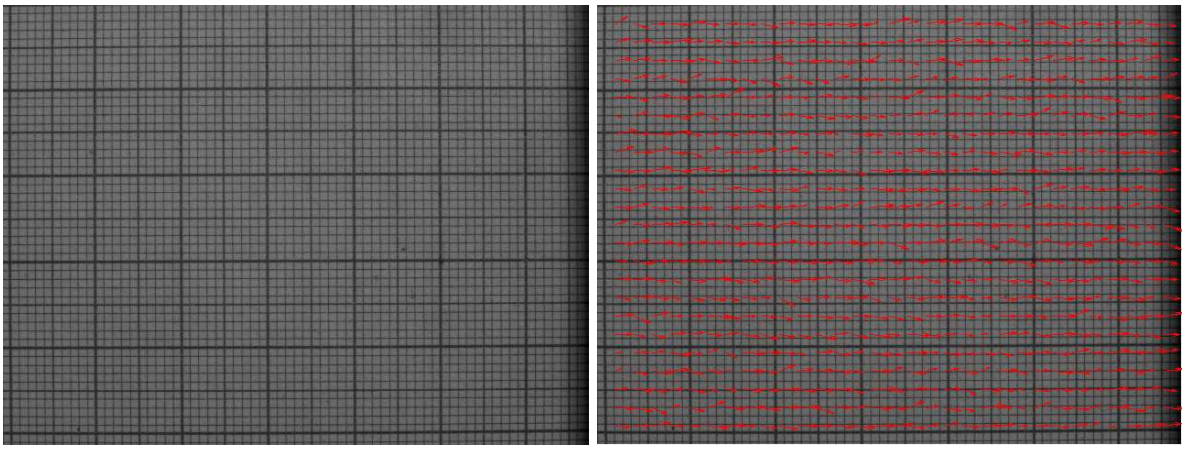


Figure 5.1: Real world test sequence and estimated flow field of a test sequence at original scale

We acquired the sequence by the stop-and-shoot method. If there are enough frames this allows for any kind of subsampling of the sequence in temporal direction in order to conduct experiments. By a proper subsampling of the frames we generated an aliased motion of the thin lines in the opposite direction of the real motion but choosing the subsampling rate such that the motion of the thick lines is not affected.

The image sequence was decomposed by the continuous wavelet transform with the Mexican hat wavelet. The Mexican hat wavelet ensures the best scale-frequency resolution in the sense of the uncertainty principle, see also section 1.4.1. Then we performed optic flow estimation with a simple local algorithm at every scale of the decomposed sequence.

