

COMPUTATIONAL IMAGING AND VISION

# Gaussian Scale-Space Theory

Jon Sporring, Mads Nielsen, Luc Florack  
and Peter Johansen (Eds.)

## Gaussian Scale-Space Theory

# Computational Imaging and Vision

---

## Managing Editor

**MAX A. VIERGEVER**

*Utrecht University, Utrecht, The Netherlands*

## Editorial Board

**OLIVIER D. FAUGERAS, INRIA, Sophia-Antipolis, France**

**JAN J. KOENDERINK, Utrecht University, Utrecht, The Netherlands**

**STEPHEN M. PIZER, University of North Carolina, Chapel Hill, USA**

**SABURO TSUJI, Osaka University, Osaka, Japan**

**STEVEN W. ZUCKER, McGill University, Montréal, Canada**

# Gaussian Scale-Space Theory

*Edited by*

**Jon Sporring**

*DIKU, Department of Computer Science,  
University of Copenhagen,  
Copenhagen, Denmark*

**Mads Nielsen**

*3D-Laboratory, School of Dentistry,  
University of Copenhagen,  
Copenhagen, Denmark*

**Luc Florack**

*Department of Computer Science,  
Faculty of Mathematics and Computer Science,  
University of Utrecht,  
Utrecht, The Netherlands*

and

**Peter Johansen**

*DIKU, Department of Computer Science,  
University of Copenhagen,  
Copenhagen, Denmark*



SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-4852-3      ISBN 978-94-015-8802-7 (eBook)  
DOI 10.1007/978-94-015-8802-7

---

*Printed on acid-free paper*

**All Rights Reserved**

© 1997 Springer Science+Business Media Dordrecht

Originally published by Kluwer Academic Publishers in 1997

No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the copyright owner.

# Contents

<b>Preface</b>	<b>xi</b>
<b>Contributors</b>	<b>xiii</b>
<b>Scale in Perspective</b>	
Jan J. Koenderink	xv
<b>I Applications</b>	<b>1</b>
<b>1 Applications of Scale-Space Theory</b>	
Bart ter Haar Romeny	3
1.1 Introduction . . . . .	3
1.2 Feature detection . . . . .	4
1.3 Shape Properties . . . . .	6
1.4 Higher Order Invariants . . . . .	8
1.5 Deblurring . . . . .	9
1.6 Ridges and Multimodality Matching . . . . .	11
1.7 Deep Structure and the Hyperstack . . . . .	12
1.8 Denoising and Edge-Preserving Smoothing . . . . .	13
1.9 Discussion . . . . .	14
<b>2 Enhancement of Fingerprint Images using Shape-Adapted Scale-Space Operators</b>	
Andrés Almansa and Tony Lindeberg	21
2.1 Introduction . . . . .	21
2.2 Shape-adapted smoothing . . . . .	23
2.3 Enhancement of ridges by shape adaptation . . . . .	24
2.4 Automatic scale selection . . . . .	26
2.5 Summary and discussion . . . . .	29
<b>3 Optic Flow and Stereo</b>	
Wiro J. Niessen and Robert Maas	31
3.1 Introduction . . . . .	31
3.2 Generalized Brightness Constraint Equation . . . . .	32

3.3	Optic flow . . . . .	34
3.4	Binocular stereo . . . . .	35
3.5	Scale selection . . . . .	37
3.6	Optic flow results . . . . .	39
3.7	Stereo results . . . . .	40
3.8	Summary . . . . .	41
<b>II</b>	<b>The Foundation</b>	<b>43</b>
<b>4</b>	<b>On the History of Gaussian Scale-Space Axiomatics</b>	
	Joachim Weickert, Seiji Ishikawa, and Atsushi Imita	45
4.1	Introduction . . . . .	45
4.2	Iijima's 1-D Axiomatic (1962) . . . . .	47
4.2.1	Motivation . . . . .	47
4.2.2	Axioms . . . . .	47
4.2.3	Consequences . . . . .	48
4.2.4	Further Results . . . . .	50
4.3	Otsu's 2-D Axiomatic (1981) . . . . .	51
4.3.1	Derivation of the Gaussian . . . . .	51
4.3.2	Further Results . . . . .	52
4.4	Relation to Other Work . . . . .	53
4.5	Discussion . . . . .	58
<b>5</b>	<b>Scale-Space and Measurement Duality</b>	
	Luc Florack	61
5.1	Introduction . . . . .	61
5.2	Pseudo-Static Image Concept . . . . .	62
5.3	Kinematic Image Concept . . . . .	66
5.4	Conclusion . . . . .	73
<b>6</b>	<b>On the Axiomatic Foundations of Linear Scale-Space</b>	
	Tony Lindeberg	75
6.1	Introduction . . . . .	75
6.2	Axiomatic formulations of linear scale-space . . . . .	78
6.2.1	Original formulation . . . . .	78
6.2.2	Causality . . . . .	78
6.2.3	Non-creation of local extrema . . . . .	79
6.2.4	Semi-group structure and continuous scale parameter . . . . .	80
6.2.5	Non-enhancement and infinitesimal generator . . . . .	81
6.2.6	Scale invariance . . . . .	83
6.3	Semi-group and causality: continuous domain . . . . .	85
6.3.1	Assumptions . . . . .	86
6.3.2	Definitions . . . . .	87
6.3.3	Necessity . . . . .	89
6.3.4	Sufficiency . . . . .	91
6.3.5	Application to scale invariant semi-groups . . . . .	91

6.4	Summary and conclusions . . . . .	91
6.5	Extensions of linear scale-space . . . . .	92
6.5.1	Relaxing rotational symmetry . . . . .	92
6.5.2	Non-linear smoothing . . . . .	95
<b>7</b>	<b>Scale-Space Generators and Functionals</b>	
Mads Nielsen		<b>99</b>
7.1	Introduction . . . . .	99
7.2	Linear scale-space generators . . . . .	101
7.3	Linear regularization . . . . .	103
7.3.1	Scale-invariant regularization . . . . .	104
7.3.2	Evolution equations . . . . .	106
7.3.3	Scale-space interpretation of regularization . . . . .	107
7.4	Truncated smoothness operators . . . . .	109
7.5	Implementation issues of truncated filters . . . . .	111
7.6	Summary . . . . .	114
<b>8</b>	<b>Invariance Theory</b>	
Alfons Salden		<b>115</b>
8.1	Introduction . . . . .	115
8.2	The Equivalence Problem . . . . .	116
8.3	Intrinsic Symmetries . . . . .	117
8.3.1	Solution of Similarity Equivalence Problem . . . . .	118
8.3.2	Solution of Affine and Euclidean Equivalence Problems . . . . .	118
8.4	Conclusion and Discussion . . . . .	127
<b>9</b>	<b>Stochastic Analysis of Image Acquisition and Scale-Space Smoothing</b>	
Kalle Åström and Anders Heyden		<b>129</b>
9.1	Introduction . . . . .	129
9.2	Image acquisition . . . . .	130
9.3	Interpolation and smoothing . . . . .	131
9.4	The random field model . . . . .	133
9.5	Applications . . . . .	135
9.6	Conclusions . . . . .	136
<b>III</b>	<b>The Structure</b>	<b>137</b>
<b>10</b>	<b>Local Analysis of Image Scale Space</b>	
Peter Johansen		<b>139</b>
10.1	Introduction . . . . .	139
10.1.1	Codimension . . . . .	141
10.2	Analysis of codimension one . . . . .	143
10.3	Analysis of codimension two . . . . .	144
10.4	Suggestions for Further Research . . . . .	146
10.5	Acknowledgement . . . . .	146

<b>11 Local Morse Theory for Gaussian Blurred Functions</b>	
James Damon	<b>147</b>
11.1 Classifying Properties of Blurred Images via Equivalences . . . . .	148
11.2 Understanding Generic Properties . . . . .	150
11.3 Genericity via Stability . . . . .	152
11.4 Weighted Homogeneity and the Heat Equation . . . . .	153
11.5 Generic Properties of Blurred Images . . . . .	157
11.6 Consequences of the Generic Properties . . . . .	158
11.7 Genericity via Transversality and Versality . . . . .	160
11.8 Conclusion . . . . .	162
<b>12 Critical Point Events in Affine Scale-Space</b>	
Lewis Griffin	<b>165</b>
12.1 Motivation for Affine Scale Space . . . . .	165
12.2 A Representation of Affine Scale Space . . . . .	166
12.3 Local Structure . . . . .	168
12.4 Mathematical Technique . . . . .	170
12.5 The Catastrophes . . . . .	172
12.5.1 Fold Catastrophe: $H(u, v; p, q, r) = pu + u^3 + v^2$ . . . . .	172
12.5.2 Cusp Catastrophe: $H(u, v; p, q, r) = pu + qu^2 + u^4 \pm v^2$ . .	174
12.5.3 Swallowtail: $H(u, v; p, q, r) = pu + qu^2 + ru^3 + u^5 + v^2$ . .	177
12.5.4 Elliptic Umbilic: $H(u, v; p, q, r) = pu + qv + r(u^2 + v^2) + u^3 - 3uv^2$ . . . . .	177
12.5.5 Hyperbolic Umbilic: $H(u, v; p, q, r) = pu + qv + r(u^2 - v^2) + u^3 + 3uv^2$ . . . . .	179
<b>13 Topological Numbers and Singularities</b>	
Stiliyan Kalitzin	<b>181</b>
13.1 Introduction . . . . .	181
13.2 Theory. Topological invariants . . . . .	183
13.3 Examples. . . . .	185
13.3.1 Second order singular points . . . . .	185
13.3.2 One dimensional case . . . . .	186
13.3.3 Two dimensional images . . . . .	186
13.4 Scale space evolution. . . . .	187
13.4.1 One dimensional fold . . . . .	187
13.4.2 Two dimensional umbilic point . . . . .	188
13.5 Summary and discussion. . . . .	189
<b>14 Multi-Scale Watershed Segmentation</b>	
Ole Fogh Olsen	<b>191</b>
14.1 Introduction . . . . .	191
14.2 Morse functions and catastrophe theory . . . . .	192
14.3 Watersheds and Catchment basins . . . . .	192
14.4 Catastrophe theory applied . . . . .	195
14.5 The segmentation tool . . . . .	197

14.6 Conclusions . . . . .	200
<b>IV Non-linear Extensions</b>	<b>201</b>
<b>15 The Morphological Equivalent of Gaussian Scale-Space</b>	
Rein van den Boomgaard and Leo Dorst	<b>203</b>
15.1 Introduction . . . . .	203
15.2 Morphological Operators . . . . .	205
15.3 Geometrical Interpretation of Morphological Operators . . . . .	208
15.4 Tangential Dilation and the Slope Transform . . . . .	209
15.5 Scale-Space from Basic Principles . . . . .	212
15.6 The Infinitesimal Generator of Scale-Space . . . . .	216
15.7 Conclusions . . . . .	219
<b>16 Nonlinear Diffusion Scale-Spaces</b>	
Joachim Weickert	<b>221</b>
16.1 Introduction . . . . .	221
16.2 The Perona-Malik model . . . . .	222
16.2.1 Basic idea . . . . .	222
16.2.2 Ill-posedness . . . . .	223
16.3 Regularized nonlinear diffusion filtering . . . . .	225
16.3.1 Basic idea . . . . .	225
16.3.2 General well-posedness and scale-space framework . . . . .	225
16.4 Semidiscrete nonlinear diffusion scale-spaces . . . . .	228
16.4.1 An example . . . . .	228
16.4.2 General well-posedness and scale-space framework . . . . .	230
16.4.3 Application to the example . . . . .	231
16.5 Fully discrete nonlinear diffusion scale-spaces . . . . .	231
16.5.1 An example . . . . .	232
16.5.2 General well-posedness and scale-space framework . . . . .	232
16.5.3 Application to the example . . . . .	233
16.5.4 A more efficient discretization . . . . .	234
16.6 Generalizations . . . . .	234
<b>Bibliography</b> . . . . .	<b>235</b>
<b>Index</b> . . . . .	<b>258</b>

# Preface

Gaussian scale-space is one of the best understood multi-resolution techniques available to the computer vision and image analysis community. It is the purpose of this book to guide the reader through some of its main aspects.

During an intensive weekend in May 1996 a workshop on Gaussian scale-space theory was held in Copenhagen, which was attended by many of the leading experts in the field. The bulk of this book originates from this workshop.

Presently there exist only two books on the subject. In contrast to Lindeberg's monograph (Lindeberg, 1994e) this book collects contributions from several scale-space researchers, whereas it complements the book edited by ter Haar Romeny (Haar Romeny, 1994) on non-linear techniques by focusing on linear diffusion.

This book is divided into four parts. The reader not so familiar with scale-space will find it instructive to first consider some potential applications described in Part I. Parts II and III both address fundamental aspects of scale-space. Whereas scale is treated as an essentially arbitrary constant in the former, the latter emphasizes the *deep structure*, i.e. the structure that is revealed by varying scale. Finally, Part IV is devoted to non-linear extensions, notably non-linear diffusion techniques and morphological scale-spaces, and their relation to the linear case.

The Danish National Science Research Council is gratefully acknowledged for providing financial support for the workshop under grant no. 9502164.

Jon Sporring, Mads Nielsen, Luc Florack, and Peter Johansen (eds.)

# Contributors

## **Andrés Almansa**

*Centro de Cálculo, Facultad de Ingeniería, Universidad de la República  
Julio Herrera y Reissig 565, piso 5. 11300 Montevideo, Uruguay*

## **Rein van den Boomgaard**

*Faculty of Mathematics & Computer Science, University of Amsterdam  
Kruislaan 403, NL-1098 SJ Amsterdam, The Netherlands*

## **James Damon**

*Department of Mathematics, University of North Carolina  
Chapel Hill, NC 27599, USA*

## **Leo Dorst**

*Faculty of Mathematics & Computer Science, University of Amsterdam  
Kruislaan 403, NL-1098 SJ Amsterdam, The Netherlands*

## **Luc Florack**

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

## **Lewis Griffin**

*INRIA-SOPHIA  
2004, Route des Lucioles, B.P. 93, Sophia Antipolis Cedex 06902, France*

## **Anders Heyden**

*Department of Mathematics, Lund University  
P. O. Box 118, S-221 00 Lund, Sweden*

## **Atsushi Imiya**

*Department of Information and Computer Sciences,  
Faculty of Engineering, Chiba University,  
1-33 Yayoi-Cho, Inage-Ku 263, Chiba, Japan*

## **Seiji Ishikawa**

*Department of Control Engineering, Kyushu Institute of Technology,  
Sensuicho 1-1, Tobata, Kitakyushu 804, Japan*

## **Peter Johansen**

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

**Stiliyan Kalitzin**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Jan J. Koenderink**

*Department of Physics and Astronomy, Utrecht University  
Buys Ballot Laboratory, Princetonplein 5, P.O. Box 80 000  
NL-3508 TA Utrecht, The Netherlands*

**Tony Lindeberg**

*Computational Vision and Active Perception Laboratory (CVAP)  
Department of Numerical Analysis and Computing Science (NADA)  
KTH (Royal Institute of Technology)  
S-100 44 Stockholm, Sweden*

**Robert Maas**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Mads Nielsen**

*3D-Lab, School of Dentistry, University of Copenhagen  
Nørre allé 25, 2200 Copenhagen N, Denmark*

**Wiro J. Niessen**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Ole Fogh Olsen**

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

**Bart ter Haar Romeny**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Alfons Salden**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Jon Sporring**

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

**Joachim Weickert**

*RWCP Novel Function SNN Laboratory,  
Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Kalle Åström**

*Department of Mathematics, Lund University  
P. O. Box 118, S-221 00 Lund, Sweden*

# Scale in Perspective

**Jan J. Koenderink**

*Department of Physics and Astronomy, Utrecht University  
Buys Ballot Laboratory, Princetonplein 5, P.O. Box 80 000  
NL-3508 TA Utrecht, The Netherlands*

Seen from a distance objects gain a certain charm because a *mystery* creeps in. According to the 16<sup>th</sup>c. Chinese(Sze, 1959) when discussing painting, “*distant faces have no eyes*” (though even without eyes they must seem to *look!*). When distance is increased *details are lost*. The image suffers a topological change.

Similar observations abound in other cultures. Lucretius(Lucretius, 1951) in his “*On the nature of the universe*” (perhaps the closest thing to a scientific treatise at the time) describes not only how things *vanish* at a distance, but also how they appear to *change*. For instance, distant square towers look rounded to him. This effect is generally known as “*blurring*”. He also observes how a pair of distant islands appears to *merge* into a single one. This is yet another type of topological change. When distance is increased *details generalize* (become rounded) and distinct details may *merge* or *vanish*.

This change of details at a distance leads to an experiential “*unfolding*” as things draw nearer and one perceives layer after layer of emerging structure. (Which has been called “*deblurring*”.) This has been described often in western literature and poetry. Here is a famous instance: In Calderón’s *The Constant Prince* a large fleet appears at the horizon. This vision is described very perceptively(Calderón, 1853):

*For, as on the coloured canvas  
 Subtle pencils softly blend  
 Dark and light in such proportions  
 That the dim perspectives end—  
 Now perhaps like famous cities,  
 Now like caves or misty capes,  
 For remoteness ever formeth  
 Monstruous or unreal shapes . . .  
 So it was, while I alone,  
 Saw their bulk and vast proportions  
 But their form remained unknown.  
 First they seemed to us uplifting  
 High in heaven their pointed towers,  
 Clouds that in the sea descended,  
 To conceive in sapphire showers  
 What they would bring forth in crystal.  
 And this fancy seemed more true,  
 As from their untold abundance*

*They, methought, could drink the blue  
 Drop by drop. Again sea monsters  
 Seemed to us the wandering droves,  
 Which, to form the train of Neptune,  
 Issued from their green alcoves.  
 For the sails, when lightly shaken,  
 Fanned by zephyrs as by slaves,  
 Seemed to us like outspread pinions  
 Fluttering o'er the darkened waves;  
 Then the mass, approaching nearer,  
 Seemed a mighty Babylon,  
 With its hanging gardens pictured  
 By the streamers fluttering down.  
 But at last our certain vision  
 Undeceived, becoming true,  
 Showed it was a great armada  
 For I saw the prows cut through  
 Foam . . .*

Notice the key observations “... *For remoteness ever formeth Monstruous or unreal shapes . . .*” and “... *But their form remained unknown . . .*”. Here the visual field of the observer appears as an infinite nested set of Chinese boxes whose contents are progressively revealed as the observer draws nearer to the scene. In this example the author rather arbitrarily limits the nesting on one side: “... *But at last our certain vision Undeceived, becoming true, . . .*”. The character simply stops the analysis at the point where vision should give rise to action. This makes perfect sense from a pragmatic viewpoint but appears rather *ad hoc* from a broader perspective.

Perhaps the most incisive and complete phenomenology of these very basic facts of sight has been presented by John Ruskin (Ruskin, 1857; Ruskin, 1860). He calls the phenomena *mystery* and insists that *never do we see anything clearly*: The mystery is not removed but only shifted to another stratum as you draw nearer to an object. For Ruskin *there is no* “certain vision” and no ultimate “truth” at all. To “see clearly” means simply:

*What we call seeing a thing clearly, is only seeing enough of it to make out what it is; this point of intelligibility varying in distance for different magnitudes and kind of things, while the appointed quantity of mystery remains nearly the same for all.*

This standpoint, which indeed captures the human condition quite well, is certainly shocking once understood and clearly at odds with our everyday common sense notions. Yet Ruskin firmly insists, for:

*We suppose we see the ground under our feet clearly, but if we try to number its grains of dust, we shall find that it is as full of confusion and*

*doubtful forms as anything else; so that there is literally no point of clear sight, and there never can be, ...*

*Every object, however near the eye, has something about it which you cannot see, and which brings the mystery of distance even into every part and portion of what we suppose ourselves to see most distinctly.*

To Ruskin we also owe the most complete descriptions of mystery, both in writing and as short sequences of sketches. Here is an example:

*Go to the top of Highgate Hill on a clear summer morning at five o'clock, and look at Westminster Abbey. You will receive an impression of a building enriched with multitudinous vertical lines. Try to distinguish one of these lines all the way down from the one next to it: You cannot. Try to count them: You cannot. Try to make out the beginning or end of any of them: You cannot. Look at it generally, and it is all symmetry and arrangement. Look at it in its parts, and it is all inextricable confusion.*

Yet Ruskin adamantly insists that the draughtsman should render this confusion veridically! If you really try Ruskin's chores you will find yourself at a loss: You are bound to be completely baffled by the "mystery" in the commonest view.

Ruskin also observed *topological changes* that to the best of my knowledge no one ever described before him. As noted earlier Lucretius already commented on the common observation that distinct details *merge* when moved away from the observer. John Ruskin remarks that the *opposite* also occurs: When you observe a leaf that is attached to a branch through a thin stem and increase your distance, you will observe how the stem becomes invisible before either the leaf or the branch. Thus at a certain distance the leaf breaks loose from the branch and hovers in the air: *Connected details may split when distance is increased!* Many people find this counterintuitive, yet Ruskin's example is easily verified in daily life if one pays a little attention.

Such observations clearly indicate that one is faced with a fundamental and important, though unfortunately ill understood, aspect of perception (or perhaps ecological optics) here. What came as a bit of a shock to me when I took a first look at the subject was that there existed essentially *no science* on the topic. The only discipline that cared about such phenomena turned out to be *cartography* (*Greenwood, 1964*). Indeed, a map at the scale of the earth is virtually useless, maps are useful exactly *because they suppress detail*. Suppressing detail is easily the most important aspect of the mapmaker's work and is known as "cartographic generalization". In real life, like in cartography summarizing or generalization is a *necessity* since you can't store or process all structure available in actuality.

Cartographic generalization is conducted largely on an intuitive basis. Oddly enough, although there is certainly a lot of science in cartography, its arguably most important aspect has always remained an *art*. Nevertheless (or perhaps because of that), one may glean some important insights from maps. First of all, each detail on a map must have a *cause* in the real world. Very often you will be able to find the cause also on a *more detailed* map, but not likely on a *coarser* map: on the contrary, the detail *itself* will be a cause, a cause for the structure

on the coarser map. Thus there is a notion of “causality” connected with atlases, a cause at some scale being found at a finer scale, structure at some scale being cause for detail on a coarser scale. This “causal direction” from fine to coarse is not unlike temporal causality in the real world, a one-way functional dependence.

Another useful hint from cartography is that it may be advantageous to look at coarser scales for structures that can't be seen at finer scales. Indeed, putting a microscope on an object may blind us to its essential structure. One reason is that increasing resolution (paging to finer maps in the atlas) typically decreases the “scope”. (And if the method doesn't do it your visual system itself will act as the bottleneck with effectively the same result.) In atlases this is clear enough: If the pagesize is fixed increased resolution means that the map *covers less terrain*. Such an effect is almost universal. Typically the scope covers a more or less fixed number of “grains”. This is definitely the case in human vision where people often complain that they “can't see the forest for the trees”. In the history of art the *locus classicus* is Roger de Piles's treatise(Piles, 1708), with his proverbial “bunch of grapes”. De Piles argues that the real painter paints the bunch whereas the mere tyro will attempt to paint the grapes. A visit to any gallery of still lifes will convince you of the sad truth of this observation. Modern painters agree with de Piles point, e.g., Jacobs(Jacobs, 1988).

*Students take a ferocious pride in showing how many details they can find. Ironically, the most difficult problem in drawing and painting is to see and render the subject in its totality.*

Likewise, in cartography the alps appear on the map of western Europe, but they can't be seen on a map of Switzerland say.

“Going coarse to fine”, or deblurring, is one way of making sure the details automatically arise in context. The method has gained some popularity in image interpretation and indeed offers a rare opportunity to substitute local for global analysis. This might—in principle—allow global shape description on the basis of local differential geometry, a possibility that has not yet been exploited to its full potential. The idea is certainly not new: Here is a (evidently tongue in cheek) quotation by Oliver Wendell Holmes of 1890:

*It's just the same thing as my plan . . . for teaching drawing nothing more. Anybody . . . can make a dot . . . Lesson No.1. Make a dot; that is, draw your man, a mile off . . . Now make him come a little nearer . . . the dot is an oblong figure now. Good. Let your scholar draw an oblong figure. It is as easy as making a note of admiration. So by degree the man who serves as a model approaches. A bright pupil will learn to get the outline of a human figure in ten lessons, the model coming five hundred feet nearer each time.*

A field where scale notions abound is of course physics and the many disciplines closely connected to it, such as for instance the earth sciences or astronomy. The ideas have been around for centuries and one could quote Francis Bacon, Leibnitz and many others. Instead of dwelling on history I recommend the very nice popular book (based on a great though short movie) “Powers of ten(Morrison and

*Morrison, 1982)*" . In physics one often *averages* over configurations of particles or subsystems, it is the major topic of statistical mechanics. Some entities such as pressure or density do indeed only *exist* at a coarse level and lose their significance at finer levels. The granularity of matter implies that many quantities of great importance in daily life exist only on a coarse enough scale, *e.g.*, retinal illuminance at low luminance levels has no meaning whatsoever for the single photoreceptor: At the visual threshold a single receptor may catch perhaps one photon every half hour! On the other hand, due to the haphazard constitution of our daily environments many quantities vary capriciously on too coarse a scale. (For instance, consider the temperature distribution over a kitchen.) Such quantities exist only in a certain limited *range* of scales. One doesn't need much modern physics to appreciate this, and indeed, most of the relevant facts can be found in early work such as Boscovitz(Boscovitz, 1763)'s.

Let us consider a very simple example, taken from meteorology: What is a cloud? This is an example based on a simple and common observation, taken from Webster's(Webster, 1955) book on partial differential equations:

*For example, if we look at a cumulus cloud from a distance, it appears to have a sharp boundary (the density of the water vapor discontinuous) but as we approach it we find it impossible to say where the boundary is, and it seems to shade off gradually, while again if we examine it with a microscope we find it is composed of minute drops, each of which we suppose to be composed of millions of atoms, each composed of many smaller parts. We have here an extreme example of the possible variety of points of view.*

Webster dismisses the point perhaps too easily. Indeed, this is not an "extreme example" of the possible variety of points of view at all, but rather *this is the generic condition in the natural sciences!* It is remarkable how Webster—who like many others clearly perceived the problem—never attempts to frame a solution. He is satisfied to state: "*We shall here adopt the hypothesis of general continuity of properties, ...* ". Apparently he is convinced that the remainder of the book will nevertheless be of use (the book is addressed to the engineer). Despite the deep cleft between the realm of smooth functions and *observed scalar fields* scientists have not been hesitant to apply the methods of differential calculus to actual data. That this actually works fine is to be considered as a tribute to common sense. However, a more principled approach is clearly desirable. So what *is* a cloud? You might be tempted to point at something white in the sky by way of an answer, but here I mean a definition in terms of the distribution of matter. A cloud is a region of the atmosphere containing condensed  $H_2O$ . If you measure  $H_2O$  density in too small a volume you find either zero (when you're in between drops) or very high (when you're inside a drop) values. If you measure in very large volumes (say  $1 \text{ km}^3$ ) you get rid of these troublesome variations, but you lose so much resolving power that you also lose the cloud. About the right scale here is  $1 \text{ m}^3$ , clouds may be defined(Mason, 1962) as atmospheric regions where the density of condensed  $H_2O$  exceeds about  $0.4 \text{ g m}^{-3}$  *at a level of resolution of about*  $1 \text{ m}^3$ . The lesson is clear: *Density* is not simply a scalar field, but a one-parameter family of scalar

fields, *scale* (of the sampling volume) being the parameter.

One feels that the understanding of the structure of such scalar fields has much progressed throughout the last decade. We now have a “field” that is now generally known as “*scale-space*” theory. Yet—given the fact that the key ideas have been around for centuries and essentially everything important was around by the end of the 19<sup>th</sup>c.—do we have anything more here than another reinvention of the wheel? I have only my personal perspective to offer. I started to be interested in the seventies and my first paper(Koenderink, 1977) on the topic (probably) dates from 1976. At that time I proposed multiscale “pyramids” as likely models for the human visual system (see also Koenderink and van Doorn(Koenderink, 1978)). Much later I started formal developments and at about the same time others (Witkin(Witkin, 1983)) developed similar interests. I remember lecturing on the diffusion equation on the first David Marr memorial workshop thrown by Whitman Richards at Cold Spring Harbor where I also met Andy Witkin and learned of the interest in the topic by computer scientists (I was approaching the subject from the angles of psychophysics and physiology). It must have been 1981 (or '82?). What was new? In retrospect probably only the formal definitions(Koenderink, 1984) of ideas that were commonplace in several sciences and with which I was well familiar. Even so, these developments turned out to be quite useful in many applications so even if the basic ideas weren't very original, at least something good came out of it. Nowadays the field is certain to remain one of the corner stones of image processing technology.

The present book contains a collection of contributions on scale-space theory that may serve as a fairly comprehensive introduction into the field for the newcomer and as a wellcome review of modern results for the professional.

# **Part I**

# **Applications**

# Chapter 1

# Applications of Scale-Space Theory

Bart ter Haar Romeny

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

## 1.1 Introduction

The attractive aspect of (linear and nonlinear) scale-space theory is its solid mathematical framework. For this reason quite some emphasis has been placed on the development of the theoretical foundations. The core of the paradigm is the notion of sampling, a physical observation through a multi-size aperture, and where the aperture size is the variable parameter 'scale'.

Applications of scale-space theory can be found in any field of analysis of sampled data, in the spatial, temporal, or whatever domain. In this chapter we will present a (necessarily much restricted) anthology of applications taken from the field of computer vision and image analysis. The examples are taken such that a variety of approaches is covered, and the reader is referred to the references for further examples and detail. The applications are taken from the following areas:

- Feature detection
- Higher order invariants
- Ridges and multimodality matching
- Segmentation by hyperstack linkages
- Computational perception theory
- Shape properties
- Deblurring
- Stellate tumor detection
- Edge preserving smoothing

## 1.2 Feature detection

Scale-space theory gives us the solution for the ill-posedness problem of differentiation of discrete data: convolution with a scaled Gaussian derivative kernel does two steps in one: observation and differentiation, as we can see from the commutation relation

$$\frac{\partial L}{\partial x} * G = L * \frac{\partial G}{\partial x}$$

Differentiation is now done by integration. The full machinery of differential geometry for  $n - D$  manifolds (see e.g. the practical book (Gray, 1993) with many Mathematica examples) is now available.

This implies that the image and its scale-space can be considered continuous after the measurement, i.e. can be differentiated to any order in any direction. High order differential structure can now be extracted robustly. Most sensible features are those independent of a particular choice of coordinates, i.e. invariant under a particular group of transformations. There is a vast number of applications of differential features published (Florack et al., 1993a; Lindeberg, 1994e; Haar Romeny et al., 1994c; Salden et al., 1992b; Fidrich and Thirion, 1995; Thirion and Gourdon, 1995; Monga et al., 1992; Danielsson and Seger, 1990; Mundy and Zisserman, 1992). We will present affine corner detection as an example. First, modern notation for differential invariants is explained in intermezzo 1.

## Intermezzo 1: Modern Notation of Geometric Structure

There are several useful notations for differential invariant properties (or features) (Haar Romeny et al., 1994c; Florack et al., 1994c):

- Cartesian notation
- Manifest invariant, tensorial or index notation
- Gauge coordinate notation

**Cartesian notation** of differential invariant features: This is the representation mostly used for physical implementation and computation. Some examples of invariant expressions in the 2D case are given in table 1.1.

**Manifest invariant notation**, well known in the physical literature: here partial derivatives are considered tensors: indices are used for the (spatial) coordinate to which the differentiation is taken: e.g.  $L_i$  denotes the vector  $(L_x, L_y, L_z)$  in 3D. Polynomial invariants are formed by full contraction (i.e. summing over the total number of dimensions: the ‘Einstein convention’) or alternation of pairs of indices, in this way precisely showing its coordinate independence. There are two constant tensors expressing the symmetric (Kronecker tensor:  $\delta_{ij}$ ) and antisymmetric (Epsilon or Levi-Civita tensor:  $\epsilon_{ij\dots}$ ) connections. The  $\epsilon$ -tensor is familiar from outer products (e.g.  $L_i \epsilon_{ij} L_j = 0$ ) and the calculation of determinants, e.g. the determinant of the Hessian matrix  $L_{ij}$  in 2D:  $\det H = \frac{1}{2} \epsilon_{ij} \epsilon_{kl} L_{ik} L_{jl} = \frac{1}{2} (L_{ii} L_{jj} - L_{ij} L_{ji})$ , where we used that the product of two  $\epsilon$ -tensors can be expressed in  $\delta$ -tensors:  $\epsilon_{ij} \epsilon_{kl} = \delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}$ . We also have:  $\epsilon_{ij} L_{ij} = 0$ .

Name	Cartesian	Manifest	Gauge
Intensity	$L$	$L$	$L$
Gradient magn.	$\sqrt{L_x^2 + L_y^2}$	$\sqrt{L_i L_i}$	$L_w$
Laplacian	$L_{xx} + L_{yy}$	$L_{ii}$	$L_{vv} + L_{ww}$
Isophote curv.	$\frac{2L_x L_y L_{xy} - L_x^2 L_{yy} - L_y^2 L_{xx}}{(L_x^2 + L_y^2)^{3/2}}$	$\frac{L_i L_j L_{ij} - L_i L_i L_{jj}}{(L_k L_k)^{3/2}}$	$-\frac{L_{vv}}{L_w}$
Flowline curv.	$\frac{L_x L_y (L_{yy} - L_{xx}) + L_{xy} (L_x^2 - L_y^2)}{(L_x^2 + L_y^2)^{3/2}}$	$\frac{L_i \epsilon_{ij} L_j}{(L_k L_k)^{3/2}}$	$-\frac{L_{vw}}{L_w}$

Table 1.1: Some examples of 2D invariants for orthogonal transformations, expressed in Cartesian, manifest invariant and gauge coordinates.

**Gauge coordinates** are local coordinates which are ‘natural’ in some sense: because they ‘fix the gauge’, they employ the extra degree of freedom, e.g. for the group of orthogonal transformations a rotation: in 2D we define  $(v, w)$  such that  $v$  is everywhere tangential to the isophote and  $w$  perpendicular to it:

$$\partial_v = \frac{L_i \epsilon_{ij} \partial_j}{\sqrt{L_k L_k}} \quad \partial_w = \frac{L_i \delta_{ij} \partial_j}{\sqrt{L_k L_k}}$$

or explicit in 2D:

$$\begin{pmatrix} v \\ w \end{pmatrix} = \frac{1}{\sqrt{L_x^2 + L_y^2}} \begin{pmatrix} L_y & -L_x \\ L_x & L_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Because  $L_v \equiv -L_i \epsilon_{ij} L_j \equiv 0$  many expressions get a much simpler form (see table 1.1).

The way gauge coordinates are constructed automatically implies that all polynomial expressions in  $(v, w)$  are invariant under orthogonal transformations.

The higher order ( $\geq 2$ ) gauge invariants suffer from being singular in points where the gradient vanishes, i.e. singular points. This problem can be solved by multiplication with an appropriate power of the gradient. In many computer vision applications only the numerator or even just its sign is important. The relation between the non-singular invariants in gauge, Cartesian and manifest invariant representation is given to third order in table 1.2.

Corners in 2D images are loosely defined as locations where the curvature of isophotes (curves of equal brightness) is relatively high. Natural coordinates on 2D isophotes are the local tangent ( $v$ ) and gradient ( $w$ ), the socalled gauge coordinates. In gauge coordinates (see the intermezzo) the isophote curvature  $\kappa$  is given as

$$\kappa = -\frac{L_{vv}}{L_w}$$

where lower coordinate indices denote differentiation. In singular points (defined as those points where the gradient  $L_w \equiv 0$ ), the curvature is not defined. Blom

Gauge	Cartesian	Manifest invariant
$L_v L_w$	$\equiv 0$	$-\epsilon_{ij} L_{ij} \equiv 0$
$L_w L_w$	$(L_x^2 + L_y^2)$	$L_i L_i$
$L_{vv} L_w^2$	$\ \nabla L\ ^{-2} (L_{xx} L_y^2 - 2L_{xy} L_x L_y + L_{yy} L_x^2)$	$L_{ii} L_j L_j - L_{ij} L_i L_j$
$L_{vw} L_w^2$	$\ \nabla L\ ^{-2} (L_{xx} L_x L_y + L_{xy} L_y^2 - L_{xy} L_x^2 - L_{yy} L_x L_y)$	$-\epsilon_{ij} L_{jk} L_i L_k$
$L_{ww} L_w^2$	$\ \nabla L\ ^{-2} (L_{xx} L_x^2 + 2L_{xy} L_x L_y + L_{yy} L_y^2)$	$L_i L_{ij} L_j$
$L_{vvv} L_w^3$	$\ \nabla L\ ^{-3} (L_{xxx} L_y^3 - 3L_{xxy} L_x L_y^2 + 3L_{xyy} L_x^2 L_y - L_{yyy} L_x^3)$	$\epsilon_{ij} (L_{jkl} L_i L_k L_l - L_{jkk} L_i L_l L_l)$
$L_{www} L_w^3$	$\ \nabla L\ ^{-3} (L_{xxx} L_x^3 + 3L_{xxy} L_x^2 L_y + 3L_{xyy} L_x L_y^2 + L_{yyy} L_y^3)$	$L_{ijk} L_i L_j L_k$

Table 1.2: Gauge, Cartesian and manifest invariant notation for some expressions of gauge invariants up to third order.

(Blom, 1992) (see also table 1.1) showed that the following expression is invariant under affine transformations:

$$L_{vv} L_w^2 = L_{xx} L_y^2 - 2L_{xy} L_x L_y + L_{yy} L_x^2$$

From  $L_{vv} L_w^2 = \kappa L_w^3$  we see that this invariant has high values at places where both the isophote curvature  $\kappa$  and the gradient  $L_w$  are large. Due to the affine (shear) invariance, the detection is independent of the angle of the corner.

Figure 1.1 shows the result for a scene of the city of Utrecht. This invariant property can also be derived from an algebraic approach to invariant properties, as we will see in section 1.4. Lindeberg added a ‘focussing’ strategy to the corner detection (Lindeberg, 1994e) to get improved localization.

The scaled differential invariant features are well suited for use in geometric reasoning. Their size can be tuned to the task at hand (scale selection) and they form a complete set of structural descriptors to choose from. An instructive example is presented by Karssemeijer (Karssemeijer, 1995), where stellate tumors are detected in mammograms. A stellate tumor is characterized by a radial pattern of lines around an centerpoint. Integrating the votes over a search area around each pixel of investigation whether there is a line present having an orientation towards this point, a very sensitive detection proved possible. The detection of the orientation of a line is done by a ‘steerable’ second order derivative filter.

### 1.3 Shape Properties

Often the local shape of 3D objects can be expressed by principal curvatures of the surface points (Koenderink, 1990). The mean and Gaussian curvatures are the respective sum and product of these principal curvatures. Table 1.3 list the

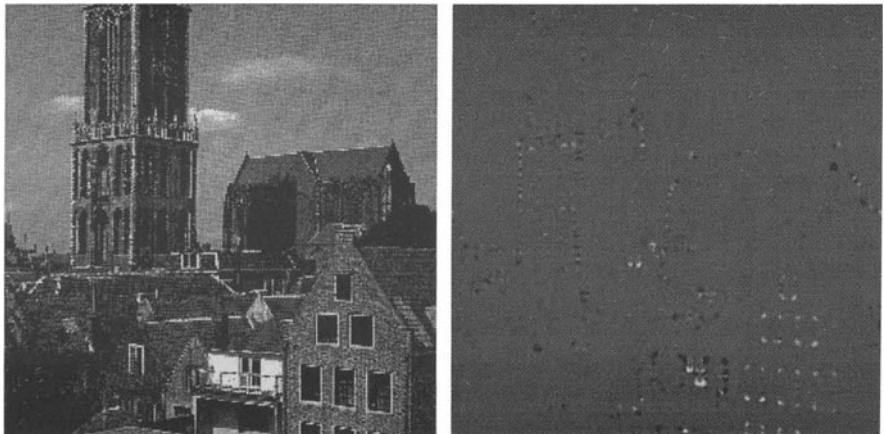


Figure 1.1: Affine corner detector  $L_{vv}L_w^2$ . Image resolution  $512^2$ , scale operator 2.44 pixels. From (Haar Romeny and Binkhuysen, 1992).

expressions (Salden et al., 1994; Salden et al., 1992a) for their calculation for iso-intensity surfaces of a 3D dataset (such as 3D Magnetic Resonance Imaging (MRI) or Computed Tomography (CT) data).

An example of an application is in oral pathophysiology, where for some investigations the shape of teeth surfaces is required to study patients with a chewing deficiency. The change in morphology of the teeth surface over time or as a result of treatment is an important parameter (Kobayashi et al., 1993). The parameters listed in table 1.3 enable a quantitative analysis. Gypsum casts were scanned mechanically with a diamond needle, with a spatial resolution in three directions of 0.1 mm. In figure 1.2 a 3D rendering and the display of the Gaussian curvature on the teeth surface is displayed.

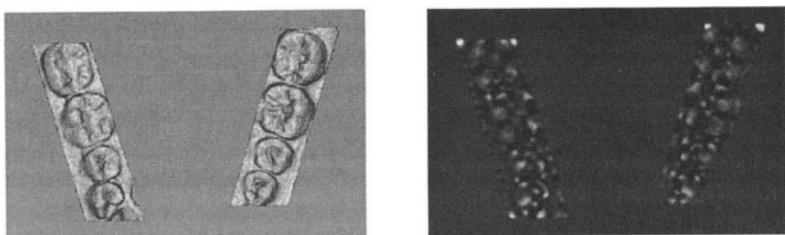


Figure 1.2: Left: 3D surface rendering of mechanically scanned gypsum cast of a patient with chewing disorders. Right: Gaussian curvature, scale operator 3 pixels. Note the hyperbolic (dark) and elliptic (bright) areas. Dataset  $512 \times 256 \times 64$  voxels. Voxelsize 0.1 mm each side. Dataset made available by B. van der Glas, Utrecht University.

Name	Manifest expression
K = Gaussian Curvature 3D	$\frac{\varepsilon_{ijk}\varepsilon_{pqr}L_iL_{jp}L_rL_{kq}}{2(L_sL_s)^2}$
H = Mean Curvature 3D	$\frac{L_iL_iL_{jj} - L_iL_{ij}L_j}{2\sqrt{L_kL_kL_lL_lL_mL_m}}$
Largest Principal Curvature	$H + \sqrt{H^2 - K}$
Smallest Principal Curvature	$H - \sqrt{H^2 - K}$

Table 1.3: Expressions for local 3D curvature shape descriptors in manifest invariant notation.

## 1.4 Higher Order Invariants

An essential observation made by Hilbert (Hilbert, 1893) states that any invariant can be represented as a polynomial function on a complete and finite set of *irreducible invariants*. This implies that these form the most fundamental and concise set of local properties (up to some spatial order) or 'primitives' that may describe all local intrinsic properties of a scalar image at a fixed level of resolution. This set is well known for second order properties (Florack et al., 1994a), and is given in manifest notation for 2D as follows:

$$L, L_iL_i, L_{ii}, L_iL_{ij}L_j, L_{ij}L_{ji}$$

For higher order ( $\geq 3$ ) there is no simple method to derive the complete irreducible set. Salden (Salden et al., 1992b) presented a method that makes use of the *algebraic* structure of the local geometric description, i.e. the binary forms in the local Taylor expansion up to the chosen order. E.g. the fourth order binary form is given by  $f(\mathbf{x}) = \frac{1}{4} L_{ijkl} x_i x_j x_k x_l$ . The main proposition of algebra states that a function is fully described by its roots. Several theorems are available relating the *measure of coincidence of roots* to geometrical symmetries. There is also a wide algebraic literature on the study of roots, defining the *discriminant*, the *resultant* which is a measure for the coincidence of roots of polynomials of different order, and *transvectants*. In table 1.4 some examples of resultants and discriminants of low order are given. These are all affine invariants. Note that  $R_{12}$  is just the corner detector  $L_{vv}L_w^2$  described in section 1.2, and  $D_2$  is the determinant of the Hessian.

The method, due to Hilbert (Hilbert, 1893), generates complete sets of irreducible invariants for affine and orthogonal transformations. The sets for third and

$R_{12}$	$L_x^2 L_{yy} - 2L_y L_x L_{xy} + L_y^2 L_{xx}$
$R_{13}$	$L_x^3 L_{yyy} - 3L_y L_x^2 L_{xyy} + 3L_y^2 L_x L_{xxy} - L_y^3 L_{xxx}$
$R_{23}$	$L_{xx}^3 L_{yyy}^2 - 6L_{xx}^2 L_{yyy} L_{yy} L_{xxy} + 6L_{xx} L_{yyy} L_{yy} L_{xxx} L_{xy} + 9L_{xx} L_{yy}^2 L_{xxy}^2$ $- 6L_{yy}^2 L_{xxy} L_{xxx} L_{xy} - 6L_{xy} L_{xyy} L_{xx}^2 L_{yyy} - 18L_{xy} L_{xyy} L_{xx} L_{yy} L_{xxy}$ $+ 12L_{xy}^2 L_{xyy} L_{yy} L_{xxx} + 12L_{xy}^2 L_{xxy} L_{xx} L_{yyy} - 8L_{xy}^3 L_{xxx} L_{yyy}$ $+ 9L_{yy} L_{xxy}^2 L_{xx}^2 - 6L_{yy}^2 L_{xxy} L_{xx} L_{xxx} + L_{yy}^3 L_{xxx}^2$
$D_1$	1
$D_2$	$-L_{xy}^2 + L_{yy} L_{xx}$
$D_3$	$-3L_{xxy}^2 L_{xyy}^2 - 6L_{xxy} L_{xyy} L_{xxx} L_{yyy} + L_{xxx}^2 L_{yyy}^2 + 4L_{xxy}^3 L_{xxx}$ $+ 4L_{yyy} L_{xxy}^3$

Table 1.4: All lower order resultants and discriminants to third order.

fourth order are increasingly more complicated, and are fully specified in (Salden et al., 1992b). Haring et al. applied these irreducibles in Kohonen neural networks for segmentation (Haring et al., 1994). In figure 1.3 (from (Salden et al., 1992b)) an example is given of the fourth order *discriminant* of the fourth order binary form, acting as a robust 4-junction detector. A successful application in object recognition was presented by Schmidt et al. (Schmid and Mohr, 1996), using the set of irreducible invariants to third order (9 members for 2D images) and a statistical analysis of the resulting high-dimensional feature space.

## Intermezzo 2: Numerical Implementation

Convolutions of a given input image with a given Gaussian derivative kernel are conveniently calculated in the Fourier domain, where a convolution is equivalent to a product ( $\mathcal{F}$  indicates the Fourier transform,  $\mathcal{F}^{-1}$  the inverse transform (Cooley and Tukey, 1965)). The  $n^{th}$  order Gaussian partial derivative of a sampled image  $L$  is calculated through

$$\frac{\partial^{n+m}}{\partial x^n \partial y^m} L * G = \mathcal{F}^{-1} \left\{ \mathcal{F}\{L\} \cdot \mathcal{F}\left\{ \frac{\partial^{n+m}}{\partial x^n \partial y^m} G \right\} \right\} \quad (1.1)$$

The Fourier transform of the derivative of a function is most conveniently calculated using

$$\mathcal{F}\left\{ \frac{\partial^{n+m}}{\partial x^n \partial y^m} G \right\} = (-i\omega_x)^n (-i\omega_y)^m \mathcal{F}\{G\} \quad (1.2)$$

where  $\omega$  is the spatial frequency. A consequence of implementation via the Fourier domain is the infinite periodicity of the image in all directions.

## 1.5 Deblurring

Gaussian blurring is a common phenomenon in physical science, where *measurements* abound. To counteract this phenomenon, many approaches have been proposed (see the references in (Haar Romeny et al., 1994b)). The problem is ill-posed,

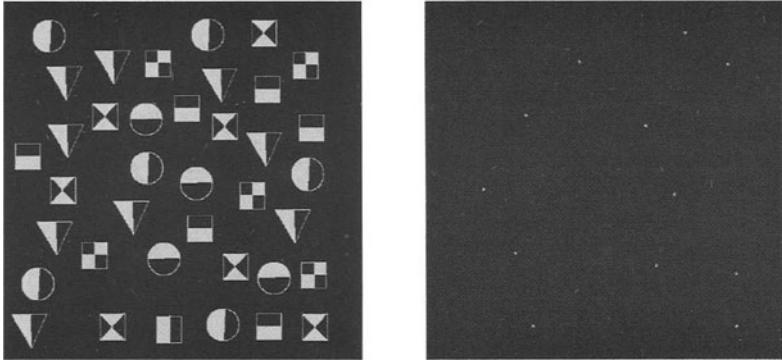


Figure 1.3: 4-junction detection: The pure fourth order invariant  $D_4 = -(\frac{1}{2}L_{iijj}L_{kkll} - L_{iijk}L_{jkll} + \frac{1}{2}L_{ijkl}L_{ijkl})^3 + (\frac{1}{8}L_{iijj}L_{kkll}L_{mmnn} - \frac{1}{8}L_{iijj}L_{klmn}L_{klmn} - \frac{1}{4}L_{iijj}L_{kklm}L_{lmnn} + \frac{1}{4}L_{iijk}L_{jklm}L_{lmnn})^2$  calculated for a set of sampled shapes. Scale of the invariant  $D4$ : 4.18 pixel units. Note the rotation invariance and that despite this high order, the fourth order property is well-represented at an appropriate scale. Resolution  $512^2$ . (From (Salden et al., 1992b)).

and a major attack to the problem is the regularization of the inverse operators, either by probabilistic methods, by Tikhonov regularization, or the decomposition of inverse operators.

The regularized Gaussian kernel makes scale-space *continuous*, and Florack *et al.* (Florack et al., 1996a) suggested the idea of inter- and extrapolation at any point in a scale-space, using the multiscale local *jet*, i.e. in this case to consider the extrapolation to smaller scales, i.e. deblurring (Haar Romeny et al., 1994b). The luminance in a neighborhood of the origin  $(x, y; t) = (0, 0; 0)$  (without loss of generality) is given by the expression

$$L_\infty(\delta\vec{x}, \delta t) = \exp \left\{ \delta t \Delta + \delta\vec{x} \cdot \vec{\nabla} \right\} L$$

which function, when developed in a series, just gives the Taylor expansion:

$$\exp(X) \equiv \sum_{n=0}^{\infty} \frac{1}{n!} X^n$$

$\Delta$  denotes the Laplacean operator,  $\vec{\nabla}$  the nabla operator, with  $\Delta = \vec{\nabla} \cdot \vec{\nabla}$ . The notation  $L_\infty$  denotes the infinite expansion, i.e. equality. The infinite expansion gives the *exact* reconstruction. In practical applications the expansion will always be truncated. The way this series must be truncated in the scale and space domain to keep exact scale-space properties is studied by Florack *et al.* in (Florack et al., 1996a; Florack, 1993) and ter Haar Romeny *et al.* in (Haar Romeny et al., 1994d). For deblurring we study the expansion along the scale axis only, and find

$$L(t + \delta t) = L(t) + \frac{\partial L}{\partial t} \delta t + \frac{1}{2!} \frac{\partial^2 L}{\partial t^2} \delta t^2 + \frac{1}{3!} \frac{\partial^3 L}{\partial t^3} \delta t^3 + \dots$$

Deblurring is then accomplished by taking a negative scale-step:

$$L(t - \delta t) = L(t) - \frac{\partial L}{\partial t} \delta t + \frac{1}{2!} \frac{\partial^2 L}{\partial t^2} \delta t^2 - \frac{1}{3!} \frac{\partial^3 L}{\partial t^3} \delta t^3 + \dots \quad (1.3)$$

The higher order scale-derivatives can be expressed in higher order Laplaceans because of the diffusion equation:

$$\begin{aligned} L_t &= L_{xx} + L_{yy} \\ L_{tt} &= L_{xxxx} + 2L_{xxyy} + L_{yyyy} \\ L_{ttt} &= L_{xxxxxx} + 3L_{xxxxyy} + 3L_{xxyyyy} + L_{yyyyyy} \\ L_{tttt} &= L_{xxxxxxxx} + 4L_{xxxxxyy} + 6L_{xxxxyyyy} + 4L_{xxyyyyyy} + L_{yyyyyyyy} \\ &\text{etc.} \end{aligned}$$

So each order of differentiation to scale gives rise to twice the order in the spatial domain. Figure 1.4 gives an example. However, the method is very sensitive to small disturbances (ill-posedness) and coarse representation of the intensity samples (see (Haar Romeny et al., 1994b)).

## 1.6 Ridges and Multimodality Matching

The sophisticated modern imaging modalities each measure the spatial distribution of a distinct (or limited set of) physical parameter(s). It can be anatomy-related, like the linear attenuation coefficient in CT or the T<sub>1</sub>, T<sub>2</sub> or proton density in MR, or function-related, as the uptake of radioactive tracers in SPECT and PET (Stokking et al., 1994), or metrical, like the distribution of e.g. thickness over the surface of a skull (Zuiderveld and Viergever, 1994). The different measurements of the same patient are done in different scanners at different times, and matching (registration) of 2D or 3D datasets is an essential requirement for integrated visualization (Zuiderveld and Viergever, 1992; Zuiderveld and Viergever, 1994) or processing.

One way to line-up 3D datasets is by extracting outline features. One can extract contour-based features, like edges, or region-based features, such as ridges, and minimize some distance function between the extracted sets. In figure 1.5 an example is given of CT/MR matching, using the ridge operator  $L_{vv}$ , i.e. the second derivative in the direction tangential to the isophote. The (global) match found is of good quality, and there is no need for external markers. Other application areas include 3D CT subtraction angiography (Haar Romeny et al., 1994a) (CTA, now feasible due to the arrival of high resolution slirping CT technology) and the matching with sophisticated 3D medical atlases. Figure 1.6 nicely illustrates the application: the integrated visualization of MR and CT data is now possible, each with their own strength: CT is the modality of choice to depict the bone, MR to depict the soft tissues, like the cortical surface.

## 1.7 Deep Structure and the Hyperstack

It is interesting to exploit the multiscale representation of image data by studying the relation between different scale levels.

De Graaf, Vincken and Koster developed the 'hyperstack' (multiscale stack of images) segmentation (Vincken et al., 1990; Vincken et al., 1994; Koster et al., 1996; Vincken et al., 1996) as a linking model based segmentation technique, originally built upon linear scale space theory. The basic idea of the hyperstack is to define relations between voxels in adjacent scale space levels, such that the levels at larger scales—containing the global information—guide the collection of voxels at the smallest scale (the original image). The entire process requires four steps: *(i)* blurring, *(ii)* linking, *(iii)* root labeling, and *(iv)* downward projection (see figure 1.7).

In the blurring phase a stack of images is created. During the linking phase voxels in two adjacent scale levels are connected by so-called child-parent linkages, with an assigned *affection* value based on heuristic and statistical features (see (Koster et al., 1996) for details). The area in which a parent in level  $n + 1$  is selected for a specific child in level  $n$  is defined by a radius  $r_{n,n+1} = k \cdot \sigma_{n,n+1}$ , where  $\sigma_{n,n+1}$  denotes the relative  $\sigma$  (that corresponds with the transition from level  $n$  to level  $n + 1$ ) and  $k$  is chosen such that only parents are considered whose intensity has been influenced significantly by the child at hand. Typically,  $k = 1.5$ . Furthermore, the linking is a bottom-up process such that only parents that have been linked to are considered children in the next linking step. This leads to the typical tree-like structure of linkages through scale space. If the linking has converged—in the sense that only few parents are left in the top level of the hyperstack—the root labeling takes place. In this phase, the children in the tree with a relatively low affection value are labeled as *roots*, each of which represents a segment in the original image. Finally, in the down projection phase the actual segments are formed by grouping all the voxels in the ground level that are linked to a single root by following the child-parent linkages downwards.

In (Vincken et al., 1996) the hyperstack is used for an in-depth comparison of four different scale space generators with respect to segmentation results. Considered are the linear (Gaussian) scale space both in the spatial and the Fourier domain, the variable conductance diffusion according to Perona & Malik, and the Euclidean shortening flow. Segmentation experiments are carried out on MR images of the brain (see figure 1.8), for which a gold standard is available. The hyperstack turns out to be rather insensitive to the underlying scale space generator. Currently we study the mathematical underpinning of this so far rather heuristic approach.

Name of flow	Image	Level set	Cons.	Flow
Linear	$\frac{\partial L}{\partial t} = \Delta L$	$\frac{\partial C}{\partial t} = \frac{-\Delta L}{ \nabla L }$	$\int L$	$\nabla L$
Variable conduct.	$\frac{\partial L}{\partial t} = \nabla \cdot (c \nabla L)$	$\frac{\partial C}{\partial t} = \frac{-\nabla \cdot (c \nabla L)}{ \nabla L }$	$\int L$	$c \nabla L$
Variable tensor	$\frac{\partial L}{\partial t} = \nabla \cdot (D \nabla L)$	$\frac{\partial C}{\partial t} = \frac{-\nabla \cdot (D \nabla L)}{ \nabla L }$	$\int L$	$D \nabla L$
Normal motion	$\frac{\partial L}{\partial t} = L_w$	$\frac{\partial C}{\partial t} = c \vec{N}$	-	-
Eucl. shortening	$\frac{\partial L}{\partial t} = L_{vv}$	$\frac{\partial C}{\partial t} = \kappa \vec{N}$	-	-
Affine shortening	$\frac{\partial L}{\partial t} = L_{vv}^{\frac{1}{3}} L_w^{\frac{2}{3}}$	$\frac{\partial C}{\partial t} = \kappa^{\frac{1}{3}} \vec{N}$	-	-
Entropy	$\frac{\partial L}{\partial t} = \alpha L_w + \beta L_{vv}$	$\frac{\partial C}{\partial t} = (\alpha + \beta \kappa) \vec{N}$	-	-

Table 1.5: Overview of the various evolution schemes. Cons.: Conserved property.

## 1.8 Denoising and Edge-Preserving Smoothing

Recently many 'geometry-driven' evolution schemes have been proposed. See for a tutorial overview (Haar Romeny, 1994)<sup>1</sup>. In Utrecht (Haar Romeny, 1996) we have implemented a general framework to generate nonlinear multi-scale representations of image data (Niessen et al., 1994a; Niessen et al., 1994b; Niessen et al., 1996a; Weickert, 1996a). The process is considered as an initial value problem with an acquired image as initial condition and a geometrical invariant as "driving force" of an evolutionary process.

The geometrical invariants are extracted using the family of Gaussian derivative operators. Stability requirements for numerical approximation of evolution schemes using Gaussian derivative operators are derived (Niessen et al., 1994b) and show that the allowed time-step is proportional to scale, and considerably larger than in nearest neighbor approximations. This approach has been used to generalize and implement a variety of nonlinear diffusion schemes. See table 1.5. Weickert (Weickert, 1996a) presents a complete scale-space theory for nonlinear diffusion filtering in the continuous, semidiscrete and discrete setting. Anisotropic diffusion with a heat conduction tensor has already been studied in (Weickert, 1994; Weickert, 1995; Weickert, 1996c).

Two approaches are distinguished: evolution of the luminance function under a flow and evolution of the isoluminance curves of the image. The duality relation between these two approaches, i.e. curve evolution and geometric diffusion, was found by Osher and Sethian (Osher and Sethian, 1988), and Lopez and Morel

---

<sup>1</sup>A number of European and US based laboratories have been collaborating in the 'DIFFUSION' program: <http://www.cv.ruu.nl/Collaborations/node1.html>. The book (Haar Romeny, 1994) is one of the fruits of this collaboration.

(Lopez and Morel, 1993): The level sets of an evolution of  $L$  constrained by

$$\frac{\partial L}{\partial t} = \nabla \cdot F(L_i, L_{ij}, \dots)$$

evolve according to

$$\frac{\partial C}{\partial t} = \frac{-\nabla \cdot F(L_i, L_{ij}, \dots)}{|\nabla L|} \vec{N}$$

and, given the curve evolution:

$$\frac{\partial C}{\partial t} = H(L_i, L_{ij}, \dots) g(\kappa) \vec{N}$$

the luminance function evolves according to

$$\frac{\partial L}{\partial t} = -H(L_i, L_{ij}, \dots) L g(\kappa)$$

Here,  $\vec{N}$  is the inward unit normal, and  $\kappa$  isophote curvature.

An example of edge preserving smoothing in medical imaging is given in figure 1.9, where a substantial noise reduction is obtained for both MR and CT images.

## 1.9 Discussion

Scale-space theory finds many applications. The regularized operators can be applied to any dimension dataset for robust extraction of a wide variety of structural information. Most applications presented have been chosen from medical imaging, but the application areas can easily be expanded to robotics, meteorology, geophysical imaging, remote sensing etc. Though promising results can now be obtained, many areas still need much further development, such as scale selection, exploitation of local orientation structure, temporal differential structure, and the important step to go from local detection to global analysis, the realm of perceptual grouping. Many of these research topics are discussed in the other chapters of this book.

## Acknowledgements

The applications mentioned in this chapter were taken from many researchers in the field, as referenced. I like to thank Luc Florack, Alfons Salden, Wiro Niessen, Robert Maas, Joachim Weickert, Twan Maintz, Koen Vincken, Andre Koster, Nico Karssemeijer, Rik Stokking, Karel Zuiderveld, Antonio Lopez, Mads Nielsen, Stiliyan Kalitzin, Hans Blom, Bas Haring and Max Viergever for their willingness to contribute part of their work to this chapter.