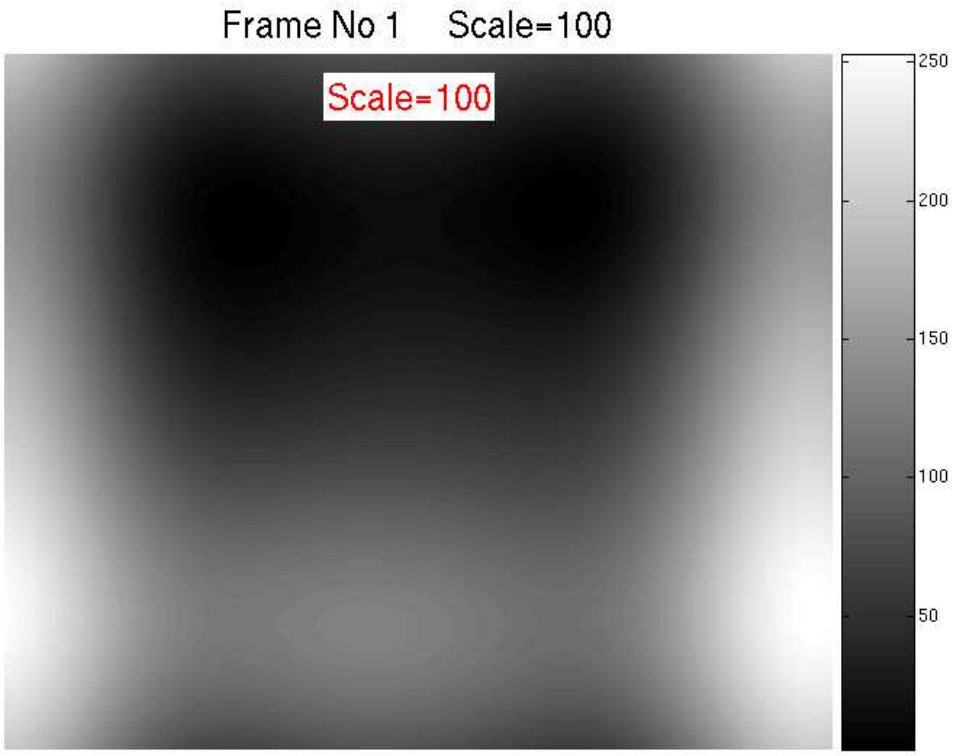


Figure 5.2: Movie: Scale decomposition of a test sequence

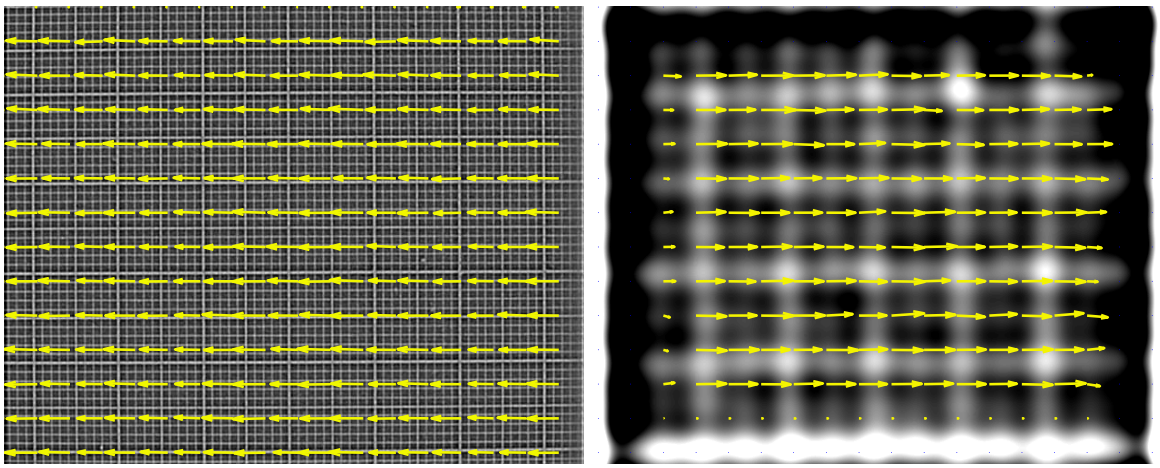


Figure 5.3: Aliasing detection, optical flow on scale 3 and scale 21

Aliasing Detection Aliasing velocity vs. wavelet scale

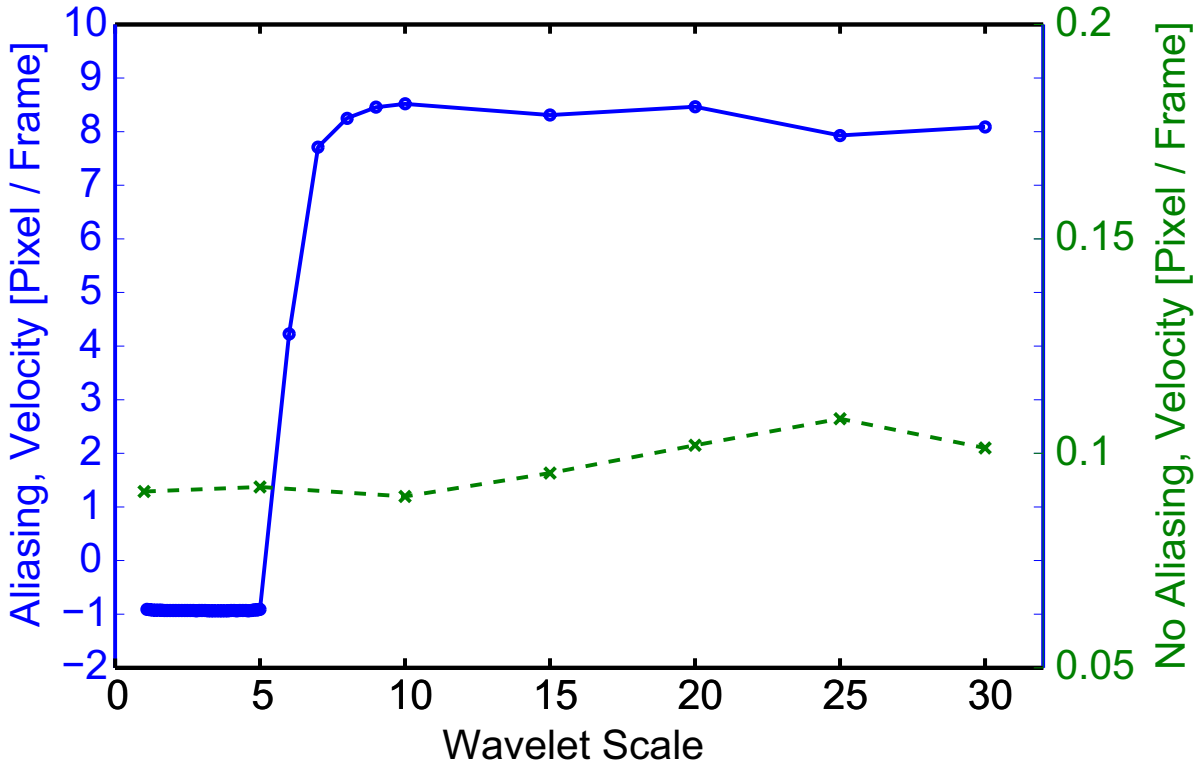


Figure 5.4: Aliasing Velosity vs. Wavelet Scale

STATEMENT 5.1. *If the estimated velocity for a given segmented object changes its direction or its order of magnitude across the scale space, then there is aliasing in the pixel.*

Unfortunately we can not detect *all* regions of aliasing by this method. It can happen, that maybe there are some objects in the image sequence, where temporal aliasing occurs, but the velocity changes neither its direction nor its order of magnitude.

5.3 Conclusion

Detection of aliasing in the scale-space is possible. By our method not all regions of temporal aliasing are detected.

The scale-space is generated by the continuous wavelet transform with rotationally symmetric wavelets. The optical flow is computed in parallel at each scale.

The underlying assumption is, that in a neighborhood of the investigated pixel there is a constant motion. This assumption is realistic in the inside of moving objects.

An open question is the comparison and cross-check of the proposed method by direct implementation of the inequality 5.1. In this case the local wavenumbers can be estimated by a transform of Cohen's class, 1.4 and the optic flow can be computed by a local method.

Chapter 6

Conclusion and outlook

The topic of this work was to reveal some new aspects of motion analysis from image sequences by scale space decomposition of the image data. The scale space generation was performed by the wavelet transform. The discrete wavelet transform is suitable for compression of multidimensional signal and image data. For analysis purposes, the continuous wavelet transform is a better tool than the discrete one, because it generates a continuous scale representation of the image sequence. In order to avoid direction dependence we used rotationally symmetric wavelets. It is easy to construct a fully isotropic multidimensional wavelet just by rotation of the one dimensional one. The Mexican hat wavelet localizes best in two dimensions in the sense of the uncertainty relation, [Dahlke, S. and Maass, P., 1995]. If there is no a priori knowledge about the image data, the data can be decomposed by the Mexican hat wavelet, because in two dimensions it is optimal with respect to minimization of the uncertainty principle, [Dahlke, S. and Maass, P., 1995].

We discussed also the issue of tensor field regularization for optic flow computation and conducted accuracy measurements for motion estimation by adaptive smoothed tensor fields. The adaptive smoothing of the tensors relies on nonlinear anisotropic diffusion with discontinuity and curvature preservation. We reached an accuracy gain under properly chosen parameters for the diffusion filter.

Another research topic was the optic flow estimation in the scale space by the energy operator. Our experiments showed, that there is a need of post-integration for the three dimensional energy operator, in order to achieve optimal average angular error of the estimated flow fields. For orientation estimation with the energy operator [Felsberg and Jonsson, 2005, page 498] reported similar results in two dimensions and applied a Gaussian post-filtering with $\sigma = 1$ and a smoothing window size of 7x7. In comparison with the structure tensor, the energy operator requires an additional computational step, this is the calculations of the Hessian of the image data for every pixel. In the numerical experiments there was an accuracy gain for a real life sequence in our measurements. In general for motion estimation purposes the energy tensor is comparable to the structure tensor, but requires less computational steps.

The energy tensor has three essential advantages:

- In comparison with the structure tensor it possesses higher spatial resolution and is auto adapted to the underlying spatial structures due to its phase invariance.

- Unlike in the case of the structure tensor, [Bigun, 2006, pages 114-117] or [Köthe, 2003, pages 26–27], there is no need of oversampling. The higher spatial resolution allows in one dimension an undersampling up to three times without significant performance reduction, [Larkin, 2005, page 8111].