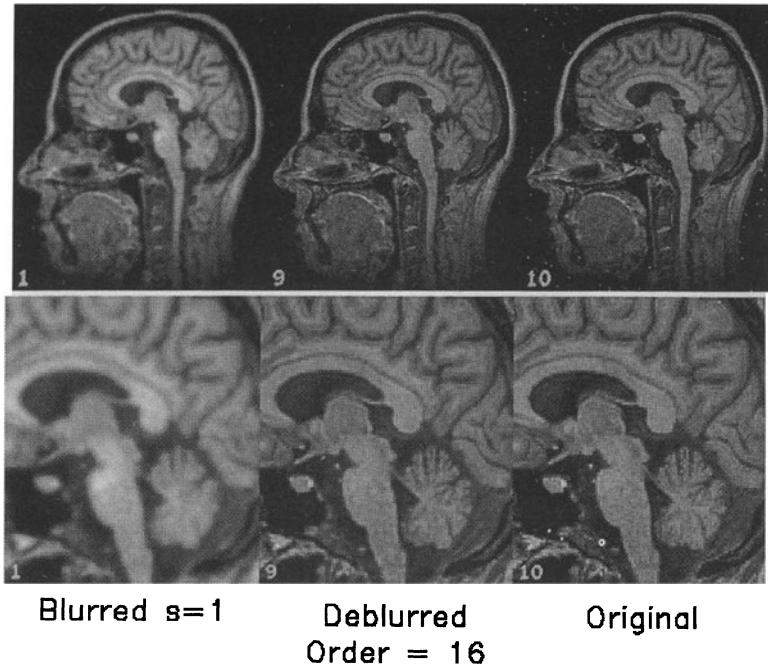


Figure 1.4: Deblurring Gaussian blur. Right image: original sagittal MR image, 12 bit resolution, spatial resolution 128x128 pixels. Left image: original blurred with Gaussian kernel, width 2.8 pixels. Middle image: Deblurred result of application of equation (1.3), truncated at 16<sup>th</sup> spatial order. Standard deviation Gaussian derivative operators 2 pixels. The bottom row shows enlargement of the top row images. From (Haar Romeny et al., 1994b).

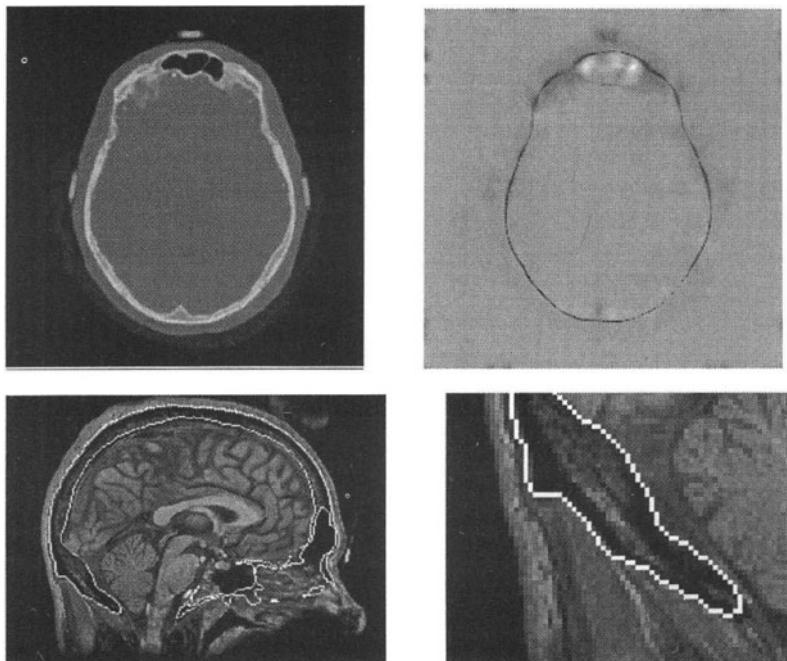


Figure 1.5: Ridge-based 3D multimodality matching of CT and MR images. Top images: CT image and the  $L_{vv}$  ridge detection. Lower images: Overlay of bone contour (extracted from the CT data) over the MR data. Note the good matching. Lower right: enlarged detail. From Maintz et al. (Maintz et al., 1994)

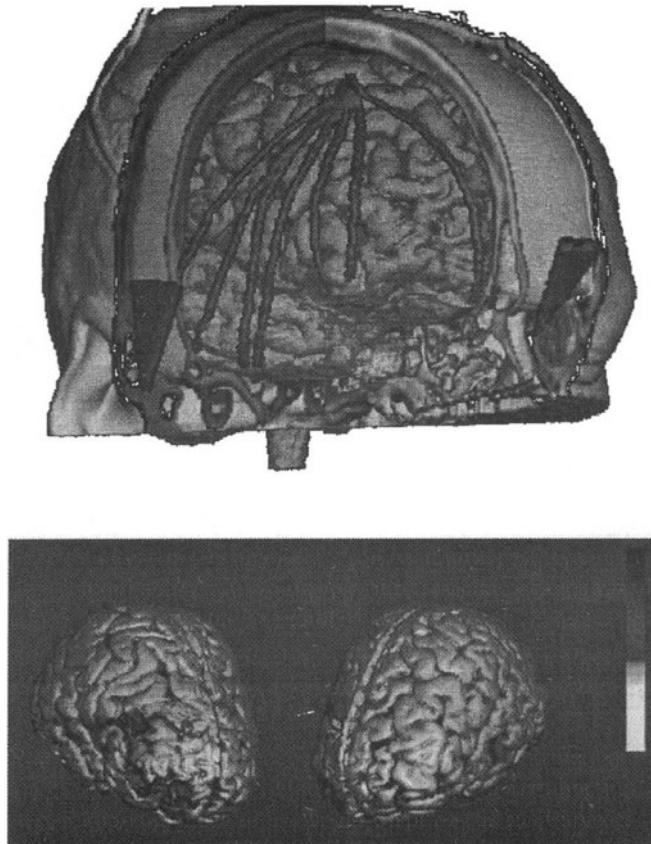


Figure 1.6: 3D Visualization of matched datasets. Top: Integrated view of CT and MR data. The cortical surface is segmented from the MR data, the bone and cortical electrodes from the CT data. From Zuiderveld (Zuiderveld, 1995). Bottom: Integrated view of SPECT and MR data: "Reverse gradient fusion" visualization; a right (A) and left (B) frontal view of a brain with a damaged region in the right frontal lobe, indicated by the atrophic area. Comparison of left- and right frontal lobe clearly yields an area of increased cerebral blood flow surrounding the damaged region. The SPECT lookup table is shown on the right hand side. From Stokking (Stokking et al., 1994). Volume rendering with the software package VROOM (Zuiderveld and Viergever, 1994).

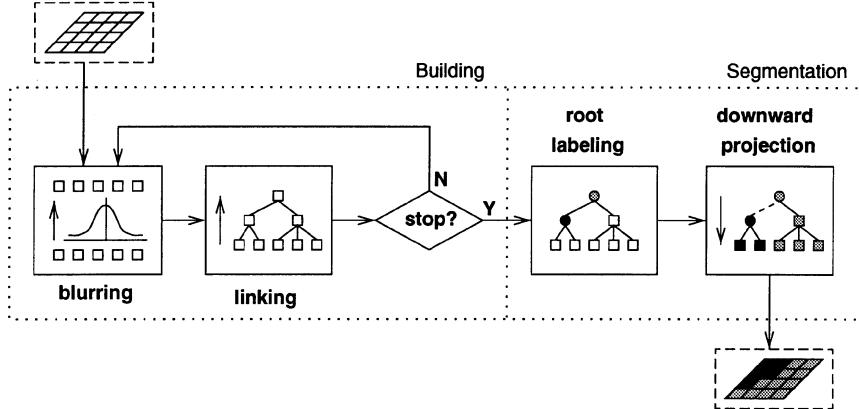


Figure 1.7: Schematic of the hyperstack segmentation process.

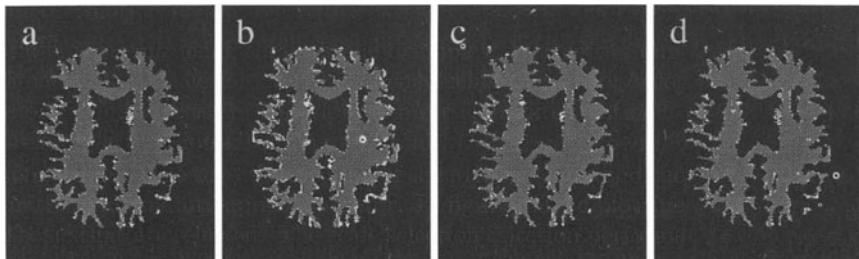


Figure 1.8: Segmentations of the white matter of a transversal MR brain image based on four different scale space generators. The grey colored areas have been segmented correctly (according to the gold standard), the white colored areas correspond to erroneously segmented pixels. The figures correspond to: (a) Linear scale space implemented in the spatial domain; (b) idem, implemented in the Fourier domain; (c) Perona & Malik anisotropic diffusion; (d) Euclidean Shortening Flow. From (Vincken et al., 1996).

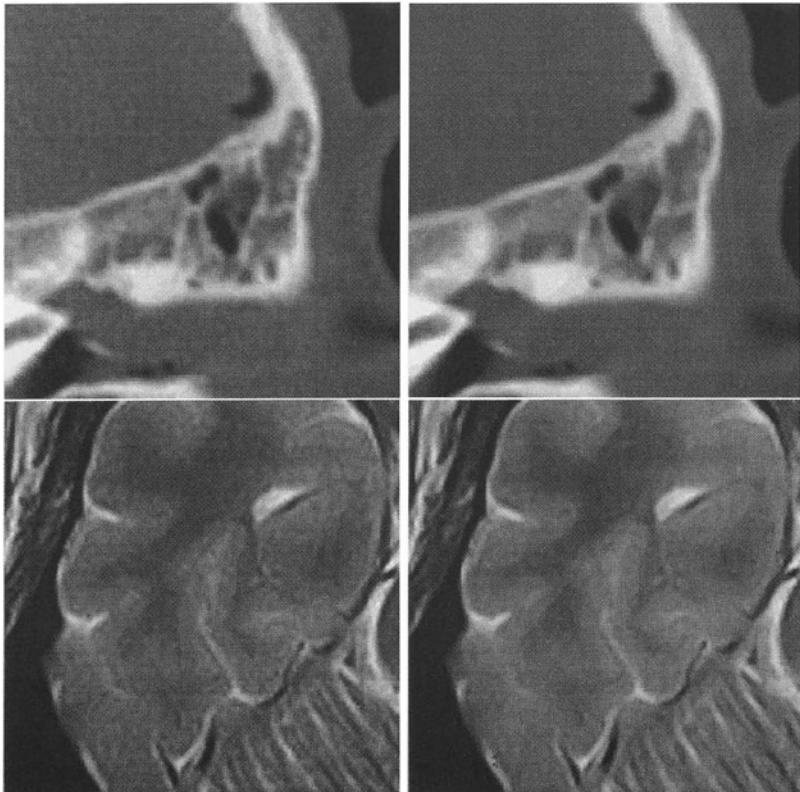


Figure 1.9: Upper figure: a part of a CT-scan showing the petrous bone and the semi-circular channels. Lower figure: part of an MR-scan showing the cerebellum and cerebrum. Both figures are taken from Elsevier's interactive anatomy atlas on CDI/CDROM edited by Hillen (Wolferen et al., 1993). To the right we see the effect (we chose a larger evolution than in the data on CDI to improve clarity) after we applied the affine invariant evolution equation  $\frac{\partial L}{\partial t} = L_{vv}^{1/3}$  which evolves the isophotes as a function of their curvature with a thresholding depending on the norm of the gradient. The process improves the visual impression of the images on the display. From (Niessen et al., 1994b).

# Chapter 2

# Enhancement of Fingerprint Images using Shape-Adapted Scale-Space Operators

**Andrés Almansa**

*Centro de Cálculo, Facultad de Ingeniería, Universidad de la República  
Julio Herrera y Reissig 565, piso 5. 11300 Montevideo, Uruguay*

**Tony Lindeberg**

*Computational Vision and Active Perception Laboratory (CVAP)  
Department of Numerical Analysis and Computing Science (NADA)  
KTH (Royal Institute of Technology)  
S-100 44 Stockholm, Sweden*

## 2.1 Introduction

Automated Fingerprint Identification Systems (AFIS) rely on the ridge structure of a fingerprint for comparing it to other fingerprints in a database. The structure of current AFIS closely resembles the manual procedure used by human experts (McCabe et al., 1992), and can be briefly described as follows:

- *Enhance* the fingerprint image. This usually involves a filtering step to suppress noise. The output from this step can be either a better grey-level image or a binary image in which the ridge structure of the original is preserved.

---

<sup>0</sup>This research has been supported by BID-CONICYT project number 96/94 and by the BITS collaboration program between KTH and the Engineering School at the University of Uruguay. We would like to thank the CeCal and CVAP staffs, especially R. Curbelo, G. Drets, O. Bergengruen and R. Oliveira for their support, and for making this project possible.

- Extract the *ridge structure* of the fingerprint, *i.e.*, build a graph-like representation of the skeletonized ridges and their incidence relations.
- *Classify* the fingerprint according to its ridge structure at a coarse scale. This reduces the search space, thus speeding up the procedure.
- Detect *minutiae*, *i.e.* singularities in the ridge structure, such as bifurcations and ridge endings. Then, build a feature vector of the fingerprint.
- Store the feature vectors in a database for later use and/or search for fingerprints with *matching* feature vectors (which will most likely correspond to the same person).

Traditional methods for fingerprint image enhancement are based on Fourier transforms and Gabor filters, dating back to the work of (Nakamura et al., 1986) and (Asai et al., 1986). Variations of such methods include (Bergengruen, 1994), (Almansa, 1994), (Candela et al., 1995), (Kaymaz and Mitra, 1992). By and large, these approaches consist of applying a suitable Gabor or similar linear filter at each point in the image, where the filter selection is usually based on local estimates of ridge orientation and width. In some cases, multi-scale processing is performed.

These methods turn out to be useful for estimating the distance between ridges. Moreover, the main orientation field (see figure 2.1), has been successfully used in classification methods (Candela et al., 1995). Such methods, however, fail to produce accurate skeletons of the fingerprint image, as needed by matching algorithms, since fairly noisy singularities are destroyed or transformed in many cases.

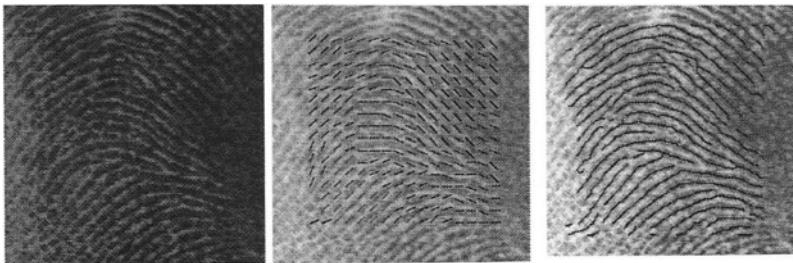


Figure 2.1: Left: Original sample image. Center: Ridge direction produced by a spectral-based ridge enhancement method (Bergengruen, 1994). Right: Skeleton produced by the same method.

The subject of this article is to present an image enhancement algorithm that can be used as part of an AFIS. Such an algorithm should provide an enhanced image and a *ridge orientation* field (*i.e.*, a coarse estimate of the main ridge direction at each point) on which either classification, ridge structure extraction, or matching algorithms should perform better than on the original, noisy image with no directional information.

## 2.2 Shape-adapted smoothing

The method we shall propose for enhancing the ridge structure in fingerprint images is based on shape adaptation of scale-space operators in the affine Gaussian scale-space representation. This approach builds upon work by (Lindeberg, 1994e; Lindeberg and Gårding, 1994) which was originally aimed at reducing the shape distortions that occur in filter based algorithms for obtaining shape from texture and disparity gradients. It also builds upon the non-linear filtering ideas proposed by (Nitzberg and Shiota, 1992), as well as the application of “tensor based” non-linear diffusion schemes to fingerprint images suggested by (Weickert, 1995). In these respects, the approach bears similarities with non-linear diffusion methods, as summarized in (Haar Romeny, 1994). The major difference is that the proposed approach will be constructed from a set of linear scale-space representations, from which interesting data are selected by non-linear scale selection and shape adaptation mechanisms (see the discussion below and in section 2.4). In this respect, the approach bears closer relationships to the open-ended suggestion by (Florack, 1993) to regard non-linear scale-space concepts as linear scale-space representations on transformed domains.

Given any image  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , and given any symmetric positive semi-definite matrix  $\Sigma_0$ , the *affine Gaussian scale-space representation* (Lindeberg, 1994e)  $L$  of  $f$  is for this  $\Sigma_0$  defined as the solution of the anisotropic diffusion equation

$$\partial_t L = \frac{1}{2} \nabla^T \Sigma_0 \nabla L \quad (2.1)$$

with initial condition  $L(\cdot; 0) = f$ . The approach we shall follow, as suggested in different forms in (Nitzberg and Shiota, 1992; Lindeberg, 1994e; Lindeberg and Gårding, 1994; Weickert, 1995), is to adapt the shape of  $\Sigma_0$  to the local image structure based on *second-moment descriptors* (Kanatani, 1984), (Förstner and Gülich, 1987), (Brown and Shvaytser, 1990), (Super and Bovik, 1992), (Bigün et al., 1991), (Rao and Schunk, 1991), (Lindeberg and Gårding, 1993; Gårding and Lindeberg, 1996):

$$\mu(\cdot; \Sigma_t, \Sigma_s) = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{12} & \mu_{22} \end{pmatrix} = g(\cdot, \Sigma_s) * (\nabla L(\cdot; \Sigma_t) \nabla L(\cdot; \Sigma_t)^T) \quad (2.2)$$

computed as the exterior product of gradient vectors by themselves at *local scale*  $\Sigma_t$  in the scale-space representation  $L$  of  $f$  and then averaged using an affine Gaussian kernel with *integration scale*  $\Sigma_s$  as weight

$$g(x; \Sigma_s) = \frac{1}{(2\pi|\Sigma_s|^{1/2})} \exp -x^T \Sigma_s^{-1} x / 2. \quad (2.3)$$

To obtain an intuitive understanding of the operation, let us expand the tensor product and introduce  $P = \mu_{11} + \mu_{22} = g(\cdot; \Sigma_s) * (L_x^2 + L_y^2)$ ,  $C = \mu_{11} - \mu_{22} = g(\cdot; \Sigma_s) * (L_x^2 - L_y^2)$ , and  $S = 2\mu_{12} = 2g(\cdot; \Sigma_s) * (L_x L_y)$ . Then, it can be seen that  $P$  is a measure of operator strength (local image contrast measured by average

gradient magnitude). The other two entities contain directional information, which can be summarized into the two *anisotropy* measures

$$Q = \sqrt{C^2 + S^2}, \quad \tilde{Q} = \frac{Q}{P}. \quad (2.4)$$

The *normalized anisotropy*  $\tilde{Q}$  is always in the range  $[0, 1]$ , and the extreme cases  $\tilde{Q} = 0$  and  $\tilde{Q} = 1$  correspond to rotationally and translationally symmetric patterns. The eigenvalues of  $\mu$  are  $\lambda_{\max,\min} = \frac{1}{2}(P \pm Q)$ .

If we interpret the symmetric positive definite matrix  $\mu$  as a quadratic form, the corresponding ellipses will be elongated in the direction of the gradient. Therefore, multiples of  $\mu^{-1}$  constitute natural choices for scale matrices, since we want a larger amount of smoothing along ridges than across them. Computation of  $\mu$ , however, requires determination of *local* and *integration scale matrices*  $\Sigma_t$  and  $\Sigma_s$ . In (Lindeberg and Gårding, 1994) it was proposed that it is natural to look for a *fixed point*  $\Sigma$  such that

$$\mu(\cdot; \Sigma_t, \Sigma_s) = \Sigma^{-1}, \quad \Sigma_t = t\Sigma, \quad \Sigma_s = s\Sigma. \quad (2.5)$$

In (Lindeberg and Gårding, 1994), it was shown that this fixed point is preserved under affine transformations. Moreover, a rapidly convergent algorithm was proposed to find this fixed point. It is the process of reaching such fixed points that is referred to as *shape adaptation*.

## 2.3 Enhancement of ridges by shape adaptation

When expressing a shape adaptation procedure in practice, it turns out that a number of technical problems have to be addressed. The fixed point condition (2.5) imposes a constraint on the *shape* of the affine Gaussian kernel (2.3), whereas its *size* is a free parameter. This leads to the need for a mechanism for automatic scale selection, which we will be treated in section 2.4. Other problems concern:

**Defining the scale of the diffusion matrices.** Assuming for a moment that the scale selection problem can be avoided by assuming that initial size information is available, a more technical problem concerns how to use such size information when the covariance matrices become successively more isotropic. A simple strategy is to keep the size of the ellipse constant over the iterations. However, how should that size be defined? Experiments on several fingerprint images showed that the *minor axis* of the corresponding ellipse constitutes a better choice than, for example, the major axis or the square root of the ellipse area. A motivation for this choice is that it ensures a certain *minimum amount of smoothing* as a mean of noise suppression (Lindeberg, 1994e). Thus, given a scalar size descriptor  $t$  and a non-isotropic matrix  $\Sigma$ , we define the *local scale matrix*  $\Sigma_t$  by  $\Sigma_t = t \frac{\Sigma}{\|\Sigma\|_{\lambda_{\min}}}$ .

**Bounding the eccentricity of the ellipses.** Another technical problem arises in highly anisotropic regions, where  $\Sigma$  becomes near singular, and shape adaptation may lead to an exaggerated amount of blurring in some direction. To avoid this effect, we use a *regularized diffusion matrix* (having the same eigenvectors)

$$\Sigma = (\mu + \varepsilon I)^{-1}. \quad (2.6)$$

From the eigenvalues of  $\mu$ , we have  $\text{cond}(\Sigma) = \text{cond}(\mu + \varepsilon I) = 1 + 2Q/(2\varepsilon + P - Q)$ . Since  $(P - Q) \geq 0$ , it follows that  $\text{cond}(\Sigma) \leq 1 + \frac{Q}{\varepsilon} = \kappa$  and the natural choice  $\varepsilon = Q/(\kappa - 1)$  ensures a uniform (user supplied) upper bound  $\kappa$  on the eccentricity of  $\Sigma$ , preventing ellipses from becoming too elongated.

**Isotropic integration scale matrix improves computational efficiency.** Concerning the practical implementation of this scheme, the second stage integration smoothing corresponds to the largest computational work. With respect to the fixed point requirement, however, the constraint on the shape of the scale matrices turns out to be important mainly for the local scale matrix  $\Sigma_t$ , and measurements of  $\mu$  are usually not strongly affected by the shape of the integration scale matrix  $\Sigma_s$ . Thus, to improve the computational efficiency, we use isotropic integration scale matrices, *i.e.*,  $\Sigma_s = s I$ , since these allow for separable smoothing.

**Ensuring size consistency.** At any stage, it is essential that the size descriptor  $\|\Sigma_s\|$  associated with an integration scale matrix  $\Sigma_s$  is greater than the size descriptor  $\|\Sigma_t\|$  associated with the local scale matrix  $\Sigma_t$ , *i.e.*, that

$$\|\Sigma_s\| \geq \gamma^2 \|\Sigma_t\|, \quad (2.7)$$

for some  $\gamma > 1$ . Failure to satisfy this relation might lead to unreliable estimates of the ridge directions when computing  $\mu$  for shape adaptation, which in turn might lead to undesirable enhancement of spurious directional fine-scale structures. Here, we have used the  $\|\cdot\|_{\det}$  norm when comparing local and integration scale matrices with different shape, and the size consistency condition then assumes the form

$$s \geq \gamma^2 t \sqrt{\kappa}. \quad (2.8)$$

A motivation for this approach is that it enforces large integration scales when the ellipses are highly eccentric, while allowing for smaller integration scales in more isotropic regions. Hence, this choice allows the second moment descriptors to adapt to rapidly varying ridge directions around minutiae

More generally, an appropriate value of  $\gamma$  for ridge enhancement (here usually between  $\sqrt{2}$  and  $2\sqrt{2}$ ) depends on the image quality (see discussion in section 2.4).

**Experimental results.** Based on the abovementioned ideas, we implemented an iterative ridge enhancement algorithm. In each iteration, the diffusion equation is solved up to a certain local scale  $t$ , and the diffusion matrix for the next iteration is adapted to the measured second moment descriptors. (Isotropic smoothing was used in the bootstrapping stage.) The local scale parameter  $t = 2.4$  was determined

from manual measurements of the distance between ridges. The safety factor  $\gamma$  and the bound  $\kappa$  on the eccentricity of the ellipses were set to 2 and 10. Then, based on (2.8), the integration scale was set to  $s = 29$ .

As can be seen in figure 2.2, the results are at least as good as for the spectral method. In particular, the ridge orientation estimates are more reliable. Furthermore, the skeleton computed by the proposed method preserves bifurcations slightly better than the spectral based method. Similar conclusions were obtained for a larger number of fingerprints taken from the NIST fingerprint database 4 (Wilson, 1992).

## 2.4 Automatic scale selection

In the scheme outlined so far, all scale levels have been determined manually. There are, however, important reasons to consider automatic ways of tuning them.

First of all, a good choice of  $t$  is always related to the characteristic distance between the ridges. This entity may take very different values for different fingerprints, and may vary substantially even within the same fingerprint. For this reason, a mechanism is needed for adapting the local scale parameter  $t$  to the local image structure. Furthermore, the integration scale should on one hand be sufficiently small, to allow rapid variations in ridge directions around minutiae to be resolved. On the other hand, it should be sufficiently large, such that these variations are slow enough and satisfy the consistency requirement (2.7). To tune these scale levels, an explicit scale selection mechanism is required.

**Selection of integration scales.** A general and powerful methodology for automatic scale selection is by maximizing a suitable normalized measure of operator strength over scales (Lindeberg, 1993c; Lindeberg, 1994e; Lindeberg, 1994b; Lindeberg, 1996a). In this study, we have found the following differential expression useful for estimating the distance between ridges:

$$\mathcal{P}_{\text{norm}} L = t (L_x^2 + L_y^2) + C t^2 (L_{xx} + L_{yy}^2 + 2L_{xy}^2 - 2(L_{xx}L_{yy} - L_{xy}^2))$$

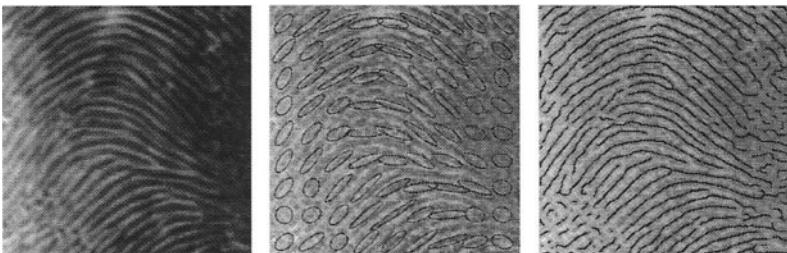


Figure 2.2: Results of shape adaptation. (a) smoothed grey-level image, (b) second moment descriptors, (c) skeleton overlaid on original image. (Parameter settings:  $t = 2.4$ ,  $\kappa = 10$ ,  $\gamma = 2$ ,  $s = 29$ , and 5 iterations.)

For a signal with one-dimensional symmetry, this entity (with  $C \approx 2/3$ ) approximates the local power spectrum (Lindeberg, 1996b), and we will use the scales  $t_{PL}$  at which this entity assumes *local maxima over scales* for selecting integration scales.

This approach agrees with the common practice in classical fingerprint enhancement methods, where the ridge distance is estimated from peaks in local power spectra. Here, however, we avoid explicit computations of local Fourier transforms (with the associated trade-offs between widow sizes and amount of overlapping between windows, efficiency and continuity of the output) and use  $\mathcal{P}_{\text{norm}} L$  instead, which can be easily computed within the scale-space framework.

Then, once  $t_{PL}$  has been calculated, we set  $s$  to the smallest possible value that satisfies the consistency relation in (2.7)

$$s = \gamma^2 \|\Sigma_{t_{PL}}\|_{\det}. \quad (2.9)$$

Figure 2.3 shows an example of applying this approach in the neighborhood of a bifurcation. Note how finer scales are selected at the singularity, which is essential for resolving the more rapid variations in the ridge directions.

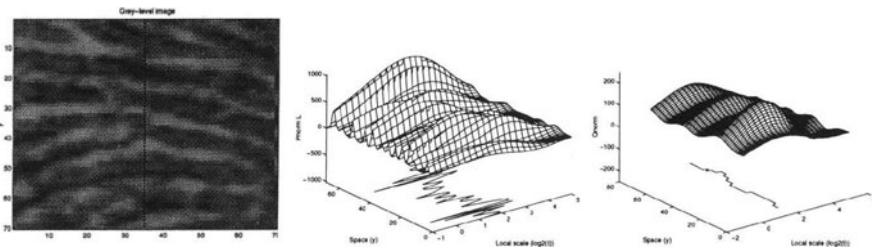


Figure 2.3: Properties of the scale selection mechanism in the neighborhood of a bifurcation. The three-dimensional graphs show scale-space signatures of  $\mathcal{P}_{\text{norm}} L$ , and  $Q_{\text{norm}}$ , respectively, computed along a vertical cross-section through the image center. The vertical axis represents the operator strength at different positions and different scales, whereas the curve in the horizontal plane shows how the selected scales vary along the cross-section (logarithmic scale axis with  $t \in [0.5, 32]$ ).

**Selection of local scales.** In the work of shape from texture and disparity gradients in (Lindeberg, 1994e; Gårding and Lindeberg, 1996), local maximization of  $\tilde{Q}$  over scales was used for automatic selection of local scales. Here, we shall multiply this anisotropy measure by the strength measure  $P$  expressed in terms of normalized derivatives  $\partial_{x_i} = \sqrt{t} \partial_{x_i}$ , and thus maximize the anisotropy measure  $Q_{\text{norm}} = t Q$  over scales.

The idea behind this approach is to select scale levels where the following criteria hold simultaneously: (i) the local image structure should be highly non-isotropic (a high value of  $Q$ ), and (ii) the scale value should reflect the size of the dominant image structures (a high value of  $P_{\text{norm}} = t P$ ).

**Measure of local image quality.** In certain regions of a fingerprint, the image quality can be poor and the ridge structure severely fragmented. To allow the second moment descriptors to capture the overall ridge direction in such regions, a larger integration scale will in general be needed than in regions where the ridge structures are approximately parallel.

To ensure that the integration scales are sufficiently large in regions with low image quality, we introduce a measure of image quality  $\mathcal{Q} \in [0, 1]$ , and let the safety factor  $\gamma$  in the size consistency condition (2.7) depend on  $\mathcal{Q}$  according to

$$\gamma^2 = \gamma_{\min}^2 \mathcal{Q} + \gamma_{\max}^2 (1 - \mathcal{Q}), \quad (2.10)$$

where  $\gamma_{\min}^2$  and  $\gamma_{\max}^2$  are minimum and maximum values of allowed safety factors (typically  $\sqrt{2}$  and  $2\sqrt{2}$ , respectively).

The actual measure of image quality we have used is closely related to the differential entity  $Q_{norm}$  used for local scale selection. Since experiments have shown that  $Q_{norm}$  often assigns too large weights to high contrast gradient information, we instead define this quality measure as

$$\mathcal{Q} = \frac{\sqrt{P} \tilde{Q}}{\sqrt{P_{\max}}} = \frac{Q}{\sqrt{P} \sqrt{P_{\max}}} \quad (2.11)$$

where  $P_{\max}$  is the maximum possible value of  $P$  at scale  $t$ , estimated as the maximum value a squared Gaussian derivative can assume at scale  $t$  (given the intensity range).

**Composed algorithm for shape adaptation and scale selection.** By combining the components described in previous sections, we obtain a scheme for iterative shape adaptation and scale selection. In the bootstrapping stage, isotropic diffusion matrices are used and  $\gamma$  is computed from the image quality using a globally selected local scale and a very large integration scale.

Then, in each iteration, local and integration scales are selected as described above, and the diffusion matrices are adapted to second moment descriptors according to section 2.3. The procedure stops when the relative updates are sufficiently small for the local and the integration scale matrices.

**Experimental results.** This algorithm was tested on 25 images from NIST fingerprint database 4. The parameters  $\gamma_{\min}$ ,  $\gamma_{\max}$  and  $\kappa$  were set to  $\sqrt{2}$ ,  $2\sqrt{2}$  and 10, respectively, and the maximum number of iterations was 3 for scale selection and 5 for shape adaptation.

Figure 2.4 illustrates how estimates of  $\mu$  are improved relative to a shape adaptation scheme at a fixed scale. The result is a much more reliable skeleton of the ridge structure, where errors due to missing or misclassified minutiae are reduced to a minimum (see figure 2.5). Similar results were obtained for other fingerprints in the database. This improvement is important for the later stages of an AFIS, since many matching procedures rely on comparisons between these minutiae.

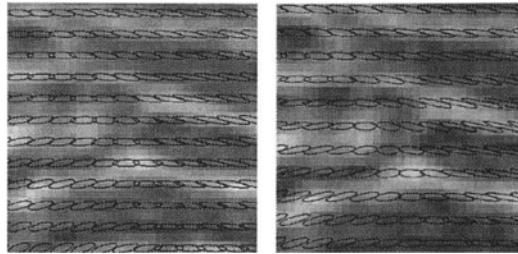


Figure 2.4: Comparison between second moment descriptors computed by the shape adaptation (a) without and (b) with a scale selection mechanism. Note the higher ability of the latter approach to adapt its estimates of  $\mu$  around minutiae.

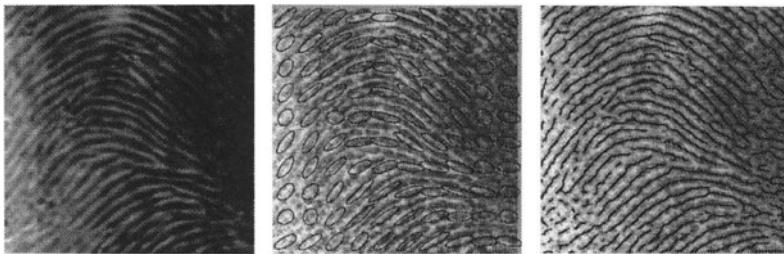


Figure 2.5: Results of composed scale selection and shape adaptation: (a) smoothed grey-level image, (b) measured second moment descriptors, (c) skeleton overlaid on original image.

**Future work.** Concerning limitations of the approach, we have observed that the performance of this method may decrease substantially in extremely noisy regions. A plausible explanation is that the current implementation of the algorithm does not correctly handle multiple local extrema over scales; it only selects the maximum over scales closest to the average scale in the bootstrapping stage. We are currently studying different ways of dealing with this issue.

## 2.5 Summary and discussion

We have presented a novel methodology for fingerprint image enhancement, and compared it to a classical spectral-based method by applying it to a well established set of sample fingerprint images. The experiments led us to the conclusion that the proposed method: (i) is in most cases less prone to miss minutiae or to confuse ridge endings and bifurcations; (ii) produces more *reliable directional fields*, by correcting in certain cases spurious ridge directions produced by scars, or fragmented ridges, that usually mislead the spectral based method; (iii) produces *directional fields* which are more useful for later stages of an AFIS since they are *continuous* both in space and magnitude; (iv) avoids the “checkerboard

effect” present in the grey-level output of the spectral based method. Thus, this shape adaptation approach addresses key weaknesses of traditional fingerprint image enhancement methods, and should be suitable for use within an AFIS. This application is currently being evaluated.

More generally, this technique provides an interesting connection between linear and non-linear scale-space approaches, in the sense that the smoothing operation is purely linear and is separated from the non-linear shape adaptation and scale selection procedures. Whereas the more general implications of this approach have not been explored in detail, this technique has the general attractive property of leading to a larger amount of smoothing along edges than across them, while presenting a more conservative behavior around non-edge areas, allowing more complex structures, such as corners or branching points, to be preserved (Nitzberg and Shiota, 1992; Lindeberg, 1994e). We believe that non-uniform smoothing schemes based on this general idea should be applicable to a large number of other purposes in image processing.

# Chapter 3

# Optic Flow and Stereo

**Wiro J. Niessen and Robert Maas**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

## 3.1 Introduction

In multi-frame or multi-camera imaging systems we aim at inferring the relation between consecutive or adjacent frames, *e.g.* for estimating object velocities or for computing disparities. We will approach these types of problems using the well-known *image brightness constraint equation* which assumes that corresponding image features in different frames have equal luminance values (Horn and Schunck, 1981).

Originally, this constraint was formulated for image isophotes. However, in practical imagery detectors have a fixed, finite resolution; isophotes as such do not exist. The measurement process essentially entails a summation of information over the region covered by a detector. The appropriate formalism to study these measurement processes is provided by linear scale space theory (Koenderink, 1984; Witkin, 1984; Florack et al., 1994c; Lindeberg, 1994e), which will be extensively addressed in the chapters in part II. It derives the one parameter family of scaled Gaussians and its derivatives as the unique filter class to perform image measurements at a certain scale. Applying the brightness conservation constraint to measurements obtained with this filter class, as first proposed by Florack & Nielsen (Florack and Nielsen, 1994; Florack et al., 1996b), we derive equations which are fundamentally different from the conventional brightness constraint equation. Intuitively the modification can be interpreted as if the conservation holds for the support window of the filter subjected to the flow.

We consider two applications of the approach, velocity estimation in image sequences and disparity estimation in binocular stereo. In the first approach the

velocity can only be determined up to a tangential component; owing to the *aperture problem* motion along isophotes can not be detected. We will use additional physical constraints to disambiguate the flow field. In a stereo system these additional constraints are supplied by the epipolar constraint. For the standard geometry there is only horizontal flow.

We emphasize the multiscale nature of the approach. The obtained estimate is scale dependent and estimates at different scales can be different, *e.g.* transparent motion is possible. In practical situations one often wants to infer the velocity of a physical object. In this case the available information should be read out in a way to most reliably extract this velocity. We will in detail discuss the problem to extract a single preferred scale and show examples for both velocity and disparity estimation.

## 3.2 Generalized Brightness Constraint Equation

We use spatiotemporal operators that are straightforward extensions of the Gaussian kernel known from classical scale space theory (Witkin, 1984; Koenderink, 1984).

**Definition 3.2.1 (Spatiotemporal Kernel).** The spatiotemporal scale space kernel  $G(\mathbf{x}, t, \sigma_s, \sigma_t)$  in D+1 dimensions is defined by the normalized Gaussian:

$$G(\mathbf{x}, t, \sigma_s, \sigma_t) = \frac{1}{\sqrt{2\pi\sigma_s^2}^D \sqrt{2\pi\sigma_t^2}} e^{-\frac{1}{2}(\frac{\mathbf{x}\cdot\mathbf{x}}{\sigma_s^2} + \frac{t^2}{\sigma_t^2})} \quad (3.1)$$

These operators apply directly to recorded sequences; for real-time applications we require adaptations for the filters to be causal.

Let  $\Psi(\mathbf{x}, t)$  denote the original or 'highest resolution' image and  $L(\mathbf{x}, t, \sigma_s, \sigma_t)$  the image obtained after convolving the image with  $G(\mathbf{x}, t, \sigma_s, \sigma_t)$ . We can now define a well-posed differentiation process:

**Definition 3.2.2 (Spatiotemporal Derivative).** The spatiotemporal derivative of an image  $\Psi(\mathbf{x}, t)$  at spatial scale  $\sigma_s$  and temporal scale  $\sigma_t$  is defined as ( $*$  denotes convolution):

$$\partial_t^m \partial_{i_1 \dots i_n} L(\mathbf{x}, t, \sigma_s, \sigma_t) = \Psi(\mathbf{x}, t) * \partial_t^m \partial_{i_1 \dots i_n} G(\mathbf{x}, t, \sigma_s, \sigma_t) \quad (3.2)$$

We modify the *brightness constraint equation* for the spatiotemporal filters defined above. The original contribution can be found in the paper by Florack & Nielsen (Florack and Nielsen, 1994). The traditional Horn & Schunck *brightness constraint equation* for scalar images states that, when co-moving with the flow, the luminance should be constant:

$$\frac{d\Psi}{dt} = \Psi_t + v^x \Psi_x + v^y \Psi_y \equiv 0 \quad (3.3)$$

Here  $v^x, v^y$  denote the  $x$  and  $y$  velocity respectively. If we access image information using some aperture function, these equations can be written *under the aperture*.

$$(\Psi_t + v^x \Psi_x + v^y \Psi_y) * G = 0 \quad (3.4)$$

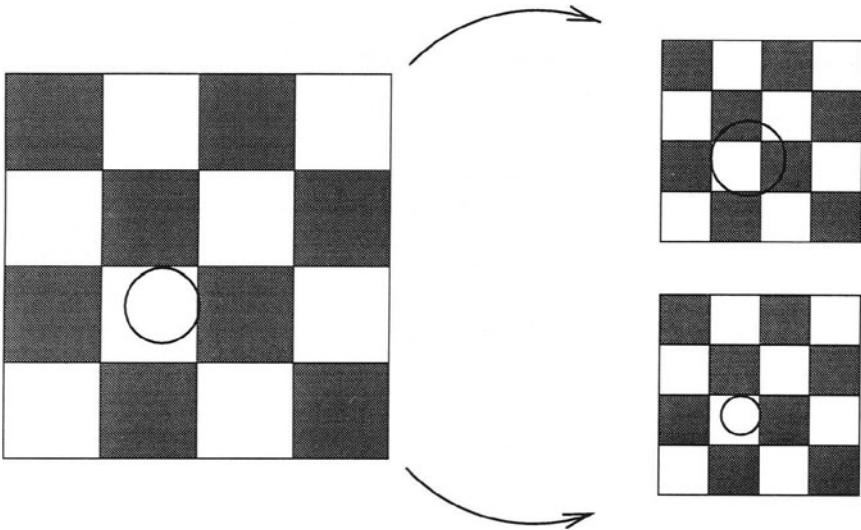


Figure 3.1: Intuitive explanation of the generalized brightness constraint equation. A fronto-parallel checkerboard is being sampled by a filter with a circular support window. If the observer moves away from the screen, while the support region remains fixed, the measurement changes owing to the divergence of the flow field within the support region (upper right). However, if we take into account this divergence (or equivalently, if we adapt the filter according to the flow, lower right) the measurement value is still conserved.

which gives (using partial integration and the fact that the Gaussian decays sufficiently fast at the boundaries)

$$-(v_x^x + v_y^y)\Psi * G + v^x\Psi * G_x + v^y\Psi * G_y + \Psi * G_t = 0 \quad (3.5)$$

Owing to the partial integration a new term appears which is absent from the original formulation. It accounts for the change of signal obtained in the Gaussian filter window. The virtue of this approach can most easily be understood using a simple example (see figure 3.1). It shows an example where an algorithm only allowing for deformations of the original iso-intensity level fails; the notion of the aperture (or scale) needs to be made explicit.

It is not trivial to extract the velocity field from equation (3.5) since the velocity field itself is present in the convolution. We therefore use the following trick: the flow field is approximated up to any order around a point  $(x, t)$  using a polynomial expansion. Let  $\xi$  denote  $(x' - x)$  and  $\eta$  denote  $(y' - y)$ . For example, up to 2nd order we have:

$$v_2^{x'} = \tilde{v} + \xi \tilde{v}_x + \eta \tilde{v}_y + \frac{1}{2}\xi^2 \tilde{v}_{xx} + \frac{1}{2}\eta^2 \tilde{v}_{yy} + \xi\eta \tilde{v}_{xy} \quad (3.6)$$

This series should not be interpreted as a truncated Taylor series; owing to the

fact that all measurements are performed at nonzero scale, increasing the order of approximation affects all components of the flow field.

We subsequently insert equation (3.6) into (3.5). Owing to the fact that polynomials times (derivatives of) Gaussians can be expressed in derivatives of Gaussians, we can express the equation in terms of the approximated velocity components and the spatiotemporal operators defined in the previous section. We therefore get an equation relating the estimated velocity components, which are now considered as given constants, and spatiotemporal image derivatives which can be obtained with the Gaussian operators.

We can differentiate the modified brightness constraint equation (3.4) to obtain additional constraint equations. However, we must assure that no velocity components arise which are not used in the approximation (3.6). This would introduce 'hidden' constraints. Except for the 1-dimensional case the approach always yields an under-determined system of equations. This is due to the *aperture problem* which we will discuss in more detail in the next section.

### 3.3 Optic flow

Owing to the *aperture problem*, *i.e.* tangential motion is possible without affecting the image data, the system of equations resulting from the image brightness equation is in general under-determined. A number of authors differentiate the optic flow equation a number of times, *apparently* fixing the aperture problem (Nagel, 1987; Tistarelli, 1994; Otte and Nagel, 1994). Such an approach can lead to a system of equations that uniquely determines (or even over-constrains!) the optic flow field, thereby violating the degrees of freedom initially present. We cannot emphasize enough that the tangential component should be fixed using physical constraints. Of course, in specific situations the implicit constraint may represent the physical situation to a good approximation in which case the resulting flow field will be in close correspondence to the true motion. In case of rigid motion this is often the case, in case of deformable motion generally not. We therefore choose to keep a strict separation between image derived information and additional constraints to fix the tangential component of the flow field. We first give two approximations of the information directly obtained from the image brightness constraint equation, and then use normal flow as an example of an additional constraint which disambiguates the flow field.

Unless otherwise stated all image derivatives ( $L_x$ ) are expressed in 'natural coordinates'  $\frac{x}{\sigma}$  (Florack et al., 1994c). An  $n$ -th order normalized derivative can be obtained from a (normal)  $n$ -th order derivative by multiplication by  $\sigma^n$ .

- In 2D the zeroth order approximation reads  $v^i = \tilde{v}^i$  which gives

$$\begin{aligned} 0 &= \Psi * (G_t + \tilde{v}^x G_x + \tilde{v}^y G_y) \\ &= L_t + \tilde{v}^x L_x + \tilde{v}^y L_y \end{aligned} \quad (3.7)$$

This is the original Horn & Schunck equation. We can not differentiate the original constraint equation since we approximated the velocity field only up to zeroth order. We therefore have one equation in two unknowns.

- In 2D the first order spatial approximation reads:  $v^i = \tilde{v}^i + \xi \tilde{v}_x^i + \eta \tilde{v}_y^i$  which gives:

$$0 = L_t + \tilde{v}^x L_x + \tilde{v}^y L_y + \tilde{v}_x^x L_{xx} + (\tilde{v}_y^x + \tilde{v}_x^y) L_{xy} + \tilde{v}_y^y L_{yy} \quad (3.8)$$

Since we now have approximated the velocity field up to first order we can differentiate the original equation up to first order with respect to  $x$  and  $y$  to obtain three equations in six unknowns. No additional information can be extracted from the image sequence without additional knowledge.

Complementary equations to the *intrinsic* optic flow equations (3.7) and (3.8) should be derived based upon physical knowledge about the observed scene. As an example we consider normal flow (which we will use in the experiments on the canine MR data in section 5).

In the zeroth order approximation we have:

$$\tilde{v}^x L_y - \tilde{v}^y L_x = 0 \quad (3.9)$$

and with equation (3.7) we can solve for  $\tilde{v}^x$  and  $\tilde{v}^y$ .

If we assume normal flow in the first order approximation we get three complementary equations that we can solve. These are most easily obtained by replacing  $\mathbf{v} = (1, v^x, v^y)$  by its dual  $\mathbf{v}^* = (0, -v^y, v^x)$  in equation (3.8) and the derived versions of it. We end up with the following system of equations:

$$\begin{pmatrix} L_x & L_y & L_{xx} & L_{xy} & L_{xy} & L_{yy} \\ L_{xx} & L_{xy} & L_{x+L_{xxx}} & L_y + L_{xxy} & L_{xxy} & L_{xyy} \\ L_{zy} & L_{yy} & L_{xxy} & L_{zyy} & L_z + L_{xyy} & L_y + L_{yyy} \\ L_y & -L_x & L_{xy} & -L_{xz} & L_{yy} & -L_{xy} \\ L_{zy} & -L_{xz} & L_{xxy} + L_y & -L_x - L_{xxx} & L_{xyy} & -L_{xxy} \\ L_{yy} & -L_{xy} & L_{xyy} & -L_{xxy} & L_y + L_{yyy} & -L_x - L_{xyy} \end{pmatrix} \begin{pmatrix} \tilde{v}^x \\ \tilde{v}^y \\ \tilde{v}_x^x \\ \tilde{v}_x^y \\ \tilde{v}_y^x \\ \tilde{v}_y^y \end{pmatrix} = \begin{pmatrix} -L_t \\ -L_{xt} \\ -L_{yt} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

### 3.4 Binocular stereo

In binocular stereo the aim is to obtain the 3-D shape of an object from two projections. If two corresponding points are known in the projections, the position in 3-space can directly be constructed. For this reason we are interested in the difference in position of two corresponding points, the so-called disparity. If we restrict ourselves to standard geometry for the camera set-up, *i.e.* two adjacent cameras with the same intrinsic parameters and optical axes pointing in the same direction, we get a simple formula for the disparity. We can use this formula to construct the position in 3-space, but it can also be used to obtain locally higher order geometric information by taking higher order derivatives of the disparity.

The epipolar constraint states that, given a point on the left projection plane, the corresponding point on the right projection plane is on the so-called epipolar line. This line is given by the intersection of the right film plane with the plane through the left and right optical centers and the point in the left film plane (Faugeras, 1994). In case of standard geometry the vertical position is equal in the left and right camera.

The disparity  $u$ , which we make a function of the central coordinate  $x = \frac{x_L+x_R}{2}$ , where  $x_L$  is the position in the left retina and  $x_R$  the position in the right retina, is defined as:

$$u(x, y) = x_R - x_L \quad (3.10)$$

If we attach a coordinate system in the middle of the interocular line with  $Z$  pointing orthogonal to the retinas and  $X$  and  $Y$  in the direction of  $x$  and  $y$  respectively, we can deduce using simple geometry:

$$\mathbf{r}(x, y) \equiv \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{l}{u(x, y)} \begin{pmatrix} x \\ y \\ f \end{pmatrix} \quad (3.11)$$

where  $f$  is the focal length of the cameras and  $l$  is the interocular distance. By differentiating  $\mathbf{r}$  with respect to  $x$  and  $y$  we can obtain higher order geometric structure (see also (Faugeras, 1994) and (Devernay and Faugeras, 1994)).

We now treat disparity estimation as an optic flow problem, by considering the two camera images as snapshots in time. To justify the optic flow approach two conditions have to be satisfied: *i*) the reflection is Lambertian, i.e. the reflection is the same in every (possible) direction and *ii*) we have a good approximation  $u_0$  of the disparity  $u$  (this is more general than the restriction that the disparity should be small).

From these assumptions we obtain an equation to calculate the (multiscale) disparity. No ambiguity is present, since the additional constraint is supplied by the camera geometry; the epipolar constraint guarantees that only horizontal flow is present, turning it into a 1-dimensional problem.

Let  $u(x, y)$  be the disparity at the central coordinate  $(x, y)$ . Using a Taylor expansion in  $\frac{u}{2}$  for fixed  $x$  and  $y$  we derive a series expansion for the disparity,

$$\begin{aligned} 0 &= \Psi^R(x + \frac{u+u_0}{2}, y) - \Psi^L(x - \frac{u+u_0}{2}, y) \\ &= \sum_{n=0}^{\infty} \left( \left(\frac{u}{2}\right)^{2n} \frac{\partial^{2n}}{\partial x^{2n}} (\Psi^R(x + \frac{u_0}{2}) - \Psi^L(x - \frac{u_0}{2})) + \right. \\ &\quad \left. \left(\frac{u}{2}\right)^{2n+1} \frac{\partial^{2n+1}}{\partial x^{2n+1}} (\Psi^R(x + \frac{u_0}{2}) + \Psi^L(x - \frac{u_0}{2})) \right) \\ &\approx \Psi_t^C(x) + u \Psi_x^C(x) \end{aligned} \quad (3.12)$$

where  $\Psi_t^C = \Psi_R(x + \frac{u_0}{2}) - \Psi_L(x - \frac{u_0}{2})$  can be interpreted as a discrete approximation to a temporal derivative and  $\Psi_x^C = (\Psi_x^R + \Psi_x^L)/2$  is a central derivative with respect to  $x$ .

Note that equation (3.12) resembles the brightness constraint equation and by virtue of this we can apply “optic flow under the aperture” to our stereo problem. The approximation of  $u$  can be up to any spatial order, but obviously in binocular stereo no higher order temporal derivatives can be obtained.

The set of equations describing the first order approximation of the (normalized) disparity function  $\tilde{u}$  becomes:

$$\begin{pmatrix} L_x^C & L_{xx}^C & L_{xy}^C \\ L_{xx}^C & L_x^C + L_{xxx}^C & L_{xxy}^C \\ L_{xy}^C & L_{xxy}^C & L_x^C + L_{xyy}^C \end{pmatrix} \begin{pmatrix} \tilde{u} \\ \tilde{u}_x \\ \tilde{u}_y \end{pmatrix} = \begin{pmatrix} -L_t^C \\ -L_{xt}^C \\ -L_{yt}^C \end{pmatrix} \quad (3.13)$$

The spatial and 'temporal' derivatives can all be *measured* at an appropriate scale (of the Gaussian derivative operators). From the above set of equations we are now able to solve for the disparity and its gradient in any point of the image, giving a *dense* disparity map.

### 3.5 Scale selection

The theoretical results described in the previous sections are valid for all possible scales. In fact, a proper (or complete) description of a spatiotemporal scene should be derived from information at all scales. A posteriori selection of a single scale in a sense destroys the multiscale nature we stressed as being essential for our approach. However, in practice one is often interested in the velocity of a specific object. In this case certain scales are better suited than others, since the image structure can not be measured at every scale with the same reliability.

The two main criteria to select a certain scale for estimating *e.g.* velocities and disparities are accuracy and robustness. The obtained estimate should be close to the *physical* velocity, and be robust to the image degrading effects that are present. These criteria are not easily reconciled; selecting higher scales often results in better noise behavior but worse accuracy.

For an extensive treatment of scale selection (mainly for image feature extraction such as blobs, edges and ridges), we refer to (Lindeberg, 1994c; Lindeberg, 1996a). In (Lindeberg, 1995; Gårding and Lindeberg, 1996) a scale selection criterion for image deformations is considered. In correlation-based binocular stereo several authors have addressed the selection of appropriate scales, *i.e.* window size (see (Okutomi and Kanade, 1992), (Fua and Leclerc, 1995) and references therein).

In this section we will propose a scale selection criterion to select the proper scale for the linear system of equations which follows from the image brightness constraint equation. One should note that the definition of a proper scale is given by the user-supplied semantics, and thus depends on the assumptions.

Consider a linear system  $\mathbf{A}_\sigma \mathbf{x}_\sigma = \mathbf{b}_\sigma$  as derived in sections 3.3 and 3.4. The index  $\sigma = (\sigma_s, \sigma_t)$ , denotes both spatial and temporal scale. The latter will be constant in the stereo case. By choosing units in physical dimensions for  $b_\sigma$  rather than in natural derivatives, we achieve that we can analyze the physical error in  $\mathbf{x}_\sigma$  (*e.g.* meter/sec).

In order to extract the scales at which we obtain robust estimates we disturb  $\mathbf{A}_\sigma$  and  $\mathbf{b}_\sigma$  and look at the change in  $\mathbf{x}_\sigma$ .

$$\begin{aligned} \|\Delta \mathbf{x}_\sigma\| &= \|(\mathbf{A}_\sigma + \Delta \mathbf{A}_\sigma)^{-1}(\mathbf{b}_\sigma + \Delta \mathbf{b}_\sigma) - \mathbf{A}_\sigma^{-1} \mathbf{b}_\sigma\| \\ &= \|(\mathbf{A}_\sigma + \Delta \mathbf{A}_\sigma)^{-1}(\Delta \mathbf{b}_\sigma - \Delta \mathbf{A}_\sigma \mathbf{x}_\sigma)\| \end{aligned} \quad (3.14)$$



**Figure 3.2:** **Left:** Input image: eighth frame of an MR sequence of a canine heart. Spatial resolution:  $100 \times 100$  pixels; 16 frames per heartbeat. **Middle:** Gradient operator ( $\sigma = 1$ ); **Right:** Temporal derivative of gradient operator ( $\sigma_s = 1, \sigma_t = 1$ ).

Since  $\Delta\mathbf{A}_\sigma$  and  $\Delta\mathbf{b}_\sigma$  are just random disturbances our first guess is that the scale selection criterion is defined by minimizing a norm of  $\mathbf{A}_\sigma^{-1}$ . For reasons that will be explained shortly we consider the Frobenius norm. The Frobenius norm of a real  $m \times n$  matrix  $\mathbf{B}$  is given by:

$$\|\mathbf{B}\|_F = \left( \sum_{i=1}^m \sum_{j=1}^n b_{ij}^2 \right)^{\frac{1}{2}} = \left( \sum_{i=1}^n \lambda_i^2(\mathbf{B}^T \mathbf{B}) \right)^{\frac{1}{2}}$$

where  $\lambda_i$  denote the singular values.

The reason for choosing the Frobenius norm is given by the following lemma:

**Lemma 3.5.1.** Let  $\mathbf{r}$  be a vector with uncorrelated random coefficients. Then

$$E(\|\mathbf{Br}\|_2) = \frac{1}{n} \|\mathbf{B}\|_F E(\|\mathbf{r}\|_2)$$

where  $E$  denotes the expected value.

In the sense of minimizing the expected error the Frobenius norm is the least committed matrix norm.

Note also that in the Frobenius of the inverse of a matrix  $\mathbf{B}$  is mainly determined by the reciprocal of the smallest singular values of  $\mathbf{B}$ :

$$\|\mathbf{B}^{-1}\|_F^2 = \sum_i \lambda_i^2(\mathbf{B}^{-1}) = \sum_i \frac{1}{\lambda_i^2(\mathbf{B})} \quad (3.15)$$

Small singular values of  $\mathbf{A}_\sigma$  imply that our linear system is sensitive to noise. In view of these properties we select the scale which minimizes the Frobenius norm of  $\mathbf{A}_\sigma^{-1}$  in the optic flow and stereo experiments in the remainder of this chapter. However, several issues that are of interest have to be addressed:

- We have to check if the system is well-defined for this choice. In general this will be the case, since we work with normalized derivatives, thus the norm of  $\mathbf{A}_\sigma$  is bounded from above.



Figure 3.3: Spatial (left) and temporal (right) scale that are selected by the minimization of the Frobenius norm for the zeroth order approximation. We use an exponential spatial and temporal scale sampling (both  $\sigma_s$  and  $\sigma_t$  range from 1 to 8 pixel units in 6 exponential steps) which are displayed from black (low scale) to white (large scale). Note that in the region of interest almost everywhere the smallest temporal scale is selected.

- We have to assure that we restrict the range of scales to those that are physically interesting. The scale selection criterion is mainly based upon stability considerations which implies that we are not likely to end up with wrong estimates owing to high frequency noise or small scale texture. However, this also results in a preference for larger scales, in which case we trade accuracy for stability. Therefore, we could add constraints which favor small scales with sufficient stability.

### 3.6 Optic flow results

We illustrate the optic flow theory using some simple examples. We restrict ourselves to velocity estimates in 2D-slices of an MR-sequence using the zeroth order approximation of the brightness constraint equation. A more elaborate treatment, including first order results, can be found in (Niessen et al., ).

Figure 3.2 shows one frame of the 2D-cross-section of the MR-sequence and the output of a purely spatial and spatiotemporal operator. In figure 3.3 we plot the temporal and spatial scale that are selected. The scale selection criterion performs well qualitatively: if edges are present a low spatial scale is selected, assuring an accurate localization. If the edge is moving, as e.g. in the region of interest, the left ventricle, a low temporal scale is selected to avoid motion blur. If at a small scale there is little or no information a larger scale is selected. We therefore get precise estimates if the image information allows it, but do not end up with wrong estimates owing to singularities. Since we usually compute the optic flow only at a discrete number of spatial and temporal scales there still can be regions in which the velocity estimate is poor. This can always be checked by considering the Frobenius norm at the selected scales. Regions in which no reliable velocity estimate can be given within a reasonable spatial and temporal scale range (to be specified for a specific application) should be discarded.

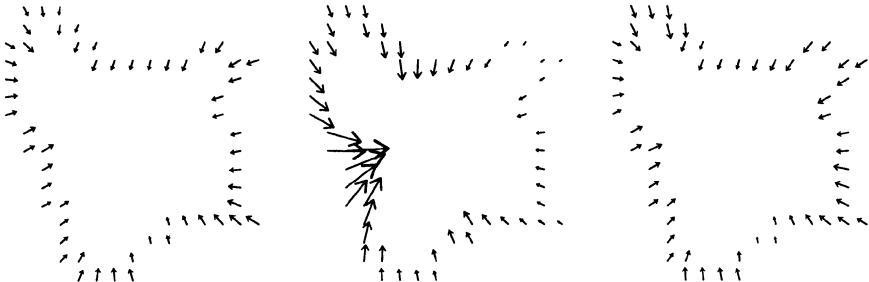


Figure 3.4: Velocity vectors calculated using the zeroth order brightness constraint equation for the eighth time-frame of the MR-sequence (figure 3.2). **Left:** vectors at fixed spatial and temporal scale ( $\sigma_s = 2, \sigma_t = 1$ ). **Middle:** vectors at fixed spatial and temporal scale ( $\sigma_s = 5.656, \sigma_t = 1.414$ ). **Right:** vectors at spatial and temporal scale selected by the scale selection criterion. Upon validating the normal component it follows that the velocity estimate in the left region of the center image is significantly off.

In figure 3.4 we show the estimated velocity fields at the endocardial wall. For a properly selected scale ( $\sigma_s = 2, \sigma_t = 1$ ) the results are qualitatively similar to the results obtained with the scale selection criterion. Quantitative validation studies are required to test improvement in these cases. For wrongly selected fixed scales the estimates are off. From a number of experiments it follows that when using fixed scales the velocity estimate is sometimes unreliable, while the scale selection criterion usually avoids these instances.

### 3.7 Stereo results

The optic flow-based stereo system should be capable of detecting the orientation of an object, provided that both the Lambertian reflection model is met and a good initial guess for the disparity is available. In a previously published study (Maas et al., 1996) these requirements were satisfied and the method was capable to extract the orientation from globally aligned synthetic images, representing a plane.

Applying the method on two real images reveals that the Lambertian reflection model should be met very accurately to get reliable results especially for the derivatives. To correct for different lighting conditions we subtract the respective (local) average from both images (*i.e.* subtracting a considerably smoothed version).