- It needs less computational steps to determine the energy tensor, if there is no need of averaging. The averaging of six tensor components is more expensive than the computation of the Hessian. To obtain the Hessian we need only three additional computational steps for the spatio-temporal derivatives of second order. For an averaging by binomial filters with equivalent standard deviation of 1 and derivative masks of size 5x5x5, in three dimensions the speed acceleration is approximately threefold or 2.9069767 times if we consider the number of multiplications as the significant time consuming computational step.

We showed also, that detection of aliasing in the scale-space is possible. By our method not all regions of temporal aliasing are detected. The scale-space is generated by the continuous wavelet transform. The optical flow is computed in parallel at each scale.

With respect to mathematical modelling of optic flow computation in this work we formulated a new regularization model for motion estimation with discontinuity and curvature preservation. We discussed the nonlocal dependence of the diffusion tensor on the image data in the mathematical sense.

Here are some open questions and comments on possible improvements of our analysis.

- Gradient computation by derivative filters, which are tuned for the wavelet scale with respect to rotation invariance.

  Due to the discretization of image data on rectangular grids, there is an angle dependent error in the derivative computation. The basic concepts of error minimization for derivative filters with cross smoothing are discussed in [Jähne, Bernd, 2004, page 416], [Scharr, 2000] and [Jähne, B. et al., 1999]. In the optimization of the filter masks with respect to better isotropy the wavenumber weighting is important for the definition of the objective function. The wavenumber weighting $w(k)$ can be set as

$$w(k) := \cos^n 2\pi k, \qquad n = 0, 1, 2, 4,$$

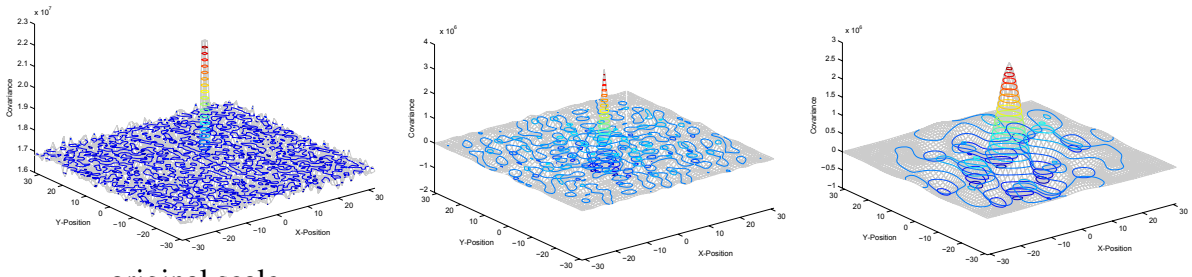
  where k is the wavenumber. Then the filter masks are computed by nonlinear optimization.

  It is still an open question, which accuracy gain can be reached by computing the tensor field with filters optimized for the wavelet and the wavelet scale.

- Tuning of derivative filters for the given wavelet scale from statistical point of view.

  Filters approximating the local gradient in image sequences can be optimized with respect to the noise characteristics of the image data, [Mester, 2003]. Experiments showed, that there is an improvement of the estimated optic flow in comparison with other methods, that do not take into account the statistical properties of the image data, [Krajsek and Mester, 2007].

  The noise in a wavelet decomposed sequence is correlated, see figure 6.1.



original scale small wavelet scale large wavelet scale

Figure 6.1: Autocorrelation of white noise across the wavelet scale. The autocorrelation increases with increasing wavelet scale. At the original scale the autocorrelation is a peak in the origin.

It is a field of further investigation, to which extent the error can be reduced, if we apply derivative filters optimized for the noise correlation originating from the wavelet decomposition.

- Adaptive integration of the tensor field by the local alignment method may improve the accuracy of the optic flow field.

  In chapter 3 we investigated the performance of a method based on nonlinear anisotropic diffusion for scale space generation of a feature extracted from the image data. The feature is the structure tensor and the diffusion flow was constrained on the manifold of matrices with given set of eigenvalues. This constraint is a good trade off between performance and computational complexity. One can investigate, whether there will be an improvement and to which extent, if the diffusion flow is constrained on the manifold of orthogonal matrices $\mathbb{SO}(3)$. This is the canonical choice in a natural way for the manifold, because the tensor field contains only orthogonal matrices. In this manner we use additional information about the feature, extracted from the image sequence. In order to avoid the formation of artificial discontinuities, the tensors should be aligned to each other in a neighborhood, [Tschumperle, 2002, page 143].

- Strategies for incorporation of the computed flow fields for the different wavelet scales.

  We showed experimentally, that in scenes with moving objects the detection of temporal aliasing is possible by computation of the optic flow in parallel at every scale. It is still an open question, how this results can be incorporated into one flow field, representing the motion in the original sequence.

- Comparison and validation of the proposed method for temporal aliasing detection on sequences with displacement ground truth by direct implementation of the inequality 5.1. In this case the local wavenumbers can be estimated by a transform of the Cohen's class, section 1.4.

- Fair comparison between different diffusion flows.

  In optic flow computation and motion estimation the accuracy of the flow field is important for practical applications. We compared the best results with respect to accuracy of two diffusion flows, on the one hand the isotropic linear diffusion, which is a Gaussian smoothing and the isotropic nonlinear diffusion on the other hand. Hence our method to compare different diffusion flows is with respect to practical results and is feature based. Maybe a better theoretical comparison of different smoothing methods will be possible

in an invariant metrics, cf. [Xavier Pennec and Nicholas Ayache, 1998; Xavier Pennec et al., 2006].

- Computation of a spatial gradient of a tensor field in Riemannian metrics.

  In chapter 3 we computed the gradient of the tensor field in the flat Euclidean metrics. Maybe we can improve the accuracy of the results, if in this step the derivatives of the tensor field are computed on the manifold of orthogonal matrices.

- How can we express the ϕ-function formalism in Rimannian metrics?

  Discontinuity preserving regularization was introduced in section 1.3.2. The idea was to replace the regularization term

  $$\int_\Omega \|\nabla u\|^2 \, dx$$

  by a discontinuity preserving term

  $$\int_\Omega \Psi(\|\nabla u\|) \, dx$$

  with suitable Ψ.

  Under regularity assumptions for Ψ we end up with

  $$\operatorname{div}\left(\frac{\Psi'(\|\nabla u\|)}{\|\nabla u\|}\nabla u\right)$$

  for the smoothing term. An interesting open question is, how can we adapt this formalism to a Riemannian or an invariant metrics, [Xavier Pennec and Nicholas Ayache, 1998; Xavier Pennec et al., 2006].