We show the results for a stereo pair of a cardboard house. The left and right images are displayed in figure 3.5. In figure 3.6 we show the obtained orientation for a certain area of the house. The points for which we compute the orientation are automatically chosen by thresholding our selection criterion. Subsequently they are subsampled to be able to visualize the results. The left image of figure 3.6 shows the needles scaled proportionally to the selected scale, whereas the right

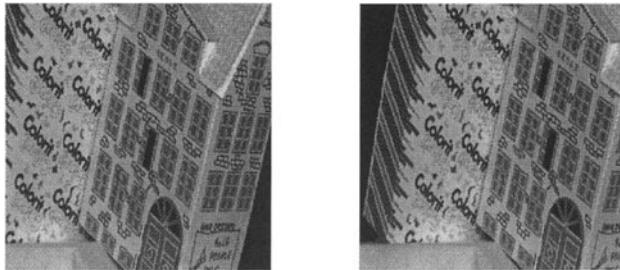


Figure 3.5: Left and right image of the stereo pair for which the surface orientation is computed.

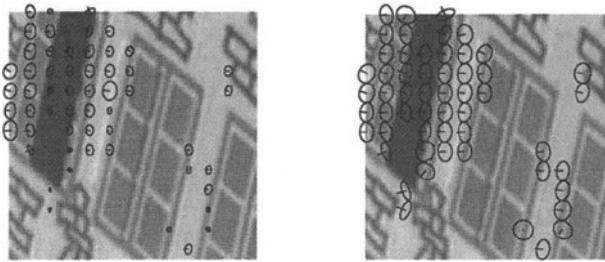


Figure 3.6: Needleplots depicting the orientation of the house in a certain area. The needles in the left image are scaled proportionally to the selected scale, while the needles have a fixed size in the right image. The points where the needles are shown are a subsampled set of automatically selected points, using a threshold on our selection criterion.

image displays the same needles on a fixed size to make some needles better visible. We use 10 exponentially spaced scales between 3 and 20 pixel units. The threshold for the Frobenius norm is set to 0.1. The images that are subtracted from the originals are computed at a scale of 25 pixel units.

## 3.8 Summary

Any image measurement necessarily has to be performed at a certain scale. The conventional image brightness constraint equation, which is often used in optic flow estimation, does not hold for measurements at a certain scale. We therefore apply a new approach which reconciles the brightness constraint equation with linear scale space theory. The new approach conserves the luminance associated with a measurement over a finite region, rather than conserving the luminance associated with a voxel (thereby simulating the physics of measurement).

For optic flow problems the scheme provides a, generally under-determined,

set of equations based on the intrinsic or image-derived knowledge. To single out a unique flow field we have to overcome the *aperture problem*. Unfortunately, more often than not, this is a difficult problem. This explains that many existing approaches implicitly "solve" the aperture problem. In certain situations this will lead to a unique, but wrong velocity field. By strictly separating the *intrinsic* and *extrinsic* optic flow equations we can refine the latter if better models become available. Furthermore, it is possible to have different additional constraints in parts of the image. This is *e.g.* interesting in medical images where different tissues have different dynamic properties. A rough segmentation could already determine whether in certain regions the rigid motion constraint (bone) or the incompressible motion constraint (blood) should be applied.

Binocular stereo can be treated as an optic flow problem provided that two assumptions are met. The reflection is Lambertian and a good initial estimate of the disparity is available. We adapt the optic flow scheme for the stereo case, in which case the additional degrees of freedom are naturally fixed by the epipolar constraint. Using this scheme we can in principle calculate the disparity and its derivatives up to predescribed order, enabling a direct extraction of higher order geometrical structure. However, higher scales are required to extract higher order derivatives with equal precision, thus compromising accuracy. Moreover, a smoothed luminance distribution is more error-prone in algorithms which rely on intensity "matching".

In optic flow and stereo we end up with a linear system of equations. We use both the well-conditionedness of this system of equations and the image content with respect to the noise level in order to select the most appropriate temporal and spatial scales. The Frobenius norm seems a good measure to test these criteria. It avoids the occurrence of singularities and moreover provides a quantitative measure of the reliability of an estimate.

We presented a study of the motion of the left ventricle which suggests that the measurements are more robust if the scale selection criterion is used, but more validation studies are required. After correcting for the violation of the assumptions, good first order results are obtained in automatically selected points of a stereo pair of a cardboard house.

## Acknowledgments

The authors would like to thank Luc Florack and Mads Nielsen for extensive discussions concerning the generalized brightness constraint equation. The 4D MR dataset was kindly made available by Dr. James Duncan, Yale University. Experiments to obtain the data were performed by Dr. Albert Sinusas, his experimental physiology research team and Dr. Todd Constable.

This research was sponsored by the Foundation for Computer Science in the Netherlands (SION) and the Life Sciences Foundation (SLW), which both are subsidized by the Netherlands Organization for Scientific Research (NWO), and by the Netherlands Foundation for Computer Vision Research (SCVR) with financial support from the industrial companies Philips Medical Systems, KEMA, Shell International Exploration and Production, and ADAC Europe.

# **Part II**

# **The Foundation**

## Chapter 4

# On the History of Gaussian Scale-Space Axiomatics

**Joachim Weickert**

*RWCP<sup>1</sup> Novel Function SNN<sup>2</sup> Laboratory,  
Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

**Seiji Ishikawa**

*Department of Control Engineering, Kyushu Institute of Technology,  
Sensuicho 1-1, Tobata, Kitakyushu 804, Japan*

**Atsushi Imiya**

*Department of Information and Computer Sciences,  
Faculty of Engineering, Chiba University,  
1-33 Yayoi-Cho, Inage-Ku 263, Chiba, Japan*

### 4.1 Introduction

A rapidly increasing number of publications, workshops and conferences which are devoted to scale-space ideas confirms the impression that the scale-space paradigm belongs to the challenging new topics in computer vision.

In scale-space theory one embeds an image  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  into a continuous family  $\{T_t f \mid t \geq 0\}$  of gradually smoother versions of it. The original image corresponds to the scale  $t = 0$  and increasing the scale should simplify the image without creating spurious structures. Since a scale-space introduces a hierarchy

---

<sup>1</sup>Real World Computing Partnership

<sup>2</sup>Dutch Foundation for Neural Networks

of the image features, it constitutes an important step from a pixel-related image description to a semantical image description.

In 1983 the scale-space idea became popular by a paper of Witkin (Witkin, 1983). He obtained a scale-space representation by convolution of the original image with Gaussians of increasing width. Koenderink (Koenderink, 1984) pointed out that this *Gaussian scale-space* is equivalent to calculating  $(T_t f)(x)$  as the solution  $u(x, t)$  of the linear diffusion process

$$\begin{aligned}\partial_t u &= \sum_i \partial_{x_i x_i} u =: \Delta u, \\ u(x, 0) &= f(x).\end{aligned}$$

This has been the starting point of a whole field of research on the linear diffusion equation in image processing. Many results have been obtained with respect to axiomatization, differential geometry, deep structure, and applications. This book gives an overview of all these aspects.

Perona and Malik (Perona and Malik, 1990) pioneered the field of nonlinear diffusion processes, where the diffusivity is adapted to the underlying image structure. Many regularized variants of the Perona–Malik filter are well-posed and reveal scale-space properties (Weickert, 1996a; Weickert, 1996b). Other important classes of nonlinear scale-space have been established as well. Some of them are continuous-scale versions of classical morphological processes such as dilation or erosion (Boomgaard, 1992b; Brockett and Maragos, 1994; Jackway and Deriche, 1996), others can be described as intrinsic evolutions of level curves (Alvarez et al., 1993; Kimia et al., 1990; Sapiro et al., 1994; Sapiro and Tannenbaum, 1993). These scale-spaces are generated by nonlinear partial differential equations (PDEs) which are designed to have affine (Alvarez et al., 1993; Sapiro and Tannenbaum, 1993) or projective invariances (Faugeras, 1993; Bruckstein and Shaked, 1993; Sapiro et al., 1994; Dibos, 1995). Overviews of nonlinear approaches can be found in (Haar Romeny, 1994; Weickert, 1996a; Deriche and Faugeras, 1996).

This diversity of scale-space approaches has triggered people to investigate which of these equations can be distinguished in a unique way from others, because they can be derived from first principles (axioms) (Koenderink, 1984; Yuille and Poggio, 1986; Babaud et al., 1986; Lindeberg, 1990; Florack et al., 1992b; Boomgaard, 1992b; Alvarez et al., 1993; Sapiro et al., 1994; Pauwels et al., 1995; Nielsen et al., 1996a; Florack, 1996) (Chapter 6). Apart from a few exceptions (Boomgaard, 1992b; Alvarez et al., 1993; Sapiro et al., 1994), all of these axiomatics use (explicitly or implicitly) one requirement: a linearity assumption. Within such a linear framework it was always possible to derive the Gaussian scale-space as the unique possibility. The fact that many of these approaches have been found recently shows that linear scale-space axiomatic is still a very active field in computer vision.

However, since the linear diffusion equation is well-established in mathematics and physics since Fourier's pioneering work in 1822 (Fourier, 1822), and image processing was already a very active field in the sixties, one might wonder whether the concept of Gaussian scale-space axiomatics is not much older as well. The goal of the present chapter is to address this question. To this end we present two

approaches which are available in Japanese only, but they are older than western ones and they resemble many of these “modern” approaches. Interestingly, the first one of them goes back to 1962, and both have been discovered in the context of optical character recognition (OCR).

The outline of this chapter is as follows: In Section 4.2 we present the basic ideas of a 1-D axiomatic for Gaussian scale-space that has been discovered by Taizo Iijima in 1962 (Iijima, 1962). Section 4.3 describes a 2-D axiomatic which has been found by Nobuyuki Otsu in 1981 (Otsu, 1981). In Section 4.4 we shall relate their results to the well-known linear axiomatics that have been established since 1984. We conclude with a short discussion in Section 4.5. In order to give the reader an impression of the spirit of both Japanese papers, we basically use the same notations as in the original work.

## 4.2 Iijima’s 1-D Axiomatic (1962)

### 4.2.1 Motivation

In 1962 Taizo Iijima wrote a paper in Japanese titled “Basic theory on normalization of a pattern” (Iijima, 1962). His intention was to give a theoretical foundation for character and speech recognition by investigating pattern normalization methods from the viewpoint of observational theory. This approach should also clarify the intrinsic nature of patterns in a mathematical way. For simplicity, he restricts himself to the 1-D case.

### 4.2.2 Axioms

Taizo Iijima considers a transformation  $\Phi$  which depends on a parameter  $\sigma$  and which transforms the original image  $g(x)$  into a blurred version  $f(x)$ . This class of blurring transformations is called “BOKE” (defocusing). He assumes that it has the structure<sup>3</sup>

$$f(x) = \Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} \phi\{g(x'), x, x'\sigma\} dx', \quad (4.1)$$

and that it should satisfy four conditions:

(I) *Linearity (with respect to multiplications):*

If the intensity of a pattern becomes  $A$  times its original intensity, then the same should happen to the transformed pattern:

$$\Phi[Ag(x'), x, \sigma] = A\Phi[g(x'), x, \sigma]. \quad (4.2)$$

(II) *Translation invariance:*

Filtering a translated image is the same as translating the filtered image:

$$\Phi[g(x'-a), x, \sigma] = \Phi[g(x'), x-a, \sigma]. \quad (4.3)$$

---

<sup>3</sup>The variable  $x'$  is a dummy variable.

(III) *Scale invariance:*

If a pattern is spatially enlarged by some factor  $\lambda$ , then there exists a  $\sigma' = \sigma'(\sigma, \lambda)$  such that

$$\Phi[g(x'/\lambda), x, \sigma] = \Phi[g(x'), x/\lambda, \sigma']. \quad (4.4)$$

(IV) *(Generalized) semigroup property:*

If  $g$  is observed at some scale  $\sigma_1$  and this observation is observed under a scale  $\sigma_2$ , then this is equivalent to observing  $g$  under a suitable scale  $\sigma_3$ :

$$\Phi[\Phi[g(x''), x', \sigma_1], x, \sigma_2] = \Phi[g(x''), x, \sigma_3], \quad (4.5)$$

where  $\sigma_3 = \sigma_3(\sigma_1, \sigma_2)$ , but not necessarily  $\sigma_3 = \sigma_1 + \sigma_2$ .

Later on we shall see that – in order to determine the Gaussian uniquely – this axiomatic has to be supplemented with a fifth requirement: preservation of positivity.

### 4.2.3 Consequences

In his Chapter 3 titled “Determination of the function  $\Phi$ ” Iijima establishes in a very systematic way four lemmas which start with the class (4.1) and confine this family by subsequently imposing one more of the conditions (I)–(IV):

(a) **Lemma 1:**

If  $\Phi$  has the structure (4.1) and satisfies axiom I (linearity), then it can be written as the integral

$$\Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} g(x') \phi(x, x', \sigma) dx'. \quad (4.6)$$

(b) **Lemma 2:**

If  $\Phi$  is given by (4.6) and satisfies axiom II (translation invariance), then it can be written as a convolution operation:

$$\Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} g(x') \phi(x - x', \sigma) dx'. \quad (4.7)$$

(c) **Lemma 3:**

If  $\Phi$  is given by (4.7) and satisfies axiom III (scale invariance), then it can be written as

$$\Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} g(x') \phi(\nu(\sigma)(x - x')) \nu(\sigma) dx', \quad (4.8)$$

where  $\nu(\sigma)$  is an arbitrary function of  $\sigma$ .

(d) **Lemma 4:**

If  $\Phi$  is given by (4.8) and satisfies axiom IV (semigroup property), then it can be written either as

$$\Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} g(x') \phi(\nu(\sigma)(x-x')) \nu(\sigma) dx' \quad (4.9)$$

with

$$\phi(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-k^{2m}\xi^{2m} + i\xi u) d\xi \quad (k \in \mathbb{R}, m = 1, 2, \dots), \quad (4.10)$$

or

$$\Phi[g(x'), x, \sigma] \equiv 0. \quad (4.11)$$

The proofs of the first three lemmas are rather short and not very complicated, whereas the longer proof of Lemma 4 involves some more sophisticated reasonings in the Fourier domain.

In a next step Iijima simplifies the result of Lemma 4. The case  $\Phi[g(x'), x, \sigma] \equiv 0$  is of no scientific interest and is not considered any further. In equation (4.9) the function  $\nu$  is eliminated by defining (without loss of generality) a new scale parameter  $\sigma$  via  $\nu(\sigma) = k/\sigma$ . Then the  $k$ -dependence in (4.9) and (4.10) immediately vanishes by means of substitution of variables. The results are summarized in his Theorem 1.2, where Iijima states that  $\Phi$  satisfying (4.1) and the axioms I–IV is given by

$$\Phi[g(x'), x, \sigma] = \int_{-\infty}^{\infty} g(x') \phi_m\left(\frac{x-x'}{\sigma}\right) \frac{dx'}{\sigma} \quad (4.12)$$

with

$$\phi_m(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-\xi^{2m} + i\xi u) d\xi \quad (m = 1, 2, \dots). \quad (4.13)$$

For this family it follows that  $\sigma'$  in (III) becomes  $\sigma' = \sigma/\lambda$ , and  $\sigma_3$  in (IV) satisfies

$$\sigma_3^{2m} = \sigma_1^{2m} + \sigma_2^{2m}. \quad (4.14)$$

For the special case  $m = 1$  equation (4.13) becomes

$$\phi_1(u) = \frac{1}{2\sqrt{\pi}} \exp\left(-\frac{u^2}{4}\right), \quad (4.15)$$

which gives

$$\Phi[g(x'), x, \sigma] = \frac{1}{2\sqrt{\pi}\sigma} \int_{-\infty}^{\infty} g(x') \exp\left(-\frac{(x-x')^2}{4\sigma^2}\right) dx'. \quad (4.16)$$

Thus,  $\Phi[g(x'), x, \sigma]$  is just the convolution between  $g$  and a Gaussian with standard deviation  $\sigma\sqrt{2}$ .

In his Theorem 1.4 he establishes that, if one requires that  $\Phi$  is *positivity preserving*, i.e.

$$\Phi[f(x'), x, \sigma] > 0 \quad \forall f(x) > 0, \quad \forall \sigma > 0, \quad (4.17)$$

then  $m = 1$  arises by necessity. The proof presents an explicit example, where the positivity is not preserved for  $m > 1$ .

This concludes his axiomatic derivation of the Gaussian kernel under five assumptions: linearity, translation invariance, scale invariance, semi-group property, and preservation of positivity.

#### 4.2.4 Further Results

Iijima's axiomatic derivation is available in Japanese only. Nevertheless, there exists an English paper by Iijima, Genchi and Mori (Iijima et al., 1972) from 1972, in which a linear diffusion equation is used as a theoretical model for generalizing figures. It is based on results from a series of eight Japanese papers by Iijima, for which English abstracts are available (Iijima, 1971a; Iijima, 1971d; Iijima, 1971e; Iijima, 1971c; Iijima, 1971f; Iijima, 1971b; Iijima, 1972a; Iijima, 1972b).

In (Iijima et al., 1972) Iijima et al. consider the 2-D case and propose to obtain a generalized image  $f(r, T)$  from an initial image  $f(r)$  by the blurring transformation

$$f(r, T) = \int_{\mathbb{R}^2} G(r - r', T) f(r') dr', \quad (4.18)$$

where  $T$  is a symmetric positive definite matrix and  $G(r, T)$  is the affine Gaussian

$$G(r, T) := \frac{1}{4\pi\sqrt{\det(T)}} \exp\left(-\frac{1}{4}rT^{-1}r\right). \quad (4.19)$$

They choose  $T$  as a function of a blurring parameter  $\tau$  such that  $T(0)$  is the zero matrix, and  $dT/d\tau$  is again symmetric positive definite. Moreover, they note that  $f(r, T(\tau))$  evolves from  $f(r)$  according to the parabolic PDE

$$\frac{\partial f(r, T(\tau))}{\partial \tau} = \operatorname{div}\left(\frac{\partial T(\tau)}{\partial \tau} \nabla f(r, T(\tau))\right). \quad (4.20)$$

Iijima calls this equation the *basic equation of figure*<sup>4</sup>. It seems that he was fully aware of the future importance of his discovery, when he wrote in (Iijima, 1971a) that the “*theory described in this paper will become the basis of the theory of visual pattern recognition*”.

---

<sup>4</sup>Evidently, this approach includes the so-called affine Gaussian scale-space, which results from (4.20) by setting  $T(\tau) := \tau\Lambda$ , where  $\Lambda$  is a symmetric positive definite matrix. If  $\Lambda$  is the unit matrix, we obtain the usual isotropic Gaussian scale-space.

Iijima considered scale-space as a first part of his theory of pattern recognition, which was mostly motivated by character recognition. He was thinking about a direct interface between humans and computers already in the sixties. His view of scale-space analysis in pattern recognition is documented in a Japanese textbook from 1973 (Iijima, 1973), in which even ideas related to wavelet analysis can be found. Apart from this work in pattern recognition, Iijima has also made significant contributions to integral equations and radiation theory. His research on scale-space techniques for OCR is still ongoing (Aoki and Iijima, 1996).

## 4.3 Otsu's 2-D Axiomatic (1981)

### 4.3.1 Derivation of the Gaussian

Another scale-space axiomatic written in Japanese has been established by Nobuyuki Otsu in 1981 (Otsu, 1981). In his Ph.D. thesis he derives the two-dimensional Gaussian scale-space in an axiomatic way by modifying the five Iijima axioms. Section 4.1 of his thesis is titled “Axiomatic derivation of the scale transformation”. There he considers some transformation of an image  $f$  into an image  $\tilde{f}$ , for which the following holds:

- (I) *Representation as a linear integral operator:*

There exists a function  $W : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$  such that

$$\tilde{f}(r) = \int_{\mathbb{R}^2} W(r, r') f(r') dr' \quad \forall r \in \mathbb{R}^2. \quad (4.21)$$

- (II) *Translation invariance:*

For all  $r \in \mathbb{R}^2$  and for all  $a \in \mathbb{R}^2$  it is required that

$$\tilde{f}(r-a) = \int_{\mathbb{R}^2} W(r, r') f(r'-a) dr'. \quad (4.22)$$

Since this is just  $\int_{\mathbb{R}^2} W(r, r'+a) f(r') dr'$ , and (I) states that  $\tilde{f}(r-a) = \int_{\mathbb{R}^2} W(r-a, r') f(r') dr'$ , it follows that the integral kernel is symmetric,

$$W(r, r'+a) = W(r-a, r'), \quad (4.23)$$

and, thus, it is a convolution kernel:

$$W(r, r') = W(r - r'). \quad (4.24)$$

- (III) *Rotation invariance (of the kernel):*

For all rotation matrices  $T_\Theta$  and for all  $r = (x, y)^T \in \mathbb{R}^2$  it is assumed that

$$W(T_\Theta r) = W(r). \quad (4.25)$$

Hence,  $W$  depends only on  $|r|$ :  $W(r) = W(x^2 + y^2)$ .

(IV) *Separability:*

There exists a function  $u : \mathbb{R} \rightarrow \mathbb{R}$  such that

$$W(r) = u(x)u(y). \quad (4.26)$$

Combining this with (III) implies after elementary manipulations that

$$W(r) = k \exp [c(x^2 + y^2)]$$

with some parameters  $k, c \in \mathbb{R}$ . In order to get  $k > 0$  and  $c < 0$ , however, additional constraints are needed.

- (V) His next requirement which he names “Normalization of energy” actually consists of two parts:

*Perservation of nonnegativity,*

$$\tilde{f}(r) \geq 0 \quad \forall f(r) \geq 0, \quad (4.27)$$

and *average grey level invariance*,

$$\int_{\mathbb{R}^2} \tilde{f}(r) dr = \int_{\mathbb{R}^2} f(r) dr. \quad (4.28)$$

This leads to  $W(r) \geq 0$  and  $\int_{\mathbb{R}^2} W(r) dr = 1$ , respectively.

Combining these results gives  $k = \frac{1}{2\pi\sigma^2}$  and  $c = -\frac{1}{2\sigma^2}$ . This yields the Gaussian kernel

$$W(r) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right) \quad (4.29)$$

and concludes the axiomatic derivation of the 2-D linear scale-space.

### 4.3.2 Further Results

Section 4.2 of Otsu’s thesis is titled “Representation of scale-space transformation and semigroup”. It is devoted to the  $N$ -dimensional Gaussian scale-space. With  $\rho := \sigma^2/2$  he defines

$$T(\rho)f(r) := \frac{1}{(4\pi\rho)^{N/2}} \exp \left( -\frac{|r|^2}{4\rho} \right) * f(r). \quad (4.30)$$

Using Fourier techniques he shows that the generator of the scale-space transformation is the Laplacean:

$$\tilde{f}(r, \rho) = T(\rho)f(r) = \exp(\rho\Delta)f(r). \quad (4.31)$$

This gives

$$\frac{\partial \tilde{f}(r, \rho)}{\partial \rho} = \Delta (\exp(\rho\Delta)f(r)) = \Delta \tilde{f}(r, \rho).$$

Thus,  $\tilde{f}$  satisfies the isotropic linear diffusion equation.

The formal inversion of the scale-space transformation by means of (4.31) is

$$f(r) = [T(\rho)]^{-1} \tilde{f}(r, \rho) = \exp(-\rho\Delta) \tilde{f}(r, \rho) = \left( I - \rho\Delta + \frac{\rho^2}{2} \Delta^2 - \dots \right) \tilde{f}(r, \rho).$$

For the case that  $\rho$  or  $\Delta^2 \tilde{f}$  is small, Otsu proposes to approximate  $[T(\rho)]^{-1}$  by  $[I - \rho\Delta]$  and to use it for recovering the original image from a blurred one<sup>5</sup>.

## 4.4 Relation to Other Work

Having sketched the basic ideas of these Japanese axiomatics, it is natural to ask about similarities and differences to other approaches. Table 4.1 gives an overview of the current axiomatics for the continuous Gaussian scale-space. These axioms and some of their relations can be explained as follows<sup>6</sup>:

- **Convolution kernel:**

There exists a family of functions  $\{k_t : \mathbb{R} \rightarrow \mathbb{R} \mid t \geq 0\}$  such that

$$(T_t f)(x) = \int_{\mathbb{R}^N} k_t(x - x') f(x') dx'.$$

In Section 4.3.1 we have already seen that this property can be derived from the two assumptions:

- **Linear integral operator:**

There exists a family of kernels  $\{k_t \mid t \geq 0\}$  with

$$(T_t f)(x) = \int_{\mathbb{R}^N} k_t(x, x') f(x') dx'.$$

Since every continuous linear functional can be written as an integral operator, it follows that Florack's topological duality paradigm (Florack, 1996) can also be interpreted as requiring the existence of a linear integral operator<sup>7</sup>.

- **Translation invariance:**

Let a translation  $\tau_a$  be defined by  $(\tau_a f)(x) := f(x - a)$ . Then,

$$\tau_a T_t = T_t \tau_a \quad \forall a \in \mathbb{R}^N, \quad \forall t > 0.$$

Since usually linearity and translation invariance are imposed in conjunction, we have summarized them under the term "convolution kernel".

---

<sup>5</sup>One should keep in mind that this is an ill-posed problem which may lead to unstable results.

<sup>6</sup>Of course, such a table can only give a "flavour" of the different approaches, and the precise description of each axiom may slightly vary from paper to paper. Several relations between the presented axioms are discussed in (Alvarez et al., 1993; Pauwels et al., 1995) and Chapter 6.

<sup>7</sup>Dirac point distributions and their derivatives are admitted as "functions under the integral".

- **Semigroup property:**

$$T_{t+s}f = T_t(T_s f) \quad \forall t, s \geq 0, \quad \forall f.$$

This property ensures that one can implement the scale-space process as a cascade smoothing which resembles certain processes of the human visual system.

- **Locality:**

For small  $t$  the value of  $T_t f$  at any point  $x$  is determined by its vicinity:

$$\lim_{t \rightarrow 0^+} (T_t f - T_t g)(x) = o(x)$$

for all  $f, g \in C^\infty$  whose derivatives of order  $\geq 0$  are identical.

- **Regularity:**

A precise definition of the smoothness requirements for the scale-space operator depends on the author:

- Since the original image creates the scale-space, it is natural to assume that it is continuously embedded, i.e.  $\lim_{t \rightarrow 0^+} T_t = I$ . In the linear convolution case, this means that  $k_t(x)$  tends to Dirac's delta distribution (Yuille and Poggio, 1986) and its Fourier transform becomes 1 everywhere (Florack et al., 1992b).
- Babaud et al. (Babaud et al., 1986) and Florack (Florack, 1996) consider infinitely times differentiable convolution kernels which are rapidly decreasing functions in  $x$ , i.e. they are vanishing at  $\infty$  faster than any inverse of polynomials.
- Lindeberg uses kernels  $k_t$  which are Borel measurable in  $t$  (Lindeberg, 1990), or kernels which converge for  $t \rightarrow 0^+$  in the  $L^1$  norm to the Dirac distribution (Chapter 6).
- Alvarez et al. (Alvarez et al., 1993) require that

$$\|T_t(f + hg) - (T_t(f) + hg)\|_\infty \leq Cht$$

for all  $h, t \in [0, 1]$ , and for all smooth  $f, g$ , where  $C$  may depend on  $f$  and  $g$ .

- Pauwels et al. (Pauwels et al., 1995) assume that the convolution kernel  $k_t(x)$  is separately continuous in  $x$  and in  $t$ .

- **Infinitesimal generator:**

The existence of

$$\lim_{t \rightarrow 0^+} \frac{T_t f - f}{t} =: A[f]$$

guarantees that the semigroup can be represented by the evolution equation

$$\partial_t u = A[u].$$

The existence of an infinitesimal generator follows from the semigroup property when being combined with regularity assumptions (Alvarez et al., 1993).

- **Causality:**

The scale-space evolution should not create new level curves when increasing the scale parameter. If this is satisfied, iso-intensity linking through the scales is possible and a structure at a coarse scale can (in principle) be traced back to the original image.

For this reason, Koenderink (Koenderink, 1984) required that at spatial extrema (with nonvanishing determinant of the Hessian) isophotes in scale-space are upwards convex; In 2-D he showed that at these extrema the diffusion equation

$$\partial_t u = \alpha(x, t) \Delta u \quad (4.32)$$

has to be satisfied. Hereby,  $\alpha$  denotes a positive-valued function.

Hummel (Hummel, 1986) established the equivalence between causality and a maximum principle for certain parabolic operators.

We may also derive the causality equation (4.32) and its  $N$ -dimensional generalizations by requiring that local extrema with positive or negative definite Hessians are not enhanced (Babaud et al., 1986) (Chapter 6). This assumption states that such an extremum in  $x_0$  at scale  $t_0$  satisfies

$$\begin{aligned} \partial_t u > 0 &\quad \text{if } x_0 \text{ is a minimum,} \\ \partial_t u < 0 &\quad \text{if } x_0 \text{ is a maximum.} \end{aligned}$$

This is just the causality requirement  $\text{sign}(\partial_t u) = \text{sign}(\Delta u)$ . Moreover, in 1-D, nonenhancement of local extrema is equivalent to the requirement that the number of local extrema does not increase (Babaud et al., 1986; Lindeberg, 1990). In higher dimensions, however, diffusion scale-spaces may create new extrema, see e.g. (Yuille and Poggio, 1986; Lifshitz and Pizer, 1990; Damon, 1995a).

- **Nonnegativity:**

If the nonnegativity of the convolution kernel,

$$k_t(x) \geq 0 \quad \forall x, \quad \forall t > 0,$$

is violated, new level crossings may appear for  $t > 0$ , such that the causality property does not hold.

Within a linear framework with spatially continuous convolution kernels, nonnegativity is equivalent to the monotony requirement (Alvarez et al., 1993)

$$f(x) \leq g(x) \quad \forall x \implies (T_t f)(x) \leq (T_t g)(x) \quad \forall x, \quad \forall t > 0$$

and the preservation of nonnegativity:

$$f(x) \geq 0 \quad \forall x \implies (T_t f)(x) \geq 0 \quad \forall x, \quad \forall t > 0.$$

- **Tikhonov regularization:**

In the 1-D case,  $u$  is called a Tikhonov regularization of  $f \in L^2(\mathbb{R})$ , if it minimizes the energy functional

$$E_f[u] = \int_{\mathbb{R}} \left[ (f - u)^2 + \sum_{i=1}^{\infty} \lambda_i \left( \frac{d^i u}{dx^i} \right)^2 \right] dx \quad (\lambda_i > 0).$$

This concept and an  $N$ -dimensional generalization has been used by Nielsen, Florack and Deriche (Nielsen et al., 1996a). The first term under the integral ensures that  $u$  remains close to  $f$ , while the second one is responsible for the smoothness of  $u$ .

- **Normalization:**

The average grey level invariance

$$\int_{\mathbb{R}^N} T_t f \, dx = \int_{\mathbb{R}^N} f \, dx \quad \forall t > 0$$

boils down to the normalization condition

$$\int_{\mathbb{R}^N} k_t(x) \, dx = 1,$$

if we consider linear convolution kernels.

In this case normalization is also equivalent to grey level shift invariance (Alvarez et al., 1993):

$$\begin{aligned} T_t(0) &= 0, \\ T_t(f + C) &= T_t(f) + C \end{aligned}$$

for all images  $f$  and for all constants  $C$ .

- **Flat kernel for  $t \rightarrow \infty$ :**

For  $t \rightarrow \infty$ , one expects that the kernel spreads the information uniformly over the image. Therefore, if the integral over the kernel should remain finite, it follows that the kernel has to become entirely flat:  $\lim_{t \rightarrow \infty} k_t(x) = 0$ .

- **Isometry invariance:**

Let  $R \in \mathbb{R}^N$  be an orthogonal transformation (i.e.  $\det R = \pm 1$ ) and define  $(Rf)(x) := f(Rx)$ . Then,

$$T_t(Rf) = R(T_t f) \quad \forall f, \quad \forall t > 0.$$

In the 1-D case with a linear convolution kernel this invariance under rotation and mirroring comes down to the symmetry condition  $k_t(x) = k_t(-x)$ .

- **Homogeneity and isotropy:**

Koenderink (Koenderink, 1984) required that the scale-space treats all spatial points equally. He assumed that the diffusion equation (4.32), which results at extrema from the causality requirement, should be the same at each spatial position (regardless whether there is an extremum or not) and for all scales. He named these requirements homogeneity and isotropy<sup>8</sup>.

- **Separability:**

The convolution kernel  $k_t(x)$  with  $x = (x_1, \dots, x_N)^T \in \mathbb{R}^N$  may be split into  $N$  factors, each acting along one coordinate axis:

$$k_t(x) = k_{1,t}(x_1) \cdots k_{N,t}(x_N).$$

- **Scale invariance:**

Let  $(S_\lambda f)(x) := f(\lambda x)$ . Then there exists some  $t'(\lambda, t)$  with

$$S_\lambda T_{t'} = T_t S_\lambda.$$

One may achieve this by requiring that, in the  $N$ -dimensional case, the convolution kernel  $k_t$  has the structure

$$k_t(x) = \frac{1}{\Psi^N(t)} \Phi \left( \frac{x}{\Psi(t)} \right)$$

with a continuous, strictly increasing rescaling function  $\Psi$ . This means that all kernels can be derived by stretching a parent kernel such that its area remains constant (Pauwels et al., 1995). It is evident that this is related to the normalization condition.

Scale invariance follows also from the semigroup property when being combined with isometry invariance and causality (Chapter 6). Moreover, scale invariance, translation invariance and isometry invariance result from the more general assumption of invariance under the spacetime symmetry group; see (Florack, 1996) for more details.

We observe that – despite the fact that all presented axiomatics use many similar requirements – not two of them are identical. Each axiomatic confirms and enhances the evidence that the others give: that Gaussian scale-space is unique within a linear framework. This theoretical foundation is the backbone of a lot of successful applications of linear scale-space theory.

Nevertheless, apart from their historical merits, the early Japanese approaches differ from the well-known axiomatics after 1984 in several aspects which make them worth being studied:

Firstly, it is interesting to note that both Japanese axiomatics require only quite a few axioms in order to derive Gaussian scale-space. Even recent approaches which intend to use a minimal set of first principles do not utilize less axioms.

---

<sup>8</sup>In our terminology, homogeneity and isotropy are much stricter requirements than translation and isometry invariance. They enable Koenderink to derive Gaussian scale-space under only one additional assumption (causality).

Iijima's one-dimensional framework is very systematic and rather modern: principles such as the semigroup property are typical for axiomatizations after 1990, and also the importance of scale invariance has been emphasized mainly in recent years (Florack et al., 1992b; Florack et al., 1994c; Pauwels et al., 1995; Florack, 1996).

Otsu's two-dimensional axiomatic is very appealing due to its simplicity: in contrast to many other approaches it does not require advanced mathematical techniques such as Fourier analysis, complex integrals, or functional analysis in order to derive the uniqueness of the Gaussian kernel. It is therefore a well-suited approach even for undergraduate courses in image processing.

## 4.5 Discussion

We have presented two axiomatics for the linear diffusion scale-space that have been unknown in the western image processing world<sup>9</sup>.

The fact that the vision community was not aware of both axiomatics had two reasons: firstly, they were written in Japanese and secondly, they came too early to be appreciated. Computer technology in the sixties was not capable of fully applying scale-space ideas to real-world problems, and so they fell into oblivion.

The last point seems to be not completely untypical for the field of partial differential equations in image processing. Another example is the fact that already in 1965 the Nobel prize winner Dennis Gabor – the inventor of optical holography and the so-called Gabor functions – proposed a deblurring algorithm based on combining mean curvature flow with backward smoothing along flowlines (Gabor, 1965; Lindenbaum et al., 1994). This long-time forgotten method is similar to modern PDE techniques for image enhancement.

Maybe this review helps a little bit so that the pioneering work of these people receives the acknowledgement that it deserves.

**Acknowledgements.** We thank Luc Florack and Alfons Salden for useful comments on a draft version of this work.

---

<sup>9</sup>To our knowledge, (Weickert et al., 1996) is the first citation of Iijima's work in the western scale-space community.

Table 4.1: Overview of continuous Gaussian scale-space axiomatics (I = Iijima (Iijima, 1962), O = Otsu (Otsu, 1981), K = Koenderink (Koenderink, 1984), Y = Yuille/Poggio (Yuille and Poggio, 1986), B = Babaud et al. (Babaud et al., 1986), L1 = Lindeberg (Lindeberg, 1990), F1 = Florack et al. (Florack et al., 1992b), A = Alvarez et al. (Alvarez et al., 1993), P = Pauwels et al. (Pauwels et al., 1995), N = Nielsen et al. (Nielsen et al., 1996a), L2 = Lindeberg (Chapter 6), F2 = Florack (Florack, 1996)).

	I	O	K	Y	B	L1	F1	A	P	N	L2	F2
convolution kernel	•	•		•	•	•	•	•	•		•	•
semigroup property	•					•	•	•	•	•	•	•
locality								•				
regularity				•	•	•	•	•	•		•	•
infinitesimal generator										•		
causality			•	•	•	•					•	
nonnegativity	•	•						•	•			•
Tikhonov regularization										•		
normalization		•			•	•		•	•			
flat kernel for $t \rightarrow \infty$				•			•					
isometry invariance		•		•	•	•	•	•	•	•	•	•
homogeneity & isotropy			•									
separability		•					•					
scale invariance	•			•	•		•		•	•		•
valid for dimension	1	2	1,2	1,2	1	1	> 1	N	1,2	N	N	N

# Chapter 5

# Scale-Space and Measurement Duality

Luc Florack

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

## 5.1 Introduction

There exists remarkably little debate about an *operational definition* of an “image” beyond computational details; the concept is usually identified with that of a “raw image” (“pixels” plus “header”, say). Raw images are normally represented as real valued functions of spacetime<sup>1</sup>, possibly discrete or quantised. However, serious conceptual and computational problems may arise if this representation is taken for granted:

- The representation is “too close to the machine” for high-level purposes; conceptually irrelevant technicalities such as quantisation and discretisation are explicit rather than “encapsulated”. This holds both for the image’s pixel format  $f_i$  as well as for any “interpolated function”  $f(x)$  intended to approximate it.
- Differentiation is ill-posed, unless one adopts an unrealistic function topology; one should not want to approximate ill-posed operators.
- Whereas raw data  $f_i$  are biased towards digitization details, the continuum form  $f(x)$  is a physical non-entity. In order to obtain numeric samples one needs a measurement<sup>2</sup> device (*detector* or *filter*), as well as a *duality principle*.

---

<sup>1</sup>“Spacetime” refers to space and/or time henceforth; the term “image” accordingly covers static images, temporal signals, as well as video sequences.

<sup>2</sup>I take “measurement” in the broad sense of an *assessment* of physical evidence.

*ple*, explaining how such samples are produced given a source-detector pair. Neither detector nor duality paradigm is deductible from raw data; they must be *postulated*.

- Finally, since an image is a *coherent* set of samples, one needs to endow it with a *spacetime topology*. This requires a *spacetime model* to be incorporated *a priori*. A “canonical” choice of sampling grid should reflect this spacetime model.

In Section 5.2 it is shown that conceptual and computational problems can be overcome by encapsulating the raw image as “read-only” data into a generalised image concept that explicitly incorporates duality, detector paradigm, and spacetime model. The resulting “pseudo-static” representation, in which samples are attributed to fixed base points in spacetime, is then reformatted into a “kinematic” one, in which orbits of invariant samples are defined by a so-called gauge field of infinitesimal generators. This “syntactical optic flow” model is explained in Section 5.3.

## 5.2 Pseudo-Static Image Concept

The approach is as follows. Firstly, I consider the duality paradigm; there is no compelling argument—at least at this level of abstraction—to single out a unique choice. I pick a suitable one and comment briefly on an alternative. Secondly, I decide on the class of detectors. In view of generality I will take it to be such that it is capable of segregating a large variety of source fields, more than one will ever need in practice. The drawback, as one might anticipate, is that it is *too* large for actual implementation. More structure is needed. I will point out an *image processing consistency requirement*, which requires an *algebraic structure* on any admissible class of detectors. The class can now be narrowed down with the help of a *point operator* compatible with the algebraic constraint. It will turn out that there is actually no leeway: the point operator is uniquely determined. At this point I have a precise and practicable definition of a *local sample*. Finally, an image can be defined by combining local samples with a spacetime model. At this stage there are again some options, out of which I take a plausible one.

It should be stressed that wherever theory does not compel us to make definite choices, the options at hand must be inspired by applications or *a priori* knowledge. An abstract approach is preferred if one cannot or does not want to anticipate final purpose, and thus has to keep an open mind regarding all possible options. By making certain choices I have chosen a compromise between flexibility and specificity. This is, however, merely for the sake of clarity. The reader is encouraged to abstract from the somewhat arbitrary choices, and to fill in the missing details of alternative ones; relations to and deviations from other models than the one pursued here are to be found “upstream” in the same general framework.

Let us turn to the details. I will collectively denote the set of all possible raw images by  $\Sigma$  and refer to it as *state space*, and the set of all admissible filters by

$\Delta$ , or *device space* (cf. Koenderink's *sensorium* (Koenderink and Doorn, 1986b; Koenderink and Doorn, 1987)). Instead of postulating a "direct" model for  $\Sigma$ , I define it as "probes of  $\Delta$ ", or "the dual of  $\Delta$ ". The reason for this is that "degrees of freedom" of a source field have no operational significance other than by way of interaction with a fiducial set of detectors ("what-you-see-is-all-you've-got").

There is more than one way to fill in "probes of  $\Delta$ ". One possibility is as follows.

**Paradigm 5.2.1 (Duality).** State space is the topological dual of device space:

$$\Sigma \stackrel{\text{def}}{=} \Delta'.$$

The topological dual of a topological vector space—indicated by the prime symbol—is the space of all real or complex valued, linear functions on that vector space, also called *linear forms*, *covectors*, or, in case of function spaces, *linear functionals*, which are *continuous*. In other words, instead of functions I consider functionals in order to model measurements. For all practical purposes such a functional can be represented in integral form as

$$F[\phi] = \int dz f(z) \phi(z), \quad (5.1)$$

in which  $f$  is now *any* plausible representation of the raw image.

Other choices are possible. One may for example replace (linear) topological duality by (nonlinear) morphological duality, defined by using probes of the type

$$F[\phi] = \sup_{z \in \mathbb{R}^n} \{f(z) + \phi(z)\}. \quad (5.2)$$

In the usual terminology,  $\phi$  is referred to as the *structuring element*. The analogy with topological duality—detailed below—can be pursued a long way, particularly if one generalises the supremum operator to a set-valued operator that returns the operand's stationary values (Boomgaard, 1992b; Dorst and Boomgaard, 1994a). (Again, the key to understand this is abstraction; *don't commit yourself!*) For more details on this morphological analogon the reader is referred to Van den Boomgaard's chapter, Chapter 15.

The following class of detectors is plausible in the context of Paradigm 5.2.1.

**Paradigm 5.2.2 (Device Space).** Let  $\mathcal{S}(\mathbb{R}^n)$  be the class of smooth functions of rapid decay, then

$$\Delta \stackrel{\text{def}}{=} \mathcal{S}(\mathbb{R}^n) \quad \text{whence} \quad \Sigma \stackrel{\text{def}}{=} \mathcal{S}'(\mathbb{R}^n).$$

It suffices here to appreciate the general features. Any filter  $\phi \in \mathcal{S}(\mathbb{R}^n)$  is smooth and "essentially compact" in the sense that all of its derivatives decrease faster as  $\|x\| \rightarrow \infty$  than any polynomial. Filter confinement is natural, since the aim is to define *local* samples. The mappings  $F \in \mathcal{S}'(\mathbb{R}^n)$  are known in the literature as *tempered distributions*. Details of this are the subject of the theory of Schwartz (Schwartz, 1951). Perhaps the most convincing argument, which at the

same time removes a widespread misconception, is that filter smoothness is a *very weak* constraint, because it is the dual space that represents our image samples. In fact,  $\mathcal{S}'(\mathbb{R}^n)$  is *much larger* than general function spaces typically employed in non-dualistic models of state space, such as  $L^1(\mathbb{R}^n)$ ! Despite this, tempered distributions are *infinitely differentiable* in a *well-posed* sense. Further motivations can be found in (Florack, 1995), as well as in the next section, were I shall be needing derivatives.

Clearly  $\mathcal{S}(\mathbb{R}^n)$  is too large even for an approximate implementation; if that is the goal, one needs some suitably constrained subspace. The essence of image processing (“image in/image out”) is the ability to cascade observations.

**Paradigm 5.2.3 (Image Processing).** A cross-section of local samples,

$$F \star \phi(x) = \int dz f(z + x)\phi(z),$$

is itself a source field that can be sampled in a cascade fashion.

The raw image may serve as an entry point to such a cascade. (Morphological counterpart of the right hand side of Equation 5.2 (Boomgaard, 1992b):  $\sup_{z \in \mathbb{R}^n} \{f(z + x) + \phi(z)\}$ , cf. Chapter 15.)

Paradigm 5.2.3 calls for additional structure on device space, known as an *algebra*. It can be shown that, together with the previous paradigms, it implies the following

**Corollary 5.2.4 (Consistency Requirement).** Device space  $\Delta$  must form a linear convolution algebra.

In other words one must satisfy the closure property  $\phi, \psi \in \Delta \Rightarrow \phi * \psi \in \Delta$ , the only nontrivial algebraic requirement (morphologically:  $\phi, \psi \in \Delta \Rightarrow \phi \oplus \psi \in \Delta$ ).

If one postulates the existence of a *point operator* as a positive filter,  $\phi \in \mathcal{S}^+(\mathbb{R}^n)$  say, then consistency with Corollary 5.2.4 demands that it generates an *autoconvolution algebra*. There is no choice: a point operator *must be a normalised Gaussian*<sup>3</sup>. (When pursuing Equation 5.2 instead, similar arguments will produce the *quadratic structuring function* as the morphological counterpart (Boomgaard, 1992b).)

Thus postulating the existence of a point operator leads to well-known scale-space theory. In turn it induces a complete family of derivative filters known as the *Gaussian family* (Koenderink and Doorn, 1990), notation:  $\Delta = \mathcal{G}(\mathbb{R}^n)$ .

**Paradigm 5.2.5 (Scale-Space).** The scale-space paradigm is defined by

$$\Delta \stackrel{\text{def}}{=} \mathcal{G}(\mathbb{R}^n) \quad \text{whence} \quad \Sigma \stackrel{\text{def}}{=} \mathcal{G}'(\mathbb{R}^n).$$

It is understood that one takes this as a *local* paradigm; it defines local samples, not images.

The missing link is a spacetime model. One usually states spacetime models in terms of *symmetries* that appear to govern the spatiotemporal behaviour of

---

<sup>3</sup>This is quite easily verified in the Fourier domain.

a system (Friedman, 1983; Misner et al., 1973). These in turn are formulated as transformations of “empty” spacetime, but in reality one always transforms *physical objects* relative to each other, *in casu* sources and detectors.

**Definition 5.2.6 (Push Forward).** Let  $\theta : M \rightarrow M : x \mapsto \theta(x)$  be a smooth spacetime automorphism. The push forward of a filter is then defined as the mapping

$$\theta_* : \Delta_x \rightarrow \Delta_{\theta(x)} : \phi \mapsto \theta_* \phi \stackrel{\text{def}}{=} J_{\theta^{-1}} \phi \circ \theta^{-1},$$

with Jacobian determinant  $J_\chi \equiv |\det \nabla \chi|$ .

The Jacobian takes care that filter normalisation is preserved despite transformations; the filter is said to “transform as a density” (a density is the geometrical object corresponding to an “ $n$ -form” (Spivak, 1965; Spivak, 1975)). One naturally “pulls back” the source field in the dual view.

**Definition 5.2.7 (Pull Back).** With the automorphism  $\theta$  and its push forward  $\theta_*$  as defined in Definition 5.2.6, the pull back of the source is defined as the mapping

$$\theta^* : \Sigma_{\theta(x)} \rightarrow \Sigma_x : F \mapsto \theta^* F \quad \text{defined by} \quad \theta^* F[\phi] \stackrel{\text{def}}{=} F[\theta_* \phi].$$

In function form:  $\theta^* f \equiv f \circ \theta$ , also referred to as “scalar field transformation”. A cross-section as defined in Paradigm 5.2.3 is an instance of a “carry-along”—push forward or pull back, depending on your view—of a local measurement sample under the translation group, which reflects homogeneity of spacetime.

The choice of group is to some extent optional, though you will agree that it must act transitively on spacetime. A natural one is the classical, *Newtonian spacetime group*, comprising spatial rotations (isotropy), spatiotemporal translations (homogeneity), and in addition one spatial and one temporal scaling (scale invariance). It should be obvious that, if one does not want to account for time (static images), one considers the the spatial subgroup, i.e. the *scale-Euclidean group*. In any case *multiple scales must be available* in order not to violate the image processing demand of Corollary 5.2.4; once given a single point operator, one must extend the spacetime group beyond mere translations. The pseudo-isotropic<sup>4</sup> Gaussian naturally generates a scale-spacetime stratification as the “default structure” upon minimal extension with the scaling group. It is important to note that this still leaves various possibilities, because the carry-along of physical objects depends on system-specific parameters, such as foveal region, present moment (Koenderink, 1988), which are susceptible to spacetime transformations. These nontrivial cases are beyond the scope of this chapter. Suffice it to conclude that if no such parameters are present, then Definition 5.2.6 or 5.2.7 combined with all paradigms constrains the point operators  $\mathcal{G}^+(\mathbb{R}^n)$  to Gaussian correlation filters (quadratic structuring functions for dilation respectively). This finally establishes an operational *image definition* in the context of linear (morphological) image processing.

---

<sup>4</sup>In Newtonian spacetime one may opt for coordinates  $(ct; \vec{x})$  that make things look isotropic;  $c$  is a free parameter relating the spatial and temporal scales  $\sigma, \tau$ :  $2s = \sigma^2 = c^2 \tau^2$ .

**Definition 5.2.8 (Image).** An image  $I$  is defined by a triple  $\{F, \phi, \Theta\}$ , comprising raw image data  $F \in \Sigma$ , a point prototype  $\phi \in \Delta$ , and a transitive spacetime symmetry group  $\Theta$ , together with a duality paradigm  $\Sigma = \text{dual } \Delta$ :

$$I \stackrel{\text{def}}{=} \langle F | \Theta | \phi \rangle \stackrel{\text{def}}{=} \{ \langle F | \theta | \phi \rangle \mid \theta \in \Theta \} ,$$

in which  $\langle F | \theta | \phi \rangle = \theta^* F[\phi] = F[\theta_* \phi]$ .

(A particular instance is obtained if  $\text{dual } \Delta$  is interpreted as  $\Delta'$ , according to Paradigm 5.2.1.) One may recognise  $\theta_* \phi$  as the (linear, respectively morphological) *propagator* of a partial differential equation, the isotropic diffusion equation

$$\partial_s u = \Delta u , \quad (5.3)$$

in the scrutinised case of topological duality, and

$$\partial_s u = \|\nabla u\|^2 , \quad (5.4)$$

in the morphological case briefly touched upon, the initial condition of which is given by the raw image  $F$  in both cases.

### 5.3 Kinematic Image Concept

The spacetime representation of an image as explained in the previous section could be called “pseudo-static” in the sense that it does not explicitly account for *kinematic* relations between local image samples. Such relations naturally arise as a consequence of apparent conservation laws. For this reason the concept of “optic flow” has been introduced by Gibson in the context of optical pilot navigation (Gibson, 1950).

One thinks of optic flow as a vector field. This reflects the desire to link corresponding points, or rather local samples—whatever these may be—separated by arbitrarily small temporal intervals. The motivation for this is of course that in the physical world such pointwise connections are actually meaningful; ideally they correspond to the true motion of material points on a physical body (or to wave phenomena).

However, it is commonly known that the induced image flow is intrinsically ambiguous due to the so-called “*aperture problem*” (Marr, 1982); one may seek to solve for the *homotopy* that links spatial iso-grey-level contours over time *as a whole*, but one cannot hope to establish any *pointwise* connections between them. A “canonical” representation of this homotopy is *normal flow*, sometimes called the optic flow field (to be distinguished from the physically induced image velocity or *motion field*) (Hildreth, 1984; Horn, 1986).

Another way of expressing the intrinsic ambiguity is to say that optic flow theory is a *gauge theory*. A gauge theory is a theory characterised by a local invariance. In the case at hand, the optic flow field can be regarded as the *gauge field*, since it contains local degrees of freedom (“tangential flow”) that do not manifest themselves in any observable way, and hence can be fixed arbitrarily.

However, an arbitrary *gauge condition* would in general single out a unique optic flow field that lacks any physical meaning. Physical considerations beyond the evidence contained in the data, i.e. the details of *image formation* and *task*, should be taken into account in order to arrive at a sensible choice of gauge (Hildreth, 1984; Niessen et al., 1995). Examples are rigidity or non-elasticity constraints for solid objects (D'Haeyer, 1986; Faugeras, 1994), (in)compressibility and continuity conditions for fluids (Amini, 1994; Devlaminck and Dubus, 1996), assertions about motion being induced by projection of smooth surface patches, etc. (Hildreth, 1984; Horn and Schunck, 1981; Tretiak and Pastor, 1984; Verri et al., 1990), or a simple geometric one, such as the normal flow constraint. Smoothness constraints are often used to fix the gauge. Typically neither one of these constraints is globally valid; the gauge is confounded with local semantics. Its validation will therefore typically require a hermeneutic kind of procedure (accept hypotheses that improve the global coherence of their semantical context). A disambiguated instance of optic flow could be called *semantical optic flow*.

In this section, the term optic flow will be assumed to entail essentially all possible flows as far as these are compatible with the data, in other words, optic flow equals normal flow modulo arbitrary tangential flow (Horn and Schunck, 1993). In this terminology, normal flow can be regarded as a canonical representation of optic flow (subject to a “canonical gauge condition” that nullifies tangential flow). I will furthermore assume that the data  $f$  represent a *scalar field*. This case is to be distinguished from that of a *density field*, which can be easily accounted for by slight adaptations of the following theory.

There are many possible approaches to optic flow measurement (Barron et al., 1994; Koenderink, 1986). Horn and Schunck's “Optic Flow Constraint Equation” (Arnspang, 1988a; Arnspang, 1988b; Arnspang, 1991; Arnspang, 1993; Barron et al., 1994; Fennema and Thompson, 1979; Horn and Schunck, 1981; Otte and Nagel, 1994; Schunck, 1984; Werkhoven, 1990) constrains the optic flow vector by means of a conservation principle; invariant grey-values are attributed to points which are dragged along the flow. Although it is quite natural to relate flow to conservation<sup>5</sup>, the traditional view is problematic, since it bypasses the role of a filter paradigm in the realisation of a grey-value sample, as explained in the previous section. The traditional OFCE, even when disambiguated by complementary gauge conditions, is itself not a recipe for actual optic flow computation; it is only a constraint on the (possibly unique) solution. In this section I explicitly account for *duality* between source data (pixels) and filter paradigm, *in casu* Gaussian scale-space theory, and propose an operational definition of optic flow based on a dualistic OFCE, assuming the gauge to be given.

The physical nature of the gauge problem, even though it is the core problem in optic flow theories, has not always been appreciated in the literature. Methods have been proposed to allow for the extraction of a unique vector field without solid motivation for the underlying gauge choice. Even worse, the choices *implicitly* imposed may very well be inconsistent! This is particularly the case in methods using the OFCE and its derivatives in combination with some *ad hoc* truncated

---

<sup>5</sup>In fact, the relation is tautological: one identifies corresponding samples by virtue of the existence of some invariant tag.

Taylor approximation of the optic flow vector and the grey-value image (Nagel, 1987; Nagel, 1992; Otte and Nagel, 1994; Tistarelli, 1994). Approximations cannot help us to find degrees of freedom that were not there in the first place. Their determination is a matter of *interpretation*. Semantical models must of course be subordinate to data evidence, but data evidence does not compel us to favour one model over another. For reasons explained above one must be willing to give up one interpretation in return for a better one. It is therefore prudent not to muddle up factual data evidence with hypothetical explanations, in other words, to manifestly separate the data-intrinsic optic flow degrees of freedom (gauge invariant optic flow: syntactical information) from data-extrinsic considerations that solve the aperture problem (gauge fixing: semantics). A clear and consistent exposition of the order-by-order approach based on differentiating and truncating the OFCE—explicitly stating the gauge choice as a *mathematical* constraint—has been developed by Arnspang (Arnspang, 1988a; Arnspang, 1988b; Arnspang, 1991; Arnspang, 1993).

Let us model the spacetime manifold as a stratification of spatial slices over absolute time (Misner et al., 1973), i.e. I adhere to the Newtonian spacetime model explained in the previous section. I will use coordinates  $x^\mu$  labeled by Greek indices for the parametrisation of space ( $\mu = 1, \dots, n - 1$ ) and time ( $\mu = 0$ ), and write  $x \equiv (x^0 = t; \vec{x})$ . Einstein’s summation convention applies to repeated indices.

Recall Definition 5.2.8; in accordance with the scale-space paradigm (Florack et al., 1994c; Koenderink, 1984; Koenderink and Doorn, 1987; Koenderink and Doorn, 1990; Lindeberg, 1994e), an image will be conceived of as a distribution over the Gaussian filter family (Florack, 1995; Florack et al., 1994b; Koenderink and Doorn, 1990). The normalised Gaussian kernel will be indicated by  $\phi(x)$ . Its centre location  $x$  and its pseudo-isotropic scale  $\sigma$  are free parameters that will not be made explicit in the notation (without loss of generality I take  $(x; \sigma) = (0; 1)$ ). Image derivatives can be defined accordingly:

**Definition 5.3.1 (Partial Image Derivatives).** Recall Definition 5.2.8, with  $\Theta$  given by the Newtonian spacetime group. Let  $\phi \in \mathcal{G}^+(\mathbb{R}^n)$  be the standard Gaussian point operator, and let  $\phi_{\mu_1 \dots \mu_k}(x)$  denote its  $k$ -th order derivative with respect to  $x^{\mu_1}, \dots, x^{\mu_k}$ . Then the corresponding  $k$ -th order image derivative is given by

$$I_{\mu_1 \dots \mu_k} \stackrel{\text{def}}{=} (-)^k \langle F|\Theta| \phi_{\mu_1 \dots \mu_k} \rangle.$$

That is,

$$I_{\mu_1 \dots \mu_k} = (-)^k \int dx f(x) \Theta_* \phi_{\mu_1 \dots \mu_k}(x).$$

I henceforth consider local samples defined at the fiducial origin  $1 \in \Theta$ .

Definition 5.3.1 can be motivated as follows. Firstly, the “classical limit”  $\sigma \downarrow 0$  will bring us back to the conventional  $k$ -th order derivative  $f_{\mu_1 \dots \mu_k}(x = 0)$ , provided of course  $f(x)$  is sufficiently smooth (simply perform a  $k$ -fold partial integration to see this). The key point, however, is that the raw data are not required to be regular in any way (in particular, smoothness of discrete data is a *contradictio in*

*terminis*). Nevertheless, its derivatives in the sense of Definition 5.3.1 are *operationally well-defined* and, moreover, *well-posed* in the sense of Hadamard<sup>6</sup>. They can be conceived of as conventional derivatives of the data  $f(x)$  after “blurring” with filter  $\phi(x)$ , although blurring and classical differentiation cannot be decoupled in practice.

Another important mathematical tool used in this section is the *Lie derivative*. Lie derivatives capture variations of spacetime quantities along the integral flow of some vector field. To take a Lie derivative, one therefore needs to know this vector field. Actually, I shall only consider the 1-st order Lie derivative of an image, which will give us a *linear* model of optic flow; this is *not* a restriction and should not be confused with the spatiotemporal differential order of the flow field one might be interested in. I will require no *a priori* restrictions on this. In the line of Horn and Schunck’s approach the vector field will, of course, be the optic flow field, which I shall denote by  $v(x) \equiv (v^0(x); \vec{v}(x))$ . Note that I include a temporal component.

The Lie derivative of a scalar function  $f(x)$  w.r.t. the vector field  $v(x)$  is given by the directional derivative  $\mathcal{L}_v f(x) = \nabla f(x) \cdot v(x)$ . However, this is ill-posed just like any classical derivative, and only works if  $f(x)$  is differentiable; it can be made well-posed by defining it on the basis of an integration against the conjugate derivative of a Gaussian kernel, as I did above.

The following assumption will be made throughout:

### Definition 5.3.2 (Temporal Gauge).

$$\forall x \in \mathbb{R}^n : \quad v^0(x) = 1.$$

This is a usual (locally weak, but globally not very realistic) assumption stating that the flow is everywhere nonvanishing and transversal to fixed-time slices, in other words, that structural details are enduring. It therefore expresses *conservation of topological detail*. It is in fact an instance of a gauge condition enforced on the basis of an *a priori* physical principle. See Figure 5.1.

A few more definitions are needed.

**Definition 5.3.3 (Filter Current Density).** Notation as in Definition 5.3.1. Define the filter current density  $j(x)$  corresponding to the filter  $\phi(x)$  and the vector field  $v(x)$  as follows:

$$j(x) = \phi(x) v(x).$$

More generally, for any order  $k \in \mathbb{Z}_0^+$ :

$$j_{\mu_1 \dots \mu_k}(x) = \phi_{\mu_1 \dots \mu_k}(x) v(x).$$

The  $k$ -th order filter current density is really a *density* in the tensorial sense, a property it inherits from the basic filter  $\phi(x)$ . Note that the  $\mu = 0$  components equal the basic Gaussian derivative filters by virtue of the temporal gauge (Definition 5.3.2).

---

<sup>6</sup>A problem is well-posed in the sense of Hadamard if it has a unique solution which depends continuously on the input data. Otherwise it is ill-posed.