Part I

Appendix

Appendix A

Tools

We used preferably open source tools to fulfill our research and to create this document. If not explicitly mentioned, all the tools listed bellow are with open source code, they are free of charge, well documented and licensed in the spirit of the GNU General Public License or the CeCILL-C license .

Type-setting

LaTeX is a standard type-setting environment based on TeX for writing out texts with mathematical formulas, creating reference lists and preparing presentations.

PDE solvers

PETSc is a free open source library for solving partial differential equations. It's a good library for implementation of partial differential equations for motion estimation by variational techniques. PETSc allows also parallel computations by Message Passing Interface.

Image processing libraries

CImg stands for *"Cool Image"* and this is an image processing library which completely deserves its name. CImg is really cool! The library offers an image class with template pixel types and a variety of image processing and matrix operations. There are also nice display possibilities for image surfaces, isosurfaces of volume data and interactive visualization implemented in the CImg library. We used the CImg library to solve tensor valued nonlinear PDE's by line integral convolution.

heurisko is a commercial closed source script language for image processing with an API for C extensions. heurisko offers a fast low-level image processing with good possibilities to control, manage, synchronize and operate on external devices for optical measurements such like cameras, light sources, translational stages etc. We used heurisko to acquire real life test sequences.

ImageJ is a program written in Java which offers basic image processing operations, geometric transformations and spatial calibration. ImageJ provides extensibility via Java plug-ins and the user can write his own applications. There are a lot of plug-ins available for ImageJ including some simple basic optical flow algorithms and average angular error measurement. ImageJ can be run either as an on-line applet or as downloadable application. We downloaded and used ImageJ to read volume data in tiff format and float precision of the pixel value.

ltilib is a template-based C++ image processing library with some motion estimation methods implemented.

vigra is a good image processing library, written in C++. The concept of the library is to develop reuseable software for image data in arbitrary dimensions.

JavaView is not exactly an image processing library, but a mathematical visualization software. JavaView is suitable for preparing interactive visualization of flow fields and incorporating it into electronic documents, which allow user interaction and on-line computations.

Wavelets libraries

YAWTB is a wavelet toolbox which provides the computation of a two or three dimensional non-separable wavelet transform for image sequences. In this manner the wavelet decomposition is at least theoretically invariant with respect to rotations and there are no preferred directions in the bandpass decomposed sequences. The separable wavelet transform introduces a systematic error in the estimated flow field because of preferred directions in the the wavelet decomposition.

Other wavelet libraries and toolboxes, we used in our research are WaveLab802 , LastWave and the Wavelet Toolbox. The Wavelet Toolbox is a commercial software.

Linear algebra and matrix computation

lapack and **blas** are standard linear algebra libraries written in FORTRAN 77 for matrix decomposition and for solving linear systems of equations and eigenvalue problems. lapack and blas are used in Expokit and levmar is a Levenberg-Marquardt optimization routine in C/C++.

Expokit is a FORTRAN package for computation of matrix exponentials. We used this package to implement the isospectral flow on the manifold of matrices with given set of eigenvalues in chapter 3.

Optimization

levmar is a Levenberg-Marquardt optimization routine in C/C++. levmar was used to search for optimal wavelet scale, wavelet bandwidth and integration size of the tensor field for the motion estimation by the energy operator in chapter 4.

Camera calibration toolboxes

Janne Heikkilä and Davide Scaramuzza and Jean-Yves Bouguet are good kamera calibration tollboxes, which were used for preliminary calibration of real life test sequences.

Bibliography

A. Bovik, J. Havlicek, M. Desai, and D. Harding . Limits on discrete modulated signals. *IEEE Transactions on Signal Processing*, 45(4):867–879, Apr 1997. citeseer.ist.psu.edu/bovik97limits.html. 60

Alan C. Bovik, Joseph P. Havlicek, and Mita D. Desai . Theorems for Discrete Filtered Modulated Signals. In *Proc. IEEE ICASSP-93*, volume 3, pages 153–156, Minneapolis, MN, USA, Apr 1993. IEEE. citeseer.ist.psu.edu/bovik93theorems.html. 59

Alfred Karl Louis, Peter Maass, and Andreas Rieder. *Wavelets*. Teubner, 1994. 23

Alvarez, L., Guichard, F., Lions, P.L., and Morel, J.M. . Axioms and Fundamental Equations of Image Processing: Multiscale Analysis and P.D.E. *Archive for Rational Mechanichs and Analysis*, 123(3):199–257, 1993. 2, 26

P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2:283–319, 1989. 4, 6

Andrés Bruhn, Joachim Weickert, and Christoph Schnörr . Combining the advantages of local and global optic flow methods. In *Proceedings of the 24th DAGM Symposium on Pattern Recognition*, volume 2449 of *Lecture Notes In Computer Science*, pages 454–462, 2002. citeseer.ist.psu.edu/553233.html. 39

G. Aubert and P. Kornprobst . A Mathematical Study of the Relaxed Optical Flow Problem in the Space $BV(\Omega)$. *SIAM J. Math. Anal.*, 30(6):1282–1308, 1999. 8, 13, 14

G. Aubert and P. Kornprobst . *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, 2002. 1, 3, 4, 9, 14, 17

G. Aubert, R. Deriche, and P. Kornprobst . Computing optical flow via variational techniques. *SIAM J. on Appl. Math.*, 60(1):156–182, 1999. 8, 14, 20

Bab-Hadiashar, A. and Suter, D.. Optic Flow Calculation Using Robust Statistics. In *CVPR*, pages 988–993, Puerto Rico, 1997. 6

Bab-Hadiashar, A. and Suter, D.. Robust Optic Flow Computation. *International Journal of Computer Vision*, 29(1):59–77, 1998. 6

Bainbridge-Smith, A. and Lane, R. G.. Determining Optical Flow Using a Differential Method. *Image and Vision Computing*, 15:11–22, 1997. 5

Barron, J. L., Fleet, D. J., and Beauchemin, S. . Performance of Optical Flow Techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994. 4

Beauchemin, S. S. and Barron, J. L. . The Computation of Optical Flow. *ACM Computing Surveys*, 27(3):433–467, 1995. 4

Eugene Belogay and Yang Wang . Arbitrarily Smooth Orthogonal Nonseparable Wavelets in $\mathbf{R}^2$. *SIAM Journal on Mathematical Analysis*, 30(3):678–697, 1999. citeseer.ist.psu.edu/belogay99arbitrarily.html. 33, 34

F. Bergholm. A theory of optical velocity fields and ambiguous motion of curves. In *Proc. IEEE ICCV*, pages 165–176, 1988. 6

Christophe P. Bernard . Discrete Wavelet Analysis for Fast Optic Flow Computation. *Applied and Computational Harmonic Analysis*, 11:32–63, 2001. http://www.cmap.polytechnique.fr/~bernard/Publications/. 28, 29

Josef Bigun. *Vision with direction*. Springer, Berlin, 2006. http://www2.hh.se/staff/josef/. 4, 62, 65, 67, 68, 73

M. J. Black. *Robust Incremental Optical Flow*. PhD thesis, Yale University, 1992. 6

Black, M. J. and Anandan, P. . The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, January 1996. 6

Bovik, Alan C. and Maragos, Petros . Conditions for positivity of an energy operator. *IEEE Transactions on Signal Processing*, 42(2):469–471, Feb 1994. http://cvsp.cs.ntua.gr/publications/jpubl+bchap/BovikMaragos_PosEnOp_ieeetSP1994.pdf. 62

Brox, Thomas, Weickert, Joachim, Burgeth, Bernhard, and Mrázek, Pavel . Nonlinear structure tensors. *Image and Vision Computing*, 24(1):41–55, January 2006. 39, 40

Bruhn, A., Weickert, J., Kohlberger, T., and Schnoerr, C. . Discontinuity-Preserving Computation of Variational Optic Flow in Real-Tome. In Ron Kimmel, Nir A. Sochen, and Joachim Weickert, editors, *Scale Space and PDE Methods in Computer Vision, 5th International Conference, Scale-Space 2005, Hofgeismar, Germany, April 7-9, 2005, Proceedings*, volume 3459 of *Lecture Notes in Computer Science*. Springer, 2005a. ISBN 3-540-25547-8. 9, 20

Bruhn, A., Weickert, J., and Schnoerr, C. . Lucas/Kanade Meets Horn/Schunk: Combining Local and Global Optic Flow Methods. *International Journal of Computer Vision*, 61(3): 211–231, 2005b. 9, 20, 39

Brian Cabral and Leith Casey Leedom . Imaging vector fields using line integral convolution. In *SIGGRAPH '93: Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 263–270, New York, NY, USA, 1993. ACM Press. ISBN 0-89791-601-8. doi: http://doi.acm.org/10.1145/166117.166151. 3

J. Canny. A Computational Approach to Edge Detection. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 8(6):679–698, Nov. 1986. 6

M. Chipot . Asymptotic behaviour of solutions to nonlinear parabolic equation with nonlocal Terms. http://www.math.unizh.ch/chipot/, December 2004. 18

W. J. Christmas. Filtering Requirements for Gradient-BasedOptical Flow Measurement. *IEEE Transactions on Image Processing*, 9(10):1817–1820, 2000. 68

Christopher R. Johnson and Charles D. Hansen. *The Visualization Handbook*. Academic Press, June 2004. 48

Chu, C. H. and Delp, E. J.. Estimating displacement vectors from an image sequence. *Journal of the Optical Society of America A*, 6(6):871–878, 1989. 6

Leon Cohen . Time-frequency distributions – a review. In *Proceedings of IEEE*, volume 77, pages 941–981, July 1989. 22, 24

Corpetti, T., Etienne, M., and Perez, P. . Dense Estimation of Fluid Flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3), March 2002. http://www.irisa.fr/vista/Themes/Demos/MouvementFluide/fluide.html. 15

Dahlke, S. and Maass, P. . The Affine Uncertainty Principle in One and Two Dimensions. *Computers Math. Applic.*, 30(3–6):293–305, 1995. 25, 36, 73

L Debnath, editor. *Wavelet Transforms and Time-Frequency Signal Analysis*. Birkhäuser, 2001. 25

Cédric Demonceaux and Djemâa Kachi-Akkouche . Fast motion estimation and motion segmentation using multi-scale approach. In *ICIP*, pages 377–380, 2004. 28, 29

O. Faugeras. On the motion of 3-d curves and its relationship to optical flow. In *Proc. 1st ECCV*, pages 107–117, Antibes, 1990. Springer-Verlag. 6

Manuela Feilner, Dimitri Van De Ville, and Michael Unser . An orthogonal family of quincunx wavelets with continuously adjustable order. *IEEE Transactions on Image Processing*, 14(4):499–510, 2005. 38

M. Felsberg and G.H. Granlund . POI Detection using Channel Clustering and the 2D Energy Tensor. In *Pattern Recognition: 26th DAGM Symposium*, volume 3175 of *LNCS*, pages 103–110, Tübingen, Germany, 2004. Springer Berlin. http://www.cvl.isy.liu.se/ScOut/Publications/Papers/DAGM2004_felsberg_granlund.pdf. 60

M. Felsberg and E. Jonsson . Energy Tensors: Quadratic, Phase Invariant Image Operators. In Kropatsch, W., Sablatnig, R., and Hanbury, A., editors, *DAGM 2005*, pages 493–500. Springer, 2005. http://www.cvl.isy.liu.se/ScOut/Publications/Papers/DAGM2005_felsberg_jonsson.pdf. 60, 64, 73

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 Mathods in Computer Vision*, volume 3459 of *LNCS*, pages 192–203. Springer, 2005. http://www.cvl.isy.liu.se/ScOut/Publications/Papers/SSVM2005_felsberg_koethe.pdf. 60

Fennema, C. and Thompson, W.. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9:301–315, 1979. 4

Fleet, D. J. and Jepson, A. D.. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5:77–104, 1990. 7

C. S. Garbe. *Measuring Heat Exchange Processes at the Air-Water Interface from Thermographic Image Sequence Analysis*. PhD thesis, University of Heidelberg, Heidelberg, Germany, December 2001. 39

Girosi, F., Verri, A., and Torre, V.. Constraints for the computation of optical flow. In Girosi, F., Verri, A., and Torre, V., editors, *Proceedings Workshop on Visual Motion, March 1989, Irvine, CA*, pages 116–124, Washington, 1989. 5

Glazer, F., Reynolds, G., and Anandan, P.. Scene Matching through hierarchical correlation. In *Proc. Conference on Computer Vision and Pattern Recognition*, pages 432–441, Washington, 1983. 4

S. Gong. Curve motion constraint equation and its applications. In *Proceedings Workshop on Visual Motion, March 1989, Irvine, CA*, pages 73–80, Washington, 1989. 7

Hampel, F. R., Ronchetti, E. M., and others. *Robust Statistics: The Approach Based on Influence Functions*. John Wiley and Sons, New York, 1986. 6

Haussecker, H., Garbe, C., and others. A Total Least Squares for Low-Level Analysis of Dynamic Scenes and Processes. In *DAGM*, pages 240–249, Bonn, Germany, 1999. Springer. 4