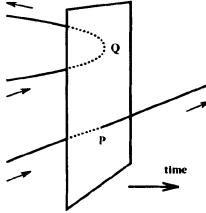


Figure 5.1: A transversal and a non-transversal flow-line, through  $\mathbf{P}$  and  $\mathbf{Q}$ , respectively (a flow-line is the integral curve of a vector field through a given point). The latter one is excluded by the transversality requirement. The upper branch of that curve has an anti-causal orientation. It is quite easy to give it a causal interpretation simply by reversing the arrow; in that case, however,  $\mathbf{Q}$  becomes an annihilation point, thus conflicting the assumption of “conservation of topological detail”. Note that if the *spatial* optic flow field vanishes, the flow-lines will be parallel to the time axis.

Finally, I shall use a *formal expansion* of  $v(x)$  near the origin, truncated at some arbitrary order  $M$ ,  $v_M(x)$ . This is a polynomial intended to approximate a finite number of local degrees of freedom of the vector field (both number as well as level of approximation depend on  $M$ ).

**Definition 5.3.4 (M-th Order Formal Expansion).** The formal expansion of order  $M$  of the vector field  $v(x)$  at  $x = 0$ , denoted  $v_M(x)$ , is an  $M$ -th order polynomial given by

$$v_M^\mu(x) = \sum_{l=0}^M \frac{1}{l!} v_{M;\rho_1 \dots \rho_l}^\mu x^{\rho_1} \dots x^{\rho_l},$$

the coefficients of which may depend on  $M$ .

The coefficients  $v_{M;\rho_1 \dots \rho_l}^\mu$  (to be defined later on so as to *approximate* the flow field’s partial derivatives  $\partial_{\rho_1 \dots \rho_l} v^\mu(x)$  at the origin), correspond exactly to the degrees of freedom I shall be looking for<sup>7</sup>. Having established notations and definitions, let us turn to the theory.

The Lie derivative of the image with respect to the optic flow vector can be made well-posed by formulating it as a distribution (cf. Definition 5.3.1).

**Proposition 5.3.5 (Lie Derivative).** See Definition 5.3.1 and Definition 5.3.3. The Lie derivative of an image (evaluated at the origin) is defined as

$$\mathcal{L}_v F[\phi] = \int dx f(x) \mathcal{L}_v^T \phi(x),$$

in which  $\mathcal{L}_v^T \phi(x)$  is given by

$$\mathcal{L}_v^T \phi(x) = -\operatorname{div} j(x).$$

---

<sup>7</sup>Other expansions are possible; it will be clear from the theory how to proceed in similar cases.

Note that  $\operatorname{div} j(x) = \partial_\mu j^\mu(x)$  is the Lie derivative of the filter<sup>8</sup>  $\phi(x)$  w.r.t.  $v(x)$ . It contains a contribution involving the divergence of the vector field, preserving the filter's volumetric normalisation. Its form is similar to the left hand side of (but should not be confused with) the “generalised motion constraint equation”, proposed by Schunck (Schunck, 1984) in the context of density data, and is well-known from fluid dynamics. Like with ordinary differentiation, Proposition 5.3.5 is motivated by virtue of correspondence with its classical counterpart in the subspace of  $C^1(\mathbb{R}^n)$ -source functions  $f(x)$  (the proof of which requires one partial integration step).

Gauge invariant optic flow can now be defined as any (causal) flow field that preserves  $F[\phi]$ :  $\mathcal{L}_v F[\phi] = 0$ . Although formulated for one sample point only, it should be understood as a globally valid identity. This can be expressed in terms of a countable set of constraints by stating that all spatiotemporal derivatives at the point of interest vanish as well.

**Definition 5.3.6 (Optic Flow).** See Definition 5.3.3 and Proposition 5.3.5. The vector field  $v(x)$  is an optic flow field if

$$\partial_{\mu_1 \dots \mu_k} \mathcal{L}_v F[\phi] = 0 \quad \forall k \in \mathbb{Z}_0^+,$$

in which the spatiotemporal derivatives of  $\mathcal{L}_v F[\phi]$  are defined as usual:

$$\partial_{\mu_1 \dots \mu_k} \mathcal{L}_v F[\phi] \stackrel{\text{def}}{=} (-)^k \int dx f(x) \mathcal{L}_v^T \phi_{\mu_1 \dots \mu_k}(x),$$

with

$$\mathcal{L}_v^T \phi_{\mu_1 \dots \mu_k}(x) = -\operatorname{div} j_{\mu_1 \dots \mu_k}(x).$$

(Note that Lie derivatives and partial derivatives do not commute; first the partial derivatives are transposed, and subsequently the Lie derivative.) Defined as such, the optic flow field  $v(x)$  contains an infinite number of degrees of freedom. This is inconvenient; moreover, most of them are irrelevant or computationally inaccessible anyway. But, depending on one's task, a 0-th order approximation is usually too restrictive. For example, in the case of real-world movies, 1-st order properties of the vector field may reveal relevant information such as qualitative shape properties, surface slant (Koenderink and Doorn, 1975), and time-to-collision (Lee, 1976; Lee, 1980). Unlike 1-st order, 2-nd order is quantitatively related to intrinsic surface properties of an object (Koenderink and Doorn, 1992b). Moreover, there is no *a priori* limit to the highest order that is still accessible and significant; this depends very much on matters such as image quality (noise and sampling characteristics), resolution of interest, etc. Last but not least, the order  $M$  of approximation introduced in Definition 5.3.4 and employed in the computational scheme below is not necessarily the same thing as order of differentiation. It will be argued that the ability to refine the order of approximation is essential.

So let us look at the  $M$ -th order case. Replacing  $v(x)$  by  $v_M(x)$  according to Definition 5.3.4 yields the following.

---

<sup>8</sup>Presuming the source  $f$  is a scalar; if  $f$  represents a density, then  $\phi$  must behave like a scalar under Lie derivation.

**Definition 5.3.7 (Optic Flow Approximation).** Recall Definitions 5.3.4 and 5.3.6. Using  $\mathcal{L}_{v_M}$  instead of  $\mathcal{L}_v$  one has

$$\partial_{\mu_1 \dots \mu_k} \mathcal{L}_{v_M} F[\phi] = - \sum_{l=0}^M v_{M; \rho_1 \dots \rho_l}^\mu \int dx f(x) \partial_\mu \Phi_{\mu_1 \dots \mu_k}^{\rho_1 \dots \rho_l}(x) = 0 \quad \forall k = 0, \dots, M,$$

in which the effective filters  $\Phi_{\mu_1 \dots \mu_k}^{\rho_1 \dots \rho_l}(x)$  are given by

$$\Phi_{\mu_1 \dots \mu_k}^{\rho_1 \dots \rho_l}(x) = \frac{(-)^k}{l!} \phi_{\mu_1 \dots \mu_k}(x) x^{\rho_1} \dots x^{\rho_l}.$$

Since the Gaussian family  $(-)^k \phi_{\mu_1 \dots \mu_k}(x)$  ( $k \in \mathbb{Z}_0^+$ ) is *complete* (Koenderink and Doorn, 1990), the set of filters  $\Phi_{\mu_1 \dots \mu_k}^{\rho_1 \dots \rho_l}(x)$  ( $k, l \in \mathbb{Z}_0^+$ ) is apparently *redundant*. Hence they can all be expressed in terms of pure Gaussian derivative filters (i.e. the case  $l = 0$ ); for technical details see (Florack and Nielsen, 1994); for an example cf. the chapter by Niessen and Maas, Chapter 3.

It is more important to note the restriction on the admissible spatiotemporal orders  $k$ . One has to take care not to introduce spurious degrees of freedom by truncating the optic flow field (it introduces a gauge condition through the back-door!). In particular, one should try to maintain gauge invariance for our *approximated* optic flow field  $v_M(x)$ . Allowing arbitrary orders of differentiation in Definition 5.3.7 would certainly break this invariance in the generic case (i.e. the usual case when  $v(x) \neq v_M(x)$ ), and may even yield an *inconsistent* system<sup>9</sup>! This is why one needs to limit the highest order to  $k = M$ . A more rigorous motivation of this can be found in (Florack and Nielsen, 1994), where it is also shown that the resulting linear equations are generically independent and indeed gauge invariant *to the same extent as the exact system*. Thus the  $M$ -th order approximation does *not* affect the intrinsic ambiguity of optic flow.

Unfortunately, the literature is abundant of instances in which a unique optic flow field is singled out by shrewd combination of truncation and differentiation; one cannot expect this to be a proper way of gauge fixing, except for coincidental cases. Gauge fixing is really a matter of data-extrinsic physical considerations independent of the OFCE or anything derived from it.

As opposed to conventional schemes based on  $M$ -fold implicit differentiation of the OFCE, *every*  $k$ -th order subset of Definition 5.3.7 contains  $M$ -th order components of the approximated optic flow field. An important thing to keep in mind is that the degrees of freedom of  $v_M(x)$ , i.e. the coefficients  $v_{M; \rho_1 \dots \rho_l}^\mu$ , depend on the order  $M$  of approximation. In other words, the polynomial approximation  $v_{M+1}(x)$  is a refinement of  $v_M(x)$  in the sense that *all* coefficients are refined. Hence, it is *not* the Taylor polynomial of  $v(x)$ ; only in the limiting case one has

$$\lim_{M \rightarrow \infty} v_{M; \rho_1 \dots \rho_l}^\mu = \partial_{\rho_1 \dots \rho_l} v^\mu(x=0),$$

so that  $v_\infty(x) = v(x)$ . Thus if the Cauchy sequence  $\|v_{k+1}^\mu - v_k^\mu\|$ , truncated at

---

<sup>9</sup>Which of course can always be “solved” in some least squares sense...

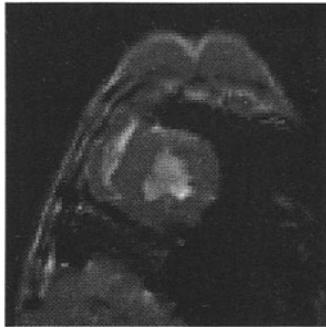


Figure 5.2: One frame taken from an MR sequence of a canine heart. Resolution:  $100 \times 100$  pixels, 16 frames per heartbeat.

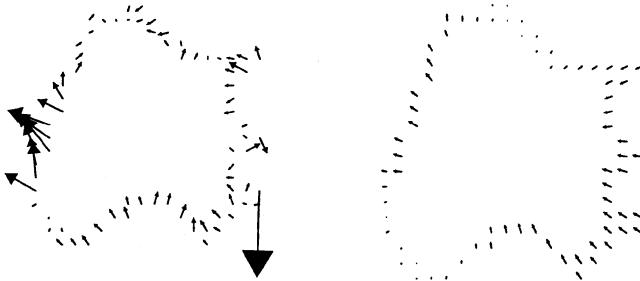


Figure 5.3: Velocity vectors calculated using the first order OFCE on the time-frame of the MR-sequence (Figure 5.2). **Left:** vectors calculated at fixed spatial and temporal scales, viz.  $\sigma_s = 2, \sigma_t = 1$ . It is obvious that in some parts of the image the velocity estimate will be unreliable if fixed scales are used. **Right:** vectors calculated at spatial and temporal scales selected using a stability criterion. For details, cf. Niessen *et al.* (Niessen et al., 1996b), and Chapter 3.

$k = M$ , say<sup>10</sup>, does not indicate a sufficient rate of convergence, then the optic flow field is effectively discontinuous.

Figures 5.2 and 5.3—reproduced with permission from Niessen *et al.* (Niessen et al., 1996b)—illustrate optic flow computation as proposed here on cine MR data.

## 5.4 Conclusion

In summary, the core message of this chapter is that raw data are meaningless without theory. In particular, *an image is, like any physical observation, “theory-*

---

<sup>10</sup>An upper limit always exists above which the required differential order  $2M + 1$  prohibits further refinement.

*laden*". An explication of this has been outlined.

The framework sketched in this chapter suggests diversification to novel kinds of multiresolution schemes, as well as unification of existing ones, e.g. based on interaction models somewhere in-between the "transparent" topological and "tactile" morphological probing mechanisms (soft-probes into surface layers of finite penetration depth). To this end it is useful to adopt an abstract viewpoint and formulate a desirable set of "axioms of coordination" (Reichenbach, 1960), akin to the notion of *a priority* in Kantian epistemology. Deductive reasoning will then reveal the details of all specific models that comply with these axioms, and a precise understanding of their mutual relations.

I have also presented conceptual and computational aspects of optic flow measurements, taking duality into account from the outset. It has been argued that the enigmatic aperture problem can be interpreted as a *gauge invariance* proper to image intrinsic optic flow. One *manifestly* separates data evidence and *a priori* knowledge when handling optic flow as a gauge field. Models come in naturally as *gauge conditions*. This segregation increases semantical flexibility, which is especially important, because any *a priori* assertion that is not supported by the data may be wrong (or ineffectual relative to one's task). An operational approach has been shown to yield a linear system of equations for optic flow measurements different from existing schemes, the coefficients of which are linear combinations of operationally defined and well-posed image derivatives. The associated spatial and temporal inner scale parameters are optional, a freedom that can be exploited by selecting those levels among a precomputed set that permit the most reliable linear inversion at each pixel or voxel (Niessen et al., ; Niessen et al., 1996b), cf. the method proposed in Chapter 3. One can solve this system after imposing appropriate gauge conditions that allow for a unique solution.

The theory can be formulated for *density images* in a straightforward way as well.

# Chapter 6

# On the Axiomatic Foundations of Linear Scale-Space

**Tony Lindeberg**

*Computational Vision and Active Perception Laboratory (CVAP)  
Department of Numerical Analysis and Computing Science (NADA)  
KTH (Royal Institute of Technology)  
S-100 44 Stockholm, Sweden*

**Overview:** Since the pioneering work by Witkin (1983) and Koenderink (1984) on the notion of “scale-space representation”, a large number of different scale-space formulations have been stated, based on different types of assumptions (usually referred to as scale-space axioms). The main subject of this chapter is to provide a synthesis between these linear scale-space formulations and to show how they are related. Another aim is to show how the scale-space formulations, which were originally expressed for continuous data on spatial domains without preferred directions, can be extended to discrete data as well as to spatio-temporal domains with preferred directions. Connections will also be pointed out to approaches based on non-uniform smoothing.

## 6.1 Introduction

One of the very fundamental problems that arises when analysing real-world measurement data originates from the fact that objects in the world may appear in different ways depending upon the scale of observation. This fact is well-known in physics, where phenomena are modelled at several levels of scale, ranging from

particle physics and quantum mechanics at fine scales, through thermodynamics and solid mechanics dealing with every-day phenomena, to astronomy and relativity theory at scales much larger than those we are usually dealing with. Notably, the type of physical description that is obtained may be strongly dependent on the scale at which the world is modelled, and this is in clear contrast to certain idealized mathematical entities, such as “point” or “line”, which are independent of the scale of observation.

In certain controlled situations, appropriate scales for analysis may be known *a priori*. For example, a desirable property of a good physicist is his intuitive ability to select appropriate scales to model a given situation. Under other circumstances, however, it may not be obvious at all to determine in advance what are the proper scales. One such example is a vision system with the task of analysing unknown scenes. Besides the inherent multi-scale properties of real world objects (which, in general, are unknown), such a system has to face the problems that the perspective mapping gives rise to size variations, that noise is introduced in the image formation process, and that the available data are two-dimensional data sets reflecting only indirect properties of a three-dimensional world. To be able to cope with these problems, an essential tool is a formal theory for describing structures at multiple scales. In image processing and computer vision, this insight has lead to the general methodology of representing data at multiple scales, using concepts such as quad-trees, pyramids, wavelets and scale-space representation. A common idea behind the creation of these concepts is that for any input image, a set of gradually smoothed or simplified images should be generated, in which fine scale structures are successively suppressed.

The subject of this chapter is to consider the axiomatic foundations of the special framework for multi-scale representation known as *linear scale-space theory* (Witkin, 1983; Koenderink, 1984; Babaud et al., 1986; Yuille and Poggio, 1986; Koenderink and Doorn, 1992a; Lindeberg, 1990; Lindeberg, 1993b; Lindeberg, 1994e; Lindeberg, 1994d; Florack et al., 1992b; Florack et al., 1994c; Pauwels et al., 1995). Assuming that no *a priori* information is available about the measurement data, the linear *scale-space representation* of any given signal is according to this theory derived from the basic constraints that it should

- be generated by linear and shift-invariant operators (convolutions),
- possess a continuous scale parameter, and
- the transformation from any fine level in the scale-space representation to any coarser level must not introduce “spurious structures”.

The aim of the last requirement is to guarantee that fine-scale structures should disappear monotonically with increasing scale, such that any coarse-scale representation in the multi-scale family can be regarded as a simplification of any finer-scale representation. Since the first scale-space formulations (Witkin, 1983; Koenderink, 1984), this condition about “non-creation of structure” has been formalized in different ways by different authors (to be reviewed in section 6.2).

A notable coincidence between the different scale-space formulations that have been stated is that the Gaussian kernel arises as a unique choice for a large num-

ber of different combinations of underlying assumptions (scale-space axioms). In summary, linear scale-space theory states that a natural way to process a given  $N$ -dimensional input signal  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  is by convolving it with Gaussian kernels<sup>1</sup>

$$g(x; \sigma^2) = \frac{1}{(2\pi\sigma^2)^{D/2}} e^{-(x_1^2 + \dots + x_D^2)/2\sigma^2} \quad (6.1)$$

and their derivatives

$$g_{x^\alpha}(x; \sigma^2) = \partial_{x_1^{\alpha_1} \dots x_D^{\alpha_D}} g(x; \sigma^2) \quad (6.2)$$

of various widths  $\sigma$ . The output from these operations can then in turn be used as a basis for expressing a large number of early visual operations, such as feature detection, matching, optic flow and computation of shape cues. A particularly convenient framework for formalizing such processes is in terms of multi-scale differential geometric invariants and singularities of these (Koenderink and Doorn, 1987; Florack et al., 1993a; Florack et al., 1994a; Lindeberg, 1993b; Lindeberg, 1994e; Johansen, 1994).<sup>2</sup>

A main purpose of this chapter is to discuss relations between the abovementioned scale-space formulations. Moreover, a complementary treatment will be given showing how a scale-space formulation previously expressed for discrete signals applies to continuous signals. Another main goal is to show how this selection of scale-space axioms, based on the assumption of a semi-group structure combined with a reformulation of the causality requirement in terms of non-enhancement of local extrema, relates to scale-space formulations based on scale invariance.

It will also be indicated how this approach can be generalized to the following types of scale-space concepts:

- affine Gaussian scale-space, ,
- spatio-temporal scale-space, and
- non-linear scale-space.

The presentation is organized as follows: Section 6.2 provides the necessary background by reviewing the notions of causality, non-enhancement of local extrema and scale invariance. Then, section 6.3 shows how assumptions about causality and a semi-group structure uniquely determine the smoothing kernel to be a Gaussian, if combined with rotational symmetry and certain regularity assumptions. In view of this result, it will be described how causality combines with scale invariance. Finally, section 6.4 gives a summary of the main results and section 6.5 points out connections to approaches based on non-uniform smoothing.

---

<sup>1</sup>Here,  $x = (x_1, \dots, x_N) \in \mathbb{R}^N$  is standard vector notation for a  $D$ -dimensional variable, and  $\alpha = (\alpha_1, \dots, \alpha_N) \in \mathbb{Z}^N$  constitutes so-called multi-index notation with  $x^\alpha = x_1^{\alpha_1} \dots x_N^{\alpha_N}$ .

<sup>2</sup>See also chapter 1 by ter Haar Romeny (1996) and chapter 8 by Salden (1996) in this volume.

## 6.2 Axiomatic formulations of linear scale-space

Since the pioneering work (Witkin, 1983; Koenderink, 1984), a large number of scale-space formulations have been stated. This section gives a brief historical review of the development of linear scale-space theory with special emphasis on the scale-space formulations that we shall later build upon.

For further overviews, the reader is referred to (Lindeberg, 1994e; Lindeberg, 1994d; Lindeberg, 1996e; Lindeberg and Haar Romeny, 1994) as well as to Chapter by Koenderink (1996), Chapter 5 and by Florack (1996) in this volume.

### 6.2.1 Original formulation

When (Witkin, 1983) introduced the term “scale-space”, he was concerned with one-dimensional signals and observed that new local extrema cannot be created under Gaussian convolution. Since differentiation commutes with convolution,

$$\partial_{x^n} L(\cdot; t) = \partial_{x^n}(g(\cdot; t) * f) = g(\cdot; t) * \partial_{x^n} f, \quad (6.3)$$

this non-creation property applies also to any  $n^{\text{th}}$ -order spatial derivative computed from the scale-space representation. Specifically, he applied this property to zero-crossings of the second derivative to construct so-called “fingerprints”.

### 6.2.2 Causality

Witkin's observation shows that Gaussian convolution satisfies certain sufficiency requirements for being a smoothing operation. The first proof of the *necessity* of Gaussian smoothing for a scale-space representation was given by (Koenderink, 1984), who also gave a formal extension of the scale-space theory to higher dimensions. He introduced the concept of *causality*, which means that new level surfaces

$$\{(x, y; t) \in \mathbb{R}^2 \times \mathbb{R}: L(x, y; t) = L_0\} \quad (6.4)$$

must not be created in the scale-space representation when the scale parameter is increased. By combining causality with the notions of *isotropy* and *homogeneity*, which essentially mean that all spatial positions and all scale levels must be treated in a similar manner, he showed that the scale-space representation must satisfy the diffusion equation

$$\partial_t L = \frac{1}{2} \nabla^2 L. \quad (6.5)$$

The technique used for proving this necessity result was by studying the level surface through any point in scale-space for which the grey-level function assumes a maximum with respect to the spatial coordinates. If no new level surface is to be created when increasing scale, the level surface should point with its concave side towards decreasing scales. This gives rise to a sign condition on the curvature of the level surface, which assumes the form (6.5) when expressed in terms of derivatives of the scale-space representation.

A similar result was given by (Yuille and Poggio, 1985; Yuille and Poggio, 1986). Related formulations have been expressed by (Hummel, 1986; Hummel, 1987).

### 6.2.3 Non-creation of local extrema

(Lindeberg, 1990) considered the problem of characterizing those kernels in one dimension that share the property of not introducing new local extrema in a signal under convolution. A kernel  $h \in \mathbb{L}_1$  possessing the property that for *any* input signal  $f_{in} \in \mathbb{L}_1$  the number of extrema (zero-crossings) in the convolved signal  $f_{out} = h * f_{in}$  is always less than or equal to the number of local extrema (zero-crossings) in the original signal is termed a *scale-space kernel*.

Such kernels must be non-negative and unimodal both in the spatial and the frequency domain. Moreover they can be completely classified using classical results by (Schoenberg, 1950; Schoenberg, 1953) (see also (Hirschmann and Widder, 1955) and (Karlin, 1968)). Besides trivial *translations* and *rescalings*, there are two primitive types of linear and shift-invariant smoothing transformations in the *continuous* case that never increase the number of extrema (zero-crossings):

- convolution with *Gaussian kernels*,

$$h(x) = e^{-\gamma x^2}, \quad (6.6)$$

- convolution with *truncated exponential functions*,

$$h(x) = \begin{cases} e^{-x/|\mu|} & x \geq 0, \\ 0 & x < 0, \end{cases} \quad h(x) = \begin{cases} e^{x/|\mu|} & x \leq 0, \\ 0 & x > 0, \end{cases} \quad (6.7)$$

In fact, this theory states that *all* continuous scale-space kernels can be decomposed into (possibly infinite) compositions of these primitive smoothing operations.

Correspondingly, in the *discrete* case, there are besides rescaling and translation, three primitive types of smoothing transformations (where  $f_{out} = h * f_{in}$ ):

- two-point weighted averaging or *generalized binomial smoothing*,

$$\begin{aligned} f_{out}(x) &= f_{in}(x) + \alpha_i f_{in}(x-1) & (\alpha_i \geq 0), \\ f_{out}(x) &= f_{in}(x) + \delta_i f_{in}(x+1) & (\delta_i \geq 0), \end{aligned} \quad (6.8)$$

- moving average or *first-order recursive filtering*,

$$\begin{aligned} f_{out}(x) &= f_{in}(x) + \beta_i f_{out}(x-1) & (0 \leq \beta_i < 1), \\ f_{out}(x) &= f_{in}(x) + \gamma_i f_{out}(x+1) & (0 \leq \gamma_i < 1), \end{aligned} \quad (6.9)$$

- *infinitesimal smoothing* (or diffusion smoothing) with the generating function of the smoothing kernel being of the form

$$H_{semi-group}(z) = \sum_{n=-\infty}^{\infty} h(n) z^n = e^{t(a z^{-1} + b z)}. \quad (6.10)$$

Among these discrete kernels, the generalized binomial kernels provide a natural basis for constructing pyramid representations (Burt, 1981; Crowley, 1981), whereas recursive filters can be used for efficient implementations of smoothing operations (Deriche, 1987). The interpretation and applications of the class of infinitesimal discrete smoothing kernels will be apparent in next section.

### 6.2.4 Semi-group structure and continuous scale parameter

A natural structure to impose on a scale-space representation is a *semi-group* structure *i.e.*, if every smoothing kernel is associated with a parameter value, and if two such kernels are convolved with each other, then the resulting kernel should be a member of the same family,

$$h(\cdot; t_1) * h(\cdot; t_2) = h(\cdot; t_1 + t_2). \quad (6.11)$$

This condition states that all scale-space transformations are of the same type. In particular, this condition ensures that the transformation from a fine scale to any coarse scale should be of the same type as the transformation from the original signal to any scale in the scale-space representation,

$$\begin{aligned} L(\cdot; t_2) &= \{\text{definition}\} = h(\cdot; t_2) * f \\ &= \{\text{semi-group}\} = (h(\cdot; t_2 - t_1) * h(\cdot; t_1)) * f \\ &= \{\text{associativity}\} = h(\cdot; t_2 - t_1) * (h(\cdot; t_1) * f) \\ &= \{\text{definition}\} = h(\cdot; t_2 - t_1) * L(\cdot; t_1). \end{aligned} \quad (6.12)$$

If this semi-group structure is combined with non-creation of local extrema, the existence of a *continuous scale parameter*, and if the kernels are required to be symmetric and normalized and to satisfy a mild degree of smoothness in the scale direction (Borel-measurability), then the family of smoothing kernels is uniquely determined to be a Gaussian (Lindeberg, 1990)

$$h(x; t) = \frac{1}{\sqrt{2\pi\alpha t}} e^{-x^2/(2\alpha t)} \quad (t > 0, \alpha \in \mathbb{R}^+). \quad (6.13)$$

If on the other hand the spatial symmetry requirements are relaxed, we obtain translated or *velocity-adapted Gaussian kernels*

$$h(x; \delta, t) = \frac{1}{\sqrt{2\pi\alpha t}} e^{-(x-\delta)^2/(2\alpha t)} \quad (t > 0, \alpha \in \mathbb{R}^+) \quad (6.14)$$

(with  $\delta$  determined by local velocity information). This family of kernels has also been derived by (Florack et al., 1992a), based on dimensional analysis (see section 6.2.6).

Corresponding arguments about a semi-group structure on a spatially symmetric discrete domain uniquely lead to the *discrete analogue of the Gaussian kernel*

$$L(x; t) = \sum_{n=-\infty}^{\infty} T(n; t) f(x - n) \quad \text{where} \quad T(n; t) = e^{-t} I_n(t), \quad (6.15)$$

where  $I_n$  are the modified Bessel functions of integer order (Abramowitz and Stegun, 1964). This scale-space family corresponds to the result of letting  $a = b$  in (6.10) and describes the solution to the *semi-discretized diffusion equation*

$$\partial_t L(x; t) = \frac{1}{2}(L(x+1; t) - 2L(x; t) + L(x-1; t)) = \frac{1}{2}\nabla_3^2 L(x; t), \quad (6.16)$$

where  $\nabla_3^2$  is the standard second-order difference operator. Indeed, the discrete analogue of the Gaussian kernel can be interpreted as the result of exponentiating the second-order difference operator

$$T(\cdot; t) = e^{\frac{t}{2}\nabla_3^2}. \quad (6.17)$$

When considering temporal image data in a real-time situation, we have to require the scale-space kernels to be *time-causal* and to not extend into the future (Koenderink, 1988; Lindeberg and Fagerström, 1996). Imposing this constraint on the infinitesimal smoothing filters in (6.10) uniquely gives rise to the *temporal scale-space* generated by convolution with the Poisson kernel (Lindeberg, 1996c)

$$L(x; t) = \sum_{n=-\infty}^{\infty} p(n; t) f(x-n) \quad \text{where} \quad p(n; \lambda) = e^{-\lambda} \frac{\lambda^n}{n!}. \quad (6.18)$$

In terms of differential equations, this scale-space family satisfies

$$\partial_\lambda L = -\delta_- L, \quad (6.19)$$

where  $\delta_-$  denotes the backward difference operator  $\delta_- L(t) = L(t) - L(t-1)$ .

Despite the completeness of these results based on non-creation of extrema (or zero-crossings), however, they cannot be extended to higher dimensions, since in two (and higher) dimensions, since there are no non-trivial kernels guaranteed to never increase the number of local extrema in a signal (see (Lifshitz and Pizer, 1990; Yuille, 1988; Lindeberg, 1990; Lindeberg, 1994e) for counterexamples).

### 6.2.5 Non-enhancement and infinitesimal generator

If the semi-group structure is combined with a strong continuity requirement with respect to the scale parameter, then it follows from well-known results in functional analysis (Hille and Phillips, 1957) that the scale-space family must have an *infinitesimal generator*. In other words, if a transformation operator  $\mathcal{T}_t$  from the input signal to the scale-space representation at any scale  $t$  is defined by

$$L(\cdot; t) = \mathcal{T}_t f, \quad (6.20)$$

then under reasonable regularity requirements there exists a limit case of this operator (the infinitesimal generator)

$$\mathcal{A}f = \lim_{h \downarrow 0} \frac{\mathcal{T}_h f - f}{h} \quad (6.21)$$

and the scale-space family satisfies the differential equation

$$\partial_t L(\cdot; t) = \lim_{h \downarrow 0} \frac{L(\cdot; t+h) - L(\cdot; t)}{h} = \mathcal{A}(\mathcal{T}_t f) = \mathcal{A}L(\cdot; t). \quad (6.22)$$

For discrete signals, (Lindeberg, 1990; Lindeberg, 1991; Lindeberg, 1994c) showed that this structure implies that the scale-space family must satisfy a (semi-discretized) diffusion equation if combined with a slightly modified formulation of Koen-derink's causality requirement expressed as *non-enhancement of local extrema*:

*Non-enhancement of local extrema:* If for some scale level  $t_0$  a point  $x_0$  is a non-degenerate local maximum for the scale-space representation at that level (regarded as a function of the space coordinates only) then its value must not increase when the scale parameter increases. Analogously, if a point is a non-degenerate local minimum then its value must not decrease when the scale parameter increases.

In summary, these conditions imply that the scale-space family  $L: \mathbb{Z}^N \times \mathbb{R}^+ \rightarrow \mathbb{R}$  of a discrete signal  $f: \mathbb{Z}^N \rightarrow \mathbb{R}$  must satisfy the semi-discrete differential equation

$$(\partial_t L)(x; t) = (\mathcal{A}_{ScSp} L)(x; t) = \sum_{\xi \in \mathbb{Z}^N} a_\xi L(x - \xi; t), \quad (6.23)$$

for some *infinitesimal scale-space generator*  $\mathcal{A}_{ScSp}$  characterized by

- the *locality* condition  $a_\xi = 0$  if  $|\xi|_\infty > 1$ ,
- the *positivity* constraint  $a_\xi \geq 0$  if  $\xi \neq 0$ ,
- the *zero sum* condition  $\sum_{\xi \in \mathbb{Z}^N} a_\xi = 0$ , as well as
- the *symmetry* requirements
  - $a_{(-\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$  and
  - $a_{P_k^N(\xi_1, \xi_2, \dots, \xi_N)} = a_{(\xi_1, \xi_2, \dots, \xi_N)}$

for all  $\xi = (\xi_1, \xi_2, \dots, \xi_N) \in \mathbb{Z}^N$  and all possible permutations  $P_k^N$  of  $N$  elements.

The locality condition means that  $\mathcal{A}_{ScSp}$  corresponds to the discretization of derivatives of order up to two. In one and two dimensions, (6.23) reduces to

$$\partial_t L = \alpha_1 \nabla_3^2 L, \quad (6.24)$$

$$\partial_t L = \alpha_1 \nabla_5^2 L + \alpha_2 \nabla_{x^2}^2 L, \quad (6.25)$$

$$\partial_t L = \alpha_1 \nabla_7^2 L + \alpha_2 \nabla_{+3}^2 L + \alpha_3 \nabla_{x^3}^2 L, \quad (6.26)$$

for some constants  $\alpha_1 \geq 0$  and  $\alpha_2 \geq 0$ . Here, the symbols,  $\nabla_5^2$  and  $\nabla_{x^2}^2$ , denote the common discrete approximations of the Laplacian operator in two dimensions (below the notation  $f_{-1,1}$  stands for  $f(x-1, y+1)$  etc.):

$$(\nabla_5^2 f)_{0,0} = f_{-1,0} + f_{+1,0} + f_{0,-1} + f_{0,+1} - 4f_{0,0}, \quad (6.27)$$

$$(\nabla_{x^2}^2 f)_{0,0} = 1/2(f_{-1,-1} + f_{-1,+1} + f_{+1,-1} + f_{+1,+1} - 4f_{0,0}). \quad (6.28)$$

whereas  $\nabla_7^2$ ,  $\nabla_{+3}^2$  and  $\nabla_{x_3}^2$  represent corresponding discrete approximations of the Laplacian operator in three dimensions:

$$\begin{aligned} (\nabla_7^2 f)_{0,0,0} &= f_{-1,0,0} + f_{+1,0,0} + f_{0,-1,0} \\ &\quad + f_{0,+1,0} + f_{0,0,-1} + f_{0,0,+1} - 6f_{0,0,0}, \\ (\nabla_{+3}^2 f)_{0,0,0} &= \frac{1}{4}(f_{-1,-1,0} + f_{-1,+1,0} + f_{+1,-1,0} + f_{+1,+1,0} \\ &\quad + f_{-1,0,-1} + f_{-1,0,+1} + f_{+1,0,-1} + f_{+1,0,+1} \\ &\quad + f_{0,-1,-1} + f_{0,-1,+1} + f_{0,+1,-1} + f_{0,+1,+1} - 12f_{0,0,0}), \\ (\nabla_{x_3}^2 f)_{0,0,0} &= \frac{1}{4}(f_{-1,-1,-1} + f_{-1,-1,+1} + f_{-1,+1,-1} + f_{-1,+1,+1} \\ &\quad + f_{+1,-1,-1} + f_{+1,-1,+1} + f_{+1,+1,-1} + f_{+1,+1,+1} - 8f_{0,0,0}). \end{aligned}$$

If the spatial symmetry requirements underlying this formulation are relaxed, then a larger class of scale-space transformations will be obtained (see section 6.5.1). The structure required from such a non-isotropic smoothing operation to satisfy non-enhancement of local extrema is that the infinitesimal scale-space generator should satisfy the locality, positivity and zero sum conditions, essentially corresponding to the discretization of a second-order differential operator.

If we would like the Fourier transform of the associated convolution kernel to be real (corresponding to a milder degree of symmetry for the smoothing kernels, such that they will be mirror symmetric along any axis through the origin), then a necessary requirement is that the filter coefficients should satisfy  $a_\xi = a_{-\xi}$ .

### 6.2.6 Scale invariance

A formulation by (Florack et al., 1992b; Florack, 1993) and continued work by (Pauwels et al., 1995) show that the class of allowable scale-space kernels can be restricted under weaker conditions, essentially by combining the earlier mentioned conditions about linearity, shift invariance, rotational invariance and semi-group structure with *scale invariance*. The basic argument is taken from physics; physical laws must be independent of the choice of fundamental parameters. In practice, this corresponds to what is known as dimensional analysis; a function that relates physical observables must be independent of the choice of dimensional units. Notably, this condition comprises no direct measure of “structure” in the signal; the non-creation of new structure is only implicit in the sense that physically observable entities subject to scale changes should be treated in a self-similar manner.

In these scale-space formulations based on scale invariance, however, a further assumption is introduced concerning the semi-group structure. In sections 6.2.1–6.2.5, the scale parameter  $t$  associated with the semi-group (see equation (6.11)) was regarded as an *abstract ordering parameter* only. *A priori*, i.e. in the stage of formulating the axioms, there was no direct connection between this parameter and measurements of scale in terms of units of length. The only requirement was the qualitative (and essential) constraint that increasing values of the scale parameter should somehow correspond to representations at coarser scales. *A posteriori*, i.e. after deriving the shape of the convolution kernel, we could conclude that this parameter is related to scale as measured in units of length, e.g. via the standard

deviation of the Gaussian kernel  $\sigma$ . The relationship turned out to be  $t = \sigma^2/2$  (up to an unessential linear reparametrization of the scale parameter) and the semi-group operation to correspond to adding of  $\sigma$ -values in the Euclidean norm.

**Restrictions from scale invariance:** In this section, we shall assume that such a relationship exists already in the stage of formulating the axioms. Let  $\sigma$  be a scale parameter of dimension length associated with each layer in the scale-space representation, and introduce a monotonically increasing transformation function

$$t = \varphi(\sigma) \quad (6.29)$$

(with  $\varphi(0) = 0$ ) such that the semi-group structure of the convolution kernel corresponds to mere adding of the scale values when measured in terms of  $t$ . For kernels parameterized by  $\sigma$ , the semi-group operation then assumes the form

$$h(\cdot; \sigma_1) * h(\cdot; \sigma_2) = h(\cdot; \varphi^{-1}(\varphi(\sigma_1) + \varphi(\sigma_2))). \quad (6.30)$$

The basic idea<sup>3</sup> is then to study the convolution operation in the Fourier domain

$$\hat{L}(\omega; \sigma) = \hat{h}(\omega; \sigma) \hat{f}(\omega). \quad (6.31)$$

where  $\hat{h}$  represents the Fourier transform of the one-parameter family of smoothing kernels  $h: \mathbb{Z}^N \times \mathbb{R}^+ \rightarrow \mathbb{R}$ . Scale invariance implies that it must be possible to write this relation in terms of dimensionless variables. Here, we choose  $\hat{L}/\hat{f} \in \mathbb{R}$  and  $\omega\sigma \in \mathbb{R}^N$  and require the following relation to hold for some  $\tilde{h}: \mathbb{R}^N \rightarrow \mathbb{R}$ :

$$\frac{\hat{L}(\omega; \sigma)}{\hat{f}(\omega; \sigma)} = \hat{h}(\omega; \sigma) = \tilde{h}(\omega\sigma). \quad (6.32)$$

If in addition, this relation is to be independent of orientation, it follows that

$$\tilde{h}(\omega\sigma) = \hat{H}(|\omega\sigma|) \quad (6.33)$$

for some function  $\hat{H}: \mathbb{R} \rightarrow \mathbb{R}$  with  $\hat{H}(0) = 1$ . The semi-group structure implies that  $\hat{H}$  must obey

$$\hat{H}(|\omega\sigma_1|) \hat{H}(|\omega\sigma_2|) = \hat{H}(|\omega\varphi^{-1}(\varphi(\sigma_1) + \varphi(\sigma_2))|) \quad (6.34)$$

for all  $\sigma_1, \sigma_2, \omega \in \mathbb{R}_+$ , and it can be shown that  $\varphi$  must then be of the form

$$\varphi(\sigma) = C\sigma^p \quad (6.35)$$

for some arbitrary constants  $C > 0$  and  $p > 0$  (where we without loss of generality can take  $C = 1$ ). Then, with  $\tilde{H}(x^p) = \hat{H}(x)$ , this constraint reduces to

$$\tilde{H}(|\omega\sigma_1|^p) \tilde{H}(|\omega\sigma_2|^p) = \tilde{H}(|\omega\sigma_1|^p + |\omega\sigma_2|^p), \quad (6.36)$$

---

<sup>3</sup>This derivation has been shortened substantially to save space. More detailed arguments showing how the assumption of scale invariance narrows down the class of smoothing kernels are presented in different forms in (Florack et al., 1992b; Pauwels et al., 1995; Lindeberg, 1994a).

which can be recognized as the definition of the exponential function ( $\psi(\xi_1)\psi(\xi_2) = \psi(\xi_1 + \xi_2) \Rightarrow \psi(\xi) = a^\xi$  for some  $a > 0$ ). In summary, for a scale-invariant rotationally symmetric semi-group, the Fourier transform must be of the form<sup>4</sup>

$$\hat{h}(\omega; \sigma) = \hat{H}(\omega\sigma) = \tilde{H}(|\omega\sigma|^p) = e^{-\alpha|\omega\sigma|^p} \quad (6.37)$$

for some  $\alpha \in \mathbb{R}$ . Requiring  $\lim_{\sigma \rightarrow \infty} \hat{h}(\omega; \sigma) = 0$  gives  $\alpha < 0$ . Moreover, we can without loss of generality let  $\alpha = -1/2$  to preserve consistency with the definition of the standard deviation of the Gaussian kernel  $\sigma$  in the case when  $p = 2$ .

**Additional conditions.** (Florack et al., 1992b) used separability in Cartesian coordinates as an additional basic constraint. Except in the one-dimensional case, this fixates  $h$  to be a Gaussian. Since, however, rotational symmetry combined with *separability per se* are sufficient to fixate the function to be a Gaussian, and the selection of orthogonal coordinate directions constitutes a very specific choice, it is illuminating to consider the effect of using other choices of  $p$ .

(Pauwels et al., 1995) showed that the corresponding multi-scale representations generated by convolution kernels of the form (6.37) have *local infinitesimal generators* (basically meaning that the operator  $\mathcal{A}$  in (6.21) is a differential operator) if and only if the exponent  $p$  is an even integer. Out of this countable set of choices,  $p = 2$  is the only one that corresponds to a *non-negative convolution kernel*<sup>5</sup> (recall from section 6.2.3 that non-creation of local extrema implies that the kernel has to be non-negative).

(Koenderink and Doorn, 1992a) carried out a closely related study, where they showed that Gaussian derivative operators are natural operators to derive from a scale-space representation, given the assumption of scale invariance.

### 6.3 Semi-group and causality: continuous domain

A main subject of this article is to extend the last two types of scale-space formulations in previous section. We shall first state explicitly how the scale-space

<sup>4</sup>In a closely related work in chapter 7 in this volume, Nielsen (1996) arrives at filters of the same form from a slightly different starting point. He considers the problem of deriving optimal smoothing filters, and formulates an optimization problem in Euclidean norm. Thereby, the solution can be expressed as a linear filter, with the filter coefficients determined by a linear system of equations. Combined with shift-invariance, this gives rise to a convolution structure, and by requiring the filters to form a semi-group and to be associated with a scale parameter of dimension length raised to some power, it then follows that the filter must have a scale invariant Fourier transform which is additive under some self-similar reparametrization of the scale parameter. In other words, the Fourier transform must be of the form (6.37). In the spatial domain, this corresponds to regularization involving infinite orders of differentiation.

If on the other hand, the regularization functional is truncated at lower orders of differentiation, then a larger class of regularization filters is obtained, including the recursive filters studied by (Deriche, 1987). It is interesting to note that these are also scale-space kernels in the sense that they are guaranteed to not increase the number of local extrema in a signal.

<sup>5</sup>This follows directly from the well-known relation  $\int_{x \in \mathbb{R}} x^n h(x) dx = (-i)^n \hat{h}^{(n)}(0)$  between moments in the spatial domain and derivatives in the frequency domain. It is straightforward to verify that the second moments of  $h$  are zero for any  $p > 2$ . Hence, the convolution kernel assumes both positive and negative values for all  $p > 2$ .

formulation in section 6.2.5 applies to continuous signals. Then, as a corollary, it follows that for the scale-space formulation in section 6.2.6 only the specific choice  $p = 2$  corresponds to a scale invariant semi-group satisfying causality requirements.

A main result we shall prove is that semi-group structure combined with the existence of a continuous scale parameter and non-enhancement of local extrema implies that the convolution kernel must be a Gaussian. This result will be obtained by an analogous way of reasoning as in a corresponding treatment for discrete signals (Lindeberg, 1990; Lindeberg, 1991; Lindeberg, 1994c; Lindeberg, 1994e). A certain number of technical modifications, however, have to be made due to the fact that the signals are continuous.

### 6.3.1 Assumptions

Given that the task is to state an axiomatic formulation of the first stages of visual processing, *the visual front-end*, a list of desired properties may be long:

linearity, translational invariance, rotational symmetry, mirror symmetry, semi-group, causality, positivity, unimodality, continuity, differentiability, normalization, nice scaling behaviour, locality, rapidly decreasing for large  $x$  and  $t$ , existence of an infinitesimal generator, invariance with respect to certain grey-level transformations, etc.

Such a list will, however, contain redundancies, as does this one. Here, a (minimal) subset of these properties is taken as axioms. In fact, it can be shown that all the other above-mentioned properties follow from the subset we shall select.

To begin, let us postulate that the scale-space representation should be generated by convolution with a one-parameter family of kernels such that  $L(x; 0) = f(x)$  and for  $t > 0$

$$L(x; t) = \int_{\xi \in \mathbb{R}^N} T(\xi; t) f(x - \xi) d\xi. \quad (6.38)$$

This form of the smoothing formula corresponds to natural requirements about *linear shift-invariant smoothing* and the existence of a *continuous scale parameter*. Specifically, the assumption about linearity implies that all scale-space properties valid for the original signal will transfer to its derivatives. Hence, there is no commitment to certain aspects of image structure, such as the zero-order representation, or its first- or second-order derivatives. The assumption of shift-invariance reflects the desire to process all spatial points identically in the absence of further information, and the requirement about a continuous scale parameter makes it unnecessary to fixate any specific scale sampling in advance.

Initially, in the absence of any information, it is natural to require all coordinate directions to be handled identically. Therefore we assume that all kernels should be *rotationally symmetric*. Let us also impose a *semi-group* condition on the family  $T$ . This means that all scale levels will be treated similarly, that is, the smoothing operation should not depend on the scale value, and the transformation from a lower scale level to a higher scale level should always be given by convolution with a kernel from the family (see (6.12)).

As smoothing criterion the *non-enhancement* requirement for local extrema is taken (see section 6.2.5). It is convenient to express it as a sign condition on the derivative of the scale-space family with respect to the scale parameter. Hence, at any non-degenerate extremum point (extremum point where the determinant of the Hessian matrix is non-zero) we require the following conditions to hold

$$\partial_t L < 0 \quad \text{at a non-degenerate local maximum,} \quad (6.39)$$

$$\partial_t L > 0 \quad \text{at a non-degenerate local minimum.} \quad (6.40)$$

qqq In the one-dimensional case, this condition is equivalent to (Babaud et al., 1986)

$$\partial_t L < 0 \quad \text{if} \quad L_{xx} < 0 \quad (6.41)$$

$$\partial_t L > 0 \quad \text{if} \quad L_{xx} > 0 \quad (6.42)$$

and since the Laplacian operator is negative (positive) at any non-degenerate local maximum (minimum) point, we can in the  $N$ -dimensional case require that

$$\operatorname{sign} \partial_t L = \operatorname{sign} \nabla^2 L \quad (6.43)$$

should hold at any extremum point where the Hessian matrix  $\mathcal{H}L$  is either positive or negative definite. To ensure a proper statement of these conditions, where differentiability is guaranteed, we shall state a series of preliminary definitions, which will lead to the desired formulation.

### 6.3.2 Definitions

Let us summarize this (minimal) set of basic properties a family should satisfy to be a candidate family for generating a (rotationally symmetric) linear scale-space.

**Definition 6.3.1.** (Pre-scale-space family of kernels)

A one-parameter family of kernels  $T: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  in  $\mathcal{L}_1$  is said to be a (rotationally symmetric) pre-scale-space family of kernels if it satisfies

- $T(\cdot; 0) = \delta(\cdot)$ ,
- the semi-group property  $T(\cdot; s) * T(\cdot; t) = T(\cdot; s + t)$ ,
- rotational symmetry

$$T(x_1, x_2, \dots, x_N; t) = T(\sqrt{x_1^2 + x_2^2 + \dots + x_N^2}, 0, \dots, 0; t)$$

for all  $x = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$  and all  $t \in \mathbb{R}_+$ , and

- the continuity requirement  $\|T(\cdot; t) - \delta(\cdot)\|_1 \rightarrow 0$  for any  $f \in \mathcal{L}_1$  when  $t \downarrow 0$ .

**Definition 6.3.2.** (Pre-scale-space representation) Given a signal  $f: \mathbb{R}^N \rightarrow \mathbb{R}$ , let  $T: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be a pre-scale-space family of kernels. Then, the one-parameter family of signals  $L: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  given by (6.38) is said to be the pre-scale-space representation of  $f$  generated by  $T$ .

Provided that the input signal  $f$  is sufficiently regular, these conditions on  $T$  guarantee that  $L$  is differentiable with respect to the scale parameter and satisfies a system of linear evolution equations.

**Lemma 6.3.3.** (A pre-scale-space representation is differentiable)

Let  $L: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  be the pre-scale-space representation of a signal  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  in  $\mathcal{L}_1$ . Then,  $L$  satisfies the evolution equation

$$\partial_t L = \mathcal{A}L \quad (6.44)$$

for some linear and shift-invariant operator  $\mathcal{A}$ .

*Proof:* If  $f$  is sufficiently regular, e.g., if  $f \in L_1$ , define a family of operators  $\{\mathcal{T}_t, t > 0\}$ , here from from  $L_1$  to  $L_1$ , by  $\mathcal{T}_t f = T(\cdot; t) * f$ . Due to the conditions imposed on the kernels, the family satisfies the relation

$$\lim_{t \rightarrow t_0} \|(\mathcal{T}_t - \mathcal{T}_{t_0})f\|_1 = \lim_{t \rightarrow t_0} \|(\mathcal{T}_{t-t_0} - \mathcal{I})(\mathcal{T}_{t_0}f)\|_1 = 0, \quad (6.45)$$

where  $\mathcal{I}$  is the identity operator. Such a family is called a strongly continuous semigroup of operators (Hille and Phillips, 1957, p. 58–59). A semi-group is often characterized by its *infinitesimal generator*  $\mathcal{A}$  defined by

$$\mathcal{A}f = \lim_{h \downarrow 0} \frac{\mathcal{T}_h f - f}{h}. \quad (6.46)$$

The set of elements  $f$  for which  $\mathcal{A}$  exists is denoted  $\mathcal{D}(\mathcal{A})$ . This set is not empty and never reduces to the zero element. Actually, it is even dense in  $L_1$  (Hille and Phillips, 1957, p. 307). If this operator exists then

$$\begin{aligned} \lim_{h \downarrow 0} \frac{L(\cdot, \cdot; t+h) - L(\cdot, \cdot; t)}{h} &= \lim_{h \downarrow 0} \frac{\mathcal{T}_{t+h} f - \mathcal{T}_t f}{h} \\ &= \lim_{h \downarrow 0} \frac{\mathcal{T}_h(\mathcal{T}_t f) - (\mathcal{T}_t f)}{h} = \mathcal{A}(\mathcal{T}_t f) = \mathcal{A}L(\cdot; t). \end{aligned} \quad (6.47)$$

According to a theorem in (Hille and Phillips, 1957, p. 308) strong continuity implies  $\partial_t(\mathcal{T}_t f) = \mathcal{A}\mathcal{T}_t f = \mathcal{T}_t \mathcal{A}f$  for all  $f \in \mathcal{D}(\mathcal{A})$ . Hence, the scale-space family  $L$  must obey the differential equation  $\partial_t L = \mathcal{A}L$  for some linear operator  $\mathcal{A}$ . Since  $L$  is generated from  $f$  by a convolution operation it follows that  $\mathcal{A}$  must be shift-invariant.  $\square$

This property makes it possible to formulate the previously indicated scale-space property in terms of derivatives of the scale-space representation with respect to the scale parameter—the grey-level value in a local maximum point must not increase with scale, whereas the grey-level value in every local minimum point must not decrease.

**Definition 6.3.4.** (Pre-scale-space property: Non-enhancement of local extrema) A pre-scale-space representation  $L: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a smooth (infinitely continuously differentiable) signal is said to possess pre-scale-space properties, or

equivalently not to enhance local extrema, if for every value of the scale parameter  $t_0 \in \mathbb{R}_+$  it holds that if  $x_0 \in \mathbb{R}^N$  is an extremum point for the mapping  $x \mapsto L(x; t_0)$  at which the Hessian matrix is (positive or negative) definite, then the derivative of  $L$  with respect to  $t$  in this point satisfies

$$\operatorname{sign} \partial_t L = \operatorname{sign} \nabla^2 L. \quad (6.48)$$

Now we can state that a pre-scale-space family of kernels is a scale-space family of kernels if it satisfies this property for *any* input signal.

**Definition 6.3.5.** (Scale-space family of kernels) A one-parameter family of pre-scale-space kernels  $T: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is said to be a scale-space family of kernels if for any smooth (infinitely continuously differentiable) signal  $f: \mathbb{R}^N \rightarrow \mathbb{R} \in L_1$  the pre-scale-space representation of  $f$  generated by  $T$  possesses pre-scale-space properties, *i.e.*, if for any signal local extrema are never enhanced.

**Definition 6.3.6.** (Scale-space representation) A pre-scale-space representation  $L: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a signal  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  generated by a family of kernels  $T: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$ , which are scale-space kernels, is said to be a scale-space representation of  $f$ .

### 6.3.3 Necessity

We shall first show that these conditions by necessity imply that the scale-space family  $L$  satisfies the diffusion equation.

**Theorem 6.3.7.** (Scale-space for continuous signals: Necessity)

A scale-space representation  $L: \mathbb{R}^N \times \mathbb{R}_+ \rightarrow \mathbb{R}$  of a signal  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  satisfies the differential equation

$$\partial_t L = \alpha \nabla^2 L \quad (6.49)$$

with initial condition  $L(\cdot; 0) = f(\cdot)$  for some  $\alpha > 0$ .

*Proof:* The proof consists of two parts. The first part has already been presented in lemma 6.3.3, where it was shown that the requirements on pre-scale-space kernels imply that a pre-scale-space family obeys a linear evolution equation where the infinitesimal generator is shift-invariant. In the second part, counterexamples will be constructed from various simple test functions to delimit the class of possible operators.

**B.1.** The extremum point condition (6.43) combined with definitions 6.3.5-6.3.6 means that  $\mathcal{A}$  must be a pure differential operator. This can be easily understood by studying the following class of counterexamples: Consider a smooth (infinitely continuously differentiable) function  $f: \mathbb{R}^N \rightarrow \mathbb{R}$  such that  $f$  has a maximum point at the origin at which the Hessian matrix is negative definite and for some  $\varepsilon > 0$   $f(x) = 0$  when  $|x| \in [\frac{\varepsilon}{2}, \varepsilon]$ . Split this function into two components

$$f = f_I + f_E \quad (6.50)$$

where

$$\nabla f_I(0) = 0, \quad (6.51)$$

$$\nabla^2 f_I(0) < 0, \quad (6.52)$$

$$f_I = 0 \quad \text{when} \quad |x| \geq \varepsilon/2, \quad (6.53)$$

$$f_E = 0 \quad \text{when} \quad |x| \leq \varepsilon. \quad (6.54)$$

Assume first that  $f_E = 0$ . Then, evaluation of  $\partial_t L = \mathcal{A}L$  at  $t = 0$  gives  $L(\cdot; 0) = f$  and  $\partial_t f = \mathcal{A}(f_I + f_E) = \mathcal{A}f_I$ . Hence, at  $(x, t) = (0, 0)$ , we must require  $\mathcal{A}f_I = C_1 < 0$ . Fixate these  $\mathcal{A}$  and  $f_I$  and consider then any  $f_E$  for which  $\mathcal{A}f_E = C_2 \neq 0$ . Then, for  $f = f_I + \beta_1 f_E$  it holds that

$$\partial_t L = \mathcal{A}f_I + \beta_1 \mathcal{A}f_E = C_1 + \beta_1 C_2. \quad (6.55)$$

Obviously, the sign of this expression can be made positive and (6.43) be violated by a suitable choice of  $\beta_1$ . Hence, for any  $\varepsilon > 0$  we have that  $\mathcal{A}f_E$  must be identically zero for all functions that assume non-zero values outside the region  $|x| < \varepsilon$ . In other words,  $\mathcal{A}$  must be a *local operator* and  $\mathcal{A}f$  can only exploit information from  $f$  at the central point. Thus, for any smooth function,  $\mathcal{A}f$  must be of the form

$$\mathcal{A}f = \sum_{\xi \in \mathbb{Z}_+^N} a_\xi L_{x^\xi} \quad (6.56)$$

where  $\xi = (\xi_1, \xi_2, \dots, \xi_N)$  is a multi-index,  $a_\xi \in \mathbb{R} \forall \xi$ , and  $L_{x^\xi} = L_{x_1^{\xi_1} x_2^{\xi_2} \dots x_N^{\xi_N}}$ .

**B.2.** The extremum point condition (6.43) also means that  $\mathcal{A}L$  must not contain any term proportional to  $L$  or derivatives of order higher than two. This can be seen by considering a test signal of the form

$$f(x) = x_1^2 + x_2^2 + \dots + x_N^2 + \beta_2 x^\eta, \quad (6.57)$$

for some  $\eta = (\eta_1, \eta_2, \dots, \eta_N) \in \mathbb{Z}^N$  with  $|\eta| = |\eta_1| + |\eta_2| + \dots + |\eta_N| > 2$ . If  $a_\xi \neq 0$  for some  $\xi \in \mathbb{Z}^N$  it is clear that a suitable choice of  $\beta_2$  can make the sign of  $\mathcal{A}f$  arbitrary and hence violate (6.43). Similarly, by considering a test signal of the form

$$f(x) = x_1^2 + x_2^2 + x_N^2 + \beta_3 \quad (6.58)$$

it follows that  $a_0 = 0$ . Thus,  $\mathcal{A}$  can only contain derivatives of order one and two.

**B.3.** Since the scale-space kernels are required to be rotationally symmetric, it follows that the contributions from the first-order derivatives as well as the mixed second-order derivatives must be zero. Moreover, the contributions from the second-order derivatives along all coordinate directions must be similar. Hence, the only possibility is that  $\mathcal{A}$  is a constant times the Laplacian operator.  $\square$

If the requirements about rotational symmetry are relaxed, then  $\mathcal{A}$  will be a linear combination of first- and second-order derivatives, for which the coefficients of the second-order derivative terms correspond to a positive definite quadratic form and the coefficients of the first-order terms are arbitrary.

### 6.3.4 Sufficiency

The reverse statement of theorem 6.3.7, *i.e.* the fact that  $L$  generated by (6.49) satisfies definition 6.3.4, is obvious, since at extremum points where  $\mathcal{H}L$  is positive or negative definite, we have  $\nabla^2 L < 0$  at maxima and  $\nabla^2 L > 0$  at minima.

### 6.3.5 Application to scale invariant semi-groups

By combining theorem 6.3.7 with the treatment in section 6.2.6, it follows that the Gaussian kernel is the only rotationally symmetric scale invariant semi-group that satisfies the causality requirement. Thus, also from this point of view, selecting  $p = 2$  in (6.37) constitutes a very special choice.

Besides the fact that the additional requirements concerning locality of the infinitesimal generator and positivity of the convolution operator can be replaced by causality, it is worth noting that causality and semi-group structure *per se* imply that the infinitesimal generator must be local and the convolution kernel must be non-negative. More importantly, if we assume a semi-group structure and combine it with causality and rotational invariance, then scale invariance arises as a consequence and is not required as an axiom.

## 6.4 Summary and conclusions

We have seen how a scale-space formulation for continuous signals can be stated based on the essential assumptions about a semi-group and a causality requirement expressed as non-enhancement of local extrema. Combined with the assumption about convolution operations and a certain regularity assumptions along the scale direction (strong continuity; *i.e.* continuity in norm) the semi-group structure implies that the multi-scale representation is differentiable along the scale direction and has an infinitesimal generator. The causality requirement then, in turn, implies that the infinitesimal generator must be local (correspond to a differential operator) and be a linear combination of derivatives of orders one and two. The essence of these results is that the scale-space representation on a spatial domain is given by a (possibly semi-discretized) parabolic differential equation corresponding to a *second-order* differential operator with respect to the spatial coordinates, and a *first-order* differential operator with respect to the scale parameter.

Rotational symmetry implies that no first-order derivatives are allowed and that the second-order derivatives must occur in a combination such that the differential operator is a constant times the Laplacian. In this case, the scale-space family is generated by convolution with rotationally symmetric Gaussian kernels.

It has also been described how the causality requirement relates to scale-space formulations based on semi-group structure combined with scale invariance. Unless additional conditions are imposed, these assumptions give rise to a one-parameter family of smoothing kernels, and scale invariance *per se* does not uniquely single out the Gaussian kernel. Previously, it has been shown that the Gaussian arises as a unique choice if the scale-space family is required to have a local infinitesimal generator and the smoothing kernel is required to be positive. As a corollary, it

follows that the Gaussian is a unique choice if the additional assumptions about locality and positivity are replaced by a causality assumption. In fact, causality implies locality and positivity. More importantly, when combined with the semi-group structure, the causality assumption gives rise to scale invariance.

Concerning scale-space representation for discrete signals, scale invariance obviously cannot be used if the discrete signal constitutes the only available data.<sup>6</sup> A perfectly scale invariant operator cannot be expressed on a discrete grid, which has a certain preferred scale given by the distance between adjacent grid points. The formulation based on non-enhancement/causality applies in both domains, provided that the definition of local maximum and the requirement about rotational symmetry are appropriately modified.

## 6.5 Extensions of linear scale-space

A natural question then arises: Does this approach constitute the *only* reasonable way to perform the low-level processing in a vision system, and are the Gaussian kernels and their derivatives the only smoothing kernels that can be used? Of course, this question is impossible to answer to without further specification of the purpose of the representation, and what tasks the visual system has to accomplish. In any sufficiently specific application it should be possible to design a smoothing filter that in some sense has a “better performance” than the Gaussian derivative model. For example, it is well-known that scale-space smoothing leads to shape distortions at edges by smoothing across object boundaries. Distortions arise also in algorithms for estimating local surface shape, such as shape-from-texture and shape-from-disparities. Hence, it should be emphasized that the linear scale-space model is rather aimed at describing the principles of the very first stages of low-level processing in an *uncommitted* visual system aimed at handling a large class of different situations, and for which no or very little *a priori* information is available.