H.J.A.M. Heijmans and J. Goutsias . Nonlinear Multiresolution Signal Decomposition Schemes–Part II: Morphological Wavelets. 9(11):1897–1913, November 2000. http://iris.usc.edu/Vision-Notes/bibliography/author/indexh.html. 30

E. C. Hildreth. Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23:309–354, 1984. 7

Matthias Holschneider and Gerd Teschke . On the existence of optimally localized wavelets. Technical report, Preprint series of the DFG priority program 1114 "Mathematical methods for time series analysis and digital image processing", April 2005. http://www.math.uni-bremen.de/zetem/DFG-Schwerpunkt/preprints/pdf/089.pdf. 25

B. K. P. Horn. *Robot Vision*. MIT Press, Cambridge, MA, 1986. 3

Horn, B. K. P. and Schunk, B. . Determining Optical Flow. *Artificial Intelligence*, 17:185–204, 1981. 4, 8

P. J. Huber. *Robust Statistics*. John Wiley and Sons, New York, 1981. 6

B. Jähne. *Spatio-Temporal Image Processing : Theory and Scientific Applications*. Lecture Notes in Computer Science. Springer-Verlag, 1993. 7

Bernd Jähne. *Digital Image Processing*. Springer-Verlag, Heidelberg, Germany, 5th edition, 2002. 1, 4, 7

Jähne, B., Scharr, H., and Körkel, S.. Principles of Filter Design. In Jähne, B., Haußecker, H., and Geißler, P., editors, *Handbook of Computer Vision and Applications*, volume 2, pages 125–151. Academic Press, 1999. 74

Jähne, B., Mester, R., Barth, E., and Scharr, H., editors. *Complex Motion*. Springer, 1st edition, 2007. 63

Jähne, Bernd. *Practical Handbook on Image Processing for Scientific and Technical Applications*. CRC Press, Boca Rota London New York Washington, D.C., 2004. 1, 74

Jähne, Bernd. *Digital Image Processing*. Springer, 2005. 60, 68

J.F. Kaiser . On a simple algorithm to calculate the 'energy' of a signal. In *International Conference on Acoustics, Speech, and Signal Processing, 1990. ICASSP-90., 1990*, pages 381–384, 1990a. 59, 62

J.F. Kaiser. On Teager's energy algorithm and its generalization to continuous signals. In *Proc. 4th IEEE Digital Signal Processing Workshop, Mohonk, NY, Sept. 1990*, 1990b. 59

J.F. Kaiser. Some useful properties of Teager's energy operators. In *Proc. International Conference on Acoustics, Speech, and Signal Processing, 1993. ICASSP-93., 1993*, volume 3, pages 149–152, Minneapolis, MN, USA, Apr 1993. IEEE. 59

Peter Kovesi . Image Features from Phase Congruency. *Videre: Journal of Computer Vision Research*, 1(3):1–26, 1999. http://mitpress.mit.edu/e-journals/Videre/001/articles/v1n3001.pdf. 60

Kai Krajsek and Rudolf Mester. Wiener-optimized discrete filters for differential motion estimation. In B. Jähne, E. Barth, R. Mester, and H. Scharr, editors, *First International Workshop on Complex Motion, Günzburg, Germany, Oct. 2004*, volume 3417 of *Lecture Notes in Computer Science*, pages 31–41, Berlin, 2007. Springer Verlag. 74

Eivind Kvedalen . *Signal processing using the Teager energy operator and other nonlinear operators*. PhD thesis, University of Oslo, 2003. http://folk.uio.no/eivindkv/ek-thesis-2003-05-12-final-2.pdf. 59, 60, 61, 62

U Köthe. *Generische Programmierung für die Bildverarbeitung*. PhD thesis, Universität Hamburg,, 2000. 60

U Köthe . Edge and junction detection with an improved structure tensor. In B. Michaelis and G. Krell, editors, *Proc. of 25th DAGM Symposium*, volume 2781 of *Lecture Notes in Computer Science*, pages 25–32, Magdeburg, 2003. DAGM. http://kogs-www.informatik.uni-hamburg.de/~koethe/papers/structureTensor.pdf. 62, 65, 73

U. Köthe and M. Felsberg . Riesz-Transforms Versus Derivatives: On the Relationship Between the Boundary Tensor and the Energy Tensor. In Kimmel,R., Sochen, N., and Weickert, J., editors, *Scale Space and PDE Methods in Computer Vision*, pages 179–191, Berlin, 2005. Springer. http://kogs-www.informatik.uni-hamburg.de/~koethe/papers/abstracts/GEToperator.html. 60

K. G. Larkin . Efficient nonlinear algorithm for envelope detection in white light interferometry. *Journal of the Optical Society of America,A*, 13(4):832–843, 1996. http://www.physics.usyd.edu.au/~larkin/Neat_Algoritm_JOSAA96.pdf. 60, 62

K. G. Larkin . *Topics in Multi-Dimensional Signal Demodulation*. PhD thesis, University of Sidney, 2001. http://www.physics.usyd.edu.au/~larkin/. 60, 62

K. G. Larkin . Uniform estimation of fringe orientation using local and nonlocal 2-D energy operators. *Optics Express*, 13(20):8097–8121, Oct 2005. http://www.opticsexpress.org/abstract.cfm?id=85762. 59, 60, 61, 62, 65, 73

Ghislain Lemaur . *On the Choice of the Wavelet Basis Function for Image Processing*. PhD thesis, Université de Mons-Hainaut, Mons, Belgium, 2003. 33, 34

Matthias Lindemann . *Approximation Properties of Non-Separable Wavelet Bases With Isotropic Scaling Matrices And Their Relation to Besov Spaces*. PhD thesis, University Bremen, 2005. 35

Little, J. J. and Verri, A.. Analysis of differential and matching methods for optical flow. In *IEEE Workshop on Visual Motion*, pages 173–180, Irvine, CA, 1989. 4

B. D. Lucas. *Generalized image matching by the method of differences*. PhD thesis, Carnegie-Mellon University, Pittsburgh, PA, 1984. 6

Lucas, B. and Kanade, T.. An Iterative Image Registration Technique with an Application to Stereo Vision. In *DARPA Image Understanding Workshop*, pages 121–130, 1981. 6

Stéphane Mallat. *A Wavelet tour of signal processing*. Academic Press, 2 edition, 1999. http://www.cmap.polytechnique.fr/~mallat/book.html. 30