Then, once initial hypotheses about the structure of the world have been generated within this framework, the intention is that it should be possible to invoke more refined processing, which can compensate for these effects and adapt to the current situation and the task at hand. These are the motivations for studying non-uniform scale-space concepts, such as affine Gaussian scale-space and non-linear diffusion techniques. From the viewpoint of such approaches, the linear scale-space model is intended to serve as a natural starting point.

### 6.5.1 Relaxing rotational symmetry

**Affine Gaussian scale-space.** A straightforward extension of the raw linear scale-space representation can be obtained by relaxing the requirement about ro-

---

<sup>6</sup>If further information is available about the image formation process, *e.g.*, such that the continuous signal can be reconstructed exactly from the sampled data, then the discrete signal can be treated as equivalent to the original continuous signal, and an equivalent discrete scale-space model be expressed for the continuous scale-space representation of the reconstructed continuous signal. Chapter 9 by Åström and Heyden (1996) in this volume exploits this idea based on the sampling theorem and the assumption of an ideally sampled band limited signal.

tational symmetry in definition 6.3.1. Then, the same way of reasoning as in sections 6.3.2–6.3.4 still applies. The only essential differences are that part B.3 in the proof of theorem 6.3.7 should be omitted and that the Laplacian operator in equation (6.49) should be replaced by an arbitrary linear and symmetric (elliptic) second-order differential operator.

In terms of convolution operations, the resulting (three-parameter) *affine Gaussian scale-space representation* is generated by non-uniform Gaussians defined by

$$g(x; \Sigma_t) = \frac{1}{(2\pi)^D/2 \sqrt{\det \Sigma_t}} e^{-x^T \Sigma_t^{-1} x / 2} \quad (6.59)$$

where  $\Sigma_t$  is a symmetric positive definite (covariance) matrix. If the covariance matrix is written  $\Sigma_t = t\Sigma_0$  for some (constant) matrix  $\Sigma_0$ , then the shape-adapted multi-scale representation satisfies the diffusion equation

$$\partial_t = \frac{1}{2} \nabla^T (\Sigma_0 \nabla L). \quad (6.60)$$

This representation, considered in (Lindeberg, 1994e), satisfies all the scale-space properties listed in sections 6.2–6.3 except those specifically connected to rotational symmetry. For example, because of the linearity of this operation, all scale-space properties transfer to spatial derivatives of the scale-space representation as well as to linear combinations of these.<sup>7</sup><sup>8</sup>

Actually, the affine Gaussian scale-space representation of any signal  $f$  is equivalent to the result of applying the rotationally symmetric linear scale-space concept to a transformed image. More specifically, it can be constructed by composing the following operations in cascade: (i) subject the original signal  $f$  to an affine transformation, (ii) apply the linear scale-space representation to the deformed signal, (iii) subject the smoothed signal to the inverse affine transformation.<sup>9</sup> By varying the affine transformations, we can in this way span the family of affine Gaussian scale-space representations.

A major motivation for considering the affine Gaussian scale-space representation is that it is *closed* under affine transformations. For two patterns  $f_L$  and  $f_R$  related by an invertible linear transformation  $\eta = B\xi$

$$f_L(\xi) = f_R(B\xi) \quad (6.61)$$

---

<sup>7</sup>For the corresponding transformation kernels from the input signal  $f$  to a linear combination of scale-space derivatives, however, the semi-group structure is replaced by a cascade smoothing property. This means that any transformation kernel  $h(\cdot; t)$  corresponds to the result of convolving some fixed kernel  $h_0$  with a Gaussian kernel, *i.e.*,  $h(\cdot; t) = h_0 * g(\cdot; t)$ . Hence, these kernels satisfy  $h(\cdot; s+t) = g(\cdot; s) * h(\cdot; t)$  where  $g$  denotes the Gaussian kernel.

<sup>8</sup>Moreover, as shown in chapter 12 by Griffin (1996), the classification of what scale-space singularities (bifurcations) can occur with increasing scale, transfers from the rotationally symmetric Gaussian scale-space to the affine Gaussian scale-space (see (Koenderink and Doorn, 1986a; Lindeberg, 1992; Lindeberg, 1994e) as well as chapter 11 by Damon (1996), chapter 10 by Johansen (1996) chapter 13 by Kalitzin (1996)). Examples of image representations depending on this *deep structure of scale-space* can be found in (Lindeberg, 1993a) and in chapter 14 by Olsen (1996).

<sup>9</sup>This essentially corresponds to the duality between transformations of image operators and image domains described in detail in chapter 5 by Florack (1996) in this volume.

the corresponding affine Gaussian scale-space representations

$$L(\cdot; \Sigma_L) = g(\cdot; \Sigma_L) * f_L(\cdot), \quad (6.62)$$

$$R(\cdot; \Sigma_R) = g(\cdot; \Sigma_R) * f_R(\cdot), \quad (6.63)$$

are related by

$$L(\xi; \Sigma_L) = R(\eta; \Sigma_R), \quad (6.64)$$

where

$$\Sigma_R = B \Sigma_L B^T. \quad (6.65)$$

Hence, for any matrix  $\Sigma_L$  there exists a matrix  $\Sigma_R$  such that the affine scale-space representations of  $f_L$  and  $f_R$  are equal (see the commutative diagram in figure 6.1), and any non-singular affine transformations can be captured exactly within this three-parameter multi-scale representation.<sup>10</sup> This extension of the linear scale-space concept is useful, for example when estimating local surface orientation from local deformations of surface patterns (Lindeberg and Gårding, 1993; Gårding and Lindeberg, 1994; Gårding and Lindeberg, 1996)) and more generally, whenever computing affine image deformations, such as in first-order optic flow and first-order stereo matching. Chapter 5 by Florack (1996) in this volume exploits this idea further, by emphasizing the equivalence between deformations of images and filters.

$$\begin{array}{ccc} L(\xi; \Sigma_L) & \longrightarrow & R(\eta; B \Sigma_L B^T) \\ \uparrow & & \uparrow \\ *g(\cdot; \Sigma_L) & & *g(\cdot; B \Sigma_L B^T) \\ | & & | \\ f_L(\xi) & \longrightarrow & f_R(\eta) \end{array}$$

Figure 6.1: Commutative diagram of the non-uniform scale-space representation under linear transformations of the spatial coordinates in the original image.

**Discrete affine Gaussian scale-space.** The discrete counterpart of the affine Gaussian scale-space is obtained by relaxing the symmetry requirements of the infinitesimal generator in (6.23), while preserving the locality, positivity and zero sum constraints. In addition, to avoid spatial shifts, the Fourier transform should be required to be real, implying  $a_{i,j} = a_{-i,-j}$ . This gives rise to the *discrete affine*

<sup>10</sup>Compared to the affine invariant level curve evolution scheme proposed by (Alvarez et al., 1993) and (Sapiro and Tannenbaum, 1993), given as a one-parameter solution to a non-linear differential equation (see equation (6.73) in section 6.5.2), an obvious disadvantage of the affine Gaussian scale-space is that it gives rise to a three-parameter variation. The advantage is that commutative properties can still be preserved within a family of linear transformations.

*Gaussian scale-space* generated by an infinitesimal generator having a computational molecule of the form

$$\mathcal{A} = \begin{pmatrix} -C/2 & B & C/2 \\ A & -2A - 2B & A \\ C/2 & B & -C/2 \end{pmatrix} + \alpha \begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}. \quad (6.66)$$

This representation can be interpreted as a second-order discretization of the diffusion equation (6.60) associated with the continuous affine Gaussian scale-space

$$\partial_t L = AL_{xx} + BL_{yy} + 2CL_{xy} \quad (6.67)$$

where  $A > 0$  and  $AB - C^2 > 0$  are necessary requirements for the operator to be elliptic. The free parameter  $\alpha$  (which controls the addition of a discretization of the mixed fourth-order derivative  $L_{xxxx}$ ) should be in the interval  $C/2 \leq \alpha \leq \min(A, B)/2$  to ensure that all non-central coefficients are non-negative.

As a general tool, the affine Gaussian scale-space constitutes a useful framework for adapting the shape of the smoothing kernels in situations where further information is available, and the process of performing a one-parameter variation of the shapes of the smoothing kernels to the image data is referred to as *shape adaptation* (Lindeberg, 1994e; Lindeberg and Gårding, 1994). Alternatively, it can be seen as a way of obtaining invariance to affine transformations by expanding the data into a three-parameter scale-space representation instead of the more common one-parameter approach.

In chapter 2 by Almansa and Lindeberg (1996), this concept is used for expressing an *adaptation of linear smoothing operations to image data*, guided by a non-linear image descriptor (a second moment matrix) and a non-linear scale selection algorithm. Non-linear smoothing approaches with a large number of structural similarities are considered in chapter 16 by Weickert (1996). An extension of the affine Gaussian scale-space concept to a more general *spatio-temporal scale-space representation*, allowing for *velocity adaptation*, is presented in (Lindeberg, 1996c).

### 6.5.2 Non-linear smoothing

Let us finally discuss how the scale-space formulation based on non-enhancement of local extrema applies to a number of non-linear scale-space approaches. Clearly, any evolution scheme of the form

$$\partial_t L = \nabla^T(c(x; t)\nabla L) \quad (6.68)$$

satisfies the non-enhancement/causality requirement as long as the conduction coefficient  $c(x; t)$  is non-negative. (Since  $\partial_t L = \nabla^T c(x; t) \nabla L + c(x; t) \nabla^2 L = c(x; t) \nabla^2 L$  at local extrema it follows that  $\partial_t L$  has the same sign as  $\nabla^2 L$ .) In contrast to linear (and affine) scale-space representation, however, this property does not necessarily extend to spatial derivatives of the scale-space representation.

**Anisotropic diffusion.** For the edge enhancing anisotropic diffusion scheme proposed by (Perona and Malik, 1990), where

$$c(x; t) = h(|\nabla L(x; t)|) \quad (6.69)$$

for some function  $h: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ , the causality violation in the first-order derivative can be illustrated as follows: Consider, for simplicity, the one-dimensional case

$$\partial_t L = \partial_x(h(|L_x|)L_x) = h_x(|L_x|)L_x^2 + h(|L_x|)L_{xx} \quad (6.70)$$

and introduce  $\phi(L_x) = h(|L_x|)L_x$ . Then, following (Whitaker and Pizer, 1993), the evolution equation can be written  $\partial_t L = \partial_x(\phi(L_x)) = \phi'(L_x)L_{xx}$  and the gradient magnitude satisfies

$$\partial_t L_x = \phi''(L_x)L_{xx}^2 + \phi'(L_x)L_{xxx}. \quad (6.71)$$

For a local gradient maximum with  $L_{xx} = 0$  and  $L_{xxx} < 0$  it holds that  $\partial_t L_x > 0$  if  $\phi'(L_x) < 0$ . For the conductivity function used by (Perona and Malik, 1990)

$$h(|\nabla L|) = e^{-|\nabla L|^2/k^2}, \quad (6.72)$$

where  $k$  is a free parameter, we have  $\phi'(L_x) < 0$  if  $L_x > k/\sqrt{2}$ . In other words, gradients that are sufficiently strong will be enhanced, and gradients that are sufficiently weak will be suppressed (no matter what are the spatial extents of the image structures). Moreover, this evolution equation depends upon an external parameter and is not scale invariant.

**Affine invariant scale-space.** Recently, (Alvarez et al., 1993) and (Sapiro and Tannenbaum, 1993) have presented an affine invariant (one-parameter) scale-space representation which essentially commutes with affine transformations. It is generated by the evolution equation

$$\partial_t L = (|\nabla L|^3 \kappa(L))^{1/3} \quad (6.73)$$

where  $\kappa(K) = (L_x^2 L_{yy} + L_y^2 L_{xx} - 2L_x L_y L_{xy})/(L_x^2 + L_y^2)^{3/2}$  is the curvature of a level curve of the intensity distribution. In (Alvarez et al., 1993), it is shown that this representation is essentially unique if semi-group structure and existence of a local infinitesimal generator are combined with invariance under monotone intensity transformations as well as affine transformations of the spatial domain.<sup>11</sup> From the affine invariance, it follows that the evolution equation is scale invariant. Regarding causality, however,  $\partial_t L$  is exactly zero at any local extremum.

Concerning other non-linear scale-space approaches, the reader is referred to (Nordström, 1990; Nitzberg and Shiota, 1992; Florack et al., 1993b; Whitaker and Pizer, 1993), the book (Haar Romeny, 1994), as well as chapter 15 by van den Boomgaard (1996) and chapter 16 by Weickert (1996) in this volume.

---

<sup>11</sup>Note the difference in terminology in (Alvarez et al., 1993), where the semi-group structure is referred to as causality.

## Acknowledgments

The support from the Swedish Research Council for Engineering Sciences, TFR, is gratefully acknowledged. An earlier version of this manuscript was presented in (Lindeberg, 1994a).

# Chapter 7

# Scale-Space Generators and Functionals

**Mads Nielsen**

*3D-Lab, School of Dentistry, University of Copenhagen  
Nørre allé 25, 2200 Copenhagen N, Denmark*

## 7.1 Introduction

In this chapter we will try to relate axiomatic formulations of scale-spaces with regularization theory. Most axiomatic formulations do not directly require regularity properties of scale-space. Nevertheless, an analytic filter is singled out so that scale-space appears to have strong regularity properties. These properties could also be stated as first principles in terms of Tikhonov regularization. In this chapter we study how scale-space properties as described in previous chapters and regularity properties interact in axiomatic formulations.

In the previous chapters, we have seen how a set of axioms regarding a scale-space representation of an image can narrow down the set of admissible scale-space operators. We call an operator  $G$  a scale-space generator if it is parametrised by a scale parameter  $t$  so that

$$I(\cdot, t) = G_t I_0(\cdot)$$

where  $I_0 : \mathbb{R}^N \mapsto \mathbb{R}$  is the input image and  $I : \mathbb{R}^N \times \mathbb{R} \mapsto \mathbb{R}$  its scale-space. The basic assumptions, the axioms, of what a scale space is, limits the class of admissible generators.

In this chapter we take as the starting point that  $G_t$  is an isotropic convolution operator. In chapter 6 we saw that this follows from the axioms of the generator being linear and not preferring any position of orientation. Linearity can be argued from the idea that no intensity value is preferred to others (Florack, 1993). Hence,

we start by the uncommitted choice of not preferring any position, orientation, or intensity<sup>1</sup>.

Koenderink (Koenderink, 1984) used the additional principle of non-creation of detail (i.e. iso-intensity curves) for increasing scale and thereby singled out the heat equation having the Gaussian kernel as Green's function. In this approach the nice scaling behaviour of the Gaussian scale-space is a bonus you get for free and not one of the first principles. The Gaussian kernel can be derived in many other ways including principles of scale invariance (Florack, 1993; Lindeberg, 1994e; Poggio and Yuille, 1985).

In this chapter we start by scale invariance of the generator and the recursivity principle which states that two successive scalings must be describable as only one. Pauwels et al. (Pauwels et al., 1995) showed that this reveals a continuous one-parameter class of scale-space filters. This is shown in Section 7.2.

One of the major reasons to generate a scale-space is that all its derivatives are well-defined at all scales larger than zero. Given a measured (digitally represented) signal in one or more dimensions, its derivatives can be defined in distributional sense (Schwartz, 1966) by convolving the signal with a smooth test function. Instead of taking the derivatives of the signal, we take the derivatives of the convolved signal by differentiation of the smooth filter prior to convolution. As a consequence the derivatives of any integrable signal are operationally defined and well-posed (see Chapter 5). In this way tools from differential geometry can be applied to measurement data, defining, for example, image features as differential invariants (Florack, 1993).

The assignment of derivatives to a digital signal can also be performed as a mapping of the signal into a space of functions possessing the desired regularity properties. This mapping can be performed as a Tikhonov regularization (Tikhonov and Arsenin, 1977), mapping an integrable function into a Sobolev space (space of functions having square integrable derivatives up to some order) of given order. This regularization is formulated as a functional minimization problem, since we seek the function in the Sobolev space which in some precise sense is closest to the original signal.

In Section 7.3 we review Tikhonov regularization and show that the scale-invariant regularizations fulfilling the recursivity principle correspond to a subclass of all linear filterings with these properties. This class is a discrete set of the one-parameter class of linear scale-spaces derived by Pauwels which is exactly the subset having local infinitesimal generators, i.e. those scale-spaces where the evolution over scale can be described by a partial differential equation (PDE).

All the scale-space regularizations are described by functionals including terms of differentiation up to infinite order. If the functional is truncated at some lower order, the recursivity principle is violated. In Section 7.4 we show, however, that truncations at certain orders implements some well-known edge detection filters such as the Castan filter (Castan et al., 1990) and the Canny-Deriche filter (Deriche, 1987). In Section 7.5 we show that the filters corresponding to any finite order Tikhonov regularization functional can be implemented recursively without

---

<sup>1</sup>Since the generator is a linear convolution operator we also denote it convolution kernel, linear filter, etc.

any other approximation than discretisation.

Many propositions and proofs are given for 1D signals to gain clarity. However, all the above results are valid for rotational invariant filters in higher dimensions. For the exact propositions and their proofs, we refer to Nielsen et al. (Nielsen et al., 1996b).

## 7.2 Linear scale-space generators

In this section we look into the admissible set of scale-space generators when we augment the convolution property with principles of scale-recursivity, scale invariance, and locality. A scale-space  $I : \mathbb{R}^N \times \mathbb{R} \mapsto \mathbb{R}$  is constructed from an image  $I_0 : \mathbb{R}^N \mapsto \mathbb{R}$  by a linear convolution with a filter  $h(\cdot, t)$ .

$$I(\cdot, t) = I_0(\cdot) * h(\cdot, t)$$

The scale-space filter is also assumed to be isotropic which in 1D is analogous to being even. We now also assume the filter property of being scale-recursive:

**Definition 7.2.1. (Scale-recursive filters)**  $h : \mathbb{R}^N \times \mathbb{R} \mapsto \mathbb{R}$  is a one-parameter filter and is said to be scale-recursive if

$$\forall s, t \geq 0 : h(\cdot, s) \circ h(\cdot, t) = h(\cdot, s \oplus t)$$

where  $\oplus$  is a general addition operation.

In the case where  $h$  is a convolution filter, we denote the filter concatenation  $*$ . In this section we will assume the scale addition operator  $\oplus$  to be the normal  $+$  operator<sup>2</sup>. Convolution filters which are also scale-recursive have Fourier transforms of the form (Rudin, 1981; Pauwels et al., 1995) (provided they exist)

$$\hat{h}(\omega, t) = e^{-r(\omega)t}$$

where  $r : \mathbb{R}^N \mapsto \mathbb{R}$  is real since  $h$  is even and even since  $h$  is real, and  $r(1) = 1$  can be accomplished by a scaling of  $t$ .

The above constraints limit  $h$  to form a *convolution algebra* as pointed out in Chapter 5. Augmented by the criteria that  $h$  is smooth, local (decays faster than polynomially), and positive, the Gaussian kernel is singled out. A smooth and local function is a Schwartz function.

In the following we augment the algebraic convolution property by scale-invariance and locality as the fundamental property of a *scale-space*. A scaling  $S_c$  maps the points of the spatial domain according to  $S_c : x \mapsto cx$ ,  $c > 0$ . We say that a filter family  $h(x, s)$  commutes with this scaling operation if for all images  $I(x)$

$$S_c(h(x, s) * I(x)) = h(S_c(x), s') * I(S_c(x)) \quad (7.1)$$

where  $s' = \phi(s, c)$  can be uniquely identified. We call this *scale-invariant filtering*.

---

<sup>2</sup>In following sections we use metric properties of the scale parameter and a redefinition of the scale parameter is necessary so that also the addition operator must be substituted.

**Proposition 7.2.2. (Scale-invariant filters)** A filter  $h$  is scale-invariant if

$$h(x, t) = \frac{1}{\psi(t)} \hat{h}\left(\frac{x}{\psi(t)}\right)$$

where  $\psi : \mathbb{R}^+ \mapsto \mathbb{R}^+$  is an increasing function and  $\hat{h}$  is the filter at unit scale. Without loss of generality we can assume  $\psi(1) = 1$ .

*Proof.* Follows directly from the scale-commuting properties, identifying  $\psi(t) = \phi(1, t)$ , where  $\phi$  is defined in relation to Eq. 7.1.  $\square$

We are now ready to make the following proposition (Pauwels et al., 1995):

**Proposition 7.2.3. (Scale-invariant scale-recursive filters)** A filter is scale-invariant and scale-recursive if

$$\hat{h}(\omega, t) = e^{-|\omega|^\alpha t}$$

where  $\hat{h}$  is the Fourier transform of  $h$  and  $\alpha \in \mathbb{R}^+$ .

*Proof.* The Fourier transform of the scale-invariant filter is

$$\hat{h} = \hat{\tilde{h}}(\psi(t)\omega)$$

Since it fulfils also the recursivity property we use Eq. 7.2 and write

$$\hat{\tilde{h}}(\psi(t)\omega) = e^{-r(\omega)t} \quad (7.2)$$

From  $\psi(1) = 1$  we find  $r(\omega) = -\log \hat{h}(\omega)$ . This and  $r(1) = 1$  imply that  $r(\psi(t)) = t$ , so we find (logarithm of Eq. 7.2)

$$r(s\omega) = r(\omega)t \equiv r(\omega)r(s)$$

where  $s \equiv r^{-1}(t) = \psi(t)$ . This implies that  $r(\omega) = |\omega|^\alpha$ ,  $\alpha \in \mathbb{R}^+$ .  $\square$

This is a one-parameter class of scale-invariant scale-recursive filters. The filters have properties according to the choice of  $\alpha$ . All filters are smooth ( $C^\infty$ ).

**Proposition 7.2.4. (Local scale-space filters).** Only for  $\alpha = 2n$ ,  $n \in \mathbb{N}$  the scale-space filters of Eq. 7.2.3 are local, i.e Schwartz functions.

*Proof.* For a function to be local (in Schwartz's sense) all moments must be finite. The moments  $M_n(h)$  are

$$M_n(h) = \left. \frac{i\partial^n}{\partial\omega^n} \hat{h} \right|_{\omega=0}$$

If  $2(n-1) < \alpha \leq 2n$  then the  $2n$ th derivative of  $\hat{h}$  is defined in zero only if  $\alpha = 2n$ . If  $\alpha = 2n$  all derivatives are bounded.  $\square$

This is proven by Pauwels et. al (Pauwels et al., 1995) although stated on the form of existence of an infinitesimal generator on differential form. That is, only for  $\alpha = 2n$  the scale-space can be generated by a PDE.

So far we have derived a one-parameter class of scale-space filters and shown that only a discrete set of these are local (i.e. Schwartz functions). In the following we look into how this class connects to regularization in the sense of Tikhonov(Tikhonov and Arsenin, 1977).

### 7.3 Linear regularization

In Tikhonov regularization the primary goal is to define well-posed derivatives by mapping into a Sobolev space. Square integrable functions are mapped into a Sobolev space of specified order. A Sobolev space of order  $N$  is the space of all functions being square integrable, and having derivatives up to order  $N$  all defined and square integrable. In the following we shall find the subset of regularizations also constructing scale-spaces.

Regularization of a signal  $g \in \mathcal{L}^2(\mathbb{R})$  can be formulated (Tikhonov and Arsenin, 1977) as the minimization with respect to the regularized solution  $f$  of an energy functional  $E$ .

**Definition 7.3.1. (Tikhonov regularization)** The Tikhonov regularized solution  $f$  of the signal  $g \in \mathcal{L}^2(\mathbb{R})$  minimises the energy functional

$$E[f] \equiv \frac{1}{2} \int dx \left( (f - g)^2 + \sum_{i=1}^{\infty} \lambda_i (\frac{\partial^i}{\partial x^i} f)^2 \right) \quad (7.3)$$

with nonnegative  $\lambda_i$ .

When  $\lambda_n \neq 0$  and  $\lambda_i = 0$  for all  $i > n$  we talk about  $n$ th order regularization. The functional is convex and a unique minimum exists. It can be found by linear filtering:

**Proposition 7.3.2. (Regularization by convolution)** Linear convolution of the signal  $g \in \mathcal{L}^2(\mathbb{R})$  by the filter  $h$ , having the Fourier transform

$$\hat{h} = \frac{1}{\sum_{i=0}^{\infty} \lambda_i \omega^{2i}} \quad (7.4)$$

yields the regularized solution of Definition 7.3.1. By definition  $\lambda_0 \equiv 1$ .

*Proof.* In the Fourier domain the energy functional (7.3) yields (according to Parseval's Theorem)

$$E[\hat{f}] = \frac{1}{2} \int d\omega \left( (\hat{f} - \hat{g})^2 + \sum_{i=1}^{\infty} \lambda_i \omega^{2i} \hat{f}^2 \right)$$

A necessary (and sufficient) condition for  $E$  to be minimised with respect to  $\hat{f}$  is that the variation of  $E$  is zero:

$$0 = \frac{\delta E}{\delta \hat{f}} = (\hat{f} - \hat{g}) + \sum_{i=1}^{\infty} \lambda_i \omega^{2i} \hat{f} \quad \Leftrightarrow \quad \hat{f} = \frac{1}{1 + \sum_{i=1}^{\infty} \lambda_i \omega^{2i}} \hat{g}$$

The optimal  $\hat{f}$  can thus be found by a linear filtering of the initial signal  $g$  with the filter  $h$  given by Eq. 7.4.  $\square$

Thus any regularization using only sums of quadratic terms of the derivatives of the solution can be reformulated as a linear convolution with the Fourier inverse (provided it exists) of  $\hat{h}$  as given by (7.4). Since regularizations perform linear filterings, we analyse the scale-space properties of these filters. We make the following proposition:

**Proposition 7.3.3. (Completeness of regularization)** Regularization on the form of (7.3) can implement any real, normalised, even, and local filter.

*Proof.* Regularization implements the linear filtering with  $\hat{h}$  of (7.4). Since the denominator of  $\hat{h}$  can be the Taylor series expansion of any analytic, even, and real function which is unity in zero, the filter  $\hat{h}$  itself can be any analytic, even, and real function which is unity in zero. Hence, the spatial filter  $h$  can be any even, real, and normalised filter with an analytic Fourier transform. Analytic Fourier transform translates to local in the spatial domain, where local is precisely defined as decaying faster than polynomially towards plus/minus infinity, i.e. having finite moments.  $\square$

### 7.3.1 Scale-invariant regularization

In the previous section we treated scale-invariant filters by explicitly demanding scaling properties of the filter but without relating the scale parameter to any metric quantity. We could apply a similar analysis for a one-parameter subclass of regularization filters, but in order to demonstrate another technique inspired by physics, we here choose the tool of dimensional analysis.

For an action to be scale-invariant it must be dimensionless, i.e. without physical units (Fourier, 1822). In the following we use dimensional analysis to insure scale invariance of the filters. The notation  $[ \cdot ]$  denotes “the dimension of”.

**Proposition 7.3.4. (Scale-invariant regularization)** A regularization implemented by a filter  $\hat{h}(\omega, s)$  is invariant to a scaling  $S : x \mapsto cx$ ,  $c > 0$  if

$$\hat{h}(\omega, s) = \frac{1}{\sum_0^{\infty} \lambda_i \omega^{2i}} \quad \text{where} \quad \lambda_i(s) = a_i s^i$$

and  $a_i \in \mathbb{R}_+$  are scale independent (dimensionless) constants.

*Proof.* If every term of the filter is dimensionless, then the filter is scale-invariant. Since  $[\omega] = \text{length}^{-1}$ , then  $[\lambda_i]$  must be  $\text{length}^{2i}$ . This is ensured since  $[s] = \text{length}^2$ .  $\square$

This construction corresponds to a special choice of the scaling function  $\psi(t)$  in Prop. 7.2.2 so that  $\psi(t) \propto t^2$ . The above construction does not limit the class of possible regularizations. It expresses how the parameters of the regularization must transform if a change of units of the spatial domain is taking place. This is similar to the mathematical definition in previous section: the filter of unit scale can be any filter, the scale invariance only reflects how it changes under a re-scaling.

We now add the recursivity constraint to the class of scale-invariant regularization filters. Here we state it in terms of the slightly stronger semi-group constraint:

**Definition 7.3.5. (Semi-group property)** A *semi-group*  $G$  is a set of compositions:  $G \times G \rightarrow G$  which is associative.

Because regularizations can be implemented as linear filters, and linear filters are associative, so are regularizations. Regularization forms a semi-group, since it also satisfies the requirement  $G \times G \rightarrow G$ , which is the recursivity constraint.

If we embed  $h(\cdot)$  into a one-parameter family  $h(\cdot, s)$ , this family is a subset of all regularizations and this subset may not satisfy the semi-group constraint. We may formulate the *semi-group* property of the one-parameter family  $h(s)$  as:

**Definition 7.3.6. (Semi-group property of filter family)** The semi-group property of the one-parameter family of linear filters  $h(\cdot, s)$  demands that

$$\forall s, t \in \mathbb{R}_+ : h(\cdot, s \oplus t) = h(\cdot, s) * h(\cdot, t)$$

in which the parameter-concatenation is a general addition:

$$s \oplus t = \gamma^{-1}(\gamma(s) + \gamma(t))$$

where  $\gamma(\cdot)$  is a monotonic function.

To satisfy a dimensional analysis the scale addition must add up quantities of same units implying that  $\gamma(\cdot)$  must be a monomial and the scale concatenation must be the  $p$ -norm addition:

$$s \oplus_p t = (s^p + t^p)^{1/p} , \quad p > 0$$

In the following we will use the terms scale-recursivity and semi-group property interchangeably, since they are identical for linear filters.

**Proposition 7.3.7. (Scale-invariant semi-group regularization)** If a one-parameter regularization filter family  $h(s)$ ,  $s \in \mathbb{R}_+$  is scale-invariant (Prop. 7.3.4) and the filter fulfils the semi-group constraint using the  $p$ -norm addition, then the filter family can be written as

$$\hat{h}_p(\omega, t) = e^{-(\omega^2 t)^p} = \frac{1}{\sum_{i=0}^{\infty} \frac{t^{ip}}{ip!} \omega^{2ip}}$$

here given in the Fourier domain, where  $t \propto s$ .

*Proof.* A direct proof of this is given earlier (Nielsen et al., 1996b). Here we use the properties of scale-invariant scale-recursive filters derived in the previous section. A scale-invariant scale-recursive filter can be written (Prop. 7.2.3)

$$\hat{h} = e^{-|\omega|^\alpha t}$$

up to a possible scale re-synchronization  $t' = \phi(t)$ . Since regularization have analytical Fourier transforms (i.e. they are local) then  $\alpha = 2n$ ,  $n \in \mathbb{N}$  (Prop. 7.2.4). This gives

$$\hat{h} = e^{-\omega^{2n}\phi(t)}$$

From the dimensional analysis  $[\omega^2] = \text{length}^2$ , using a scale parameter  $t$  of dimension length squared, we identify the dimensionless exponent so that

$$\hat{h} = e^{-\omega^{2n}t^n}$$

To this filter the  $p$ -norm of scale addition applies if  $n = p$ . The corresponding constants  $\lambda_i$  are found by a Taylor series expansion of the analytical quantity  $1/\hat{h}$ .  $\square$

We see that we obtain the well-known Gaussian for  $p = 1$ . That is if scale parameters of dimension length is added up by Pythagoras formula.

All  $\hat{h}_p$ ,  $p \in \mathbb{N}$  are Fourier-invertible. For  $p = 1$  it is the Gaussian, for  $p$  increasing towards infinity, it converges towards the ideal low-pass filter with cut-off frequency  $\omega_c = 1/\sqrt{t}$ .

The above filters were derived from the class of regularization filters, but constitute the full class of all local scale-invariant scale-recursive filters as it was derived in the previous section. This is, of course, due to the completeness of the class of regularization filters.

In Chapter 5 the filters were assumed to be Schwartz functions. The Schwartz functions are local and analytical, i.e. they are local  $C^\infty$  functions having a convergent Taylor series. The regularization filters constitute a larger set of filters since they are local, but only analytical if regularization of infinite order is applied. In general an  $n$ th order regularization filter is only  $C^n$ . Only if we imply that for all  $i$  there exists a  $j > i$  such that  $\lambda_j > 0$ , then the corresponding regularization filter is  $C^\infty$ . The general scale-invariant semi-group filters are taken from another super-set of the Schwartz functions. They are  $C^\infty$  since all moments of the Fourier transform are finite, but they are only local if  $\alpha = 2n$ . If  $2(n-1) < \alpha < 2n$  then the even moments of order  $2n$  or larger are not finite. All the scale-invariant semi-group filters of  $\alpha = 2n$  are Schwartz functions, but only for  $\alpha = 2$  they are positive. The intersection of the filter spaces of Schwartz functions, Tikhonov regularizations, and scale-space filters is illustrated in Fig. 7.1.

### 7.3.2 Evolution equations

In the above we adjusted the weighting of the derivatives in Tikhonov regularization to implement Gaussian filtering ( $p = 1$ ). We can also reverse the argumentation: Gaussian scale-space can be implemented as the minimization of an energy

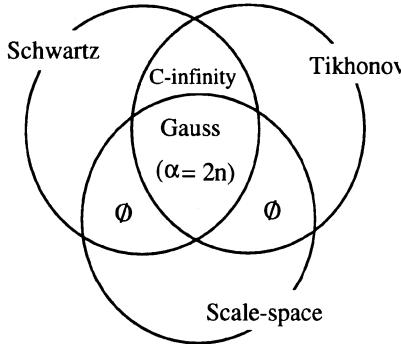


Figure 7.1: Intersections of the filter spaces: Schwartz functions, Tikhonov regularization filters, and scale-invariant semi-group filters. Of regularizations some are of infinite order and then also Schwartz functions, a few of these are also scale-space filters.

functional (using a smoothness term of infinite order). The corresponding energy functional is

$$E[f] = \frac{1}{2} \int dx (f - g)^2 + \sum_{i=1}^{\infty} \frac{t^i}{i!} \left[ \frac{\partial^i}{\partial x^i} f \right]^2$$

A necessary condition for  $f$  to minimize this is that the first variation yields zero. This is the Euler-Lagrange formulation and yields for the above functional

$$0 = (f - g) + \sum_{i=1}^{\infty} \frac{(-t)^i}{i!} \frac{\partial^{2i}}{\partial x^{2i}} f \equiv e^{-t\Delta} f - g$$

where  $\Delta$  is the Laplacean and its exponentiation is defined by its Taylor series expansion. This differential equation is an equation using only spatial derivatives, but having the same solutions as the Heat Equation. The solution can formally (only formally since  $g$  is not necessarily differentiable) be written as

$$f = e^{t\Delta} g$$

By differentiation with respect to  $t$ , we obtain the Heat Equation:

$$\begin{cases} \frac{\partial}{\partial t} f &= \Delta f \\ f|_{t=0} &= g \end{cases}$$

The same derivation can be made for the other scale-invariant semi-group filters under the mapping  $\Delta \mapsto \Delta^p$ , where  $\Delta^p$  denotes the Laplacean applied  $p$  times.

### 7.3.3 Scale-space interpretation of regularization

We have seen that infinite order regularization with appropriate weights of the different orders boils down to linear scale-space. We can reformulate this result

as follows: as time evolves, the signal governed by the Heat Equation travels through the minima of a one-parameter functional. In general, we would like to be able to construct a PDE travelling through the minima of a given one-parameter functional. Here, we look into the case of Tikhonov regularization. We want to find the function  $f$  minimising the functional

$$E[f] = \int dx \left( (f - g)^2 + \lambda \sum_{i=1}^n \lambda_i (\frac{\partial^i}{\partial x^i} f)^2 \right)$$

where  $\lambda_i \geq 0$  are arbitrary. We find the Euler-Lagrange equation by setting the first variation of this to zero

$$(f - g) + \lambda \sum_{i=1}^n \lambda_i (-\Delta)^i f = 0$$

where  $(-\Delta)^i$  denotes minus the Laplacean applied  $i$  times. We perceive of  $f$  as being a function of space  $x$  and evolution parameter  $\lambda$  and find by differentiation with respect to  $\lambda$  the following differential equation:

$$(1 + \lambda \sum_{i=1}^n \lambda_i (-\Delta)^i) \frac{\partial}{\partial \lambda} f = - \sum_{i=1}^n \lambda_i (-\Delta)^i f$$

This is a differential equation governing  $f$ , so that it travels through the minima of the above functional. In Fourier space it yields

$$(1 + \lambda \sum_{i=1}^n \lambda_i \omega^{2i}) \frac{\partial}{\partial \lambda} \hat{f} = \sum_{i=1}^n \omega^{2i} \hat{f}$$

In case of first order regularization (i.e.  $\lambda_1 = 1$  and all other  $\lambda_i = 0$ ) this is the Heat Equation, but with the following remapping of scale

$$d\lambda = (1 + \lambda \omega^2) ds \quad \Leftarrow \quad s(\lambda; \omega) = \frac{\log(1 + \lambda \omega^2)}{\omega^2}$$

In this way first order regularization can be interpreted as Gaussian scale-space with a frequency-dependent remapping of scale. Note that this remapping of scale is not Fourier invertible. Hence, a first order regularization cannot in general be constructed as a Gaussian scale-space with a spatially varying scale.

So far, we have discussed 1-dimensional regularization. In more dimensions mixed derivatives show up complicating the picture. However, imposing isotropy (Hilbert, 1893) on the functional  $E$  leads to:

**Proposition 7.3.8. (Isotropic regularization)** The most general translationally and rotationally invariant functional used for regularization in  $D$  dimensions has only one independent parameter  $\lambda_i$  per order of differentiation.

The proof is given in (Nielsen et al., 1996b). Using this translationally and rotationally invariant regularization, we can easily generalize the results from the previous sections to higher dimensions. All what is needed is to substitute  $\omega$  by  $|\omega|$  in the filters. Also the semi-group property using the 1-norm addition leads trivially to the Gaussian filter.

## 7.4 Truncated smoothness operators

To be consistent with scale-space, Tikhonov regularization must be of infinite order. Nevertheless, low-order regularization is often performed (Grimson, 1981). We might perceive of low-order regularization as a regularization using a truncated Taylor-series of the smoothness functional. Here we list the linear filters  $h_z$  found by Fourier inversion. The notation  $h_z$  indicates the filter where all  $\lambda_i$  are zero except those mentioned by their index in the set  $z$  (E.g. in  $h_{1,2}$  all  $\lambda_i$  are zero except  $\lambda_1$  and  $\lambda_2$ ).

$$h_1(x) = \frac{1}{\sqrt{2\lambda}} e^{-\frac{|x|}{\sqrt{\lambda}}} \quad (7.5)$$

$$h_2(x) = \frac{\pi}{\lambda^{1/4}} \cos\left(\frac{\sqrt{2}|x|}{\lambda^{1/4}}\right) e^{-\frac{|x|}{\sqrt{2\lambda^{1/4}}}} \quad (7.6)$$

In the case of mixed first and second order regularization with  $\lambda = \lambda_1$  and  $\lambda_2 = \lambda_1^2/2$  corresponding to the truncation to second order of the energy functional implying Gaussian filtering, we find

$$h_{1,2}(x) = \frac{\pi(|x| + \sqrt{\lambda})}{2\lambda} e^{-\frac{|x|}{\sqrt{\lambda}}}$$

We notice that the first order regularization filter is always positive, while the second order filter is similar apart from a multiplication with an oscillating term. The latter explains why second order regularization might give inadequate and oscillating results (so-called overshooting or ringing effects). The second order truncated Gaussian filter  $h_{1,2}$  is always positive, and will in general give more reasonable results than one without the first order term. Higher order non-oscillating regularization filters can be constructed by successive first order regularizations. An analysis of positiveness of regularization filters can be found in (Nielsen et al., 1996b).

Canny proposed three criteria of optimality of a feature detection filter. The feature is detected in the maxima of the linear filtering. In the case of a general feature  $e(x)$  with uncorrelated Gaussian noise, the measures of signal-to-noise ratio  $\Sigma$ , localization  $\Lambda$ , and uniqueness  $\Upsilon$  are:

$$\Sigma[f] = \frac{|\int dx e(x)f(x)|}{(\int dx f^2(x))^{1/2}}, \quad \Lambda[f] = \frac{|\int dx e'(x)f'(x)|}{(\int dx f'^2(x))^{1/2}}, \quad \Upsilon[f] = \frac{\left(\int dx f'^2(x)\right)^{1/2}}{\left(\int dx f''^2(x)\right)^{1/2}}$$

where all integrals are taken over the real axis. We now *define* the feature to be a symmetric step edge:

$$e(x) = \int_{-\infty}^x dt \delta_0(t)$$

The symbol  $\delta_0$  denotes the Dirac delta functional. Canny tries first simultaneously to maximize  $\Sigma$  and  $\Lambda$  and finds the box filter as the optimal solution on a finite domain. To avoid the box-filter he then introduces the uniqueness measure. After

this, a simultaneous optimization of all three measures using Lagrange multipliers is performed. Deriche87 (Deriche, 1987) finds the optimal solution on an infinite domain. In the following we show that the uniqueness criterion can be omitted on the infinite domain leading to a conceptually simpler result, but lacking simplicity of the edges. Simplicity means they are zero-crossings of a differentiable function.

In order to find the optimal smoothing filter  $h$  to be differentiated so as to give the optimal step edge detector, we substitute  $h'(x) = f(x)$  in the measures and try to find  $h$ . Assume,  $h$  is symmetric and normalised. We use one of the factors as optimality criterion and the others as constraints to find a composite functional  $\Psi$ . To obtain symmetry in this formulation we multiply all factors by an arbitrary Lagrange multiplier  $\lambda_i$ .

$$\Psi[h] = \int dx \lambda_1 e h' + \lambda_2 h'^2 + \lambda_3 e' h'' + \lambda_4 h''^2 + \lambda_5 h'''^2$$

Omitting the uniqueness constraint corresponds to  $\lambda_5 = 0$ . A necessary condition for an optimal filter is a zero variation:

$$\frac{\delta \Psi}{\delta h} = -\lambda_1 \delta_0 - 2\lambda_2 h^{(2)} + \lambda_3 \delta_0^{(2)} + 2\lambda_4 h^{(4)} - 2\lambda_5 h^{(6)} = 0$$

Here, a parenthesised superscript denotes the order of differentiation. By Fourier transformation we find

$$\hat{h}(\omega) = \frac{\lambda_1 + \lambda_3 \omega^2}{2(\lambda_2 \omega^2 + \lambda_4 \omega^4 + \lambda_5 \omega^6)}$$

Normalised in the spatial domain implies  $\lambda_1 = 0$  and  $\lambda_3 = 2\lambda_2$ :

$$\hat{h}(\omega) = \frac{1}{1 + \alpha \omega^2 + \beta \omega^4}$$

where  $\alpha = 2\lambda_4/\lambda_3$  and  $\beta = 2\lambda_5/\lambda_3$ . This is a second order regularization filter. Deriche87 (Deriche, 1987) chooses  $\alpha = 2\sqrt{\beta}$ , and fixes in this way a one-parameter family of optimal filters. This choice corresponds well with a dimensional analysis and is the second order truncated Gaussian  $h_{1,2}$ .

Another way to select a one-parameter family is to select  $\beta = 0$ . By omitting the uniqueness criteria we find the optimal step edge detection filter to be the derivative of

$$\hat{h}(\omega) = \frac{1}{1 + \alpha \omega^2}$$

This is a first order regularization filter. It seems intuitively correct that the optimal first order filter should project into a Sobolev space of first order. We do not in general require the second derivative of the signal to exist when we consider only first order characteristics. In the spatial domain, we find the filter

$$h'(x) = \frac{\pi}{\sqrt{\alpha}} \frac{x}{|x|} e^{-|x|/\sqrt{\alpha}}$$

This is the edge detection filter proposed by Castan et al. (Castan et al., 1990). The derivative of this filter is not well-defined, and we see that we indeed have projected into a Sobolev space of first order.

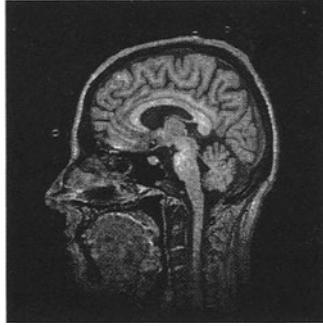


Figure 7.2: Brain image used in the following examples.

The above results can be generalized to optimal detection of higher order image features (Nielsen et al., 1996b; Nielsen et al., 1994). This also implies regularization but of higher order.

## 7.5 Implementation issues of truncated filters

The regularization using quadratic stabilisers can be implemented in various ways. It can be implemented by a gradient descent algorithm. This has the advantage that the natural boundary conditions can be implemented directly, leading to sensible results near the boundaries. The disadvantage is slow convergence for wide filters.

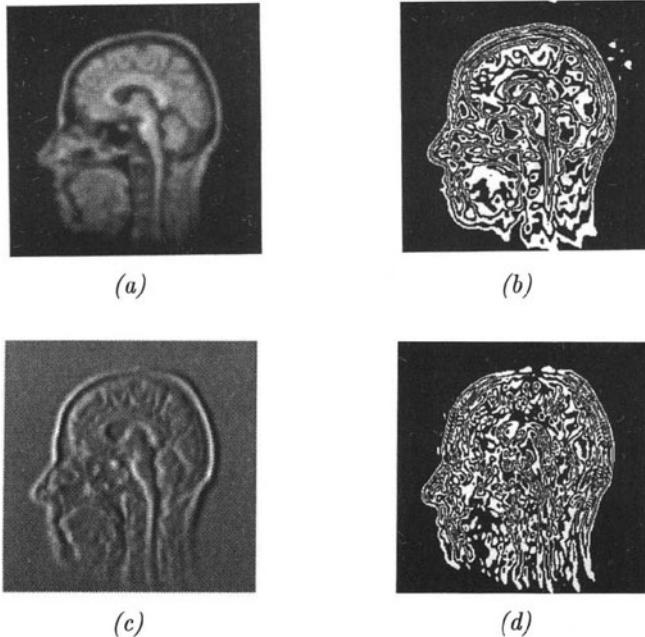
Implementation can be done by convolution in the spatial domain. In this case the boundaries could be handled by cutting off the filters and re-normalisation. The computational complexity will be  $O(MN\lambda^2n)$ , where the size of the image is  $M \times N$  pixels and  $n$  is the order of regularization.

The filtering can also be implemented in the Fourier domain as a multiplication, using the Discrete Fourier Transform (DFT). In this case the image is assumed to be cyclic, which may imply strange phenomena near the boundaries. These effects may be reduced by embedding the image into a large trivial image from which the interaction is minimal. The computational complexity is  $O(MN \log M \log N)$  independently of the order of regularization and the values of  $\lambda_i$ .

Finally, the regularization can be implemented as a recursive filtering. In this case, the boundaries can be handled by cutting off like in the case of convolution in the spatial domain. The computational complexity is  $O(MNn)$ . In most practical cases, this will be the fastest implementation.

In order to deal with the recursive system, we need to reformulate the regularization on a discrete grid. We define the energy

$$E(f) \equiv \sum_x \left( (f - g)^2 + \sum_{i=1}^N \lambda_i (d_i * f)^2 \right)$$



**Figure 7.3:** First order regularized brain ( $h(\omega) = \frac{1}{1+\lambda_1\omega^2}$ ). (a) is the filtered image ( $\lambda_1 = 6$  pixels), (b) is level sets of (a), generated by dividing the intensities into intervals and letting every second interval become black and every second white, (c) is the horizontal first derivative of (a), and (d) is level sets of (c). Notice that the first derivative is not regular in theory. It is contained in a Sobolev space of order 0, but not necessarily in one of order 1.

where  $d_i$  is the  $i$ th order difference filter. We recall the definition of the  $z$ -transform

$$\hat{h} = \sum_{x=-\infty}^{\infty} h(x)z^{-x}$$

and make the following proposition:

**Proposition 7.5.1. (Space-recursive filtering)** The discrete  $N$ th order regularization can be implemented by recursive filtering using no more than  $2N$  multiplications and additions per output element.

*Proof.* By following the argumentation in the continuous case (and substituting the Fourier transform by the  $z$ -transform) or the discrete formulation given by Unser et al. (Unser et al., 1991), we find that the minimization is implemented by convolution with the filter

$$\hat{h}(z) = \frac{1}{1 + \sum_{i=1}^N \lambda_i \hat{d}_i(z) \hat{d}_i(z^{-1})}$$

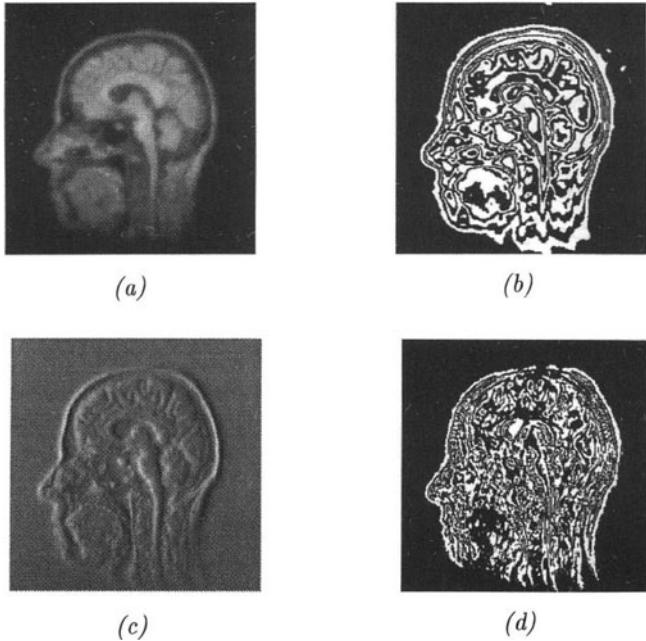


Figure 7.4: Second order regularized brain ( $h(\omega) = \frac{1}{1+\lambda_2\omega^4}$ ). (a) is the filtered image ( $2\sqrt{\lambda_2} = 6$  pixels), (b) is level sets of (a), (c) is the horizontal first derivative of (a), and (d) is level sets of (c). Notice the ringing effects.

where the hat indicates the  $z$ -transform. The transform of the difference operator is given by

$$\hat{d}_i(z) = z^{-i/2}(1-z)^i$$

implying that  $\hat{d}_i(z)\hat{d}_i(z^{-1})$  is a  $N$  order polynomial in  $z$  multiplied by a  $N$  order polynomial in  $z^{-1}$ . Because the transfer function is symmetric in  $z$  and  $z^{-1}$  the  $2N$  roots in the denominator will appear in pairs of  $z$  and  $z^{-1}$ . The transfer function can be decomposed as  $\hat{h}(z) = \hat{h}^+(z)\hat{h}^+(z^{-1})$ , where

$$\hat{h}^+(z) = \frac{c}{\prod_{i=1}^N(1-z_iz^{-1})} = \frac{c}{1-a_1z^{-1}-\dots-a_Nz^{-N}}$$

and  $z_i$  are the roots in the denominator ordered by length. This means that the regularization can be implemented by the forward and backward recursive filters with identical coefficients:

$$\begin{aligned} f^+(x) &= g(x) + a_1f^+(x-1) + \dots + a_Nf^+(x-N) \\ f^-(x) &= f^+(x) + a_1f^-(x+1) + \dots + a_Nf^-(x+N) \\ f(x) &= c^2f^-(x) \end{aligned}$$

□

We have proven here that regularization in the discrete 1D case can be implemented (without any other approximation than the discrete implementation) as recursive filtering. The proof follows the lines of the proof given by Unser et al. (Unser et al., 1991) for the case where a single order stabilizer is used.

## 7.6 Summary

We have shown that scale-invariant filters possessing the algebraic structure of a semi-group is a one-parameter class of  $C^\infty$  filters. Only a discrete set of these are local filters giving a PDE formulation of scale-space. This discrete set is the same set as the scale-invariant semi-group filters implemented as Tikhonov regularizations since all Tikhonov regularizations can be formulated as PDEs by the Euler-Lagrange formalism.

The discrete set can be ordered according to the scale addition norm. The first filter is then the Gaussian. The last is the ideal low-pass filter. The only of these filters which is positive is the Gaussian. The only linear scale-space with local generator fulfilling the principle of causality (Koenderink, 1984), or even weaker, preserving positivity is thus Gaussian scale-space.

We have shown that the Canny-Deriche87 filter is a second order regularization filter corresponding to the truncation of the regularization functional leading to Gaussian scale-space. The Canny-Deriche87 filter is optimised for detection of zero order step edges.

Furthermore, we have pointed out that regularization in a higher dimensional space in its most general form is a rotation of the 1D regularization, when isotropy is imposed. This means that all of the above results can be generalized to any dimension, under assumption of isotropy.

The regularizations of finite order cannot implement scale-spaces that are recursive in scale and simultaneously scale-invariant. However, the finite order regularizations can be implemented as space-recursive filters. The infinite order regularizations can implement any scale-recursive scale-space filter, but cannot be implemented by spatially recursive filtering.

The above is a concentrate of several papers (Pauwels et al., 1995; Nielsen et al., 1996b; Nielsen et al., 1996a; Nielsen et al., 1994), where the omitted proofs and further results can be found.

# Chapter 8

## Invariance Theory

**Alfons Salden**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

### 8.1 Introduction

A linear scale space of a 2-dimensional grey-valued image consists of a set of spatial positions and scales (see chapter 6). As the input image is normally modeled with respect to a fixed Cartesian frame a Euclidean movement of the frame within the image plane induces a local transformation of the linear scale space preserving the metrical relations. If the linear scale space is subjected to a similarity transformation, which is a particular scaling of the spatial coordinates, the scale parameter and the input image simultaneously, then certain similarity relations are preserved. The question arises which are the intrinsic invariants of a linear scale space that are invariant under the group of Euclidean movements and under the similarity group. Normally such an equivalence problem is tackled by means of differential and integral geometry (Salden et al., 1995a). Intrinsic invariants of a linear scale space can be read out by means of differential geometric slot-machines such as the torsion and curvature tensor or by means of integral geometric slot-machines such as circuit integrals measuring Burgers and Frank vectors. The integral geometric slot-machines can in turn be used to define topological currents, which are true physical observables. In this chapter, however, after stating the equivalence problem for a linear scale space, the group of Euclidean movements and the similarity group, a complete and irreducible set of intrinsic invariants of a linear scale space is derived by means of invariance theory.

## 8.2 The Equivalence Problem

First the definitions of a linear scale space and of its extended  $r$ -jet are recalled. Next the transformation groups that are consistent with the intrinsic invariants of a linear scale space are specified. Finally, a so-called equivalence problem for an extended  $r$ -jet and its transformation groups is stated. A precise definition of the equivalence problem will shortly follow.

On our Euclidean image domain the linear scaling paradigm is governed by an isotropic linear diffusion equation with an initial-boundary condition. A linear scale space defined on Euclidean space  $E^2$  can thus be defined as follows:

**Definition 8.2.1.** A linear scale space is a one-parameter family of images  $L$  on Euclidean space  $E^2$  governed by the following linear Cauchy problem:

$$\frac{\partial L}{\partial s} - \Delta L = 0, \quad L(x, 0) = L_0(x), \quad s \in \mathbb{R}_0^+.$$

As the solutions to the above Cauchy problem are just convolutions of the input image  $L_0$  with a normalised Gaussian  $G$  of width  $\sigma = \sqrt{2}s$ , and since convolutions of the input image with ordinary spatial derivates of these Gaussians are also solutions to the Cauchy problem, the extended  $r$ -jet of a linear scale space is readily obtained.

**Definition 8.2.2.** The extended  $r$ -jet of a linear scale space, Definition 8.2.1, is defined by:

$$j^\infty(L_0) = \{x, s, L_{\vec{n}} | \vec{n} \in \mathbb{Z}_0^+ \times \mathbb{Z}_0^+\}, \quad L_{\vec{n}} = L_0 * G_{\vec{n}}, \quad G_{\vec{n}} = \frac{\partial^n G}{\partial x^{i_1} \dots \partial x^{i_n}}, \quad |\vec{n}| \leq r.$$

Here  $*$  denotes the ordinary convolution operator,  $\vec{n}$  is shorthand for the  $n$ -tuple  $(i_1, \dots, i_n)$ ,  $G_{(0,0)} = G$  and  $|\vec{n}|$  denotes the order of derivative of  $G$  symbolised by  $\vec{n}$ .

Now one is normally interested in those intrinsic invariants of the extended  $r$ -jet that are invariant under the similarity group (Salden et al., 1995b). Note that the similarity group is a subgroup of the symmetries of the diffusion equation.

**Definition 8.2.3.** The group of similarity transformations  $\Sigma$  is a group of transformations whose elements all commute with the symmetries of the linear isotropic diffusion equation, and is defined by the one-parameter scaling group of point transformations:

$$(\tilde{x}, \tilde{s}, \tilde{L}) = (\lambda x, \lambda^2 s, \lambda^{2a} L), \quad \lambda \in \mathbb{R}^+, \quad a = -\frac{n+D}{2}, \quad n = 0, 1, \dots$$

As our image domain geometry is Euclidean the other natural intrinsic symmetry of a linear scale space concerns invariance under the group of Euclidean movements. Because in the algebraic invariant solution of the equivalence problem the group of affine transformations will appear crucial, both the group of affine transformations and that of Euclidean movements are defined below.

**Definition 8.2.4.** The group of affine transformations  $A(2, \mathbb{R}) = Gl(2, \mathbb{R}) \triangleright T(2, \mathbb{R})$ , i.e. the semi-direct product of the group of real general linear transformations  $Gl(2, \mathbb{R})$  and the group of translations  $T(2, \mathbb{R})$  on the affine space  $A^2 = (O, \mathbb{R}^2)$  with origin  $O$ , is given in terms of the following transformation rules for fixed Cartesian coordinates  $x$ :

$$x' = ax + t, \quad a \in Gl(2, \mathbb{R}), \quad t \in T(2, \mathbb{R}).$$

**Definition 8.2.5.** The group of Euclidean transformations  $E(2, \mathbb{R}) = SO(2, \mathbb{R}) \triangleright T(2, \mathbb{R})$ , i.e. the semi-direct product of the group of the special orthogonal transformations  $SO(2, \mathbb{R})$  and the group of translations  $T(2, \mathbb{R})$  on the Euclidean space  $E^2 = (O, \mathbb{R}^2; (\cdot, \cdot))$  with standard inner product  $(\cdot, \cdot)$ , is given in terms of the following transformation rules for fixed Cartesian coordinates  $x$ :

$$x' = ax + t, \quad a \in SO(2, \mathbb{R}), \quad t \in T(2, \mathbb{R}).$$

Applying a similarity group action and a Euclidean movement to the input image  $L_0$  one may ask oneself which properties of the extended  $r$ -jet are simultaneously invariant under both transformation groups. Finding such properties requires first of all a clear definition of an invariant of the extended  $r$ -jet.

**Definition 8.2.6.** A function  $I$  of the extended  $r$ -jet Definition 8.2.2 satisfying:

$$I(H(j^r L_0)) = m(H, I)I(j^r L_0), \quad H \in \Sigma \times E(2, \mathbb{R}),$$

is called an invariant or covariant of the  $r$ -jet under the product of the similarity group Definition 8.2.3 and the group of Euclidean movements Definition 8.2.5. If  $m(H, I)$  is equal to the identity operator, then  $I$  is an absolute invariant (covariant), else a relative invariant (covariant).

Note that a covariant is a property that remains the same under the transformation group, but is dependent on the variables  $x$  and  $s$ . Furthermore, note that both group actions commute, enabling us to treat invariance under those actions separately.

The extended  $r$ -jet, the product of the similarity group and the group of Euclidean movements and the definition of an invariant under those groups now provide a natural basis for stating the equivalence problem for the extended  $r$ -jet.

**Definition 8.2.7.** The equivalence problem for an extended  $r$ -jet given the similarity group and the group of Euclidean movements is the problem of finding a complete and irreducible set of invariants necessary and sufficient to describe any other invariant  $I$  of the extended  $r$ -jet unaffected by those groups.

## 8.3 Intrinsic Symmetries

The equivalence problem Definition 8.2.7 stated in the previous section is solved firstly for equivalence under the similarity group (in subsection 8.3.1) and secondly for equivalence under the group of Euclidean movements (in subsection 8.3.2). The division of the equivalence problem 8.2.7 into equivalence problems for the separate groups is allowed, because the group actions commute.

### 8.3.1 Solution of Similarity Equivalence Problem

Consider the extended  $r$ -jet under the similarity group action Definition 8.2.3. It is easily shown that the system states

$$\xi = \frac{x}{\sqrt{s}}, \quad \Lambda = s^{-a} L,$$

are invariant under the similarity group. Here the word state is used to emphasise the fact that it is physically really existing entity and no mathematically constructed entity. The first state  $\xi$  expresses the fact that the Euclidean distance  $|x|$  has only meaning in terms of a unit length  $\sqrt{s}$  relative to some underlying topology realised by a vision system. Scaling both the distance and as well as the unit of distance has no effect on metric relations. This finding prohibits the use of a Euclidean metric on the product topology  $E^2 \times \mathbb{R}_0^+$  in a linear scale space. The second state  $\Lambda$  says that a local measurement of the luminance field is just the total flux through an area, and that this flux does not change if both the unit of area and that of the density are subjected to a similarity scaling. To summarise, the invariants of the extended  $r$ -jet under the similarity group are given by a so-called similarity jet (Salden et al., 1995b):

**Definition 8.3.1.** The similarity jet of the extended  $r$ -jet is defined by:

$$j^\infty(\Lambda_0) = \left\{ \xi, \Lambda_{\vec{n}} \mid \xi = \frac{x}{\sqrt{s}}, \quad \Lambda_{\vec{n}} = s^{\frac{n+2}{2}} L_0 * G_{\vec{n}}, \quad \vec{n} \in \mathbb{Z}_0^+ \times \mathbb{Z}_0^+ \right\}.$$

Note that the system states  $\Lambda_{\vec{n}}$  are dimensionally equivalent so, that they can be compared quantitatively. This property of the similarity jet is useful whenever a certain aspect of the extended  $r$ -jet degenerates, for example the first order jet at a critical point. This definition of the similarity jet also implies that the "natural derivative operators" lack a crucial factor, namely  $s$ . This additional factor expresses the fact that a total flux through a fixed area expressed in terms of scale parameter  $s$  and not a density as such is involved in any measurement.

### 8.3.2 Solution of Affine and Euclidean Equivalence Problems

In this section the Euclidean equivalence problem, i.e. the equivalence problem Definition 8.2.7 restricted to invariance under the group of Euclidean movements Definition 8.2.5, is solved and examples are given for the extended  $r$ -jet up to fourth order. Firstly, the affine equivalence problem is tackled, that is the Euclidean equivalence problem restricted to invariance under the affine group Definition 8.2.4. Subsequently, the found solution method is extended to the Euclidean equivalence problem. Before presenting these solutions and examples, some basic notions and facts from algebraic invariance theory should be introduced and the connection with the extended  $r$ -jet should be clarified.

Considering the extended  $r$ -jet of an image it is clear that the partial derivatives of the map  $L$  can be locally coordinatised by so-called *binary forms*.

**Definition 8.3.2.** A binary form  $Q_n$  of  $n$ -th order is a homogeneous polynomial in the coordinate functions  $x$  with (symmetric) coefficients  $Q_{i_1 \dots i_n}$ :

$$Q_n(x) = \frac{1}{n!} Q_{i_1 \dots i_r} x^{i_1} \dots x^{i_n},$$

where the Einstein convention with respect to index pairs is used.

In the context of the affine and Euclidean equivalence problems above this coordinatisation simply means that the  $r$ -jet 8.2.2 of a grey valued image  $L$  can be described locally in terms of a truncated Taylor series with respect to a Cartesian coordinate frame, which is just a sum of all the binary forms up to  $r$ -th order. In the sequel the  $r$ -jet is represented by systems of binary forms  $\{(Q_0, \dots, Q_r)\}$  up to  $r$ -th order over the image domain.

On systems of binary forms several processes can be defined, for example the *transvection* and the *resultant* (Weitzenböck, 1923).

**Definition 8.3.3.** The  $k$ -th order transvection  $[ \cdot, \cdot ]^k$  of two binary forms  $P$  and  $Q$  of arbitrary orders is defined by:

$$[P, Q]^k(x) = \lim_{x \rightarrow y} \prod_{l=1}^k \epsilon^{i_l j_l} \frac{\partial}{\partial x^{i_l}} \frac{\partial}{\partial y^{j_l}} P(x) Q(y),$$

in which  $\epsilon^{ij}$  is the parity of the ordered pair  $(ij)$ .

It is easily shown that the transvection is a relative (absolute) invariant operation under the group of affine transformations (the group of Euclidean movements).

*Example 8.3.4.* A simple example consists of the first order transvection, also known as Cayley's omega process or the Poisson bracket, applied to two arbitrary binary forms  $P$  and  $Q$ :

$$[P, Q]^1 = \frac{\partial P}{\partial x^1} \frac{\partial Q}{\partial x^2} - \frac{\partial P}{\partial x^2} \frac{\partial Q}{\partial x^1}.$$

This operation tells us whether the gradients of  $P$  and  $Q$  at the same point in the image domain are parallel or not.

In the sequel the transvection will appear to be one of the basic operations in constructing a complete and irreducible set of integral and algebraically independent (multi-)local invariants of systems of binary forms, that is in solving the affine and Euclidean equivalence problems, respectively. Note that one does not encounter the notion of multi-local invariant in algebraic invariance theory. The closest notion to it is that of a joint invariant of systems of binary forms. Furthermore, in the sequel integrally and algebraically independent invariants are simply referred to as independent invariants. Integral invariants are in this context polynomial invariants of the coefficients of the binary forms, whereas algebraically independent invariants are those polynomial invariants for which no specific algebraic identities hold.

The desired complete and irreducible set of invariants will primarily be generated by considering the resultant of two binary forms of equal order. In order to

define such a resultant first of all the *polar of a binary form* and the *Cayley-form* of two binary forms of equal order should be introduced (Gordan, 1871):

**Definition 8.3.5.** The  $\kappa$ -th order polar  $Q_{y^\kappa}$  of a binary form  $Q$  of order  $n$  is defined by:

$$Q_{y^\kappa} = \binom{n}{\kappa}^{-1} \prod_{p=1}^{\kappa} y^{i_p} \frac{\partial Q(x)}{\partial x^{i_p}}.$$

**Definition 8.3.6.** The Cayley-form  $F$  of two binary forms  $Q^1$  and  $Q^2$  of order  $n$  is defined by:

$$F = \frac{Q^1 Q_{y^n}^2 - Q_{y^n}^1 Q^2}{(xy)}, \quad (xy) = x^1 y^2 - x^2 y^1.$$

Using the fact that the product  $Q^i Q_{y^n}^j$  can be expressed in terms of products of polars of transvectants of  $Q^i Q^j$  and of powers of  $(xy)$ :

$$Q^i Q_{y^n}^j = \sum_k \frac{\binom{n}{k} \binom{n}{k}}{\binom{2n-k+1}{k}} [Q^i, Q^j]_{y^{n-k}}^k (xy)^k,$$

it appears that the Cayley-form  $F$  can be written as (Gordan, 1871):

$$F = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} c_{ij} (x^1)^i (x^2)^{n-1-i} (y^1)^j (y^2)^{n-1-j},$$

with  $c_{ij}$  a symmetric matrix function  $c$  on the coefficients  $Q_{i_1 \dots i_n}^1$  and  $Q_{j_1 \dots j_n}^2$  of the binary forms  $Q^1$  and  $Q^2$ , respectively. On the basis of these definitions and the simplification of the Cayley-form, the resultant between two binary forms  $Q^1$  and  $Q^2$  of order  $n$  can be defined as the determinant of the matrix  $c$  (Gordan, 1871):

**Definition 8.3.7.** The resultant  $R$  of two binary forms  $Q^1$  and  $Q^2$  of order  $n$  is defined by:

$$R(Q^1, Q^2) = \det(c),$$

with matrix  $c$  as in Definition 8.3.6.

This definition of the resultant is to be preferred above the standard definition as a so-called Sylvester determinant of a non-symmetric matrix (Waerden, 1940). The determinant of a matrix  $c$  can be written as a sum of traces of powers of  $c$ , but now in a clearly manifest invariant manner.

Because our approach for finding complete and irreducible sets of invariants hinges on the method proposed by Hilbert it is necessary to know for both the affine and Euclidean equivalence problem which systems of binary forms to consider. The

reason for this is that Hilbert investigates the invariant properties of forms under the general linear group. Klein has pointed out in his "Erlangen Programm" (Klein, 1893) that the invariance theory for forms under special transformation groups can be viewed as particular manifestations of projective invariance problems for forms. This result encouraged Weitzenböck (Weitzenböck, 1923) to state affine equivalence for a system of forms in a projective manner. For binary forms this equivalence reads as follows:

**Theorem 8.3.8.** A complete set of integral and rational affine invariants of the system of binary forms  $Q^1, \dots, Q^r$  is equivalent to a complete set of projective invariants of the system of equations:

$$q^i = Q^i(\xi, 1) = 0, \quad \xi = \frac{x^1}{x^2} \in \mathbb{RP}^1,$$

under the fractional action  $\xi \rightarrow \frac{p\xi+q}{r\xi+s}$  on the projective line  $\mathbb{RP}^1$  induced by the general linear group  $GL(2, \mathbb{R})$  on  $\mathbb{R}^2$ .

*Proof.* See (Weitzenböck, 1923) □

Because the  $r$ -jet Definition 8.2.2 can be identified with all the binary forms up to order  $r$  over the image domain the affine equivalence problem can be translated into a projectively equivalent problem. Note that the translation group does not affect the equivalence problem, for the binary forms are defined with respect to a local Cartesian coordinate frame. Because the group of Euclidean movements is just a sub-group of the affine group Weitzenböck also formulates Euclidean equivalence for a system of forms in a projective manner (Weitzenböck, 1923). For binary forms this equivalence reads:

**Theorem 8.3.9.** A complete set of integral and rational Euclidean invariants of the system of binary forms  $Q^1, \dots, Q^r$  is equivalent to a complete set of projective invariants of the system of equations:

$$q^i = Q^i(\xi, 1) = 0, \quad \xi = \frac{x^1}{x^2} \in \mathbb{RP}^1, \quad \xi^2 + 1 = 0,$$

under the fractional action  $\xi \rightarrow \frac{p\xi+q}{r\xi+s}$  on the projective line  $\mathbb{RP}^1$  induced by the linear group  $GL(2, \mathbb{R})$  on  $\mathbb{R}^2$ .

*Proof.* See (Weitzenböck, 1923) □

These results of the "Erlangen Programm" in combination with Hilbert's method (to be discussed below) make explicit the systems of binary forms actually to consider in order to solve the affine and Euclidean equivalence problems, respectively. Especially, adjoining the additional algebraic equation  $\xi^2 + 1 = 0$  in the projective formulation of the Euclidean equivalence Definition 8.3.9 is crucial for applying Hilbert's method effectively to the Euclidean equivalence problem. The appearance of that algebraic equation should not surprise us, for it just captures the distance (and angle) preserving property of the group of Euclidean movements Definition 8.2.5.

**The Solution of the Affine Equivalence Problem** In the following the affine equivalence problem is solved and illustrated up to fourth order. Firstly, all this is done for a single binary form. Secondly, it is extended for a system of binary forms at one position in the image domain. Thirdly, it is extended for systems of binary forms over the entire image domain.

According to Theorem 8.3.8 the affine equivalence problem is a special case of a projective problem for binary forms. For this projective problem there always exists a complete system of a finite number of invariants (Weitzenböck, 1923). Now Hilbert addressed the problem of finding a complete set of independent invariants for a system of (binary) forms that is invariant under the general linear group  $GL(n, \mathbb{R})$ . His construction method for such a set is based on the concept of *null forms* (Hilbert, 1893).

**Definition 8.3.10.** A null form is a (binary) form all invariants of which vanish.

In the case of a single binary form the question arises how such a null form looks like and which invariants have to vanish in order for the binary form to be a null form. To this end Hilbert proves the following:

**Theorem 8.3.11.** If all the invariants of a binary form of order  $n = 2h + 1$ , respectively  $n = 2h$  are zero, then the binary form possesses an  $(h + 1)$ -fold linear factor. Conversely, if it possesses an  $(h + 1)$ -fold linear factor, then all invariants are equal to zero.

*Proof.* See (Hilbert, 1893) □

Thus a binary form of order  $n = 2h + 1$  or  $n = 2h$  is a (canonical) null form if it has a  $(h + 1)$ -fold linear factor and therefore is parametrised by  $n - h$  coefficients.

**Definition 8.3.12.** A canonical null form  $N_n$  of order  $n = 2h + 1$  or  $n = 2h$  is defined as a product of a  $(h + 1)$ -fold linear factor  $x_1^{h+1}$  and a binary form  $q_{n-(h+1)}$  of order  $(n - (h + 1))$ :

$$N_n(x) = (x^1)^{h+1} q_{n-(h+1)}(x).$$

For example,  $N_4 = x^1(x^2)^3$  is a null form of fourth order, whereas  $q_4 = (x^1 + x^2)^2(x^1 - x^2)^2$  is not.

Now the proof of Theorem 8.3.11 supplies a method for constructing a set of invariants of a binary form, through which all others can be expressed as an integral algebraic function of that set. Let us follow closely and elaborate on the first part of the proof presented by Hilbert (Hilbert, 1893).

In order to explicitly state which invariants of a single binary form  $Q_n$  of order  $n = 2h + 1$ , respectively  $n = 2h$ , have to vanish, such that it is a null form, Hilbert starts off with the construction of the set  $Z_n$  of transvectants  $[Q_n, Q_n]^k$ :

$$Z_n \equiv \{Q_n, [Q_n, Q_n]^2, \dots, [Q_n, Q_n]^{2g}\}, \quad g = \begin{cases} h & \text{if } n = 2h + 1 \\ h - 1 & \text{if } n = 2h \end{cases}.$$

Next he actually uses the following theorem to define the desired invariants:

**Theorem 8.3.13.** The necessary and sufficient condition for a binary form  $Q_n$  of order  $n = 2h + 1$ , respectively  $n = 2h$ , to have a root of multiplicity  $h + 1$ , is equivalent to requiring the set of transvectants to have one linear factor in common.

*Proof.* See (Hilbert, 1893), (Gordan, 1871) and (Waerden, 1940).  $\square$

The latter condition on the set of transvectants,  $Z_n$ , is now formulated in terms of the vanishing of the resultant of two linear independent combinations  $U$  and  $V$  of powers of the transvectants that have the same order, namely  $M$ . Here  $M$  is defined to be the least common multiple of the numbers  $n, \dots, 2(n-2g)$ , such that the powers  $m_0, m_1, \dots, m_g$  of the transvectants are related to this least common multiple  $M$  and the order  $n$  of the binary form  $Q_n$  as follows:

$$M = m_0n = 2m_1(n-2) = \dots = 2m_g(n-2g),$$

where  $g$  is  $h$  or  $h-1$ , depending on whether  $n = 2h+1$  or  $n = 2h$ , respectively. On the basis of these numbers the two forms  $U$  and  $V$  with indeterminate parameters  $u = u_0, u_1, \dots, u_g$  and  $v = v_0, v_1, \dots, v_g$ , are given by:

$$U = \sum_{k=0}^g u_k([Q_n, Q_n]^k)^{m_k}, \quad V = \sum_{k=0}^g v_k([Q_n, Q_n]^k)^{m_k}.$$

Now Theorems 8.3.11 and 8.3.13 hold, if and only if the resultant  $R$  of the forms  $U$  and  $V$  vanishes:

$$R(U, V) = \sum_{\mu} J_{\mu} P_{\mu} = 0,$$

where each  $P_{\mu}$  is a product of powers of the parameters  $u$  and  $v$  above and the invariants  $J_{\nu}$  or the invariants  $J_{\mu}$  and  $[Q_n, Q_n]^n$  depending on whether  $n$  is odd or even, respectively. The latter invariants form the necessary and sufficient conditions for a binary form  $Q_n$  of order  $n$  to be a null form  $N_n$ . Furthermore, all other invariants of the binary form  $Q_n$  of order  $n$  can be written as an integral algebraic function of these invariants. The invariants  $J_{\mu}$  are readily seen to be resultants of suitable powers of transvectants defined in Definition 8.3.3.

*Example 8.3.14.* The binary forms  $Q_n$  of order  $2 \leq n \leq 4$  are null forms if the following invariants vanish:

$$\begin{aligned} d_2 &= -[Q_2, Q_2]^2, \\ d_3 &= [[Q_3, Q_3]^2, [Q_3, Q_3]^2]^2, \\ i_4 &= [Q_4, Q_4]^4, \\ j_4 &= [[Q_4, Q_4]^2, Q_4]^4. \end{aligned}$$

The vanishing of invariant  $d_2$ , i.e. the discriminant of the inhomogeneous second order polynomial  $q_2 = Q_2(\xi, 1)$  (up to a factor), means that the binary form  $Q_2$  can be put in a null form  $N_2$ . Similarly, the vanishing of invariant  $d_3$ ,

i.e. the discriminant of the inhomogeneous third order polynomial  $q_3 = Q_3(\xi, 1)$  (up to a factor), indicates that the binary form  $Q_3$  can be put in a null form  $N_3$ . If the invariants  $i_4$  and  $j_4$  of the binary form  $Q_4$  vanish simultaneously, then the inhomogeneous fourth order polynomial  $Q_4$  can be put in a null form  $N_4$ .

Note that  $Q_0$ , i.e. the zero-th order jet component, is not an affine invariant and that from  $Q_1$  cannot be constructed locally an affine invariant. Below they will appear to be constituting ones under the Euclidean group of movements.

Now the affine equivalence problem for one binary form is extended to that of a system of binary forms at one position in the image domain. Its solution concerns finding a complete and irreducible set of independent invariants of a local  $r$ -jet. This is equivalent to deriving the conditions for a system of binary forms  $(Q_1, \dots, Q_r)$  necessary and sufficient to put this system simultaneously in a system of canonical null forms  $(N_1, \dots, N_r)$  (see also (Hilbert, 1893)). Studying the latter canonical system one immediately observes that all the null forms have a linear factor in common. Exploiting the above solution of the affine equivalence problem for a single binary form a complete and irreducible set of simultaneous independent invariants for a system of two binary forms readily follows through application of the following theorem.

**Theorem 8.3.15.** A complete and irreducible set of simultaneous independent invariants of a system of two binary forms  $(Q_p, Q_q)$  of orders  $p$  and  $q$ , respectively, is generated firstly by determining the least common multiple  $M_{p,q}$  of the orders  $p$  and  $q$  of the binary forms, and secondly by computing the simultaneous invariants of the new system of two binary forms  $\left(Q_p^{\frac{M_{p,q}}{p}}, Q_q^{\frac{M_{p,q}}{q}}\right)$  through calculation of the independent invariants of a linear combination of this new pair of binary forms of order  $M_{p,q}$ .

*Proof.* On the basis of Theorem 8.3.11 a binary form is a null form if it has an  $(h+1)$ -fold linear factor. This means in the case of a system of two binary forms  $(Q_p, Q_q)$  of orders  $p = 2h_p + 1$  or  $p = 2h_p$  and  $q = 2h_q + 1$  or  $q = 2h_q$ , respectively, that the constituting binary forms have a linear factor of multiplicity  $h_p + 1$  and  $h_q + 1$ , respectively. Let  $M_{p,q}$  and  $\left(Q_p^{\frac{M_{p,q}}{p}}, Q_q^{\frac{M_{p,q}}{q}}\right)$  be defined as in Theorem 8.3.15. Assume the binary forms  $Q_p$  and  $Q_q$  to be put separately in null forms  $N_p$  and  $N_q$ . If, however, the binary forms can be brought simultaneously in null forms such that their linear factor of multiplicity  $h_p + 1$  and that of multiplicity  $h_q + 1$  are coinciding, then only the conditions for this happening should be derived. Denoting with  $h_{p,q} + 1 = \text{Mod}(M_{p,q}/2) + 1$  the multiplicity of a linear factor in a null form  $N_{M_{p,q}}$  for a binary form  $Q_{M_{p,q}}$  then the following inequalities between this multiplicity and those of  $\left(N_p^{\frac{M_{p,q}}{p}}, N_q^{\frac{M_{p,q}}{q}}\right)$  hold:

$$h_{p,q} + 1 \leq \frac{h_p + 1}{p} M_{p,q}, \quad h_{p,q} + 1 \leq \frac{h_q + 1}{q} M_{p,q}.$$

For a single binary form a complete and irreducible set of independent invariants were automatically generated by proving Theorem 8.3.11. So for a single binary

form  $Q_{M_{p,q}}$  and also for an arbitrary linear combination of the two binary forms  $Q_p^{\frac{M_{p,q}}{p}}$  and  $Q_q^{\frac{M_{p,q}}{q}}$  the same procedure can be exploited. As the above inequalities always hold for any orders  $p$  and  $q$ , all the necessary and sufficient conditions for the system of two binary forms  $(Q_p, Q_q)$  to be a system of null forms  $(N_p, N_q)$  are captured by the vanishing of the invariants of the two parameter binary form  $\lambda_p Q_p^{\frac{M_{p,q}}{p}} + \lambda_q Q_q^{\frac{M_{p,q}}{q}}$  of order  $M_{p,q}$ . Of course, the invariants that involve only integral and algebraic functions of the coefficients of one of the binary forms, say  $Q_p$ , can be neglected, for they can be expressed by those of  $Q_p$ . Furthermore, some of the simultaneous invariants will be identical to zero due to the typical multiplicities of the new binary forms.  $\square$

Considering all possible systems of two binary forms  $(Q_p, Q_q)$  of a system of binary forms  $(Q_1, \dots, Q_r)$  and calculating the simultaneous invariants of pairs of binary forms is necessary and sufficient to solve the local affine equivalence problem for the local  $r$ -jet.

*Example 8.3.16.* The system of binary forms  $(Q_1, Q_2)$  can be put in a system of null forms  $(N_1, N_2)$ , if and only if, the following invariants vanish:

$$R_{12} = [Q_1^2, Q_2]^2, \quad d_2 = [Q_2, Q_2]^2,$$

where the invariant  $R_{12}$  is the resultant of the linear form  $Q_1$  and the quadratic form  $Q_2$ . Note that one usually applies this invariant as a "corner" detector.

Now let us conclude by solving and giving examples of the affine equivalence problem for systems of binary forms  $\{(Q_1, \dots, Q_r)\}$  over the image domain, i.e. the extended  $r$ -jet. The solution is readily seen to be an extension of the solution for the local problem for the system of binary forms  $(Q_1, \dots, Q_r)$  to a bi-local problem concerning two systems of binary forms  $\{(Q_1, \dots, Q_r), (P_1, \dots, P_r)\}$ , to a tri-local, etc. Because the elements of the translation group  $T(2, \mathbb{R})$  and the group of scale shifts intertwine with the local partial spatial derivative operators it is allowed to transport parallelly two systems to the origin of one of them and apply a scale shift in order to solve the affine equivalence problem for the system of binary forms  $(Q_1, \dots, Q_r, P_1, \dots, P_r)$ . This problem is solved by consecutively adjoining forms  $P$  to the system  $(Q_1, \dots, Q_r)$  and solving the problem for the extended problem. Transporting, subsequently, more and more systems to the origin, adjoining the additional forms to the previous ones and solving the affine equivalence problem for the extended system gives a solution to the corresponding problem for the  $r$ -jet. These multi-local invariants one also encounters in the differential and integral geometry of linear scale spaces (Salden et al., 1995a): computing the difference of connections and the related curvature and torsion aspects or the Burgers and Frank vectors involve just these multi-local invariants.

*Example 8.3.17.* The systems of binary forms  $\{(Q_1), (P_1)\}$  can be put in the system of binary forms  $\{N_1(Q), N_1(P)\}$ , if and only if, the following invariant vanishes:

$$R_{1,1} = [Q_1, P_1].$$

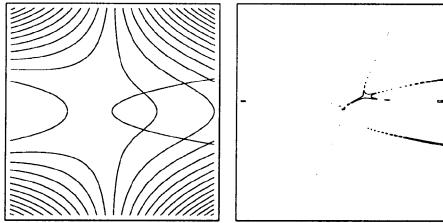


Figure 8.1: To the left level contours of  $z(x,y) = x^2 + xy^2$  of a surface with a cusp at the origin together with its parabolic curve projected onto the  $(x,y)$ -plane. To the right level contours on which invariant  $R_{1,1}$  is approximately zero. The disconnected curve segments are clearly close to the projected parabolic curve.

This image property is just the resultant of both linear forms of the systems of binary forms. In the sequel this invariant is used to trace the parabolic curves on a surface  $S$ , i.e. curves on a surface on which at least one of the principal curvatures of the surface is vanishing.

Analogous to (Koenderink and Doorn, 1980) assume a fixed illumination vector  $\vec{s}$ , the surface to be Lambertian with constant albedo and no vignetting. Then the surface radiance  $E$  is equal to  $\vec{s} \cdot \hat{n}$  where  $\hat{n}$  is the unit normal vector field to the surface. Applying pseudo-orthographic projection or the weak perspective condition planar images  $I$  of the radiance field from different viewing directions are just affine transformed versions of each other. Having a fixed Cartesian reference frame  $(O, (\hat{e}_x, \hat{e}_y, \hat{e}_z))$  the surface can be described in terms of a vector function  $X = (x, y, z(x, y))$ . Consequently, the unit normal vector field is simply  $\hat{n} = \frac{X_x \times X_y}{\|X_x \times X_y\|}$ . Having a foliation of the surface  $S$  given by isophotes  $I = cst$  in the image of the radiance field  $E$  one is interested which image properties are related to those of objects on the surface. Taking the image gradient field  $\nabla I = (I_x, I_y)$ , computing the second fundamental form  $\Xi = X_{xx} \cdot \hat{n} dx^2 + 2X_{xy} \cdot \hat{n} dxdy + X_{yy} \cdot \hat{n} dy^2$  and performing a rotation around the  $z$ -axis such that  $X_{xy} \equiv 0$  shows that the gradient vector field has constant direction along parabolic curves irrespective the direction of illumination vector  $\vec{s}$  or the viewing direction. This photometric invariant object of the surface retrievable in the image together with the notion of multi-local invariance of  $R_{1,1}$  will enable us to detect the parabolic curves on the surface  $S$  from one viewpoint on the basis of two or more different source directions  $\hat{s}$  and  $\hat{s}'$ . In this context a parabolic curve divides the surface in elliptic and hyperbolic regions, where the sign of the Gaussian curvature, equal to that of  $X_{xx}X_{yy} - X_{xy}^2$ , is positive and negative, respectively.

In figure 8.1 the invariant  $R_{1,1}$  is computed on the basis of the image gradient fields  $\nabla I$  and  $\nabla I'$  of two orthographic images  $I$  and  $I'$ , respectively, of a Lambertian surface illuminated from two different source directions  $\hat{s}$  and  $\hat{s}'$ , but viewed from one viewpoint. If those image gradient fields have to be parallel along parabolic curves obviously the invariant  $R_{1,1}$  has to vanish. Thus one finds photometric invariant objects of the pair of images connected to parabolic curves on the surface.

This example indicates that the parabolic curves form so to speak the generators for the image gradient field as function of the source direction and the surface description. Perturbing the source direction isotropically one can find the projected parabolic curves and the flow on the image gradient field, which forms a dynamic realisation of the interaction between surface and illumination field.

**The Solution of the Euclidean Equivalence Problem** The solution of the Euclidean equivalence problem for a  $r$ -jet is not difficult to guess on the basis of the projective formulation of the Theorem 8.3.9. Adjoin the quadratic binary  $S_2 = \frac{1}{2}x^i x^i$  to the previous systems and extend Hilbert's method for deriving a complete and irreducible set of independent invariants for the extended system. Instead of presenting a dull repetition of the previous section let's elaborate on two examples, namely of the additional invariants of a local two jet, and of two first order jets at different positions. In both examples the trivial invariants are the zero-th order binary forms  $Q_0$

*Example 8.3.18.* The binary forms  $Q_n$  of order  $1 \leq n \leq 2$  are identically zero if besides the invariants listed in Example 8.3.16 up to that order the following invariants also vanish:

$$l_1 = [Q_1^2, S_2]^2, \quad l_2 = [Q_2, S_2]^2,$$

respectively. The invariants  $l_1$  and  $l_2$  are proportional to the squared length of the image gradient and to the Laplacian of the image, respectively.

*Example 8.3.19.* The systems of first order binary forms  $\{(Q_1), (P_1)\}$  are identically vanishing if besides the invariant in Example 8.3.17 also the following invariants are zero:

$$l_1^q = [Q_1^2, S_2]^2, \quad l_1^p = [P_1^2, S_2]^2,$$

where the superscripts  $p$  and  $q$  express the fact that the first order binary forms are computed at different locations in the image domain.

One could proceed formulating the necessary and sufficient conditions for the systems of binary forms to have more and more multiple roots (Gordan, 1871). The construction method for such conditions relies again on the computation of transvectants of covariants. Of course, these conditions can be rewritten in those already found above.

## 8.4 Conclusion and Discussion

Summarising, the equivalence problem consistent with a linear scale-space has been solved by restricting the problem firstly to finding a complete and irreducible set of invariants under the similarity group. Secondly, the problem has been solved by restricting the search to a complete and irreducible set of independent invariants under the group of Euclidean movements. A complete and irreducible set of independent set of invariants that guarantee both the invariance under the group

of Euclidean movements and the similarity group are readily found by combining the above solutions. As the spatial positions and the scale parameter are independent physical entities it is allowed to perform Euclidean movements and scale shifts, and correlating the local similarity jets over scale and space.

Recently our invariant theoretic approach to image analysis has been successfully applied to object recognition problems (Schmid and Mohr, 1996). Our approach is readily extended to linear scale spaces of spatial input images of higher dimensions by just following Hilbert's paper (Hilbert, 1893). The latter theory and application is readily extendable to finding the intrinsic invariants of nonlinear scale spaces (Salden et al., 1995a). For example, the affine shortening flow of a 2-dimensional image can be described in terms of a complete and irreducible set of (multi)-local affine invariants found in this chapter if invariance under monotonic grey-value transformations is not assumed to be applicable or necessary. It also allows multi-local image analysis by means of differential and integral geometry. For example, the variation of the similarity jets of partly overlapping subimages along a closed circuit in the image domain operationalises the curvature of the similarity jet. Normally, one unconsciously measures such entities as coherencies in the similarity jet over the image domain (Salden, 1996).

# Chapter 9

# Stochastic Analysis of Image Acquisition and Scale-Space Smoothing

Kalle Åström and Anders Heyden

*Department of Mathematics, Lund University  
P. O. Box 118, S-221 00 Lund, Sweden*

## 9.1 Introduction

Low level image processing is often used to detect and localise features such as edges and corners. It is also used to correlate or match small parts of one image with parts in another. Methods for doing this have been developed for some time, see (Canny, 1986; Michelli et al., 1989; Deriche, 1987; Faugeras, 1994; Hildreth and Marr, 1980; Marr, 1982; Nalwa and Binford, 1986; Torre and Poggio, 1986). However, the stochastic analysis of these algorithms have often been based upon poorly motivated stochastic models. In particular, the effects of image discretisation, interpolation and scale-space smoothing is often neglected or not analysed in detail.

In this chapter, image acquisition, interpolation and scale-space smoothing are modelled into some detail. *Image acquisition* is viewed as a composition of blurring, ideal sampling and added noise, similar to (Pratt, 1978). The discrete signal is analysed after *interpolation*. This makes it possible to detect features on a sub-pixel level. Averaging or *scale-space smoothing* is used to reduce the effects of noise. To understand feature detection in this framework, one has to analyse the effect of noise on interpolated and smoothed signals. In doing so a theory is obtained that connects the discrete and continuous scale-space theories.

The chapter is organised as follows. Section 9.2 treats the image acquisition model. In Section 9.3 a method is proposed where the discrete scale-space is induced from the continuous scale-space theory. The stochastic properties of the intensity error field are discussed in Section 9.4. A short introduction to stationary random fields is given and some important results relevant for our model are demonstrated. The ideas are verified with numerical experiments on real images. The analysis of sub-pixel correlation and edge detection is commented upon briefly in Section 9.5. A more detailed description can be found in (Åström, 1996; Åström and Heyden, 1996a; Åström and Heyden, 1996b).

## 9.2 Image acquisition

To model the image acquisition process, the intensity distribution that would be caught by an ideal camera is first affected by aberrations in the optics of the real camera, e.g. blurring caused by spherical aberration, coma and astigmatism. Other aberrations deform the image, like Petzval field curvature and distortion, see (Hecht, 1987). Such distortion can typically be handled by geometric considerations in mid-level vision and will not be commented upon here. One way to model camera blur is to convolve the ideal intensity distribution with a kernel corresponding to the smoothing caused by the camera optics. This process attenuates high frequency components in the image.

In a video-camera, the blurred image intensity distribution is typically measured by a CCD array. One can think of each pixel intensity as the weighted mean of the intensity distribution in a window around the ideal pixel position. Taking the weighted mean around a position is equivalent to first convolving with the weighting kernel and then ideal sampling. Finally, due to quantisation and other errors, stochastic errors are introduced.

Led by this discussion we will use the following image acquisition model:

$$W_{\text{ideal}} \xrightarrow{\text{blur}} W \xrightarrow{\text{sampling}} w_0 \xrightarrow{\text{noise}} v_0 , \quad (9.1)$$

where *upper case letters*,  $W$ , denote signals with *continuous* parameters, whereas *lower case letters*,  $w$ , denote *discrete* signals, see Fig. 9.1. Here, and often in the sequel, we use the word signal synonymously with function, and discrete signal synonymously with sequence or function defined on  $\mathbb{Z}^n$ , for some  $n$ . These three steps of blurring, sampling and noise will now be discussed in a little more detail.

*Blurring* is modelled as an abstract operator  $h$ , such that  $W = h(W_{\text{ideal}})$ . We assume that no aliasing effects are present, when the function  $W$  is sampled at integer positions.

**Assumption 9.2.1.** All energy in the high spatial frequencies is cancelled before sampling. The function  $W$  is band-limited, i.e.

$$W \in \mathcal{B}(\mathbb{R}^n) = \{W \in L_2(\mathbb{R}^n) \mid \text{supp } \mathcal{F}W \subset (-1/2, 1/2)^n\} ,$$

where  $\mathcal{F}$  denotes the Fourier transform and  $\text{supp } f$  denotes the support, i.e. the smallest closed set containing the points where  $f$  is non-zero. ■

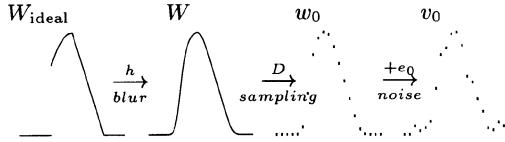


Figure 9.1: Illustration of the image acquisition model.

In the definition of the Fourier transform, we use the formula

$$\mathcal{F}W(f) = \int_{\mathbb{R}^n} W(\tau) e^{-i2\pi f \cdot \tau} d\tau , \quad (9.2)$$

where  $f \cdot \tau$  denotes scalar product.

The *sampling* is assumed to be ideal. Introduce the *sampling* or *discretisation* operator,  $D : \mathcal{B} \rightarrow l_2$ ,

$$w(i, j) = (DW)(i, j) = W(i, j), \quad i, j \in \mathbb{Z} . \quad (9.3)$$

The sampling operator maps a continuous signal  $W$  onto a discrete signal  $w$ .

Finally, *noise* is assumed to be an additive stationary random field, i.e. we ignore the quantisation effects. Experimentally, it is verified that the errors in individual pixel intensities often can be modelled as independent random variables with similar distribution.

These assumptions will serve as an initial model. Further improvements can be made by a more detailed camera acquisition model. Nevertheless, these assumptions will help us to model and analyse the next stage, namely estimating the continuous image intensity distribution from the discrete image. Obviously, it is impossible to reconstruct the original intensity distribution  $W_{\text{ideal}}$  without prior knowledge. It is, however, reasonable to try to estimate the blurred and distorted intensity distribution  $W$ , or to estimate an even more blurred version.

### 9.3 Interpolation and smoothing

Scale-space theory and its application to computer vision has obvious advantages. In the continuous case, smoothing with the *Gaussian kernel*

$$G_b(x) = \frac{1}{\sqrt{2\pi b^2}} e^{-|x|^2/2b^2} \quad (9.4)$$

is very natural. Under some consistency conditions (symmetry, semi-group property, non-enhancement of local extrema), see Chapter 4–7, the Gaussian kernel is the only choice that gives a consistent scale-space theory, cf. (Babaud et al., 1986; Koenderink, 1984; Lindeberg, 1994e; Witkin, 1983). The *smoothing* operator  $S_b$  represents convolution with the Gaussian kernel  $G_b$ . A signal  $W$  is represented at scale  $b$  by its smoothed version  $W_b$ :

$$W_b = S_b(W) = G_b * W . \quad (9.5)$$

The signal  $W_b$  is called the *scale-space representation* of  $W$ , at scale  $b$ . In the sequel subscripts are used to denote different scales.

Scale-space theory in the discrete time case has been investigated in (Lindeberg, 1994e). It turns out that just by sampling a continuous scale-space kernel, one obtains a discrete scale-space kernel. However, in doing so one does not obtain a scale-space theory with all the nice features of the continuous scale-space theory. There are difficulties with fine scales. In particular it is difficult to define higher order derivatives at fine scale levels. For the same reason it is difficult to define local extremum and zero crossings for fine scales. The semi-group property is lost. However, a purely discrete theory is introduced in (Lindeberg, 1994e) using generalised Bessel functions. This solves some of the difficulties above, but it is only possible to calculate derivatives at integer positions.

## Interpolation and smoothing

The main idea of our approach is to induce the discrete signal, the scale spaces, etc. from the associated interpolated quantities. By an *interpolation* or *restoration* method we mean an operator that maps a discrete signal,  $w$ , to a continuous one,  $W$ . The following types of interpolation operators  $I_F$  will be used:

$$W(s) = (I_F w)(s) = \sum_i F(s - i)w(i) . \quad (9.6)$$

The ideal low-pass interpolation operator  $I = I_{\text{sinc}}$  is of special interest. (Ideal in the sense that it is the inverse operator to the sampling operator on signals that are band limited below the Nyquist frequency, (Nyquist, 1928).) We propose to use this operator,  $I$ , with discretisation  $D$  as mappings between the continuous and discrete signals to solve the restoration and discrete scale-space problems. In other words we relate the discrete and continuous signals through the operations of discretisation and interpolation.

Note that if the camera induced blur cancels the high frequency components in  $W$  as in Assumption 9.2.1, the deterministic restoration  $W_0 = I(w_0)$  is equal to  $W$  by the Nyquist Theorem.

Using these definitions, the discrete and continuous scale-space representations can be defined simultaneously and consistently. We propose the following:

1. If the primary interest is the interpolated continuous signal, then *restore* the scale-space smoothed continuous signal  $W_b$  from the discrete signal  $w_0$  first using ideal interpolation and then continuous scale-space smoothing.
2. If the primary interest is a discrete scale-space representation, then use the induced representation from the continuous scale-space representation.

The procedure is illustrated by the diagram:

$$\begin{array}{ccc} W_0 & \xleftarrow{I} & w_0 \\ s_b \downarrow & & \downarrow s_b \\ W_b & \xrightarrow{D} & w_b \end{array} \quad (9.7)$$

Thus, from the discrete signal  $w_0$ , the *continuous* scale-space smoothed signal  $W_b$  is obtained as  $W_b = S_b(I(w))$ . The *discrete* scale-space signal  $w_b = s_b(w_0)$ , is induced from the continuous scale-space signal, i.e.

$$w_b = s_b(w_0) \stackrel{\text{def}}{=} D(S_b(I(w_0))) , \quad (9.8)$$

where  $s_b$  is introduced as the discrete scale-space smoothing operator. Notice that  $s_b$  is a convolution with a kernel  $g_b$ ,

$$g_b = D(G_b * \text{sinc}) . \quad (9.9)$$

The differences between this approach and others, like the sampled Gaussian approach, is very small for large scales but significant for small scales. In fact it can be shown that

$$\| \text{sinc} * G_b - G_b \|_2^2 \leq \frac{1}{b\sqrt{\pi}} \Phi(-\pi b\sqrt{2}) , \quad (9.10)$$

where  $\Phi$  is the normal cumulative distribution function. Notice that the right hand side is small when  $b$  is large. The sampled Gaussian approach is also equivalent to using interpolation with the delta distribution followed by Gaussian smoothing. The main motivation for using ideal low-pass interpolation is, however, that the approach is well suited for stochastic analysis as will be shown later. Observe that the interpolated signal  $W$  is smooth. Thererfore, there is no difficulty in defining higher order derivatives.

This scale-space theory has several theoretical advantages: It works for all scales. The semi-group property,  $s_{\sqrt{a}}s_{\sqrt{b}} = s_{\sqrt{a+b}}$ , holds. The coupling to continuous scale-space theory gives a natural way to interpolate in the discrete space. There are no difficulties in defining derivatives at arbitrary scales. It is possible to calculate derivatives at arbitrary interpolated positions. Operators which commute in the continuous theory automatically commute in the discrete theory. The effect of additive stationary noise can easily be modelled. It makes it possible to compare the real intensity distribution with the interpolated distribution.

In practice this scale-space theory is difficult to use for small scale parameters, because of the large tail of the function in (9.9). However, the function  $\text{sinc} * G_b$  has a very small tail for larger scales. In practise one may use the approximation  $\text{sinc} * G_b \approx G_b$  for large scales, according to (9.10). This simplifies implementation substantially.

## 9.4 The random field model

The discrete image  $v_0 = w_0 + e_0$  is analysed directly or through scale-space smoothing, as illustrated by the diagram:

$$\begin{array}{ccc} W_0 + E_0 & \xleftarrow{I} & w_0 + e_0 \\ s_b \downarrow & & \downarrow s_b \\ W_b + E_b & \xrightarrow{D} & w_b + e_b \end{array} \quad (9.11)$$

Note that all operations are linear. The stochastic and deterministic properties can, therefore, be studied separately and the final result is obtained by superposition. Thus with an a priori model on  $W_{\text{ideal}}$ , for example an ideal edge or corner, it is possible to predict the deterministic parts  $W_b$  and  $w_b$ . The stochastic properties of the error fields  $e_0$ ,  $e_b$ ,  $E_0$  and  $E_b$ , will now be studied.

## Stationary random fields

The theory of random fields is a simple and powerful way to model noise in signals and images. Stationary or wide sense stationary random fields are particularly easy to use. Denote by  $\mathcal{E}$  the expectation value of a random variable.

**Definition 9.4.1.** A random field  $X(t)$  with  $t \in \mathbb{R}^n$  is called *stationary* or *wide sense stationary*, if its *mean*  $m(t) = m_X(t) = \mathcal{E}[X(t)]$  is constant and if its *covariance function*  $r_X(t_1, t_2) = \mathcal{E}[(X(t_1) - m(t_1))(X(t_2) - m(t_2))]$  only depends on the difference  $\tau = t_1 - t_2$ . A random field  $X(t)$  with  $t \in \mathbb{R}^n$  is called *strictly stationary* if for all  $(t_1, \dots, t_n)$  and all  $\tau$  the stochastic variable  $(X(t_1), \dots, X(t_n))$  has the same probability distribution as  $(X(t_1 + \tau), \dots, X(t_n + \tau))$ . ■

For stationary fields we will use  $r_X(s, t)$  and  $r_X(s - t)$  interchangeably as the *covariance function*. The analogous definition is used for a stationary field in discrete parameters. The notion of *spectral density*

$$R_X(f) = (\mathcal{F}r_X)(f) = \int r_X(\tau) e^{-i2\pi f \cdot \tau} d\tau \quad (9.12)$$

is also important. Again the same definition can be used for random fields with discrete parameters  $s \in \mathbb{Z}^n$ , but whereas the spectral density for random fields with continuous parameters is defined for all frequencies  $f$ , the spectral density of discrete random fields is only defined on an interval  $f \in [-1/2, 1/2]^n$ . Introductions to the theory of random processes and random fields are given in (Adler, 1985; Cramér and Leadbetter, 1967; Cressie, 1991). In these books you will find that convolution, discretisation and derivation preserves stationarity.

$$\begin{aligned} w &= D(W) &\Rightarrow r_w &= D(r_W) \\ Y &= h * X &\Rightarrow R_Y &= R_X |\mathcal{F}h|^2 \\ Y &= \frac{\partial}{\partial t_i} X &\Rightarrow r_Y &= -\frac{\partial^2}{\partial \tau_i^2} r_X \end{aligned}$$

It is shown in (Åström, 1996) that ideal interpolation  $I$  preserves stationarity as well.

**Theorem 9.4.2 (Interpolation of a random field).** Let  $e(i_1, \dots, i_n)$  be a stationary discrete random field with zero mean and covariance function

$$r_e((i_1, \dots, i_n), (j_1, \dots, j_n)) = r_e(i_1 - j_1, \dots, i_n - j_n) ,$$

such that

$$r_e \in l^p, \quad \text{for some } p < \infty ,$$

where  $l^p$  denotes the space of p-summable sequences. Then the ideal interpolation of the discrete random field,

$$E = I(e) , \quad (9.13)$$

is a well defined random field in quadratic mean and  $E$  is stationary with covariance function

$$r_E(\tau) = I(r_e)(\tau) . \quad (9.14)$$

Thus, all operations in the commutative diagram (9.11) preserve stationarity. This simplifies the modelling of errors in scale-space theory. The effects of the operators  $I$ ,  $D$ ,  $S_b$  and  $s_b$  on covariance  $r$  and spectral density  $R$  are thus known.

It is often convenient to assume that the discrete noise  $e_0$  can be modelled as white noise, i.e.

$$r_e(k) = \begin{cases} \epsilon^2, & \text{if } k = 0, \\ 0, & \text{if } k \neq 0. \end{cases}$$

It can then be shown that the covariance function of the interpolated and smoothed error field is

$$r_{E_b} = \epsilon^2 \operatorname{sinc} * G_{b\sqrt{2}} . \quad (9.15)$$

**Remark.** The restored image intensity distribution  $V_b$  is a sum of a deterministic part  $W_b$  and a stationary random field  $E_b$ . Notice that the restoration and the residual are *invariant* of the position of the discretisation grid. The effect of discretisation is thus removed. ■

## 9.5 Applications

Analysis of sub-pixel edge-detection and sub-pixel correlation are two application of our scale-space theory.

In *sub-pixel correlation* the idea is that, at least locally, the images only differ by an unknown translation  $\rho$ . Denote by  $V = W + E$  and  $\bar{V} = \bar{W} + \bar{E}$  the restored intensity fields in two images for a fixed scale  $b$ . The deterministic functions are identical except for a translation. For a fixed translation  $\rho_0 = (\rho_1, \rho_2)$ , we thus have

$$W(t) = \bar{W}(t + \rho_0), \quad \forall t .$$

To determine the translation  $\rho$  with sub-pixel accuracy a least squares integral is minimised,

$$F(\rho) = \int_{t \in \Omega} (V(t) - \bar{V}(t + \rho))^2 dt .$$

Furthermore, the residual field  $V(t) - \bar{V}(t + \rho)$  can be used to empirically study the stochastic properties of the camera noise  $e_0$ . The quality of the estimated sub-pixel translation,

$$\hat{\rho} = \operatorname{argmin} F(\rho) ,$$

can be analysed using the statistical model given above. Let  $X = \hat{\rho} - \rho_0$  be the error in estimated translation. By linearising the function  $F$  it can be shown, see (Åström, 1996; Åström and Heyden, 1996a), that the probability distribution of  $X$  can be approximated with a normal distribution with zero mean and covariance matrix given by

$$C = \mathcal{C}[X] \approx A^{-1}BA^{-T}, \quad (9.16)$$

$$A = 2 \int_{t_1 \in \Omega} (V\bar{W}\bar{W}^T)(t_1)dt_1, \quad (9.17)$$

$$B = \int_{t_1 \in \Omega} ((V\bar{W}) * r_{E-\bar{E}})(t_1) (V\bar{W})(t_1)dt_1. \quad (9.18)$$

Similarly, *sub-pixel edge detection* can be analysed by modelling the edge  $W_{\text{ideal}}$ , the blur  $h$ , the noise  $e_0$  and analysing its effect on a sub-pixel edge detector. This is studied in (Åström and Heyden, 1996b). For low levels and low curvature edges it can be shown that the deviations in the normal direction of the detected curve is approximately a stationary random process with respect to the arc-length of the curve.

## 9.6 Conclusions

In this chapter we have modelled the image acquisition process, taking into account both the deterministic and stochastic aspects. In particular the discretisation process is modeled in detail. This interplay between the continuous signal and its discretisation is very fruitful and the increased knowledge sheds light on scale-space theory, feature detection and stochastic modelling of errors.

The relation between the continuous signal and its discretisation is used to obtain an alternative scale-space theory for discrete signals. It is also used to derive methods of restoring the continuous scale-space representation from the discrete representation. This enables us to calculate derivatives at any position and of any scale.

Furthermore, the stochastic errors in images are modelled and new results are given that show how these errors influence the continuous and discrete scale-space representations and their derivatives. This information is crucial in understanding the stochastic behaviour of scale-space representations as well as fundamental properties of feature detectors. In particular, we have analysed a simple sub-pixel edge detector and a sub-pixel correlator in detail.

In order to validate the theory, experiments and simulations both on real and simulated data have been presented. Good agreement with the theoretical model is achieved.

The work can be extended in several directions. Edges were modelled as straight ideal step edges. It would be interesting to study the effect (the bias) caused by other types of edges and the effect of high curvature edges. The model of image acquisition, interpolation and scale space smoothing can also be used to analyse other feature detectors.

The authors thank Georg Lindgren for inspiration and valuable discussions.

# **Part III**

# **The Structure**

# Chapter 10

## Local Analysis of Image Scale Space

Peter Johansen

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

### 10.1 Introduction

Scale space theory is a framework which permits one to analyze an image at different resolutions. The crucial feature is to let the resolution of an image vary continuously rather than discrete. When extrema and saddle points of an image are tracked with increasing scale we know that they will have disappeared at sufficiently large scale. In this chapter we investigate what happens to the extrema and saddles that disappear. It turns out that extrema and saddles also appear for two dimensional images when scale is increased. Two dimensional images are more complicated than one dimensional signals, and some results generalize from two to higher dimensions. In this chapter we consider two dimensional images as a compromise between generality and simplicity. One generic event is a saddle and an extremum merging and annihilating. Prior to annihilation they approach each other from opposite directions. The other generic event is the creation of a saddle-extremum pair. It will be shown that these are the only generic events. Studying scale space events for images in an image sequence or for a symmetric image, one discovers that more complicated interactions between extrema and saddles take place when scale is increased. The situations which arise when extrema and saddles disappear and appear can be classified by their codimension. The codimension is a measure of complexity. Generic events are the least complex and have codimension one. We also analyse events of codimension two. The mathematical tool which will be used for computing the local structure of top points is differential geometry

of curves.

A digital image is routinely represented as a rectangular lattice of integer numbers. This representation is inadequate for pattern recognition. A slight modification of the imaged object in one of two identical images (for instance by translation, rotation or scaling of the imaged object) changes the representation drastically. Scale space is intended to provide a better representation (Koenderink, 1984) (Lindeberg, 1994e) (Witkin, 1983). This alternative representation should be uncommitted in the sense that it must be possible in principle to reconstruct the lattice representation.

The idea in this chapter is to select the generic events in scale space as the atoms of a representation. They can be described as follows. A *critical point* is a point in which the gradient is zero. The critical points are the extrema (maxima and minima) and the saddles of the image's grey level function, which is denoted  $L$ . A *critical curve* is a curve in scale space consisting of critical points. The Hessian  $H$  is the matrix whose elements are the partial derivatives of second order

$$H(x, y) = \begin{Bmatrix} L_{xx}(x, y) & L_{xy}(x, y) \\ L_{yx}(x, y) & L_{yy}(x, y) \end{Bmatrix}$$

Partial derivatives are denoted by lower indices, for example  $L_{xx}$  denotes  $\frac{\partial^2 L}{\partial x^2}$ . The directions of maximum and minimum curvature are the eigenvectors of the Hessian. A critical point is a *toppoint* if at least one eigenvalue is zero, i.e. if the Hessian does not have full rank. It will be shown that the topoints are precisely the points at which the generic events take place. The critical curves are solutions to the equations which expresses that the gradient is zero:

$$\begin{aligned} L_x(x, y, t) &= 0 \\ L_y(x, y, t) &= 0, \end{aligned} \tag{10.1}$$

where  $x$  and  $y$  denote the spatial coordinates,  $\sigma$  is scale and  $t = 2\sigma^2$ . In order to study this system of curves, it is useful to find an expression for the tangent at a given point. As is usual in differential geometry, we let  $s$  be a monotonically varying parameter along the curve. Thus  $x$  and  $y$  are functions of  $s$  and in this way  $L$  and its partial derivates with respect to  $x$ ,  $y$  and  $t$  are implicit functions of  $s$ . It follows from equation (10.1), that for all integer  $n$

$$\begin{aligned} \frac{d^n}{ds^n} L_x(x(s), y(s), t(s)) &= 0 \\ \frac{d^n}{ds^n} L_y(x(s), y(s), t(s)) &= 0. \end{aligned} \tag{10.2}$$

In what follows we shall use dots to denote differentiation with respect to  $s$ , such that for instance  $\ddot{x}$  denotes  $\frac{d^2 x}{ds^2}$ . For  $n = 1$  we get two equations using the chain rule.

$$\begin{aligned} L_{xx}(x, y, t)\dot{x} + L_{xy}(x, y, t)\dot{y} + L_{xt}(x, y, t)\dot{t} &= 0 \\ L_{xy}(x, y, t)\dot{x} + L_{yy}(x, y, t)\dot{y} + L_{yt}(x, y, t)\dot{t} &= 0, \end{aligned}$$

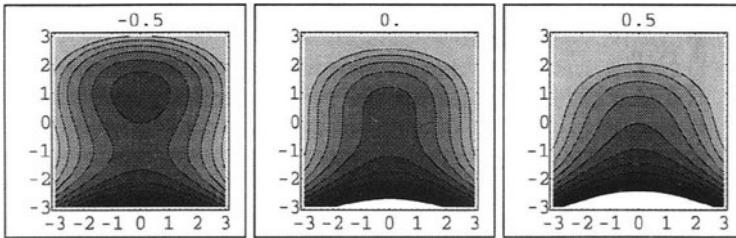


Figure 10.1:  $L = 2x^2 + y^3 + 4t + 6yt$ . The convex generic event. An extremum and a saddle merge when scale is increased. (left)  $t = -4/10$ . (center)  $t = 0$ . (right)  $t = 4/10$ .

from which the directions of the tangent  $(\dot{x}, \dot{y}, \dot{t})$  can be found at any point on the curve.

We shall analyse the curve at a particular point  $(x_0, y_0, t_0)$  by computing the tangent and its derivatives. In the following we will omit reference to  $(x_0, y_0, t_0)$  and by for instance  $L_{xx}$  understand  $L_{xx}(x_0, y_0, t_0)$ . We observe that if the Hessian has full rank,  $\dot{t}$  can be used as a parameter along the tangent, meaning that the tangent is not horizontal, which in turn has the geometrical interpretation that the critical point in the image does *not* disappear when the scale is infinitesimally increased. On the other hand, if the Hessian has rank 1 (still assuming the equations to be independent), the tangent is horizontal. The geometrical interpretation is that two critical points merge and disappear when the scale is infinitesimally increased.

The event just described is unavoidable for Gaussian smoothing. Critical points must disappear with increasing scale and this event is the generic way in which it happens. Figure 10.1 shows an example. It is visualized by the contour curves of the image at scales below, at, and above the scale of the top point.

The term generic means that if the image is changed slightly, the event may change position in scale space, but it will still be present. Perhaps surprisingly, also another event is generic: the *appearing* of two critical points. Below we shall derive the conditions for this to happen. These two events are the only generic ones, and are among the primitives proposed for image representation.

### 10.1.1 Codimension

Even if the two above events are sufficient to describe a typical image, in three situations it is required to study degenerate cases in which the image can not be said to be typical. These include *local symmetry*, *quantization* of scalar and spatial resolution, and temporal *image sequences*. In all three cases the additional image constraints have the effect that more than two points may be involved in a generic scale space event. Thus, it will be necessary to include degenerate events as description primitives. Below tools for describing degenerate events are put forward. In order to understand the algebra involved, let us rewrite the equations

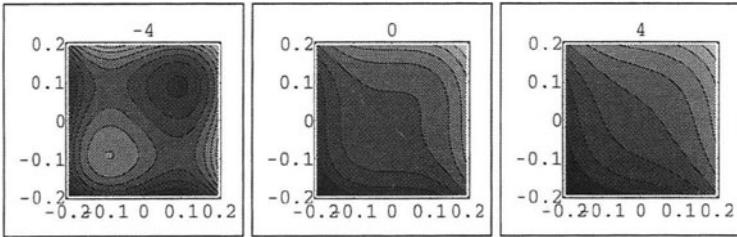


Figure 10.2:  $L = x^3 + y^3 + 6(x + y)t$ . Two extrema and two saddles disappear simultaneously when scale is increased. The two saddles merge from opposite directions simultaneously with two extrema (a maximum and a minimum) merging from the perpendicular direction. Extrema enter from north-east and south-west and saddles from north-west and south-east. (left)  $t = -4/1000$ . (center)  $t = 0$ . (right)  $t = 4/1000$ .

(10.2) in which the partial derivatives have been evaluated at  $(x_0, y_0, t_0)$  for  $n = 1$  and  $n = 2$ .

$$\begin{aligned} L_{xx}\dot{x} + L_{xy}\dot{y} + L_{xt}\dot{t} &= 0 \\ L_{xy}\dot{x} + L_{yy}\dot{y} + L_{yt}\dot{t} &= 0 \end{aligned} \quad (10.3)$$

$$\begin{aligned} L_{xxx}\dot{x}^2 + L_{xyy}\dot{y}^2 + L_{xtt}\dot{t}^2 + 2L_{xxy}\dot{x}\dot{y} + 2L_{xxt}\dot{x}\dot{t} + 2L_{xyt}\dot{y}\dot{t} + L_{xx}\ddot{x} + L_{xy}\ddot{y} + L_{xt}\ddot{t} &= 0 \\ L_{yxx}\dot{x}^2 + L_{yyy}\dot{y}^2 + L_{ytt}\dot{t}^2 + 2L_{xyy}\dot{x}\dot{y} + 2L_{xyt}\dot{x}\dot{t} + 2L_{yyt}\dot{y}\dot{t} + L_{xy}\ddot{x} + L_{yy}\ddot{y} + L_{yt}\ddot{t} &= 0. \end{aligned} \quad (10.4)$$

If we impose one algebraic constraint, then we say that the event described has *codimension* one. We choose the constraint that the rank of  $H$  is one. An immediate consequence is that the tangent is horizontal (use equations (10.3)). From the same equations the tangent direction can be found.

If we impose an additional constraint the codimension becomes two. Codimension two events are determined by the additional constraint that the Hessian has rank zero. In this case equations (10.3) still imply that a tangent must be horizontal but they are no longer sufficient to determine the tangent direction. We must invoke equations (10.4). Since the Hessian has rank zero,  $\dot{t}$  can be eliminated to get a quadratic equation in the unknowns  $\dot{x}$  and  $\dot{y}$ . It turns out that for Gaussian scale space there are two horizontal perpendicular tangents in a codimension two event. The geometric interpretation is that two critical curves intersect, and the solutions provide the direction of the tangents at the point of intersection. Figure 10.2 shows two extrema and two saddles which disappear simultaneously.

Please note that the computations are undertaken projectively. Apparently we have one equation too few for the tangent and for its higher derivatives. If we

insist on computing the exact size of the tangent and its derivatives, we can use a set of equations which expresses that the hypothetical point traversing the curve has constant velocity. This means that the parameter  $s$  measures arclength. The first few equations are

$$\begin{aligned} \dot{x}^2 + \dot{y}^2 + \dot{t}^2 &= 1 \\ 2\dot{x}\ddot{x} + 2\dot{y}\ddot{y} + 2\dot{t}\ddot{t} &= 0 \\ 2\ddot{x}^2 + 2\ddot{y}^2 + 2\ddot{t}^2 + 2\dot{x}\ddot{x} + 2\dot{y}\ddot{y} + 2\dot{t}\ddot{t} &= 0. \end{aligned}$$

## 10.2 Analysis of codimension one

Let the rank of  $H$  be one.  $(L_{xx}, L_{xy})$  and  $(L_{xy}, L_{yy})$  cannot vanish simultaneously. Let  $(L_{xx}, L_{xy}) \neq (0, 0)$ .  $\dot{x}$  and  $\dot{y}$  can be eliminated from equation (10.3) to get

$$\left| \begin{array}{cc} L_{xx} & L_{xt} \\ L_{xy} & L_{yt} \end{array} \right| \dot{t} = 0,$$

and we may infer that  $\dot{t} = 0$ , which implies horizontal tangent. The tangent direction can then be computed from equation (10.3).

In order to know if the curve is convex (two points disappear with increasing scale) or concave (two points appear with increasing scale) we must compute  $\ddot{t}$ . In the convex case it is negative and it is positive in the concave case. We wish further to know which types (saddle or extrema) are involved. To this end we need to know the sign of the Hessian determinant on the curve. To find  $\ddot{t}$  we can eliminate  $\ddot{x}$  and  $\ddot{y}$  from equation (10.4) to get

$$\left| \begin{array}{cc} L_{xxx} & L_{xx} \\ L_{xxy} & L_{xy} \end{array} \right| \dot{x}^2 + 2 \left| \begin{array}{cc} L_{xxy} & L_{xx} \\ L_{xyy} & L_{xy} \end{array} \right| \dot{x}\dot{y} + \left| \begin{array}{cc} L_{xyy} & L_{xx} \\ L_{yyy} & L_{xy} \end{array} \right| \dot{y}^2 + \left| \begin{array}{cc} L_{xt} & L_{xx} \\ L_{yt} & L_{xy} \end{array} \right| \ddot{t} = 0.$$

Depending on the values of the partial differential quotients at the topoint,  $\ddot{t}$  can be positive or negative. This means that the curve may be convex or concave. An example of a generic convex event is shown in figure 10.1, and figure 10.3 shows an example of a generic concave event.

Next we want to know the type of points involved. An approximation to first order in  $s$  of the Hessian is computed on each of the two tangents. This is done by approximating the partial derivatives on the curve to first order in  $s$ .

$$\begin{aligned} \left| \begin{array}{cc} L_{xx}(x, y, t) & L_{xy}(x, y, t) \\ L_{yx}(x, y, t) & L_{yy}(x, y, t) \end{array} \right| &= \left| \begin{array}{cc} L_{xx} + s(L_{xxx}\dot{x} + L_{xxy}\dot{y}) & L_{xy} + s(L_{xxy}\dot{x} + L_{xyy}\dot{y}) \\ L_{yx} + s(L_{xxy}\dot{x} + L_{xyy}\dot{y}) & L_{yy} + s(L_{xyy}\dot{x} + L_{yyy}\dot{y}) \end{array} \right| \\ &= \mathcal{O}(s). \end{aligned}$$

This means that the sign changes as  $s$  passes through the topoint. The geometric interpretation is that a saddle and an extremum are involved. This is true for both a convex and a concave curve.

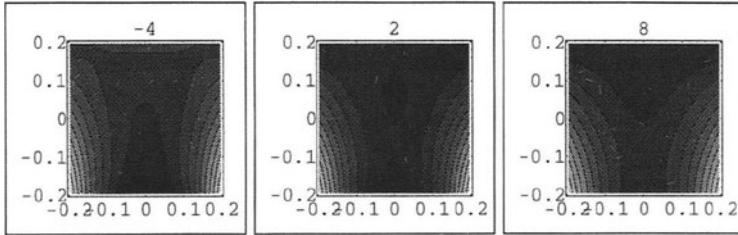


Figure 10.3:  $L = 2x^2 - 12x^2y + 2y^3 - 12yt + 4t$ . The concave generic event. An extremum and a saddle split. (left)  $t = -4/1000$ . (center)  $t = 2/1000$ . (right)  $t = 8/1000$ .