P. Maragos and A. C. Bovik . Image demodulation using multidimensional energy separation. *Journal of the Optical Society of America A*, 12:1867–1876, September 1995. http://cvsp.cs.ntua.gr/publications/jpubl+bchap/MaragosBovik_ImageDemodMultidimEnergSepar_JOSA1995.pdf. 60

Maragos, P., Bovik, A. C., and Quatieri, T. F. . A multidimensional energy operator for image processing. In *SPIE Conference on Visual Communications and Image Processing*, pages 177–186, 1992. 60

Maragos, P., Kaiser, J.F., and Quatieri, T.F. . On amplitude and frequency demodulation using energy operators. *IEEE Transactions on Signal Processing*, 41(4):1532–1550, Apr 1993. 59, 62

Markandey, V. and Flinchbaugh, B. E.. Multispectral constraints for optical flow computation. In *Proc. of ICCV*, pages 38–41, Osaka, Japan, 1990. 5

Marr, D. and Hildreth, E. C.. Theory of edge detection. *Proc.Royal Soc.London Ser.B*, 207: 187–217, 1980. 6

Martin Lefébure and Laurent D. Cohen . Image registration, optical flow and local rigidity. *Journal of Mathematical Imaging and Vision*, 14:131–147, 2001. 28

McCane, B., Novins, K., and others . On Benchmarking Optical Flow. *CVIU*, 84(1):126–143, October 2001. http://www.cs.otago.ac.nz/research/vision/Research/OpticalFlow/opticalflow.html. 19, 20

Rudolf Mester. A New View at Differential and Tensor-Based Motion Estimation Schemes. In Bernd Michaelis, editor, *Pattern Recognition 2003*, Lecture Notes in Computer Science, Magdeburg, Germany, September 2003. Springer Verlag. 74

M. Meyer . *Wavelets and Operators*. Cambridge University Press, 1992. 35

Michael Moore, Reinhard Bernstein, and Sanjit Mitra . A Generalization of the Teager algorithm. In *Proceedings IEEE Workshop on Nonlinear Signal and Image Processing NSIP 97*, Ann Arbor, Michigan, Sept 1997. IEEE. http://citeseer.ist.psu.edu/cache/papers/cs/18678/http:zSzzSzverdi.ece.ucsb.eduzSzpublicationszSz97NSIP.pdf/moore97generalization.pdf. 60

MITICHE, A. AND BOUTHEMY, P.. Computation and analysis of image motion: a synopsis of current problems and methods. *International journal of computer vision, IJCV*, 19(1):29–55, July 1996. 4

MITICHE, A., WANG, Y. F., AND AGGARWAL, J. K.. Experiments in computing optical flow with the gradient-based, multiconstraint method. *Pattern Recognition*, 20(2):173–179, 1987. 5

MITRA, S.K., LI, H., LIN, I.-S., AND YU, T.-H. . A new class of nonlinear filters for image enhancement. In *International Conference on Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991*, volume 4, pages 2525–2528, Toronto, Ont., Canada, Apr 1991. IEEE. 60

FERNANDO A. MUJICA, JEAN-PIERRE LEDUC, ROMAIN MURENZI, AND MARK J. T. SMITH. A new motion parameter estimation algorithm based on the continuous wavelet transform. *IEEE Transactions on Image Processing*, 9(5):873–888, 2000. 28, 29

H. H. NAGEL. Displacement vectors derived from second-order intensity variations in image sequences. *Computer Graphics and Image Processing*, 21:85–117, 1983. 4, 5

H. H. NAGEL. On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33(3):299–324, 1987. 5

J. V. NEUMANN. *Mathematische Grundlagen der Quantenmechanik*, volume 38 of *Grundlehren der math. Wissenschaften*. Springer, Heidelberg, 1968. 22

PETROVIC, A., DIVORRA ESCODA, O., AND VANDERGHEYNST, P. . Multiresolution segmentation of natural images: From linear to non-linear scale-space representations. *IEEE Transactions on Image Processing*, 13(8):1104–1114, 2004. http://ltspc4.epfl.ch/%7Epetrovic/publications.htm. 17

PRASSAD, L. AND IYENGAR, S.S.. *Wavelet analysis with applications to image processing*. CRC Press LLC, Boca Raton, Florida, 1997. 21

PROESMANS, M., VAN GOOL, L., AND OTHERS . Determination of Optical FLow and its Discontinuities using Non-Linear Diffusion. In *ECCV*, pages 295–304, 1994. 8

R. HAMILA, J. ASTOLA, M. GABBOUJ, AND M. RENFORS. Teager Energy and the Ambiguity Function, 1997. citeseer.ist.psu.edu/hamila99teager.html. 25

R. HAMILA, J. ASTOLA, F. ALAYA CHEIKH, M. GABBOUJ, AND M. RENFORS . Teager Energy and the Ambiguity Function. *IEEE Transactions on Signal Processing,*, 47(1):260–262, Jan 1999. citeseer.ist.psu.edu/hamila99teager.html. 25, 62

RIOUL, OLIVIER AND FLANDRIN, PATRICK . Time-scale energy distributions: A general class extending wavelet transforms. *IEEE Transactions on Signal Processing*, 40(7):1746–1757, 1992. 24

ROUSSEEUW, P. J. AND LEROY, A.. *Robust regression and outlier detection*. Wiley, 1987. 6

ROUSSEEUW, P. J. AND VAN AELST, S.. Positive-Breakdown robust methods in computer vision. *Computing Science and Statistics*, 31:451–460, 1999. 6

BALU SANTHANAM . Generalized energy demodulation for large frequency deviation and wideband signals. *IEEE Signal Processing Letters*, 11(3):341–344, Mar 2004. 60

Hanno Scharr. *Optimale Operatoren in der Digitalen Bildverarbeitung*. PhD thesis, University of Heidelberg, Heidelberg, Germany, 2000. 62, 63, 74

Scharr, H. and Weickert, J.. An Anisotropic Diffusion Algorithm with Optimized Rotation Invariance. In *DAGM*, pages 460–467, Kiel, Germany, September 2000. 62

Scharr, H., Körkel, S., and Jähne, B.. Numerische Isotropieoptimierung von FIR-Filtern mittels Querglättung. In Wahl, F. M. and Paulus, E., editors, *DAGM*, pages 199–208, 1997. 62, 63

C. Schnörr . Determining Optical Flow for Irregular Domains by Minimizing Quadratic Functionals of a Certain Class. *International Journal of Computer Vision*, 6(1):25–38, 1991. 8

C. Schnörr . Segmentation of Visual Motion by Minimizing Convex Non-Quadratic Functionals. In *ICPR94*, pages A:661–663, 1994. 8