### 10.3 Analysis of codimension two

Let us next assume that the rank of the Hessian is 0, which means that  $L_{xx} = L_{xy} = L_{yy} = 0$ . It follows from equation (10.3) that  $\ddot{t} = 0$ , which implies that the tangents are horizontal. I use the plural here since there are actually two tangents. This can be seen from equations (10.4) from which  $\ddot{t}$  has been eliminated.

$$\begin{vmatrix} L_{xxx} & L_{xt} \\ L_{xxy} & L_{yt} \end{vmatrix} \dot{x}^2 + 2 \begin{vmatrix} L_{xxy} & L_{xt} \\ L_{yyy} & L_{yt} \end{vmatrix} \dot{x}\dot{y} + \begin{vmatrix} L_{yyy} & L_{xt} \\ L_{yyy} & L_{yt} \end{vmatrix} \dot{y}^2 = 0. \quad (10.5)$$

We notice that in the case of gaussian smoothing (use the relations  $L_{xt} = L_{xxx} + L_{xxy}$  and  $L_{yt} = L_{xxy} + L_{yyy}$  which are induced by the heat equation)

$$\begin{vmatrix} L_{xxx} & L_{xt} \\ L_{xxy} & L_{yt} \end{vmatrix} + \begin{vmatrix} L_{xxy} & L_{xt} \\ L_{yyy} & L_{yt} \end{vmatrix} = 0.$$

Since the coefficients of  $\dot{x}^2$  and  $\dot{y}^2$  have the same magnitude but different sign, the discriminant of the quadratic equation is positive and two real solutions exist for the tangent directions. A further consequence is that the two directions are perpendicular to each other. This follows by observing that the product of the roots is  $-1$ .

$\ddot{t}$  can be found from equation (10.4). It is further seen that  $|H| = \mathcal{O}(s^2)$  and we infer that along each of the two intersecting curves two critical points of the same type approach each other. In order to gain more insight in the structure we now orient the coordinate system such that the tangents point along the  $x$ -axis and the  $y$ -axis. This has an equivalent formulation in terms of a constraint between the partial derivatives. It can be found from equation (10.5) which was solved for the two tangent directions. The coefficient of the quadratic term must be zero. This means that  $L_{xxx}L_{yyy} - L_{xxy}L_{yyy} = 0$ . The Hessian computed on the tangent which points along the  $y$ -axis becomes

$$H_1 = H(0, s) = (L_{xxy}L_{yyy} - L_{xxy}^2)s^2.$$

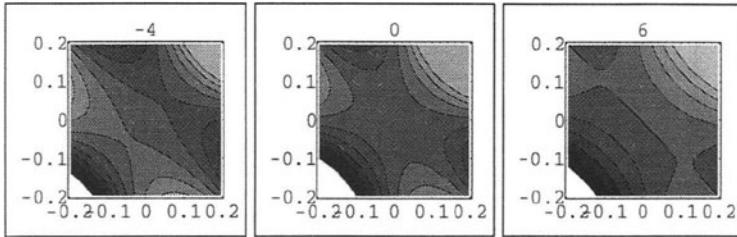


Figure 10.4:  $L = x^2y + y^2x + 2(x+y)t$ . Two saddles merge from opposite directions and separate in the perpendicular direction. They enter from north-east and south-west and separate in directions north west and south east. (left)  $t = -4/1000$ . (center)  $t = 0$ . (right)  $t = 6/1000$ .

$H$  on the tangent directed along the  $x$ -axis is

$$H_2 = H(s, 0) = (L_{xxx}L_{yyy} - L_{xxy}^2)s^2.$$

Notice that

$$L_{xxy}H_1 + L_{yyy}H_2 = 0.$$

We shall also need  $\ddot{t}$  on the two tangents. Let  $\ddot{t}_1 = \ddot{t}(0, s)$  be  $\ddot{t}$  on the tangent directed along the  $y$ -axis, and let  $\ddot{t}_2 = \ddot{t}(s, 0)$  be  $\ddot{t}$  on the tangent directed along the  $x$ -axis.

From equation 10.4 we get

$$\ddot{t}_1 = -L_{yyy}/(L_{xxy} + L_{yyy})$$

and

$$\ddot{t}_2 = -L_{xxy}/(L_{xxy} + L_{yyy}).$$

Notice that

$$\ddot{t}_2 H_1 + \ddot{t}_1 H_2 = 0. \quad (10.6)$$

Case 1a:  $L_{xxy}L_{yyy} > 0$ .  $L_{xxy}$  and  $L_{yyy}$  have same sign. Both  $\ddot{t}_1$  and  $\ddot{t}_2$  are negative, which means that both curves are convex. Equation (10.6) implies that  $H_1$  and  $H_2$  have opposite signs, which means that one curve consists of extrema and the other curve consists of saddles. Figure 10.2 shows an example.

Case 1b:  $L_{xxy}L_{yyy} < 0$ .  $L_{xxy}$  and  $L_{yyy}$  have opposite signs. It is immediately seen that  $H_1$  is negative.  $\ddot{t}_1\ddot{t}_2 < 0$  which means that one curve is convex and the other concave. Equation 10.6 implies that  $H_2$  is also negative. Consequently both curves consist of saddles. Figure 10.4 shows an example.

Two critical curves may intersect with non-horizontal tangents (Johansen, 1994). An example is shown in figure 10.5.

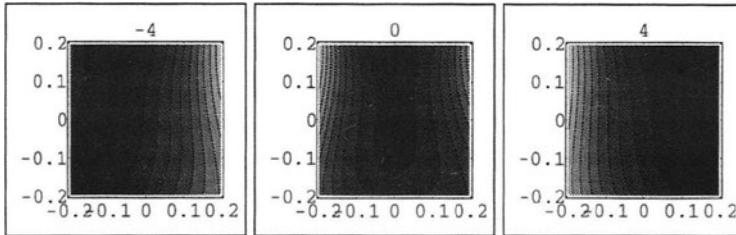


Figure 10.5:  $L = x^2 - x^3 + 3x^2y - y^3 - x^4y + 2t - 6tx - 12t^2y - 12tx^2y$ . Two non-horizontal intersecting tangents. For negative  $t$  a saddle approaches from the second quadrant at the same time as a maximum from the third quadrant. For  $t = 0$  they form a shoulder. For positive  $t$  it is split into a saddle running up into the first quadrant and a maximum into the fourth quadrant. The two critical curves intersect. (left)  $t = -4/10$ . (center)  $t = 0$ . (right)  $t = 4/10$ .

## 10.4 Suggestions for Further Research

**Local Symmetry** Describe the degenerate toppoints by a symmetry group. A generic top point has a  $180^\circ$  symmetry when simultaneously interchanging extremum and saddle. Codimension two events have larger symmetry groups. Can one assign a unique group to a topoint? Which groups are they? Does a family of groups characterize gaussian scale space?

**Image Representation** Is it possible and useful to represent an image by its topoints? In one dimension a representation by topoints has been proposed (Johansen et al., 1986). Knowledge of the position of topoints in scale space is sufficient. It is not immediately transferable to 2 dimensions because concave events are generic. Convex topoints are unavoidable. Are concave topoints also unavoidable? How frequent are they?

## 10.5 Acknowledgement

Luc Florack gave me comments on the final draft.

## Chapter 11

# Local Morse Theory for Gaussian Blurred Functions

**James Damon**

*Department of Mathematics, University of North Carolina  
Chapel Hill, NC 27599, USA*

When Gaussian blurring is applied to an intensity function  $f_0(x)$ , it yields a family  $f(x, t)$  of intensity functions parametrised by  $t$ , which is a solution to the heat equation

$$\frac{\partial f}{\partial t} = \Delta(f) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} \quad \text{on } R^n \times R_+ \quad (11.1)$$

and satisfying initial conditions  $f(x, 0) = f_0(x)$  for  $f_0 : R^n \rightarrow R$ .

We would like to know what specific properties we can expect  $f$  to have, beyond those basic ones it has as a result of being a solution to the heat equation (such as being a smooth function of  $(x, t)$  for  $t > 0$ ). We know from singularity theory that it is unrealistic to expect to be able to give specific properties for all functions, even smooth ones. For example, we generally expect that for a smooth function  $f : R^n \rightarrow R$ ,  $f^{-1}(0)$  is a smooth  $n - 1$ -dimensional manifold in  $R^n$ , except possibly on a smaller dimensional subset of points at which it exhibits some type of singular behavior (such as a smooth curve in  $R^2$  crossing through itself). However, for any closed subset  $A \subset R^n$ , Whitney has shown that there is a smooth function  $f$  with  $f^{-1}(0) = A$ . However, in a well-defined sense, this example is quite exceptional, and is not the type of example we “generally expect” to encounter.

Thus, realistically, we can only hope to describe properties which “generally hold”. We shall see that these are properties that hold for “almost all” functions, and we will refer to such properties as “generic properties”. We still have to make precise what exactly these terms mean; however, as one consequence, it will follow that a non-generic function can be perturbed by an arbitrarily small amount

to yield a generic function. Also, although it doesn't follow from genericity, we might also hope that once we've reached a generic function, a sufficiently small perturbation will not destroy the genericity; this is a form of stability which we shall also establish. Furthermore, we would like to have simple criteria in terms of low order derivative information to check for genericity. Finally, in the case of blurred functions  $f$  we are interested in the properties  $f$  will generally have for "almost all" initial intensity functions  $f_0$ . We also wish to make sense of such a statement.

From among the many possible properties of blurred functions to be investigated, we shall concentrate on two specific properties, namely, the behaviour of critical points of  $f$  (viewed as a function of  $x$ , with  $t$  as a parameter) and the interaction of regions of given intensity in the course of the blurring. However, other properties such as the generic behaviour of "Canny edges" (Rieger, 1995) or the generic properties of "cores" (introduced by Pizer-Eberly (Eberly et al., 1994; et al, 1994)) have also been determined using these methods (Damon, 1995b).

Briefly we proceed by first explaining how classical Morse theory classifies critical points and introduce the modified notions of equivalence which capture the properties of interest to us. Then, we explain the various ideas associated with genericity as it will apply to these properties. To establish the genericity of these properties, we make extensive use of the weighted homogeneity of the heat equation and the associated weight decomposition of the solutions. With this information, we are able to both list the generic local forms exhibiting the generic local behaviour, and to describe the local properties. We also indicate the type of generic behavior occurring in one parameter families, as would be expected in a one parameter deformation of an intensity function. Lastly, we provide a brief indication of the ideas of transversality, versality of unfoldings, and criteria for stability which underly the results we describe.

## 11.1 Classifying Properties of Blurred Images via Equivalences

A standard method for studying the behaviour of functions near critical points involves changing coordinates so that the function can be put into a particularly simple form, referred to as a "normal form", from which the properties may be determined. For solutions to the heat equation, the transformations which preserve the heat equation lead to a classification which is too sensitive. Any quadratic function  $f(x) = \sum_{i=1}^n a_i x_i^2$  with  $\sum_{i=1}^n a_i = 0$  is a solution to the heat equation; however, the only linear transformations on  $R^n$  which preserve the heat equation are orthogonal ones and for these the eigenvalues  $a_i$  are invariant. By contrast, the structure of the level sets  $f^{-1}(y)$  only depend on the signs of the  $a_i$ .

In the opposite direction, suppose  $f$  has a critical point at a point  $x_0$  but has terms of arbitrarily high order. Now we allow composition with a local nonlinear change of coordinates  $x = \varphi(x')$ . Provided the *critical point is nondegenerate*, which means the Hessian matrix of  $f$  at  $x_0$  is nonsingular, then a basic lemma of

Morse theory states there is such a  $\varphi$  so that

$$f \circ \varphi = \sum_{i=1}^n \delta_i x_i'^2 + c \quad \text{with each } \delta_i = \pm 1 \quad (11.2)$$

*Example 11.1.1.* The function  $f(x, y) = x^2 - y^2 + x^3 - 3xy^2$  has a critical point at 0. By the inverse function theorem, the transformation

$$(x', y') = \varphi(x, y) = (x\sqrt{1+x}, y\sqrt{1+3x})$$

defines a local nonlinear change of coordinates near 0, and  $f \circ \varphi^{-1}(x', y') = x'^2 - y'^2$ .

However, in using classical Morse theory, we may leave the space of solutions to the heat equation. Furthermore, the presence of  $t$  requires an extension to functions with parameters as introduced by Thom (Thom, 1972) in his catastrophe theory. However, there are several significant problems in trying to directly apply Morse theory to solutions to the heat equation. First, it is not clear that generic solutions to the heat equation must be generic in the Morse theory sense. Second, standard local models for Morse critical points and their annihilation and creation do not satisfy the heat equation. How must these models be modified? Third, there is the question of what constitutes generic behaviour. This depends upon what notion of local equivalence one uses between solutions to the heat equation.

We shall allow local (nonlinear) changes of coordinates in  $R^{n+1}$ , keeping track of two properties which reflect the special nature of the solutions: 1) the blurring parameter  $t$  as a distinguished variable, and 2) relative intensities of the critical points to understand the changes in the isointensity surfaces. Also, as the heat equation is invariant under translation, it is enough to define the notions of local equivalence near 0 for mappings  $f : R^{n+1} \rightarrow R$  sending  $f(0) = 0$ . We denote such a local mapping by  $f : R^{n+1}, 0 \rightarrow R, 0$ . Under local equivalence, we are only interested in  $f$  in some small neighbourhood of 0; then  $f$  is referred to as a *germ of a function (or mapping) at 0*. A local smooth mapping  $\varphi : R^{n+1}, 0 \rightarrow R^{n+1}, 0$  with nonsingular Jacobian matrix at 0 defines a local nonlinear change of coordinates (usually called a *local diffeomorphism* or a *germ of a diffeomorphism*).

Since we will be principally interested in the behaviour for  $t > 0$ , where the solutions to the heat equation are smooth, the local genericity results will be stated for smooth germs. Thus, *unless there is an explicit statement to the contrary, all mappings and germs will be smooth*.

For solutions to the heat equation,  $t$  has a distinguished role. We incorporate this distinguished role into our first notion of equivalence, called *H-equivalence* ("H" to remind us it preserves this feature for the heat equation). We will use local coordinates  $(x, t) = (x_1, x_2, \dots, x_n, t)$  for  $R^{n+1}$  and  $y$  for  $R$ .

**Definition 11.1.2.** Germs  $f, g : R^{n+1}, 0 \rightarrow R, 0$  will be called *H-equivalent* if there is a local change of coordinates  $\varphi : R^{n+1}, 0 \rightarrow R^{n+1}, 0$  of the form  $\varphi(x, t) = (\varphi_1(x, t), \varphi_2(t))$  with  $\varphi'_2(0) > 0$  and a germ  $c(t) : R, 0 \rightarrow R, 0$  so that

$$g(x, t) = f \circ \varphi(x, t) + c(t).$$

**Note :** By the form of  $\varphi$ , it follows that  $\varphi_1$  defines a local change of coordinates in  $R^n$  parametrised by  $t$ ; however,  $\varphi_2$  only depends on  $t$  and defines a change of coordinates for  $t$ . Thus, the  $t$ -coordinate of  $\varphi(x, t)$  only depends on  $t$ , preserving its distinguished role. Finally we saw in the Morse lemma that a constant  $c = f(x_0)$  enters into (11.1); as  $t$  varies; we allow the constant to vary with  $t$ . This form of equivalence does not keep track of the intensity level.

*Example 11.1.3.* The examples in (11.1.1), viewed as functions of  $(x, y, t)$  are  $H$ -equivalent to each other and also to a germ such as  $g(x, y, t) = 2x^2 - y^2 + t^2y^3 + 2t$  using

$$(x', y', t') = \varphi(x, y, t) = \left( \sqrt{2}x, y\sqrt{1-t^2}y, t \right) \quad \text{and} \quad c(t) = -2t$$

so that  $g \circ \varphi^{-1}(x', y', t') + c(t') = x'^2 - y'^2$ .

We can further refine this notion of equivalence to an “Intensity-Sensitive” equivalence which keeps track of both the local changes of an isointensity surface as it undergoes a transition and also the intensity level of that critical point. In so doing, we replace the addition of the constant  $c(t)$  depending on  $t$  with a  $t$ -dependent change of coordinates of  $R$  in the target (preserving the sense of orientation of  $R$ ).

**Definition 11.1.4.** We say two germs  $f, g : R^{n+1}, 0 \rightarrow R, 0$  are *IS-equivalent* if there is a local change of coordinates  $\varphi : R^{n+1}, 0 \rightarrow R^{n+1}, 0$  as in definition (11.1.2), but now we replace the addition of  $c(t)$  depending on  $t$  with a local change of coordinates  $\psi : R^2, 0 \rightarrow R^2, 0$  of the form  $\psi(y, t) = (\psi_1(y, t), t)$  with  $\frac{\partial \psi_1}{\partial y}(0, 0) > 0$  and  $\psi_1(0, t) = 0$  for all  $t$ , so that

$$g(x, t) = \psi_1(f \circ \varphi(x, t), t).$$

When we consider stability of the germs under deformations we wish to allow the target to move so we will add a nonzero  $c$ , depending on the parameters and which may be nonzero for nonzero parameter values.

## 11.2 Understanding Generic Properties

Underlying the notion of genericity and its consequences is the role played by a topologyfunction space topology on the space of functions. Initially we consider a function defined on some open set  $U \subseteq R^n$  for some  $n > 0$ . As a special case, we have the blurred functions defined on some open set  $U \subseteq R^n \times R_{++}$  where  $R_{++} = (0, \infty)$ . We consider a compact subset  $C \subseteq U$ , by which we mean a closed bounded subset of  $R^n$ . Important examples for blurred functions are “compact viewing areas”, which are subsets  $C \subseteq U$  of the form  $C = C_1 \times [\epsilon, R]$  where  $0 < \epsilon < R < \infty$  and  $C_1$  is a compact subset of  $R^n$ . Such a subset is a compact subset of  $U$  and any compact subset of  $U$  is contained in a finite union of such compact sets  $C$ . In the case  $U = R^n \times R_{++}$ , any compact subset is contained in a single “compact viewing area”.

If  $f, g \in C^k(C)$  are two  $C^k$  functions defined on such a compact set  $C$ , we can measure how close they are, using as a measure of distance the maximum of the (absolute) difference of their partial derivatives of order  $\leq k$  at all points of  $C$ .

$$d_k(f, g) = \max |D^\alpha(f)(x) - D^\alpha(g)(x)| \quad \text{taken over } |\alpha| \leq k, \text{ all } x \in C.$$

Here  $\alpha = (\alpha_1, \dots, \alpha_n)$  is a multi index,  $D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$ , and  $|\alpha| = \sum_{i=1}^n \alpha_i$ .

Then, a subset  $\mathcal{U} \subseteq C^k(C)$  is *open* in the usual metric sense if any  $f \in \mathcal{U}$  has an  $\epsilon > 0$  so that all  $g$  within  $\epsilon$  of  $f$  belong to  $\mathcal{U}$ . We can also speak of *dense* subsets  $A \subseteq C^k(C)$  which have the property that given any  $g \in C^k(C)$  and  $\epsilon > 0$ , there is an  $f \in A$  within distance  $\epsilon$  of  $g$ . Note that open dense subsets  $\mathcal{U} \subseteq C^k(C)$  capture the notions that interest us. Openness insures that if  $f \in \mathcal{U}$ , then there is an  $\epsilon > 0$  such that any perturbation  $g$  of  $f$  which has distance less than  $\epsilon$  to  $f$  will still belong to  $\mathcal{U}$ ; and by denseness, given any  $h$  and a distance  $\epsilon > 0$ , there is an  $f \in \mathcal{U}$  whose distance to  $h$  is less than  $\epsilon$ .

For example, suppose  $C$  is a compact subset of  $R^n$  (e.g. consisting of the set of  $x \in R^n$  of distance  $\leq R$  from 0). Suppose  $f$  has only a single critical point at 0 with nonsingular Hessian there. Then, it is possible to find  $\epsilon > 0$  so that any  $g$  whose derivatives of order  $\leq 2$  differ from those of  $f$  at each point of  $C$  by less than  $\epsilon$  will still have at most one critical point and at that point the Hessian will still be nonsingular (and in fact will have the same  $\delta_i$  in (11.2)). Hence, there is an open set about  $f$  in  $C^k(C)$  consisting of functions exhibiting the same kind of critical behaviour as  $f$ , i.e. the behaviour is stable (in fact, the statement of the stability can be made much more precise (Mather, 1969)).

Then, for a set of functions possessing a given property, denseness corresponds to genericity of the property and openness to stability. Actually denseness by itself is not enough, for the intersection of two dense sets can be empty. However, the intersection of a finite number of open dense sets is again open and dense. In fact,  $C^k(C)$  is a complete metric space and the Baire category theorem implies that the intersection of a countably infinite collection of open dense sets is still dense (although it may no longer be open). Such a set is called “residual”. It follows that the intersection of a countably infinite collection of residual sets is still residual.

This is the basis for genericity. However, there is an important point we have overlooked. Namely, in general we are considering smooth functions on  $U$ , so neither the compact set  $C$  nor the integer  $k$  are fixed. Instead we must allow both to vary. Then, instead of obtaining a measure of distance between functions, we obtain a topology made up from the open sets from all possible  $C$  and  $k$ . Such a topology is called the regular  $C^\infty$ -topology. Importantly, for this topology, the residual sets still dense. Also, if  $U \subseteq R^n \times R_{++}$ , we can consider the space

$$\mathcal{H}(U) = \{f \in C^\infty(U) : f \text{ is a solution to the heat equation on } U\}.$$

Then, residual subsets of  $\mathcal{H}(U)$  are still dense, a consequence of  $\mathcal{H}(U)$  being a Baire space, see e.g. (Damon, 1995a, prop. 5.5). This leads us to define genericity in general.

**Definition 11.2.1.** A local property of functions in  $C^\infty(U)$  (resp.  $\mathcal{H}(U)$ ) is generic if:

- i) the set of functions having the property at each point of  $U$  is residual; and
- ii) for any compact subset  $C \subset U$ , there is an open dense set of functions in  $C^\infty(U)$  (resp.  $\mathcal{H}(U)$ ) which have the property at each point of  $C$ .

**Remark :** Often genericity is stated only requiring i), but in fact, the stronger condition is usually true as in our case (see (Damon, 1996, §1)).

For example, the property of functions on  $U \subset R^n$  having only nondegenerate critical points is a generic property, and such functions are called *Morse functions*. What the generic properties are for  $\mathcal{H}(U)$  is less obvious. Before answering this, we conclude with a consequence for the initial intensity functions.

Convolution with the Gaussian kernel  $K_t(x)$  defines a continuous mapping

$$L^2(R^n, \mu) \longrightarrow C^\infty(R^n \times R_{++})$$

where  $L^2(R^n, \mu)$  denotes functions square-integrable over  $R^n$  for the exponentially decreasing measure  $\mu$  defined by  $e^{-\|x\|} \cdot dx$ . It is a Hilbert space and we can define genericity for it using the standard  $L^2$ -metric

$$\|f - g\|_{L^2} = \left( \int |f - g|^2 \cdot e^{-\|x\|} \cdot dx \right)^{1/2}$$

Significantly,  $L^2(R^n, \mu)$  contains polynomial functions, hence (by applying (Damon, 1995b, prop. 4.12)) we can conclude genericity for initial functions.

*Lemma 11.2.2. Genericity property for initial intensity functions :* if  $\mathcal{P}$  is a generic property for solutions to the heat equation on  $R^n \times R_{++}$  then initial intensity functions will generically have an associated blurred function with property  $\mathcal{P}$ .

### 11.3 Genericity via Stability

We describe the generic properties of solutions to the heat equation captured by *H* and *IS*-equivalence using the corresponding notions of stability (which, as we shall see in §12.7 is related to both transversality and versality). For these notions we also need to consider the equivalence of deformations. A *deformation* of a germ  $f_0 : R^{n+1}, 0 \rightarrow R, 0$  is a germ  $f : R^{n+1+q}, 0 \rightarrow R, 0$  such that  $f(x, t, 0) = f_0(x, t)$ . We let  $u$  denote local coordinates for  $R^q$ , which we can view as the parameters by which we deform  $f_0$ .

**Definition 11.3.1.** We say a germ  $f_0 : R^{n+1}, 0 \rightarrow R, 0$  is *H-stable* if for any deformation  $f$  there is a local change of coordinates  $\varphi : R^{n+1+q}, 0 \rightarrow R^{n+1+q}, 0$  of the form  $\varphi(x, t, u) = (\varphi_1(x, t, u), \varphi_2(t, u), u)$  with  $\frac{\partial \varphi_2}{\partial t}(0, 0) > 0$  and a germ  $c(t, u) : R^{1+q}, 0 \rightarrow R, 0$  so that

$$f_0(x, t) = f \circ \varphi(x, t, u) + c(t, u).$$

Likewise the germ  $f_0$  is called *IS-stable* if for any deformation  $f$  of  $f_0$ , there are local changes of coordinates  $\varphi$  as above and  $\psi : R^{2+q}, 0 \rightarrow R^{2+q}, 0$  of the form  $\psi(y, t, u) = (\psi_1(y, t, u), t, u)$  with  $\frac{\partial \psi_1}{\partial y}(0, 0, 0) > 0$  and  $\psi_1(0, t, u) = 0$  for all  $(t, u)$ , and a germ  $c : R^q, 0 \rightarrow R, 0$  so that

$$f_0(x, t) = \psi_1(f \circ \varphi(x, t, u), u, t) + c(u).$$

**Remark :** The form of stability used here is “stability under deformations”. We explain in §12.7 how it is equivalent via transversality to stability in the sense of topology that we mentioned earlier. In the parlance of singularity theory,  $f_0$  is *H-stable* (resp. *IS-stable*) if  $f_0$  is its own “versal unfolding”, i.e. every small perturbation of  $f_0$  is *H-equivalent* (resp. *IS-equivalent*) to  $f_0$ .

Let  $U$  be an open subset of  $R^n \times R_{++}$ . Then the basic theorem describing the local structure of solutions to the heat equation are given by the following.

**Theorem 11.3.2.** For mappings  $f \in \mathcal{H}(U)$ , being *H-stable* (resp. *IS-stable*) at each point of  $U$  is a generic property.

As a result of this theorem, it is natural to define  $f : U \rightarrow R$  to be *H-generic* (resp. *IS-generic*) if  $f$  is *H-stable* (resp. *IS-stable*) at each point of  $U$ . Then, by (11.2.2), we can conclude for initial intensity functions.

**Theorem 11.3.3.** Having a blurred intensity function  $f(x, t)$  which is *H-generic* or *IS-generic* is a generic property for initial intensity functions  $f_0 \in L^2(R^n, \mu)$ .

As a consequence of these theorems, we conclude

**Corollary 11.3.4.** For a compact viewing area  $C = C_1 \times [\varepsilon, R] \subset R^n \times R_{++}$ .

- There is an open dense set of solutions to the heat equation on  $R^n \times R_{++}$  which are *H-generic* (resp. *IS-generic*) on  $C$ .
- There is an open dense set of initial intensity functions  $f_0 \in L^2(R^n, \mu)$ , for which the blurred intensity function  $f$  is *H-generic* (resp. *IS-generic*) on  $C$ .

This leads us to ask whether we can determine the local properties of *H-stable* and *IS-stable* germs. We will do much more; namely, we shall give a list of local “normal forms” for such germs. For this we need further information about the solutions to the heat equation.

## 11.4 Weighted Homogeneity and the Heat Equation

To understand properties of solutions to the heat equation, we take a different tack from the usual approach of viewing solutions as arising via convolution with the

Gaussian kernel. Instead, we consider a weight decomposition of the space of polynomial germs in such a way that the heat operator preserves the decomposition. This allows us to determine the local properties of solutions to the heat equation by using the weighted terms in their Taylor expansion. We consider smooth solutions; since the heat equation is invariant under translation, we may assume we are considering germs at the origin.

Let the heat operator  $\Delta - \frac{\partial}{\partial t}$  be denoted by  $D$ . Also, we assign weights to  $(x, t)$  so that  $wt(x_i) = 1$  for all  $i$ , and  $wt(t) = 2$ . We use the standard notation for monomials  $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$  where  $\alpha = (\alpha_1, \dots, \alpha_n)$  and  $|\alpha| = \sum_{i=1}^n \alpha_i$ . Then, we define the weight of the monomial  $x^\alpha t^b$  to be  $wt(x^\alpha t^b) = |\alpha| + 2b$ . A polynomial is *weighted homogeneous of weight m* if the monomials appearing in it all have weight  $m$ . For example, the polynomial  $2x^5 - 7xy^2t + 6yt^2$  consists of monomials, each of weight 5, so it is weighted homogeneous of weight 5.

To decompose the space of polynomials according to weight, we let

$$W_k = \{\text{weighted homogeneous polynomials in } (x, t) \text{ of weight } k\}$$

and also let

$$\mathcal{P}_k = \{\text{homogeneous polynomials in } x \text{ of degree } k\}.$$

Then,  $W_k$  can be expressed as a direct sum

$$W_k = \mathcal{P}_k \oplus t \cdot \mathcal{P}_{k-2} \oplus t^2 \cdot \mathcal{P}_{k-4} \oplus \cdots \oplus t^{[k/2]} \cdot \mathcal{P}_{k-2[k/2]} \quad (11.3)$$

To describe the polynomial solutions to the heat equation we make an elementary observation which is key to the algebraic analysis.

*Lemma 11.4.1. Reduction to Properties on the Weighted Summands*

- i) If  $u : R^{n+1}, 0 \rightarrow R$  is a solution to the heat equation  $Du = 0$ , and has Taylor series  $\sum u_i$  with  $u_i \in W_i$  then,  $Du_i = 0$  for all  $i$ ;
- ii)  $D$  preserves the decomposition by the subspaces  $W_k$  and decreases weight by 2.

Thus, it suffices to examine the behaviour of  $D$  on  $W_k$ . We can give the form of a solution  $g \in W_k$  to  $Du = 0$  as follows. Let  $g_0(x) = g(x, 0)$ , then

$$g(x, t) = \exp(t\Delta)(g_0) = \sum_{j=0}^{[k/2]} \Delta^j(g_0) \left( \frac{t^j}{j!} \right) \quad (11.4)$$

Because the heat equation defines a flow on the space of functions, this is the standard method to integrate such a flow with initial condition provided  $\exp(t\Delta)(g_0)$  converges. However, more instructive is the way the algebraic decomposition (11.3) yields (11.4).

We see that in fig. 11.1,  $\frac{\partial}{\partial t}$  acts along the vertical arrow and  $\Delta$  along the horizontal arrow; each preserving the decomposition (11.3). Thus, if  $g = \sum g_j(x) \left( \frac{t^j}{j!} \right)$

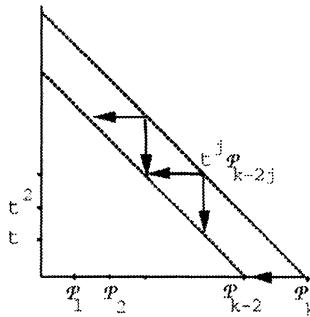


Figure 11.1: Decomposition of weighted homogeneous polynomials

with  $wt(g_j) = k - 2j$  and  $D(g) = 0$ , then

$$\Delta \left( g_j(x) \frac{t^j}{j!} \right) = \frac{\partial}{\partial t} \left( g_{j+1}(x) \frac{t^{j+1}}{(j+1)!} \right) \quad \text{or} \quad \Delta(g_j(x)) = g_{j+1}(x)$$

Conversely, we can begin with a polynomial  $g$  and add terms to it to obtain a solution to the heat equation. To see how this works, we recall a basic property of the Laplacian  $\Delta$  acting on polynomials in  $x$ . Let  $r^2 = \sum_{i=1}^n x_i^2$ ; then, we have the fundamental property (see e.g. (Folland, 1976, chap. 2)).

*Lemma 11.4.2.*  $\Delta$  viewed as an operator  $\Delta : \mathcal{P}_k \rightarrow \mathcal{P}_{k-2}$  has image all of  $\mathcal{P}_{k-2}$  and induces a vector space isomorphism  $\Delta : r^2 \cdot \mathcal{P}_{k-2} \xrightarrow{\sim} \mathcal{P}_{k-2}$ .

We denote the space of harmonic polynomials of degree  $k$  by

$$H_k = \{g \in \mathcal{P}_k : \Delta(g) = 0\}.$$

A number of properties of harmonic polynomials are given in (Folland, 1976, chap. 2). Then, (11.4.2) yields the decomposition.

*Lemma 11.4.3.* *Decomposition of  $\mathcal{P}_k$ :*

- i)  $\mathcal{P}_k = H_k \oplus r^2 \cdot \mathcal{P}_{k-2}$  ; and hence,
- ii)  $\mathcal{P}_k = H_k \oplus r^2 \cdot H_{k-2} \oplus r^4 \cdot H_{k-4} \oplus \cdots \oplus r^{2[k/2]} \cdot H_{k-2[k/2]}$  .

For weight  $k$  the harmonic polynomials  $H_k$  are already solutions to the heat equation. Using the decomposition (11.4.3), we construct a basis for the complementary subspace of solutions using the harmonic polynomials. We begin by defining an operation for generating solutions starting with a monomial  $x^\alpha \frac{t^j}{j!}$  and adding to it terms of the same weight. Then, the usual role played by monomials for giving “normal forms” in singularity theory will be played instead by these canonical solutions. The precise way this construction is carried out is given by the following lemma.

**Lemma 11.4.4.** There is a unique linear operator  $E$  on the space of polynomials  $R[\mathbf{x}, t]$  which preserves weights and satisfies:

i)  $DE(g) = 0$  for  $g \in R[\mathbf{x}, t]$  ;

ii)

$$E\left(x^\alpha \frac{t^j}{j!}\right) = \sum_{i=0}^{[\lvert \alpha \rvert / 2]} \Delta^i(x^\alpha) \left( \frac{t^{j+i}}{(j+i)!} \right) \bmod \left( \sum_{m,\ell} t^{j-m} \cdot r^{2\ell} \cdot H_{\lvert \alpha \rvert + 2m - 2\ell} \right)$$

**Remark:** If  $g(x) \in W_k$  then  $E(g)$  is given by (11.4). Also, if  $g$  is a polynomial with at most one nonzero monomial of any given weight, then for any nonzero monomial of  $g$ , it follows that the monomial appears in  $E(g)$  with the same coefficient.

Generally to see what form  $E(g)$  takes, it is enough by the linearity of  $E$  to see how it is defined for monomials  $x^\alpha \cdot \frac{t^{j+i}}{(j+i)!}$ . For this monomial, (11.4) shows us what terms with higher powers of  $t$  to add. For lower powers of  $t$  we must solve  $\Delta^k(h) = x^\alpha$ . However, by (11.4.3) there is a unique right inverse to  $\Delta$  defined as

$$\Delta^{-1} : \mathcal{P}_{k-2} \xrightarrow{\sim} r^2 \cdot \mathcal{P}_{k-2} \quad (11.5)$$

**Note :** this inverse is not multiplication by  $r^2$ . However, it is uniquely defined. To apply the operator  $E$ , we can take advantage of the especially simple form for the action of  $\Delta$  on  $r^{2j} \cdot p$  with  $p \in H_k$

$$\Delta(r^{2j} \cdot p) = c(j, k)r^{2j-2} \cdot p \quad \text{where } c(j, k) = 2j(n+2(k+j-1)) \quad (11.6)$$

**Example 11.4.5.** Then, to apply  $E$  to  $t^2$  in the case  $n = 2$ , we use (11.6) to compute  $\Delta^{-1}(1) = (1/4)r^2$  and  $\Delta^{-1}(r^2) = (1/16)r^4$ . Thus, by Lemma 11.4.4 and 11.4

$$E(t^2) = t^2 + 2(1/4)r^2t + 2(1/4)(1/16)r^4 = t^2 + (1/2)r^2t + (1/32)r^4$$

Also, by (11.4.3) and (11.6) it follows that a basis for the subspace of solutions in  $W_m$  to the heat equation is given by

$$\{E(t^\ell \cdot p_k) : p_k \text{ belongs to a basis for } H_k; 2\ell + k = m \text{ and } \ell, k \geq 0\} \quad (11.7)$$

Alternately,  $E(t^\ell \cdot p_k)$  can be replaced by  $E(r^{2\ell} \cdot p_k)$  in (11.7).

**Example 11.4.6.** Thus, a basis for  $W_k$  can be constructed using (11.7), beginning with the harmonic polynomials  $H_k$ . For example, for the lowest weights the generators for solutions on  $R^3$  with coordinates  $(x, y, t)$  are given in Table 11.1.

Having provided a weight decomposition of the space of polynomial solutions to the heat equation, we can use the weighted homogeneous solutions as components in the normal forms used for describing the generic behaviour of solutions.

wt k	$H_k$	remaining basis
0	1	
1	$x, y$	
2	$x^2 - y^2, xy$	$E(t) = t + (1/4) \cdot r^2$
3	$x^3 - 3y^2x, y^3 - 3x^2y$	$x^3 + 6tx, y^3 + 6ty$ or e.g. $E(xt) = xt + (1/8) \cdot xr^2, E(yt)$
4	$x^3y - xy^3, x^4 + y^4 - 6x^2y^2$	$x^4 + 12tx^2 + 12t^2, xy^3 + 6txy$ $x^2y^2 + 2tr^2 + 4t^2$ , or e.g., $E((x^2 - y^2)t), E(xyt)$ , and $E(t^2) = t^2 + (1/2) \cdot tr^2 + (1/32) \cdot r^4$

Table 11.1: Some basis functions for  $W_k$ 

## 11.5 Generic Properties of Blurred Images

**Theorem 11.5.1.** The list I is a complete list, up to the corresponding equivalence , for the sets of H-stable, respectively IS-stable, germs  $R^{n+1}, 0 \rightarrow R, 0$  subject to the following conditions:

- Two normal forms within the same case are  $H$ -equivalent provided the quadratic forms  $\sum a_i x_i^2$  have the same index.
- For  $IS$ -equivalence the normal forms must satisfy the condition  $c \neq 0$ . Then, two normal forms in the same case are  $IS$ -equivalent provided the quadratic forms  $\sum a_i x_i^2$  have the same index and  $c$  have the same sign.

**Remark :** As a consequence of the appropriate theorems from singularity theory, it follows that for the germs on either of the lists,  $E(g)$  may be replaced by  $g$  without changing the  $H$ -equivalence or  $IS$ -equivalence class or the stability (although the modified germ will then fail to satisfy the heat equation ).

**List I: H-stable and IS-stable germs** For each of the following classes, all  $a_i \neq 0$ .

case	normal form	conditions/description
0)	$x_1$	submersion
1)	$\sum_{i=1}^n a_i x_i^2 + 2ct$	$c = \sum_{i=1}^n a_i$
1')	$\sum_{i=1}^n a_i x_i^2 + \varepsilon E(t^2)$	$\sum_{i=1}^n a_i = 0, \quad \varepsilon = \pm 1$
2)	$x_1^3 + 6tx_1 + \sum_{i=2}^n a_i x_i^2 + 2ct$	$c = \sum_{i=2}^n a_i$
3)	$x_1^3 - 6tx_1 - 6x_1 x_2^2 + \sum_{i=2}^n a_i x_i^2 + 2ct$	$c = \sum_{i=2}^n a_i$

**Remark :** It follows from the classification (Damon, 1996) (but we do not address it here) that there are specific criteria in terms of derivatives of

order  $\leq 3$  for identifying the specific local form. Furthermore, the especially nice form we obtain is a direct result of the weighted homogeneity. However, surprisingly, if we turn instead to nonlinear blurring, e.g. using an equation of Perona–Malik type (Perona and Malik, 1990), then it is still possible to identify the local generic behaviour, now using results which apply to more general partial differential equations (Damon, 1996). Now, we cannot expect the local normal forms to be polynomial as above; but, somewhat unexpectedly, we can obtain normal forms by adding higher order terms to some of the local forms already given. Because this does not change the germ up to equivalence, it follows that we can still identify the generic behaviour that occurs (Damon, 1996, §6).

*Example 11.5.2.* We saw that the germs  $x^2 - y^2$ ,  $x^3 - 3xy^2 + x^2 - y^2$  and  $2x^2 - y^2 + t^2y^3 + 2t$  are all  $H$ –equivalent. Thus, the structure of the level sets near the critical points are the same, even taking into account the special role of  $t$ . However, for  $IS$ –equivalence, the first two are not equivalent to the third. For them, the intensity level of the critical point remains constant for all  $t$  (a highly unstable situation); while for the third it varies with  $t$ . In fact, the first two have “infinite codimension” for  $IS$ –equivalence. This essentially means that there are an infinite number of different ways to deform these germs so that the isointensity surfaces undergo distinct changes.

This notion of codimension enters when we consider problems in which other parameters are present. The codimension of a germ is the number of parameters that are needed to describe all possible behaviour that can occur in small perturbations. Given such a germ  $f_0(x, t)$  of codimension  $q$ , there is a well-defined way to construct a  $q$ –parameter unfolding  $f(x, t, v)$  (i.e. deformation) of  $f_0$  which, in a very well-defined sense, contains all possible small perturbations of  $f_0$ . Such unfoldings are called *versal unfoldings*, and those involving  $\leq q$  parameters determine which transitions occur generically in  $q$ –parametrised families.

For example, generic one–parameter families of solutions to the heat equation are  $IS$ –equivalent to versal unfoldings of germs of  $IS$ –codimension 1, where the extra parameter is denoted by  $v_1$ . The versal unfoldings are given below.

#### List II: Versal Unfoldings for $IS$ –codimension 1 germs

For all of the following  $\sum a_i = 0$ , all  $a_i \neq 0, \varepsilon = \pm 1$ ;

- 2)  $x_1^3 + 6tx_1 + \sum_{i=2}^n a_i x_i^2 + \varepsilon E(t^2 + v_1 t);$
- 3)  $x_1^3 - 6tx_1 - 6x_1 x_2^2 + \sum_{i=2}^n a_i x_i^2 + \varepsilon E(t^2 + v_1 t);$
- 4)  $\sum_{i=1}^n a_i x_i^2 + \varepsilon E(t^3 + v_1 t);$

Here  $v_1$  is treated as a constant by the operator  $E$ .

## 11.6 Consequences of the Generic Properties

Let’s examine some of the consequences of the classification of stable germs for the generic behaviour of solutions to the heat equation. We are especially interested in

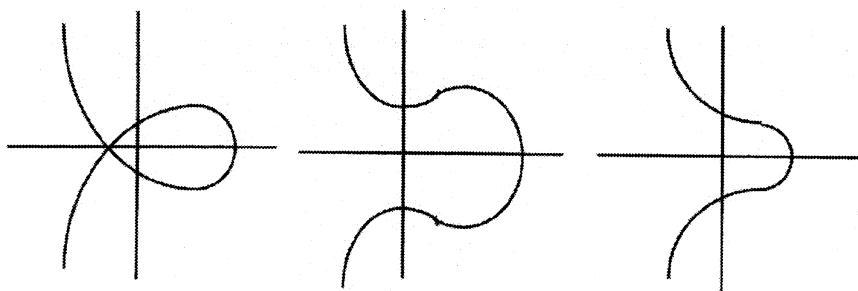


Figure 11.2: Isointensity curve evolution in Morse theory

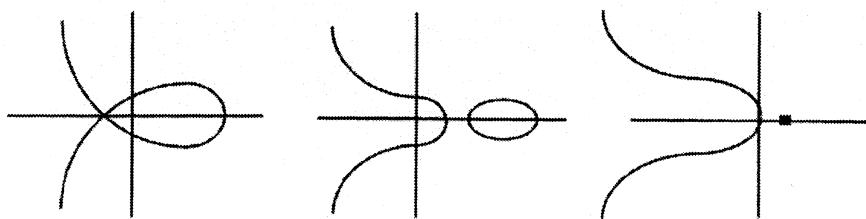


Figure 11.3: The corresponding (11.2) evolution in the Heat Equation

how regions of given intensity interact in the course of the blurring. For example,  $f(x, t)$  represents the amount of heat at the point  $x$  at time  $t$ . Since heat diffuses, we would expect that the region where the heat content locally exceeded a given amount would expand so that: if two different regions joined they should merge and further expand; and moreover, no new regions of relative higher intensity would be created. In fact, both of these intuitive expectations are false; however, we can use the classification to determine exactly what happens generically.

*Example 11.6.1. Annihilation of critical points:*  $x^3 + 6tx + y^2 + 2t$  is of type 2) with  $n = 2$ . It is the example considered by Lifshitz (Lifshitz, 1987) for the annihilation of a pair of critical points. Both critical points must approach the annihilation intensity level from the same side. In Morse theory the critical points may approach from opposite sides so that at a near-by level the isointensity curves evolve as in fig. 11.2.

However, this cannot happen for solutions to the heat equation. Instead the curve surrounding the local maximum must break off into a separate component and then this component shrinks to a point and disappears, as in fig. 11.3.

*Example 11.6.2. Creation of critical points:*  $x^3 - 6tx - 6xy^2 + y^2 + 2t$  is of type 3) with  $n = 2$ . It was expected based on the reasoning explained earlier that for solutions to the heat equation, no critical points, especially local maxima, would be created. A heuristic argument for the possibility of critical points being created in certain circumstances was given by Steve Pizer and John Gauch (Pizer et al., 1988). This example shows that such a creation of critical points, including local

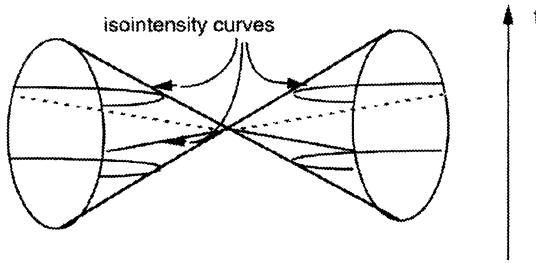


Figure 11.4: The merging and splitting of regions for the germ  $x^2 - y^2 + E(t^2)$

maxima, not only can occur but, in fact, does so generically for solutions to the heat equation.

*Example 11.6.3. Merging and separating of regions:*  $x^2 - y^2 + E(t^2)$  is of type 1') with  $n = 2$ . Since isointensity curves separate regions of greater intensity from those of lower intensity, one would expect that if two regions bounded by isointensity curves of the same intensity level join then they will continue to grow together. In fact, this example shows that two such regions can join and then break apart again. By example (11.4.5), this germ is given by the equation

$$x^2 - y^2 + t^2 + (1/2)t(x^2 + y^2) + (1/32)(x^2 + y^2)^2$$

In fact, this germ is *IS*-equivalent to the simpler germ  $g(x, y, t) = x^2 - y^2 + t^2$ . Hence, using this simpler germ, we can see in fig. 11.4 that the level curves for the intensity level 0 join together and then break apart again.

This merging and separating of regions can become even more complicated for higher codimension germs which appear generically in families, even involving codimension 1 germs.

*Example 11.6.4. Merging and separating of regions continued:*  $x^2 - y^2 + \varepsilon E(t^3 + v_1 t)$  is of type 4) with  $n = 2$  (note that the conditions on the  $a_i$  prevent all but 4) from occurring for  $n = 2$ ). This exhibits even more complicated behaviour regarding the merging and separating of regions. Again, it is *IS*-equivalent as an unfolding to  $x^2 - y^2 + \varepsilon(t^3 + v_1 t)$ . When  $v_1 < 0$  with  $\varepsilon = 1$ , the region of intensity  $> 0$  for  $x^2 - y^2 + t^3 + v_1 t$  undergoes a series of separations and mergers. With the intervals as indicated in the graph of  $t^3 + v_1 t$  in fig. 11.5, we see that in interval *A* there is a single region as indicated. In interval *B* the region separates into two parts. These regions rejoin in *C*; only to undergo a final separation in *D*.

## 11.7 Genericity via Transversality and Versality

We conclude this chapter by explaining how genericity is established and normal forms are obtained by: using transversality and deducing local stability via the “unfolding” theorem, and applying the “finite determinacy theorem” which together with the classification yields the normal forms as a consequence. These

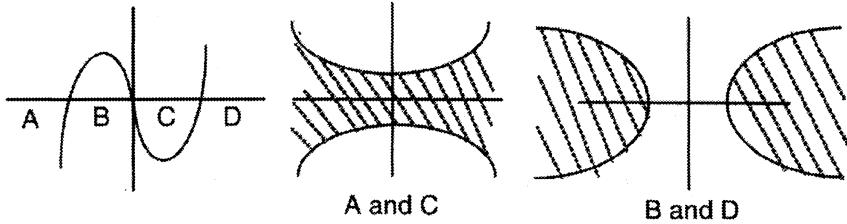


Figure 11.5: The merging and splitting of regions for the germ  $x^2 - y^2 + \varepsilon E(t^3 + v_1 t)$

results depend upon extensions of the Thom transversality theorem and of the method of Mather originally introduced for another equivalence called  $\mathcal{A}$ -equivalence (Mather, 1970).

For an open subset  $U \subset R^n \times R_{++}$ , the jet space  $J^\ell(U, R) = U \times J^\ell(n+1, 1) \times R$ , where  $J^\ell(n+1, 1)$  consists of the set of  $\ell$ -th order Taylor expansions without constant term at a fixed point. For a smooth function,  $f : U \rightarrow R$ , there is the “jet extension map”  $j^\ell(f) : U \rightarrow J^\ell(U, R)$  sending  $(x, t) \mapsto j^\ell(f)(x, t) = \ell$ -th order Taylor expansion of  $f$  at  $(x, t)$ . Singular behaviour of functions can be described by conditions given in terms of smooth submanifolds of  $W \subset J^\ell(U, R)$ . For example, the *IS*-stable points of type 1) (in List I) are points  $((x, t), f, y) \in J^2(U, R)$  at which  $\nabla_x f(x) = 0$ ,  $\frac{\partial f}{\partial t}(x, t) \neq 0$ , and the Hessian  $H_x(f)$  is nonsingular (here ‘ $\nabla_x$ ’ and  $H_x$  denote the operations with respect to the  $x$ -coordinates).

*Transversality of  $j^\ell(f)$  to  $W$*  means that at any point  $x \in U$ , either  $j^\ell(f)(x) \notin W$ , or the tangent space to  $W$  at  $j^\ell(f)(x)$  together with the image of the tangent space  $T_x U$  under the derivative  $dj^\ell(f)(x)$  span the tangent space of  $J^\ell(U, R)$  at  $j^\ell(f)(x)$ . If  $j^\ell(f)$  is transverse to  $W$ , then it follows that the set of points  $(x, t)$  where  $f$  exhibits the singular behaviour described by  $W$  is a smooth submanifold of the same codimension as  $W$  (or is empty, as it must be if  $\text{codim}(W) > n+1$ ). Thus, transversality ensures that singular behaviour is exhibited in an especially regular way and guarantees that behaviour of codimension  $> n+1$  does not occur.

The Thom transversality theorem describes how such singular behaviour for  $f \in C^\infty(U, R)$  should generically occur in terms of the transversality of  $j^\ell(f)$  to submanifolds describing this behaviour (see e.g. (Golubitsky and Guillemin, 1974, Chap. 2)). Because the subset of functions guaranteed to be transverse to a given  $W$  may entirely miss the space of solutions to the heat equation  $\mathcal{H}(U)$ , the Thom transversality theorem may provide no information. However, there is an extension of the theorem where in place of  $J^\ell(U, R)$  we use  $\mathcal{H}^\ell(U)$ , consisting of  $\ell$ -jets of solutions to the heat equation. Then, for  $f \in \mathcal{H}(U)$ , the  $\ell$ -jet extension map factors  $j^\ell(f) : U \rightarrow \mathcal{H}^\ell(U)$ . Furthermore, the  $W$  which we consider satisfy a certain type of Whitney regularity condition for the manifolds in their closure. By  $f$  being *completely transverse* to  $W$ , we shall mean that it is also transverse to these submanifolds in the closure. Then, the version of the Thom transversality theorem for the space  $\mathcal{H}(U)$  becomes (see (Damon, 1995a) together with (Damon, 1996)).

**Theorem 11.7.1.** Suppose that  $W$  is a smooth submanifold of  $\mathcal{H}^\ell(U)$ , then the set of mappings  $f \in \mathcal{H}(U)$  for which  $j^\ell(f)$  is completely transverse to  $W$  on a compact subset  $C \subset U$  is an open dense set for the regular  $C^\infty$ -topology

Now, we can identify the relevant submanifolds. We know that the manifolds  $\mathcal{O}_i$  in  $J^\ell(U, R)$  which describe  $H$  and  $IS$ -classes in List I are not contained in  $\mathcal{H}(U)$ . Nonetheless, they are transverse to  $\mathcal{H}^\ell(U)$  so that their intersection with  $\mathcal{H}^\ell(U)$  are smooth submanifolds of  $\mathcal{H}^\ell(U)$  (this crucially uses the weighted homogeneous decomposition and the classification). We decompose  $\mathcal{H}^3(U)$  into the submanifolds coming from the  $\mathcal{O}_i$  together with a finite union of submanifolds  $\Gamma_i$  of codimension  $> n+1$ . Then, by Theorem 11.7.1 and the Thom transversality theorem, the set of mappings in either  $\mathcal{H}(U)$  or  $C^\infty(U, R)$  that are completely transverse on a compact subset  $C \subset U$  to all of the  $\mathcal{O}_i$  and  $\Gamma_i$  form an open dense set. These functions will miss the  $\Gamma_i$  and only intersect the  $\mathcal{O}_i$ , i.e. only exhibit the type of behaviour captured by transverse intersection with  $\mathcal{O}_i$ .

Thus, we conclude: the genericity of the transverse appearance of certain type of singular behaviour is guaranteed by the appropriate transversality conditions. There remains the question of how this translates into specific normal forms. Here we use a line of reasoning similar to that used by Mather.

The tangent spaces appearing in the transversality conditions have an algebraic structure far more refined than just that of being vector spaces (they are “modules over systems of rings”). Because of this sophisticated algebraic structure, arguments using the Malgrange preparation theorem “almost miraculously” allow us to transform the transversality condition into an infinitesimal condition for stability under deformations,

$$T\mathcal{G}_e \cdot f = C^\infty(n+1, 1) \quad (11.8)$$

where  $T\mathcal{G}_e \cdot f$  denotes the “deformation theoretic tangent space” to the orbit of  $f$  for  $\mathcal{G} = H$  or  $IS$ -equivalence, and  $C^\infty(n+1, 1)$  denotes the space of germs  $R^{n+1}, 0 \rightarrow R$ . Then, by the unfolding theorem in singularity theory, which holds for quite general notions of equivalence (see (Damon, 1984)), this infinitesimal condition (11.8) actually implies that the germ  $f$  is stable under deformations for the appropriate equivalence (i.e. it is its own versal unfolding). Third, this condition implies by the finite determinacy theorem, yet another theorem from singularity theory (see (Damon, 1984) again generalising another important theorem of Mather (Mather, 1968)), that such a stable germ is equivalent to a finite part of its Taylor expansion. The exact part depends upon additional arguments and the specific germ; but it yields one of the normal forms given earlier.

## 11.8 Conclusion

Hence, via this line of reasoning, we have first established the genericity of properties through transversality, and then transformed the transversality conditions first into stability conditions, then into versality conditions, and finally into specific local normal forms exhibiting the behaviour. Because the arguments are so

general they apply to many other properties, situations, etc and have only just begun to fulfill their potential use in computer imaging.

# Chapter 12

## Critical Point Events in Affine Scale-Space

Lewis Griffin

INRIA-SOPHIA

2004, Route des Lucioles, B.P. 93, Sophia Antipolis Cedex 06902, France

### 12.1 Motivation for Affine Scale Space

There are (at least) three lines of argument motivating the interest in scale space: ecological, physiological and pragmatic. The **ecological** line runs thus. It is desirable for an agent that the environmental information it can extract by visual means should depend as little as possible upon the agent's location and pose. This is confounded by several factors: occlusion, limited field of view and physical limitations on the measurements it can make of the luminance incident upon its receptors. Consideration of physical limitations inspires scale space (Koenderink, 1984). It is acknowledged that the apertures of physically plausible measurements must be of non-zero size. Hence there will be an unavoidable loss of information as the agent recedes from a scene. If the agent produces a scale space (i.e. it measures the luminance with apertures at all sizes not smaller than its smallest set) then it achieves a semi-invariance to viewing distance: it ensures that *although it loses information by receding, it never loses by approaching*. The **physiological** line is to observe that receptive fields in mammalian visual systems have a range of sizes; to realize that this means that the retinal image is represented at multiple scales; and then to seek a formal framework to describe and investigate such a representation (Marr, 1982). The **pragmatic** line is the computer scientist's. He observes that image objects come in a range of sizes and that operations that can extract small objects fail with larger objects and *vice versa*. Rather than design his operators separately for each scale of interest he conceives the plan of using

one set of operators and changing the scale of his image (Witkin, 1983).

Affine scale space can also be motivated along these three lines. The **ecological** line is to continue the quest for invariance to viewpoint but to supplement the requirement of invariance to viewing distance with the requirement of invariance to viewing angle (known as slant). Again, physical limitations on the measurement apertures prevent this goal being attained, but just as scale space uses larger apertures to deduce what would be seen at greater distance, affine scale space uses elliptical apertures of varying size, eccentricity and orientation to deduce what would be seen at more extreme distances and viewing angles. For an agent that computes an affine scale space, *although it loses information when receding or when the slant of a surface is increased, it loses nothing when approaching or when the slant is decreased*. The **physiological** line is to note that not only are receptive fields of different sizes but from area V1 onwards they are predominantly elliptical rather than circular with varying eccentricities and orientations (Jones and Palmer, 1987). Again, a framework is sought to describe such a representation (Daugman, 1990). The **pragmatic** line is to observe that image structures can have different scales in different directions (Lindeberg, 1994e). Examples are not restricted to slanted textured surfaces, for instance printed text: lines of text can be extracted as distinct entities far more effectively using anisotropic, rather than isotropic, smoothing.

## 12.2 A Representation of Affine Scale Space

An affine scale space is realized by convolving the incident luminance with apertures having the form of elliptical Gaussians. For the present analysis the following representation of the family of apertures is convenient:

$$G_{i,j,k}(x,y) = \frac{\sqrt{2ik - j^2}}{4\pi ik} e^{-\frac{kx^2 + \sqrt{2}jxy - iy^2}{2\sqrt{2ik}}}, \quad i, k \geq 0, \quad 2ik \geq j^2$$

The affine scale space ( $L_{i,j,k}$ ) is defined as the result of convolution of the apertures with the incident luminance ( $\mathcal{L}$ ):  $L_{i,j,k} = G_{i,j,k} \otimes \mathcal{L}$ . This representation is only valid in a restricted portion of  $\mathbb{R}^3$  (figure 12.1 left) i.e. the boundary and interior of a (single mantle) right cone. Images on the boundary of the cone are generated by slit-like apertures, those within the cone by fully 2-D apertures. Only along the cone axis are the apertures isotropic; the images along this line form an ordinary scale space. Figure 12.1 middle shows three ways that affine scale space can be foliated: by constant aperture orientation (planes containing the cone axis); by constant aperture aspect ratio (cones sharing axis and origin with the full space); and by constant aperture area (one sheet of a hyperboloid of two sheets).

The use of Gaussian apertures for ordinary scale space allows characterization of the space by an equation linking spatial and scale derivatives i.e.  $\frac{\partial L}{\partial s} = \nabla^2 L$ . A

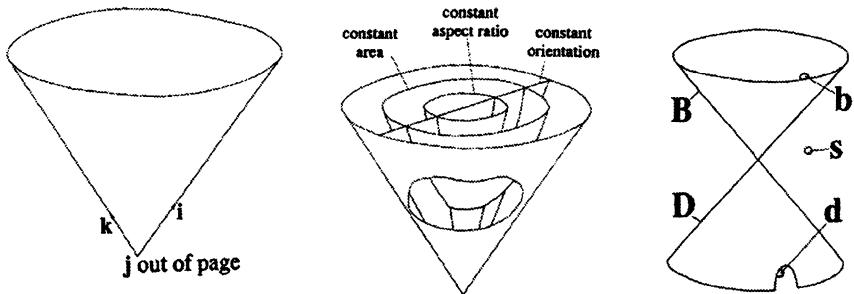


Figure 12.1: Global and Local Representations of Affine Scale Space

similar relationship holds for affine scale space:

$$\vec{Q} = \frac{2ik}{(2ik - j^2)^2} \begin{pmatrix} \frac{1}{\sqrt{2}} \frac{\partial^2 L}{\partial x^2} \\ \frac{\partial^2 L}{\partial x \partial y} \\ \frac{1}{\sqrt{2}} \frac{\partial^2 L}{\partial y^2} \end{pmatrix}$$

$$\frac{\partial L}{\partial i} = \begin{pmatrix} 2(ik - j^2) \\ -j^3/i \\ -j^2k/i \end{pmatrix} \cdot \vec{Q}$$

$$\frac{\partial L}{\partial j} = \begin{pmatrix} ij \\ 2ik + j^2 \\ jk \end{pmatrix} \cdot \vec{Q}$$

$$\frac{\partial L}{\partial k} = \begin{pmatrix} -ij^2/k \\ -j^3/k \\ 2(ik - j^2) \end{pmatrix} \cdot \vec{Q} \quad (12.1)$$

If only a single image is being considered then by choice of gauge coordinates it may be brought about that  $j = 0$ , which simplifies the equations. I will write

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial i} \\ \frac{\partial}{\partial j} \\ \frac{\partial}{\partial k} \end{pmatrix} \text{ and } \square = \begin{pmatrix} \frac{1}{\sqrt{2}} \frac{\partial^2}{\partial x^2} \\ \frac{\partial^2}{\partial x \partial y} \\ \frac{1}{\sqrt{2}} \frac{\partial^2}{\partial y^2} \end{pmatrix}.$$

So, assuming the gauge condition is satisfied, affine scale space may be characterized by:  $\nabla L = \square L$ . The gauge coordinates work by rotating the  $x, y$ -system causing the  $i, j, k$ -system to rotate around the cone axis. The rotation which makes  $j = 0$  at the image of interest is such that the new  $x$  and  $y$  axes are aligned with the axes of the aperture associated with that image.