Shizawa, M. and Mase, K. . Simultaneous Multiple Optical Flow Estimation. In *ICPR*, pages 274–278, Atlantic City, NJ, USA, June 1990. 19

E. P. Simoncelli. *Distributed analysis and representation of visual motion*. PhD thesis, MIT, Cambride, MA, USA, January 1993. 5, 28

Peter Singer . Uncertainty inequalities for the continuous wavelet transform. *IEEE Transactions on Information Theory*, 45(3):1039–1042, 1999. 25

M. A. Snyder . On the mathematical foundations of smoothness constraints for the determination of optical flow and for surface reconstruction. In *Proceedings Workshop on Visual Motion*, pages 107–115, Irvine, CA, March 1989. 8

Sobey, P. and Srinivasan, M. V.. Measurement of optical flow by a generalized gradient scheme. *J.Opt.Soc.Am.A*, 8(9):1488–1498, 1991. 5

M. V. Srinivasan. Generalized gradient schemes for the measurement of two-dimensional image motion. *Biological Cybernetics*, 63:421–431, 1990. 5

Fridtjof Stein. Efficient Computation of Optical Flow Using the Census Transform. In Carl Edward Rasmussen, Heinrich H. Bülthoff, Bernhard Schölkopf, and Martin A. Giese, editors, *DAGM-Symposium*, volume 3175 of *Lecture Notes in Computer Science*, pages 79–86. Springer, 2004. ISBN 3-540-22945-0. http://springerlink.metapress.com/openurl.asp?genre=article{\&}issn=0302-9743{\&}volume=3175{\&}spage=79. 7

Ingo Stuke . *Algorithmen zur Bestimmung überlagerter Bewegungen und Orientierungen in Bilddaten*. PhD thesis, Universität zu Lübeck, 2006. 18

Stuke,I., Aach,T., Mota,C., and Barth,E. . Estimation of multiple motions: regularization and performance evaluation. In *Image and Video Communication and Processing, Proceedings of SPIE*, volume 5022, pages 75–86, May 2003. 19

Tretiak, O. and Pastor, L.. Velocity estimation from image sequences with second order differential operators. In *Proc. 7th International Conference on Pattern Recognition*, pages 20–22, 1984. 4, 5

D. Tschumperle . *PDE's based regularization of multivalued images and applications*. PhD thesis, Université de Nice-Sophia, 2002. 9, 15, 39, 40, 41, 45, 46, 47, 75

D. Tschumperle . Fast anisotropic Smoothing of multi-valued images using curvature preserving PDE's. Technical report, Equipe Image/GREYC, UMR CNRS 6072, 6 Bd du Maréchal Juin, 14050 Caen Cedex, France, 2005. 9, 15, 16, 17, 39, 40, 47

Tschumperle, D. and Deriche, R. . Vector-valued image regularization with PDEs: a common framework for different applications. *IEEE Transactions on PAMI*, 27(4):506–517, 2005. 9, 15, 39, 40

Uras, S., Girosi, F., and others. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988. 5

Van Huffel, S. and Vandewalle, J.. *The Total Least Squares Problem: Computational Aspects and Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, 1991. 6

Verri, A. and Poggio, T.. Against Quantitative Optical Flow. In *Proc. of 1st International Conference on Computer Vision*, pages 171–180, London, 1987. 3

Verri, A. and Poggio, T.. Motion field and optical flow: Qualitative Properties. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(5):490–498, 1989. 3

Waxman, A. M. and Wohn, K.. Contour evolution, neighborhood deformation, and global image flow: planar surfaces in motion. *International Journal of Robotics Research*, 4(3): 95–108, 1985. 6, 7

Waxman, A. M., Wu, J., and Bergholm, F.. Convected activation profiles and receptive fields for real time measurement of short range visual motion. In *Proc. Conf. Comput. Vis. Patt. Recog*, pages 771–723, Ann Arbor, 1988. 6

Weber, J. and Malik, J.. Robust Computation of Optical Flow in a Multi-Scale Differential Framework. *International Journal of Computer Vision*, 14(1):67–81, 1995. 5

J. Weickert. *Anisotropic Diffusion in Image Processing*. Dissertatio, Faculty of Mathematics, University of Kaiserslautern, 1996. 41

J. Weickert . On Discontinuity-Preserving Optic Flow. In Orphanoudakis, S., Trahanias, P., and Crowley, J.and Katevas, N., editors, *Proc. Computer Vision and Mobile Robotics Workshop*, pages 115–122, September 1998. 8

Weickert, J. and Schnörr, C. . Räumlich-zeitliche Berechnung des optischen Flusses mit nichtlinearen flußabhängigen Glattheitstermen. In *DAGM*, pages 317–324. Springer, 1999. 8

Weickert, J. and Schnörr, C. . A theoretical Framework for Convex Regularizers in PDE-Based Computation of Image Motion. *International Journal of Computer Vision*, 45(3): 245–264, 2001a. 9, 11, 12, 14

Weickert, J. and Schnörr, C. . Variational Optic Flow Computation with a Spatio-Temporal Smoothness Constraint. *Journal of Mathematical Imaging and Vision*, 14:245–255, 2001b. 9

Weickert, J., Romeny, B. M. T. H., and Viergever, M. A. . Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, March 1998. 8

Elke Wilczok . New uncertainty principles for the continuous Gabor transform and the continuous wavelet transform. *Documenta Math.*, 5, 2000. http://www.mathematik.uni-bielefeld.de/documenta/vol-05/08.pdf. 25

R. J. Woodham. Multiple light source optical flow. In *Proc. of ICCV*, pages 42–46, Osaka, Japan, 1990. 5

Wu, J., Brockett, R., and Wohn, K.. A contour-based recovery of image flow: Iterative method. In *Proc. IEEE CVPR*, pages 124–129, San Diego, 1989. 6

Xavier Pennec and Nicholas Ayache . Uniform distribution, distance and expectation problems for geometric features processing. *Journal of Mathematical Imaging and Vision*, 9(1): 49–67, July 1998. doi: 10.1023/A:1008270110193. http://www.springerlink.com/openurl.asp?genre=article&issn=0924-9907&volume=9&issue=1&spage=49. A preliminary version appeared as INRIA Research Report 2820, March 1996. 56, 76

Xavier Pennec, Pierre Fillard, and Nicholas Ayache . A Riemannian Framework for Tensor Computing. *International Journal of Computer Vision*, 66(1):41–66, 2006. http://citeseer.ist.psu.edu/702078.html. 56, 76

E. Zeidler . *Nonlinear Functional Analysis and its Applications*, volume II b. Springer, 1990a. 12

E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume I. Springer, 1990b. 11