Isotropic Gaussians have a concatenation property: i.e. convolving two together produces a third ( $G_s \otimes G_t = G_{s+t}$ ); elliptical Gaussians have the same property

but the expression for the concatenation is complicated. In an ordinary scale space this property has the consequence that for any two images in the space one of them can be produced from the other by convolution and it is simple to decide which can be derived from which: the coarser scale can be derived from the finer scale but not *vice versa*. In affine scale space it is no longer true that any two images are related by a convolution. So for a given image  $(i', j', k')$ , the space can be divided into three regions: images reachable by blurring, images from which the target can be reached by blurring, and the rest. The surface which separates these regions is a complicated warped double cone but in an infinitesimal neighbourhood of the target image it is simple (figure 12.1 right). The neighbourhood is divided by a right cone of two mantles with axis parallel to that of the full space i.e.  $2(i - i')(k - k') = (j - j')^2$ . Images within the upper mantle ( $b$  for blur) can be reached by 2-D blurring; images on the boundary of the upper mantle ( $B$ ) can be reached by a 1-D blur. From images within the lower mantle ( $d$  for de-blur) the target image can be produced by a 2-D blur; from images on the boundary of the lower mantle ( $D$ ) by a 1-D blur. Images outside the cone mantles ( $s$ ) can neither be reached by blurring nor can the target image be reached from them by blurring. In terms of the ecological motivation of affine scale space: sub-space  $b$  consists of images that would be seen at greater viewing distance and slant;  $B$  at greater slant;  $d$  at lesser viewing distance and slant, and  $D$  at less slant.

### 12.3 Local Structure

Almost all images of an affine scale space are Morse functions. This means that the local structure is highly restricted. Almost all points of a Morse function have non-zero gradient (slope points) and so the pattern of isophotes in the neighbourhood is diffeomorphic to parallel lines (figure 12.2 first column). Diffeomorphic is an equivalence relation that holds between two local structures if the domain of either can be smoothly warped so that the two patterns of isophotes are identical. The non-slope points of a Morse function are critical points (zero gradient), they will be isolated and have the form of an extremum or a saddle point. They will always be diffeomorphic to the symmetrical canonical examples shown in the second column of figure 12.2.

At *versal* (i.e. non-generic) images of an affine scale space other local structures will occur as a consequence of extrema and saddle points merging in various combinations. To produce a catalogue of these critical point events it would be nice to make use of Thom's theorem (Gilmore, 1981) which gives exactly that for a generic continuous space of continuous functions (as long as the space of functions has dimension five or less). Unfortunately an affine scale space is not a generic space of functions (it satisfies  $\nabla L = \square L$ ). Counter-examples show the danger of applying Thom's theorem without satisfying its conditions. For example, at one extreme, a space of constant functions will never display any catastrophes. At the other extreme the gradient magnitude of an ordinary scale space displays catastrophes that it ought not to according to a misapplication of Thom's theorem (i.e. a Hopf bifurcation (Griffin and Colchester, 1995)). Damon (Chapter 11) has

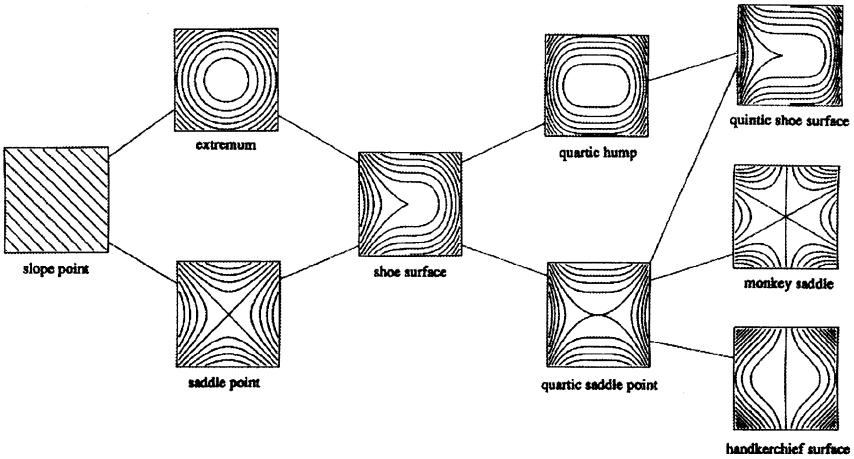


Figure 12.2: Local point types encountered in an affine scale space

shown that in the case of ordinary scale space the list of catastrophes suggested by a misapplication of Thom's theorem is identical with the catastrophes that do occur. In this paper I will show that all the catastrophes that Thom's theorem suggests can occur in affine scale space actually do. I hypothesize that this list is in fact complete.

Thom's theorem presents catastrophes as a local germ and an unfolding. The germ is the local structure at the point of the catastrophe and the unfolding describes the local structures in a neighbourhood of the function space. For example in a 1-D family of functions we can encounter the shoe surface (figure 12.2 column 3) which means that there will be local structure diffeomorphic to  $u^3 + v^2$ . This is a critical point but it is not diffeomorphic to either the saddle point or extremum. This is apparent because the cusp that occurs in the isophote through the point can never be removed by smooth deformation of the domain. The unfolding is given by  $pu + u^3 + v^2$  where the control parameter  $p$  is a smooth function of the indexing parameter of the space. Several points need to be made to clarify this. Firstly, the diffeomorphism that makes the function look like the canonical example of the unfolding will change (smoothly) with  $p$ . Secondly, the extent of validity is only some open neighbourhood not only of  $x$  and  $y$  but also of the indexing parameter. Thirdly, a constant term (varying with  $p$ ) will also normally be needed.

The fold catastrophe can occur in a 1-D family of images. Affine scale space is a 3-D family and therefore a given fold catastrophe will occur in every image in a 2-D manifold of images. That is to say there will be a fold surface in  $i, j, k$ -space, every image lying in this fold surface will contain a shoe surface and the mapping of the unfolding onto the actual affine space will vary smoothly across the fold surface. The next local structures in terms of rarity (figure 12.2 fourth column) are the quartic hump ( $u^4 + v^2$ ) and the quartic saddle ( $u^4 - v^2$ ). These are both germs of

the cusp catastrophe which occurs in 2-D families of images. In affine scale space it will occur as a curve in  $i, j, k$ -space, every image on the curve experiencing the catastrophe. Finally, there are three germs that occur in 3-D spaces of functions (the fifth column of figure 12.2); as such they will occur at isolated images of an affine scale space. They are the quintic shoe surface ( $u^5 + v^2$ ) associated with the swallowtail catastrophe; the monkey saddle ( $u^3 - 3uv^2$ ) associated with the elliptic umbilic catastrophe; and the handkerchief surface ( $u^3 + 3uv^2$ ) associated with the hyperbolic umbilic catastrophe.

In the remainder of this chapter I describe the catastrophes in more detail and address two questions: which of these catastrophes can occur and what restrictions are there on their embeddings in affine scale space? To clarify the second question, an example of a restriction on the embedding of a catastrophe is that for an ordinary scale space of 1-D images, the control parameter  $p$  of the fold catastrophe always runs in the direction of increasing scale. In affine scale space the significant questions about the embeddings of the catastrophes do not refer to the  $i, j$  and  $k$  directions but to the double cone diagram on the right of figure 12.1. For example, can the  $p$  direction of the fold catastrophe lie within the  $d$  region of the neighbourhood of the catastrophe?

## 12.4 Mathematical Technique

Let  $H : \mathbb{R}^2 \times \mathbb{R}^3 \mapsto \mathbb{R}$ ,  $(u, v; p, q, r) \mapsto H$  describe a critical point catastrophe. I wish to characterize the statement that  $L_{i,j,k}$  exhibits the catastrophe  $H$  at some point. I will assume without loss of generality that the point is  $j = x = y = 0$ . The  $i$  and  $k$  parameters cannot be set to zero by rotating the  $x, y$ -system like the  $j$  parameter can; but, for simplicity, in what follows I will assume  $i = k = 0$ ; this makes no difference to the following argument and avoids a lot of  $(i - i')$  and similar terms. The occurrence of the catastrophe is characterized by there being functions  $P, Q, R, U, V, C \in \mathbf{C}^\infty$ :

- A) Three control parameters  $P, Q, R : \mathbb{R}^3 \mapsto \mathbb{R}$ ,  $(i, j, k) \mapsto P, Q, R$  where:
  - 1.  $P(0, 0, 0) = Q(0, 0, 0) = R(0, 0, 0) = 0$
  - 2. The initial directions of the control parameters ( $\vec{p} = [\nabla P]_{i=j=k=0}$  and  $\vec{q}$  and  $\vec{r}$  defined similarly) span  $i, j, k$ -space i.e.  $\vec{p} \cdot (\vec{q} \times \vec{r}) \neq 0$
- B) An  $i, j, k$ -family of local diffeomorphisms mapping  $(x, y)$  to  $(u, v)$  described by  $U, V : \mathbb{R}^2 \times \mathbb{R}^3 \mapsto \mathbb{R}$ ,  $(x, y; i, j, k) \mapsto U, V$  where:
  - 1.  $U(0, 0; 0, 0, 0) = V(0, 0; 0, 0, 0) = 0$
  - 2. The Jacobian is non-zero:  $J = \left[ \frac{\partial U}{\partial x} \frac{\partial V}{\partial y} - \frac{\partial U}{\partial y} \frac{\partial V}{\partial x} \right]_{x=y=i=j=k=0} \neq 0$
- C) A function  $C : \mathbb{R}^3 \mapsto \mathbb{R}$ ,  $(i, j, k) \mapsto C$ . I define  $\vec{c} = [\nabla C]_{i=j=k=0}$

Such that, in an open  $\mathbb{R}^2 \times \mathbb{R}^3$ -neighbourhood of  $(0, 0; 0, 0, 0)$ ,  $U, V$  and  $P, Q, R$  warp the catastrophe unfolding to match  $L$  i.e.

$$\begin{aligned} L_{i,j,k}(x, y) = & H(U(x, y; i, j, k), V(x, y; i, j, k); P(i, j, k), Q(i, j, k), R(i, j, k)) \\ & + C(i, j, k) \end{aligned} \quad (12.2)$$

Since  $j = 0$ ,  $\nabla L = \square L$  can be applied to equation 12.2. From this it is deduced that for the image at  $i = j = k = 0$ :

$$\begin{aligned} \vec{0} = & \vec{c} + \vec{p} \frac{\partial H}{\partial p} + \vec{q} \frac{\partial H}{\partial q} + \vec{r} \frac{\partial H}{\partial r} + [(\nabla - \square) U] \frac{\partial H}{\partial u} + [(\nabla - \square) V] \frac{\partial H}{\partial v} \\ & - \left( \begin{array}{c} \frac{1}{\sqrt{2}} \left( \frac{\partial U}{\partial x} \right)^2 \\ \frac{\partial U}{\partial x} \frac{\partial U}{\partial y} \\ \frac{1}{\sqrt{2}} \left( \frac{\partial U}{\partial y} \right)^2 \end{array} \right) \frac{\partial^2 H}{\partial u^2} - \left( \begin{array}{c} \sqrt{2} \frac{\partial U}{\partial x} \frac{\partial V}{\partial x} \\ \frac{\partial U}{\partial x} \frac{\partial V}{\partial y} + \frac{\partial U}{\partial y} \frac{\partial V}{\partial x} \\ \sqrt{2} \frac{\partial U}{\partial y} \frac{\partial V}{\partial y} \end{array} \right) \frac{\partial^2 H}{\partial u \partial v} \\ & - \left( \begin{array}{c} \frac{1}{\sqrt{2}} \left( \frac{\partial V}{\partial x} \right)^2 \\ \frac{\partial V}{\partial x} \frac{\partial V}{\partial y} \\ \frac{1}{\sqrt{2}} \left( \frac{\partial V}{\partial y} \right)^2 \end{array} \right) \frac{\partial^2 H}{\partial v^2} \end{aligned} \quad (12.3)$$

Next, I note that since  $U, V \in \mathbf{C}^\infty$  and the equivalence described by equation 12.2 is only required to hold in an open neighbourhood around the catastrophe,  $U$  and  $V$  can be described by Taylor series i.e.

$$\begin{aligned} U(x, y; i, j, k) &= U_x x + U_y y + U_i i + U_j j + U_k k + U_{xx} x^2 + U_{ix} ix + \dots \\ V(x, y; i, j, k) &= V_x x + V_y y + V_i i + V_j j + V_k k + V_{xx} x^2 + V_{ix} ix + \dots \end{aligned}$$

These power series can be inserted into equation 12.3 yielding a vector, equal to the zero vector, whose terms are polynomials in  $x$  and  $y$  ( $i, j, k$  being zero). The fact that the vector is identically zero can be used to determine  $\vec{p}, \vec{q}, \vec{r}$  in terms of the coefficients of the power series of  $U$  and  $V$  (the constant function  $C$  is irrelevant). It could turn out that condition A.2 fails to be satisfied. If this occurred it would show that the catastrophe could not occur in affine scale space. As will be shown, this does not occur with the catastrophes suggested by Thom's theorem. Although equation 12.3, with the power series substituted in, will contain coefficients such as  $V_i$  and  $V_{jx}$  these can be eliminated and will not figure in the expressions for  $\vec{p}, \vec{q}, \vec{r}$  which will be entirely in terms of the coefficients for terms such as  $x^n y^m$ . Thus  $\vec{p}, \vec{q}, \vec{r}$ , the initial  $i, j, k$ -directions of the control parameters of the catastrophe unfolding, will be described in terms of the  $x, y$ -warping of the catastrophe germ onto the actual catastrophe. There are no restrictions on this warping but it will be seen that the expressions for  $\vec{p}, \vec{q}, \vec{r}$  are such that these directions are restricted in terms of the double cone diagram.

## 12.5 The Catastrophes

In the following sections I will give expressions for  $\vec{p}$ ,  $\vec{q}$  and  $\vec{r}$  for five catastrophes. To allow compact presentation I pre-define some terms:

$$\begin{aligned}\vec{a}_1 &= 3\sqrt{2} \begin{pmatrix} U_x^2 \\ \sqrt{2}U_xU_y \\ U_y^2 \end{pmatrix} \\ \vec{a}_2 &= \frac{2\sqrt{2}}{J} \begin{pmatrix} V_x(2V_yV_{xx} - V_xV_{xy}) \\ \sqrt{2}(V_y^2V_{xx} - V_x^2V_{yy}) \\ V_y(V_yV_{xy} - 2V_xV_{yy}) \end{pmatrix} \\ \vec{a}_3 &= \frac{V_{xy}^2 - 4V_{xx}V_{yy}}{J^2} \begin{pmatrix} \sqrt{2}V_x^2 \\ 2V_xV_y \\ \sqrt{2}V_y^2 \end{pmatrix} \\ \vec{a}_4 &= \frac{V_y^2U_{xx} - V_xV_yU_{xy} + V_x^2U_{yy}}{J^2} \vec{a}_2 \\ \vec{a}_5 &= \frac{2}{J^2} \begin{pmatrix} \sqrt{2}V_x(3V_y^2V_{xxx} - 2V_xV_yV_{xxy} + V_x^2V_{xyy}) \\ 3V_y^3V_{xxx} - V_xV_y^2V_{xxy} - V_x^2V_yV_{xyy} + 3V_x^3V_{yyy} \\ \sqrt{2}V_y(3V_x^2V_{yyy} - 2V_xV_yV_{xyy} + V_y^2V_{xxy}) \end{pmatrix} \\ \vec{a}_6 &= \frac{V_y^2V_{xx} - V_xV_yV_{xy} + V_x^2V_{yy}}{J^3} \begin{pmatrix} 2\sqrt{2}U_x(2V_yV_{xx} - V_xV_{xy}) \\ 4(U_yV_yV_{xx} - U_xV_xV_{yy}) + 2JV_{xy} \\ 2\sqrt{2}U_y(V_yV_{xy} - 2V_xV_{yy}) \end{pmatrix} \\ \vec{a}_7 &= \frac{3\sqrt{2}}{J} \begin{pmatrix} 2(U_xV_y + U_yV_x)U_{xx} - 2U_xV_xU_{xy} \\ \sqrt{2}(2U_yV_yU_{xx} - 2U_xV_xU_{yy}) \\ -2(U_xV_y + U_yV_x)U_{yy} + 2U_yV_yU_{xy} \end{pmatrix}\end{aligned}$$

### 12.5.1 Fold Catastrophe: $H(u, v; p, q, r) = pu + u^3 + v^2$

As already noted, the fold catastrophe has a 1-D unfolding and so takes the form of a surface in  $i, j, k$ -space. Figure 12.3 shows a patch of fold surface. The vector  $\vec{p}$  is a surface normal and  $\vec{q}$  and  $\vec{r}$  lie in the tangent plane. The vector  $\vec{p}$  corresponds to the  $p$  parameter of the unfolding while the  $\vec{q}$  and  $\vec{r}$  vectors correspond to dummy control parameters along which nothing happens. The inset contour plots show canonical examples of the local structure at different stages of the catastrophe. The two plots attached to small spheres occur over a 3-D region of affine scale space; while the third plot, attached to the fold surface, occurs at every point within the fold surface (that is to say at some  $x, y$ -point within every image whose  $i, j, k$ -position lies in the fold surface). From the contour plots it should be clear that, read in the direction of increasing  $p$ , the catastrophe describes an extremum and a saddle point approaching and mutually annihilating; read in the  $-p$  direction it describes an extremum and a saddle point being created and moving apart.

The mathematical technique described in section 12.4 can be applied to the fold catastrophe in order to express the vector  $\vec{p}$  in terms of the diffeomorphism that warps the actual shoe surface on to the canonical shoe surface. To apply the technique I calculate the various derivatives of  $H$  and then eliminate terms involving  $p$ ,  $q$  or  $r$  as these will be zero at  $i = j = k = 0$  (property A.1). This

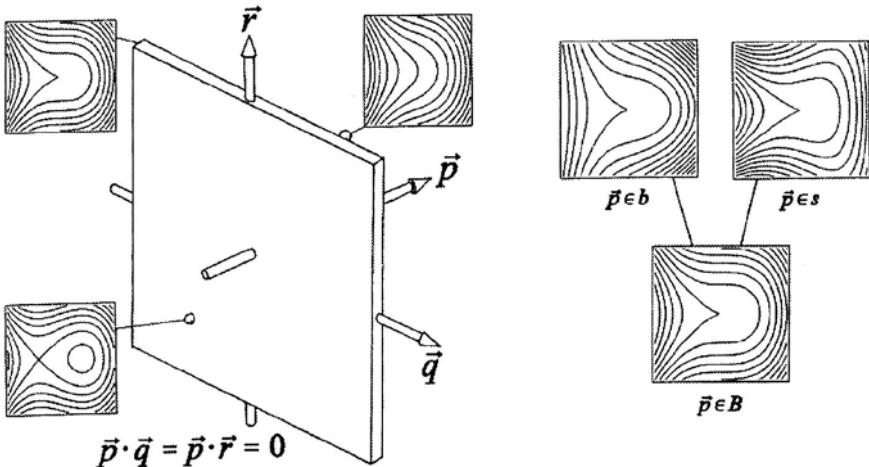


Figure 12.3: Fold Catastrophe: embedding and example germs

results in  $\frac{\partial H}{\partial p} = u$ ,  $\frac{\partial H}{\partial q} = \frac{\partial H}{\partial r} = 0$ ,  $\frac{\partial H}{\partial u} = 3u^2$ ,  $\frac{\partial H}{\partial v} = 2v$ ,  $\frac{\partial^2 H}{\partial u^2} = 6u$ ,  $\frac{\partial^2 H}{\partial u \partial v} = 0$  and  $\frac{\partial^2 H}{\partial v^2} = 2$ . These are then inserted into equation 12.3;  $u$  and  $v$  terms are replaced by power series; and terms are balanced to make everything disappear. For the fold catastrophe only the coefficients of the terms  $x$  and  $y$  need to be considered and the result is:  $\vec{p} = \vec{a}_1 + \vec{a}_2$ .

Since  $\frac{\partial H}{\partial q} = \frac{\partial H}{\partial r} = 0$  this method fails to determine the directions  $\vec{q}$  and  $\vec{r}$ . In fact, since they are associated with dummy variables there is no point in trying to determine them. What is required is to show that  $\vec{p}$  is normal to the fold surface. To show this, consider a derivative in a direction  $\vec{d}$  lying in the tangent plane of the fold surface:  $\frac{\partial H}{\partial \vec{d}} = (\vec{d} \cdot \nabla U)3u^2 + (\vec{d} \cdot \nabla V)2v + (\vec{d} \cdot \vec{p})u = 0$ . This derivative should be 0 since nothing happens within the fold surface. The vectors  $\nabla U$  and  $\nabla V$  can be expressed as power series in  $u$  and  $v$ . If this is done then the only  $u$  term (rather than  $u^2$  or  $uv$  etc.) in the expression will be  $(\vec{d} \cdot \vec{p})u$  and hence for the derivative to be zero it must be the case that  $\vec{d} \cdot \vec{p} = 0$  and so, as required,  $\vec{p}$  is shown to be normal to the fold surface.

From the expression for  $\vec{p}$  it can be shown that  $\vec{p} \in b \cup B \cup s$ . Thus, locally a fold surface will be one of three types: generically  $\vec{p} \in b$  or  $\vec{p} \in s$ , or versally  $\vec{p} \in B$ . Generically (with respect to the incident luminance), the versal points ( $\vec{p} \in B$ ) will be curves between regions of the two generic types. Alternative configurations, such as isolated points or crossings, are not stable. Examples of germs that will produce fold catastrophes of these three subtypes are shown in figure 12.3. All three examples are diffeomorphic to the canonical shoe surface of figure 12.2, it is the details of the diffeomorphism which determine the embedding of the fold surface.

Finally, I comment on how this relates to the fold catastrophe in ordinary scale space (Chapter 10). In an infinitesimal  $i, j, k$ -neighbourhood of the image

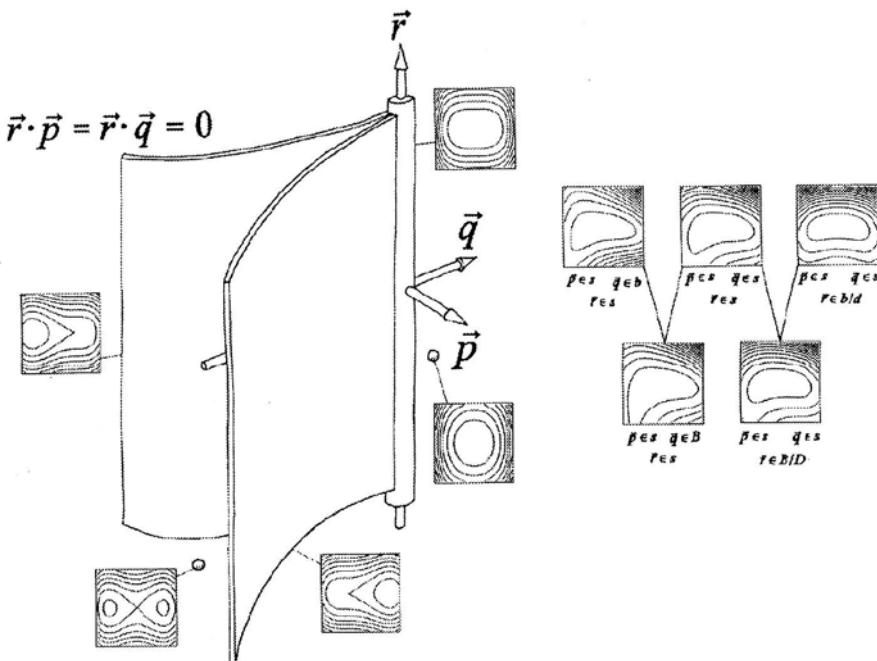


Figure 12.4: Extremum Cusp Catastrophe: embedding and example germs

containing the shoe surface, isotropic blurring occurs along the axis of the double cone. If  $\vec{p}$  is in  $b$  or the top half of  $s$  then for the ordinary scale space along this line, the fold catastrophe will be experienced as an annihilation; alternatively, if  $\vec{p}$  is in the lower half of  $s$  as a creation.

### 12.5.2 Cusp Catastrophe: $H(u, v; p, q, r) = pu + qu^2 + u^4 \pm v^2$

As shown in figure 12.2 there are two varieties of germ for the cusp catastrophe. I refer to  $u^4 + v^2$  as a quartic hump and its associated catastrophe as an extremum cusp; to  $u^4 - v^2$  as a quartic saddle and associated with it a saddle cusp catastrophe. The unfolding of the two catastrophes are shown in figures 12.4 and 12.5. The figures show that the  $i, j, k$ -structure of the catastrophes are identical, the difference lies in the  $x, y$ -structures. Since in both cases the catastrophe has two control parameters the locus of points containing the catastrophe germ forms a curve in  $i, j, k$ -space. The curve is the meeting, with equal normal, of two fold surfaces. The contour plots show the nature of the unfolding. The same conventions as in figure 12.3 are used with the addition that when a contour plot is connected to a curve it is indicated that the local structure shown occurs in every image whose  $i, j, k$ -position lies in the curve. In both types of catastrophe three critical points are involved. The extremum cusp catastrophe is the coming together of two extrema and a saddle point resulting in an extremum; in the saddle cusp

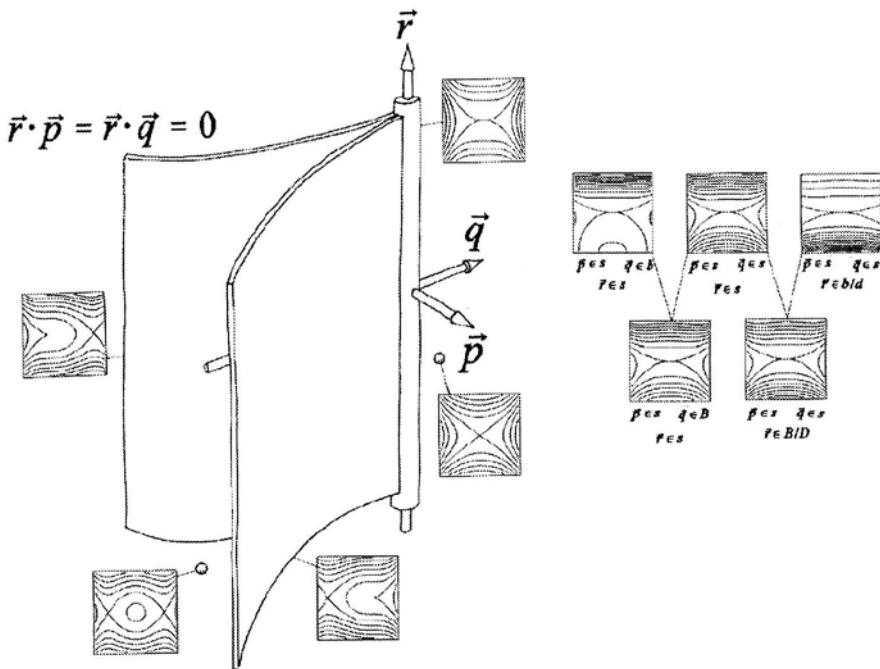


Figure 12.5: Saddle Cusp Catastrophe: embedding and example germs

catastrophe, two saddle points and an extremum produce a saddle point.

The result of applying the technique of section 12.4 is:  $\vec{p} = \pm \vec{a}_2$ ,  $\vec{q} = 2\vec{a}_1 \pm \vec{a}_3 \mp \vec{a}_4 \pm \vec{a}_5 \pm \vec{a}_6$ . The argument that was used for the fold catastrophe to show that the vector through the surface was a surface normal can be used to show that  $\vec{r}$  is normal to both  $\vec{p}$  and  $\vec{q}$ . So  $\vec{r} = \vec{p} \times \vec{q}$ .

The possible directions of  $\vec{p}$ ,  $\vec{q}$  and  $\vec{r}$  relative to the double cone diagram, can be determined. It turns out that although the expressions for the two types of cusp are different they have the same possible embeddings. Germs having the different possible embeddings are shown in figures 12.4 and 12.5. Note that I describe the direction of  $\vec{r}$  as  $b/d$ , etc. rather than  $b$  or  $d$  as the orientation of the vector is arbitrary; the same is true for  $\vec{p}$  but the issue does not arise as it only ever points into  $s$ . There are three generic cases (i)  $\vec{p} \in s$ ,  $\vec{q} \in b$ ,  $\vec{r} \in s$  (ii)  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in s$  and (iii)  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in b/d$ ; and two versal cases  $\vec{p} \in s$ ,  $\vec{q} \in B$ ,  $\vec{r} \in s$  and  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in B/D$ . The versal cases will occur at isolated points on a cusp catastrophe curve separating stretches of the generic cases. The first versal case must lie between generic types (i) and (ii); the second between (ii) and (iii). Generic cases (i) and (iii) cannot be adjacent on a cusp catastrophe curve.

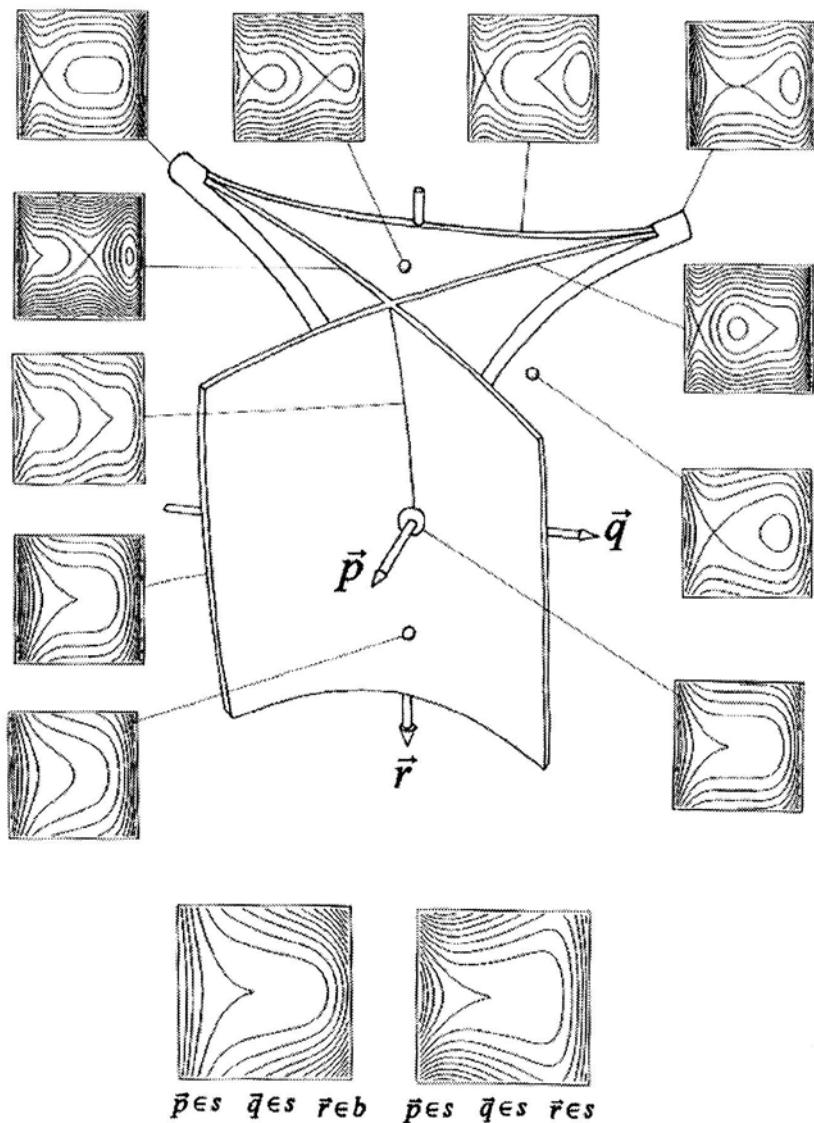


Figure 12.6: Swallowtail Catastrophe: embedding and example germs

### 12.5.3 Swallowtail: $H(u, v; p, q, r) = pu + qu^2 + ru^3 + u^5 + v^2$

The swallowtail catastrophe has three control parameters and thus occurs at an isolated point of  $i, j, k$ -space. The location of the catastrophe is shown in figure 12.6 as ball. The catastrophe point lies at a meeting of fold surfaces and three curves: an extremum cusp curve!cusp curve, a saddle cusp curve and an intersection between two fold surfaces. To read the catastrophe it is easiest to start with the plot with the most critical points (i.e. the one for the region shaped like a downwards pointing pyramid) and work out from there.

Again, the technique of section 12.4 can be applied to determine the initial directions of the control parameters. This gives  $\vec{p} = \vec{a}_2$ ,  $\vec{q} = \vec{a}_3 - \vec{a}_4 + \vec{a}_5 + \vec{a}_6$ . I have not given the expression for  $\vec{r}$  as I have only been able to derive it with the aid of software and the result is too long to reproduce.

From these expressions, the possible embeddings may be determined; there are two stable possibilities  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in b$  and  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in s$ , examples of germs with these embeddings are shown in figure 12.6. Comparison of this figure with figure 12.3 suggests that  $\vec{r} \in B$  is possible. It is, but is not stable because the catastrophe is just a single point, so perturbing the incident luminance that generates the affine scale space would change the catastrophe to one of the two types shown.

### 12.5.4 Elliptic Umbilic:

$$H(u, v; p, q, r) = pu + qv + r(u^2 + v^2) + u^3 - 3uv^2$$

The elliptic umbilic, like the swallowtail, is a catastrophe of three control parameters. Unlike the preceding catastrophes, it only occurs for functions whose domain is 2-D. Figure 12.7 shows the unfolding. The catastrophe occurs at the meeting of fold surfaces and saddle cusp curves. Reading the catastrophe in the  $\vec{r}$  direction one sees three saddle point surrounding a minimum; as the catastrophe is approached these merge to produce a monkey saddle; out the other side, three saddle points re-appear but now surrounding a maximum. Applying the technique of section 12.4 results in:

$$\begin{aligned}\vec{p} &= 2\sqrt{2} \begin{pmatrix} U_x^2 - V_x^2 \\ \sqrt{2}(U_x U_y - V_x V_y) \\ U_y^2 - V_y^2 \end{pmatrix} \\ \vec{q} &= -6 \begin{pmatrix} \sqrt{2} U_x V_x \\ U_x V_y + U_y V_x \\ \sqrt{2} U_y V_y \end{pmatrix} \\ \vec{r} &= \vec{a}_7 + \frac{3\sqrt{2}}{J} \begin{pmatrix} 2(U_x U_y - V_x V_y)V_{xx} + (V_x^2 - U_x^2)V_{xy} \\ \sqrt{2}((U_y^2 - V_y^2)V_{xx} + (V_x^2 - U_x^2)V_{yy}) \\ -2(U_x U_y - V_x V_y)V_{yy} - (V_y^2 - U_y^2)V_{xy} \end{pmatrix}\end{aligned}$$

Two embeddings are possible:  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in b/d$  and  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in s$ ;

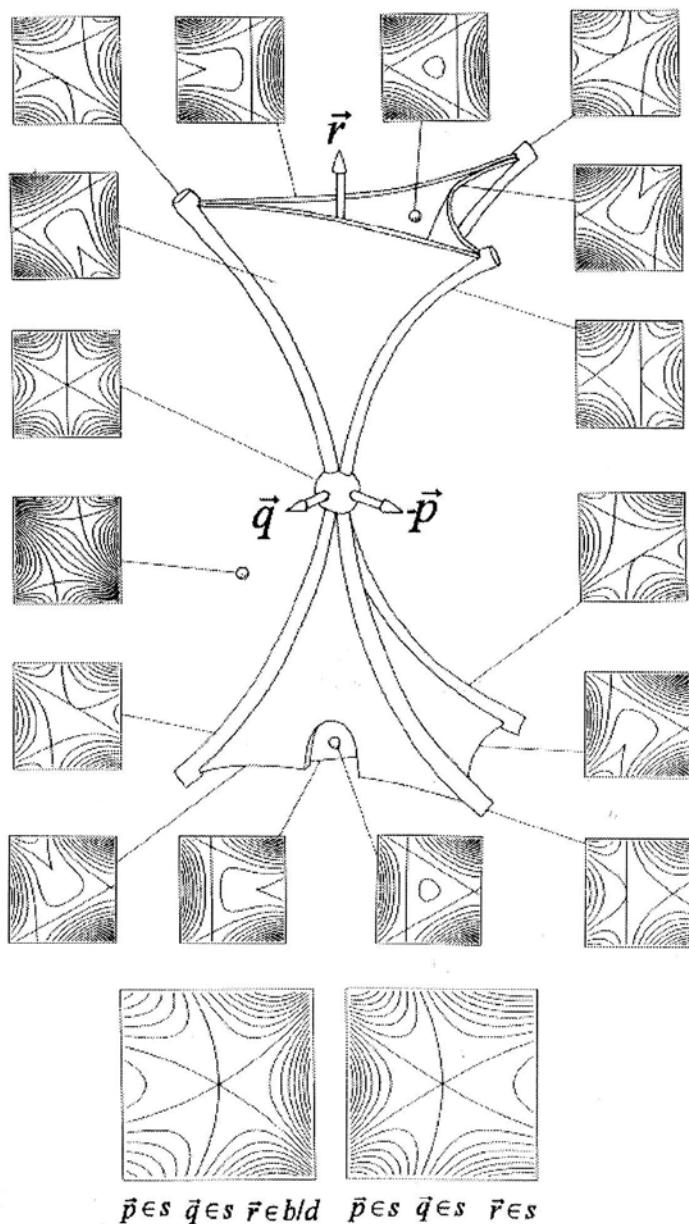


Figure 12.7: Elliptic Umbilic Catastrophe: embedding and example germs

examples of germs with these embeddings are shown at the bottom of figure 12.7. Like the swallowtail,  $\vec{r} \in B$  is possible but unstable.

### 12.5.5 Hyperbolic Umbilic:

$$H(u, v; p, q, r) = pu + qv + r(u^2 - v^2) + u^3 + 3uv^2$$

The final catastrophe (figure 12.8) also has three control parameters and can only occur with functions which have a 2-D domain. The catastrophe occurs at the meeting of fold surfaces and four curves: two saddle cusps and two intersections of distinct fold surfaces. The contour plot with the most critical points is for the region nearest the viewer. It has two saddles, a maximum and a minimum. Passing straight through the catastrophe causes these four critical points to mutually annihilate leaving no critical points. Another way of reading the unfolding is to read the sequence as one moves, via the catastrophe, between the two tapering regions. Read like this, a maximum approaches a saddle point and passes through it turning into a minimum during the passage. Applying the technique of section 12.4 results in:

$$\begin{aligned}\vec{p} &= 2\sqrt{2} \begin{pmatrix} U_x^2 + V_x^2 \\ \sqrt{2}(U_x U_y + V_x V_y) \\ U_y^2 + V_y^2 \end{pmatrix} \\ \vec{q} &= 6 \begin{pmatrix} \sqrt{2} U_x V_x \\ U_x V_y + U_y V_x \\ \sqrt{2} U_y V_y \end{pmatrix} \\ \vec{r} &= \vec{a}_7 + \frac{3\sqrt{2}}{J} \begin{pmatrix} 2(U_x U_y + V_x V_y)V_{xx} - (V_x^2 + U_x^2)V_{xy} \\ \sqrt{2}((U_y^2 + V_y^2)V_{xx} - (V_x^2 + U_x^2)V_{yy}) \\ -2(U_x U_y + V_x V_y)V_{yy} - (V_y^2 + U_y^2)V_{xy} \end{pmatrix}\end{aligned}$$

There are two possible embeddings  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in b/d$  and  $\vec{p} \in s$ ,  $\vec{q} \in s$ ,  $\vec{r} \in s$ ; examples of germs with these embeddings are shown in figure 12.8. Again a third embedding is possible but unstable.

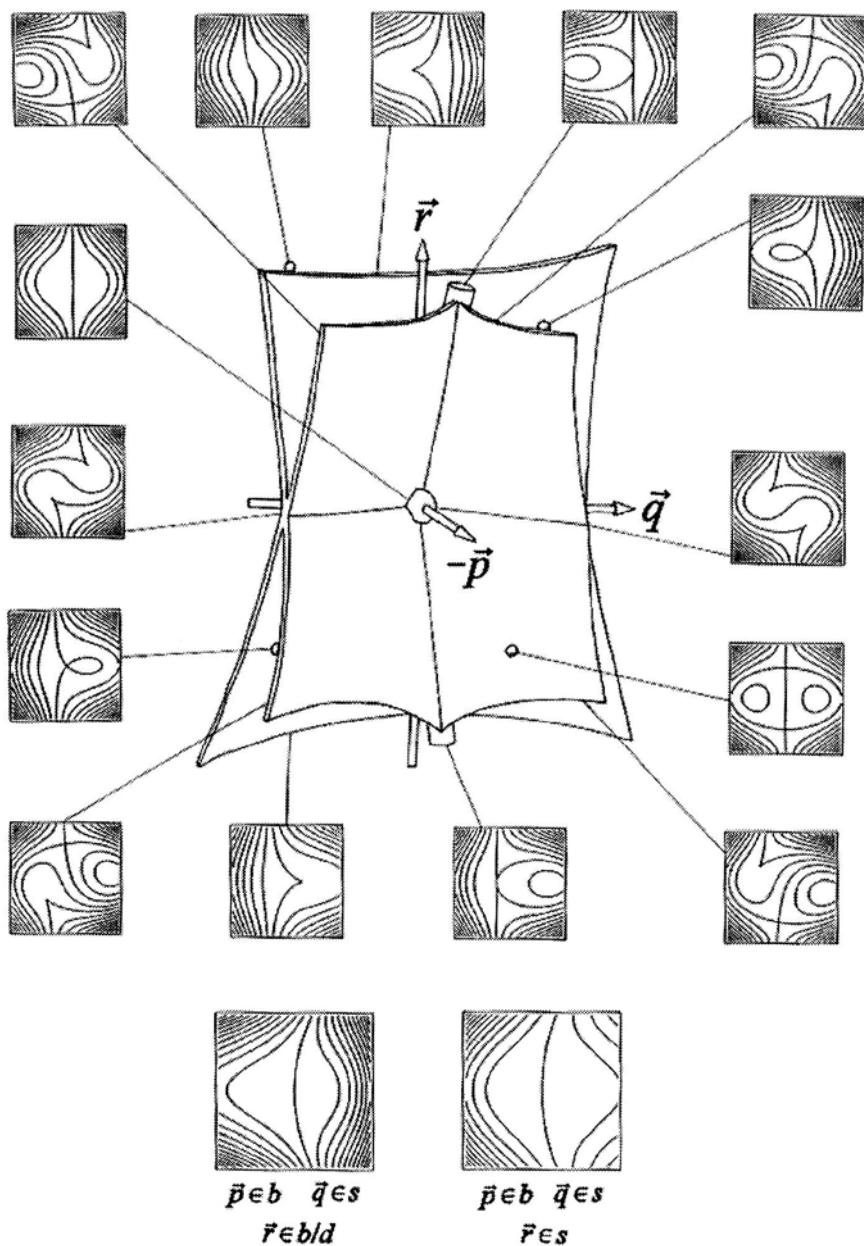


Figure 12.8: Hyperbolic Umbilic Catastrophe: embedding and example germs

# Chapter 13

# Topological Numbers and Singularities

**Stiliyan Kalitzin**

*Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

## 13.1 Introduction

Scale space approach provides a tool for studying a given image at all scales simultaneously. Features that can be detected at large scales can provide clues for tracing more detailed information at fine scales. For this ideology to be constructive, one needs to investigate which *local* properties (or local operators) are appropriate for quantifying the desired features and how these properties are changing from scale to scale. In various applications (Kass et al., 1987; Koenderink, 1990), specially those concerned with oriented structures, singular points are of particular interest. Singular points, or simply singularities below, are those points in a gray-scale image, where the gradient vector field vanishes (Lindeberg, 1992; Johansen, 1994). Examples of such points in two dimensional images are extrema, saddle points, “monkey” saddles etc. Singular points can be characterized by their order. The order of a singular point is the lowest non-vanishing power in the Taylor expansion of the image field  $L(x_1, x_2, \dots, x_d)$  around this point. Obviously the order must be greater than 1, since the gradient vector  $L_i = \partial_i L(x_1, x_2, \dots, x_d)$  is equal to zero in this point.

In this chapter we consider an additional topological characteristic, or number, which can classify singularity points of an image. Its definition is completely non-perturbative in the sense that it makes no reference to any higher order differential structures at the singularity (Salden et al., 1996) (Chapter 8). Instead, this number is computed on a hyper-surface closely surrounding the singular point. For such

a definition to be well-posed, we must ensure that our image field contains only isolated singularities. If the image is convolved with a Gaussian kernel, then all derivatives are well defined (Florack et al., 1994b) and it is well known that the image field will have only isolated singular points (Milnor, 1963). The amount of blurring, or the scale, will be treated according to the scale-space paradigm as an independent “coordinate”  $t$ . The image field then obeys the linear diffusion equation (Koenderink, 1984)

$$\partial_t L(x, t) = \Delta L(x, t). \quad (13.1)$$

Here we assume  $d$  spatial dimensions so, unless otherwise stated, we use the abbreviation  $x = (x_1, \dots, x_d)$ . When the scale is considered fixed at some value and the dependence on it is not explicitly used, the  $t$  variable will be omitted in our formulas.

The topological number considered in this chapter is the natural quantity that describes singular points in all image dimensions. For one dimensional signals this quantity reduces to a binary number indicating whether a point is a local maximum or minimum. For two dimensional signals this number (known also as the winding number for the gradient vector field) extends the definitions of extrema and saddle points that characterize the non degenerate singularities, for the degenerate, higher order cases. Winding numbers have been introduced for analyzing singular points in two dimensional oriented patterns (Kass et al., 1987). Here we give a more general approach including any order of singularity and any number of image dimensions.

By computing the topological number in every (scale) space point, we can define in a consistent way a (scale) space *distribution*, or density, of the topological “charge”. This density provides a practical tool for *localization* of the singularity points and thus contains both spatial information and information about the *type* of the singularities.

When scale space evolution is considered, singularity points “drift” and eventually interact with each other in the so-called catastrophe points (Johansen, 1994) (see also the Chapters 10–12 and 14 in this book). Our topological number has a conservation property which enables us to analyze the outcome of these interactions.

The organization of the chapter is as follows. In Section 13.2 we give a self-contained theoretical description of the topological numbers and their major properties. Section 13.3 is devoted to some low-dimensional or low-order applications of the general definitions.

In section 13.4 we (re)introduce the scale-space dynamics. Evolution of the singular points and conservation properties of their topological numbers are considered. Examples of typical one and two dimensional top points are presented as illustrations of the topological current conservation property.

Finally, section 13.5 summarizes the properties of the topological quantity considered in this chapter and we discuss some possibilities for further research.

## 13.2 Theory. Topological invariants

Now we introduce some brief notations from the theory of homotopy groups (Nakahara, 1989). They provide the natural basis for the introduction of a topological number associated with any singular image point.

Suppose  $P$  is a (singular or regular) point in the image and  $S_P$  is a closed oriented hyper-surface around the point  $P$ , topologically equivalent to a  $d - 1$  dimensional sphere. Suppose there are no singular points on  $S_P$ , then the normalized gradient vector field

$$\xi_i = \frac{\partial_i L}{(\partial_j L \partial_j L)^{1/2}} \quad (13.2)$$

is well defined on the surface  $S_P$  (summation over repeated indices is assumed).

The space of all unit-length  $d$ -dimensional Euclidean vectors is isomorphic to the  $d - 1$  dimensional sphere of unit radius  $S_1^{(d-1)}$ . Therefore the vector field  $\xi_i$  defines a mapping

$$S_P \rightarrow S_1^{(d-1)}$$

But recalling that  $S_P$  is a manifold homotopic to a  $d - 1$  dimensional sphere, we see that the above mapping can be classified by an element of the homotopy group  $\pi_{(d-1)}(S^{(d-1)})$ . This group comprises all homotopically non-equivalent mappings between two  $d - 1$  dimensional spheres. It is known that  $\pi_{(d-1)}(S_1^{(d-1)}) \cong \mathbf{Z}$  which is the commutative group of all integer numbers (where addition is the group operation). Just to illustrate the last equivalence, consider the mapping  $S^1 \rightarrow S^1$  between two unit circles (embedded, for example, in the complex plane). In angular parameterization, we have  $\psi_2 = f(\psi_1)$  where  $\psi_1$  and  $\psi_2$  are the arc-lengths of the two circles and  $f(\psi)$  is a smooth function. Then the periodicity condition imply  $f(\psi_1 + 2\pi) \equiv 2\nu\pi$  for some integer  $\nu$ . Smooth re-parameterizations can not change this number and therefore, it labels non-equivalent classes of functions  $f(\psi) \equiv f_\nu(\psi)$ .

In general, we can characterize the unit vector field  $\xi_i$  on the surface  $S_P$  (taken around the chosen point  $P$ ) by the homotopy number  $\nu_P \in \mathbf{Z}$  it defines.

To make the above construction explicit, we first give the operational definition of the quantity  $\nu$ .

**Definition 13.2.1.** Let  $L(x, t) : R^d \times R^+ \rightarrow R$  be a differentiable  $d$ -dimensional scalar image (represented by its gray-scale field) with at most isolated singularity points. At a *non-singular* point  $A = (x_1 \dots x_d)$  we define a  $d - 1$  form:

$$\begin{aligned} \Phi(A) &= \xi_{i_1} d\xi_{i_2} \wedge \dots \wedge d\xi_{i_d} \epsilon^{i_1 i_2 \dots i_d} \\ \epsilon^{i_1 i_2 \dots i_d} &= (-1)^{\text{Perm}(i_1 i_2 \dots i_d)} \end{aligned} \quad (13.3)$$

Let  $S_P$  be a closed ( $\partial S_P = 0$ ), oriented hyper-surface around a given image point  $P$ . If there are no singular points on  $S_P$  then we define the quantity:

$$\nu_{S_P} = \oint_{A \in S_P} \Phi(A). \quad (13.4)$$

The integral above is the natural integral of a  $(d-1)$  form over a  $(d-1)$  dimensional manifold without border.

This definition is applicable in any coordinates as long as *covariant* external differentiation is assumed. For simplicity, Cartesian coordinates will be assumed for the rest of this chapter.

An important property of the  $d-1$  form  $\Phi$  is that it is a closed form, i.e.  $d\Phi(A) = 0$ . For the proof of the above propositions see (Kalitzin et al., 1996). The last property of the form (13.3) is essential for the applications of the topological quantity (13.4). From the generalized Stoke's theorem (Eguchi et al., 1980; Boothby, 1975):

$$\oint_{\partial W} \Phi(x) = \int_W d\Phi(x) \equiv 0. \quad (13.5)$$

We see that the topological quantity (13.4) is invariant under smooth deformations of the surface  $S_P$  as long as no singular points are crossed by  $S$  in the process of this deformation. The last property justifies the term “topological” that we assign to the quantity  $\nu_S$ . It depends on the properties of the image in the region where  $S_P$  is placed, but in general not on the surface  $S$  itself. More precisely, the topological number depends only on the number and type of singularities surrounded by the surface  $S_P$ . Therefore, for any image point  $P$  we can define an unique topological number.

**Definition 13.2.2 (Topological number).** The topological number of an image point  $P$  is

$$\nu_P = \nu_{S_P},$$

where  $S_P$  is *any* closed oriented hyper-surface taken closely around  $P$ . The surface  $S$  must be close to  $P$  in order to ensure that no other singularities are surrounded.

To give a more intuitive description of Definitions 13.2.1 and 13.2.2 we can present the two-dimensional case where the one-form  $\Phi \equiv \epsilon^{ij} \xi_i d\xi_j$  gives just the angle between the normalized gradients in two neighboring points. Integrating this angle along a closed contour we find the *winding number* (see also Section 13.3.3) associated with this contour.

An important property of the topological number in Definition 13.2.2 is the following “conservation law” (for proof see (Kalitzin et al., 1996)):

**Proposition 13.2.3.** Let  $\nu_S$  be defined as in Definition 13.2.2 and let  $L(x, \lambda)$ ,  $\lambda \in [0, 1]$  be a one-parameter family of images, smoothly depending on the deformation parameter  $\lambda$ , such that  $L(x, 0) = L(x)$ . If the new field  $L(x, \lambda)$  has no singular points on the hyper-surface  $S$  for any value of  $\lambda \in [0, 1]$ , then quantity  $\nu_S(\lambda)$  given in (13.3) and (13.4) for the field  $L(x, \lambda)$ , is the same for all  $\lambda \in [0, 1]$ .

The last Proposition is of fundamental importance when smooth image evolution is considered. Such evolution can be induced, for example, by the diffusion equation (13.1) or by any other smooth flow or deformation.

Topological numbers  $\nu_P$  can be associated with every point of the image. It is clear from the closeness of (13.3) however, that the topological number of a non-singular point is zero. If we plot the value of  $\nu_P$  in every point of an image, we will obtain a map of the singular points of the image representing their topological “charge”. We can go one step further and define a *scalar density field*,  $\nu(x_1, x_2, \dots, x_d)$ , that gives the *distribution* of the topological singularities in a given image. This distribution can be obtained as (Kalitzin et al., 1996):

$$\nu(x) = \lim_{A_P \rightarrow 0} \frac{\nu_{S_P}}{A_P} \quad (13.6)$$

where  $\nu_{S_P}$  is the same as in Definition 13.2.1 and  $A_P$  is the area inside the contour  $S_P$ . For an image with at most discrete set of singularity points the above formula gives:

$$\nu(x) = \sum_P \nu_P \delta(x - x_P). \quad (13.7)$$

Here  $\nu_P$  are the topological numbers from Definition 13.2.2,  $\delta(x)$  is the  $d$ -dimensional Dirac delta distribution, and the sum is over all *singular* points (at locations  $P$ ) in the image. In practice, formula (13.7) provides a tool for localizing the singularity points by computing the distribution  $\nu(x)$  from the image data.

## 13.3 Examples

In this section we present some particular cases, where the above general theory can be illustrated and the topological number can be computed analytically.

### 13.3.1 Second order singular points

Non-degenerate, or Morse, singular points are those points where the gradient vector vanishes, but the determinant of the Hessian is non-zero. For these cases the topological number represents the sign of this determinant.

**Proposition 13.3.1.** Suppose  $P$  is a non-degenerate (Morse) singular point of second order, so that locally

$$L(x) = \frac{1}{2!} x^i x^j L_{ij} + \dots; \\ L_{ij} \equiv \partial_i \partial_j L(P); \quad \det(L_{ij}) \neq 0, \quad (13.8)$$

then  $\nu_P \sim \text{sign}(\det L_{ij})$

This proposition shows that the topological number in Definition 13.2.2 can be considered also as an extension of the notion of the sign of the determinant of the Hessian for image points where the last quantity vanishes.

### 13.3.2 One dimensional case

For one-dimensional signals  $L(x)$ , the topological number 13.2.2 in a point  $P = x$  is ultimately simple:  $\nu_P \equiv \text{sign}(L_x)_B - \text{sign}(L_x)_A$  for any  $A, B : A < P < B$  in the close vicinity of  $P$ . In other words, the topological number of a point is reduced to the difference between the signs of the image derivative taken from the left and from the right. Obviously,  $\nu_P = 2$  for local minima,  $\nu_P = -2$  for local maxima and  $\nu_P = 0$  for regular points or in-flex singularities.

Although this case is trivial, it shows that the general notion of topological quantity (13.4) reduces to the natural concept of local minimum and maximum in one dimensional signals. For higher dimensional images, the topological number does not reduce to a local extremum identifier, but it provides a more subtle information for the behavior of the image around its singular points.

### 13.3.3 Two dimensional images

In the two dimensional case the topological quantity of Definition 13.2.2 labels the equivalent class of mappings between two unit circles (see the beginning of the Section 2). This label is known also as the *winding number*. The winding number represents the number of times that the normalized gradient turns around its origin, as a test point circles around a given contour (hyper-surface in  $d = 2$ ).

Clearly, the winding number of any closed contour must be an integer multiple of  $2\pi$ . We can compute this number directly from the Definitions 13.2.1 and 5.3.5, but in two dimensions it is more convenient to use a complex-number representation of the image space  $z = x + iy, \bar{z} = x - iy$ .

In accord with (13.3), we can present the form  $\Phi(A)$  in complex notation as:

$$\begin{aligned}\Phi(A) &= \frac{(L_x dL_y - L_y dL_x)}{(L_x L_x + L_y L_y)} = \\ \text{Im}(dW/W) &\equiv \text{Im}(d\ln(W)).\end{aligned}\tag{13.9}$$

Application of (13.9) for homogeneous polynomial structures are considered in (Kalitzin et al., 1996). Here we give only two examples.

Singular point of order  $n=2k$  given by the local field:  $L^{(n)}(z, \bar{z}) = z^k \bar{z}^k$  is a higher order extremum and one finds that  $\nu = +2\pi$  for all  $k$ . This generalizes the case  $k=1$  of a local extremum (positive determinant of the Hessian) for a Morse singularity.

As another example, let us consider the so called “monkey saddles”. A non-degenerate  $n^{th}$  order symmetric saddle point (for simpler notations assumed at  $x_1 = x_2 = 0$ ) is defined locally by a scalar field of the form  $L^{(n)}(z, \bar{z}) = z^n + \bar{z}^n$ . General analysis in (Kalitzin et al., 1996) gives  $\nu = 2\pi(-n + 1)$ . When  $n = 2$ , we have a non-degenerate saddle point of second order with a winding number  $-2\pi$ .

## 13.4 Scale space evolution

Here we will reintroduce the explicit dependence of the image field on the scale parameter  $t$ . Singular points of each scale now will group into curves in scale space, the so-called *singularity strings*. These singularity strings may occasionally interact with each other: scatter, annihilate etc. The events of interactions between the singularities appear as catastrophes (in Thom's sense) (Morse and Cairns, 1969; Lindeberg, 1994e; Johansen, 1994) (see also Chapters 10–12 and 14 in this book) for the image scalar field.

Topological numbers from Definitions 13.2.1 and 13.2.2 play an essential role in understanding the evolution of singularities across scales. To establish this, let us recall the “conservation law” given in general in Proposition 13.2.3.

A first implication of this proposition for the case where scale evolution is considered, is that all singular points preserve their topological numbers while drifting across scales as long as they do not come infinitely close to other singularities.

When interactions occur, or in other words, when two or more singularity points are colliding at some critical scale, the *total* topological number is conserved. Therefore, the sum of topological numbers before a catastrophe is equal to the sum of topological numbers after the catastrophe.

The above conservation laws justify the notion of *topological current*, that can be attached to each singularity string. It represents the flow of the topological number of the corresponding singular point for any given fixed scale. This current is constant along the string (therefore correctly defined) and obeys the Kirchoff summation law at the catastrophic points in scale space. Exact formulation of these conservation laws are given in (Kalitzin et al., 1996).

To illustrate the conservation property of the topological current we present below some typical one and two dimensional examples of singularity interactions in scale space.

### 13.4.1 One dimensional fold

Consider the following solution of the diffusion equation (13.1):

$$L(x, t) = x^3 + 6xt. \quad (13.10)$$

Singular points, defined as the points where  $L_x = 0$  are

$$\begin{aligned} t < 0 : \quad x &= \pm(-2t)^{1/2} \\ t > 0 : \quad &\text{none.} \end{aligned} \quad (13.11)$$

The two singular points for any  $t < 0$  carry topological numbers +2 and -2. In one dimensional signals, this means that they correspond to a minimum and a maximum. The two singularity points annihilate at  $t=0$ , as permitted by the topological current conservation.

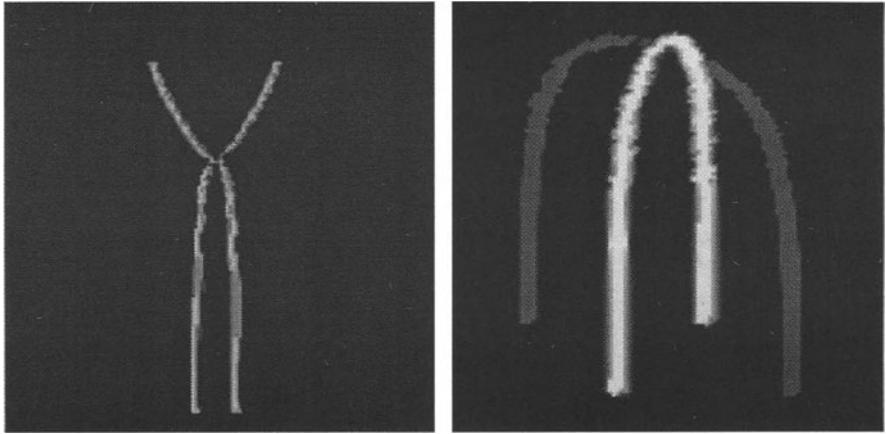


Figure 13.1: Three dimensional plots of scatter (left frame) and annihilation (right frame) of singularities in scale space.

### 13.4.2 Two dimensional umbilic point

In the next example no annihilation occurs. The two singular strings carry both current of  $-2\pi$ . Therefore these singularities can only scatter in the scale space as the total current cannot vanish.

Consider the following solution of (13.1) in two dimensions:

$$L(x, y, t) = x^2y - y^3 - 4yt. \quad (13.12)$$

Singular points, ( $\partial_x L = \partial_y L = 0$ ) are

$$\begin{aligned} t < 0 : \quad &x = 0; y = \pm(-4t/3)^{1/2} \\ t > 0 : \quad &y = 0; x = \pm t^{1/2}. \end{aligned} \quad (13.13)$$

So both for  $t > 0$  and  $t < 0$  there are two singular points. The topological numbers are easily computed and are in this case  $\nu = -2\pi$  for all singular points at  $t \neq 0$ . Therefore no annihilation is possible, the singularity strings scatter (see fig.13.1, left frame) at  $t = 0$  and change their plane of “polarization” from  $x = 0$  to  $y = 0$ . At the catastrophe point  $t = 0$ , the two singularities merge into a single one, with a winding number  $-4\pi$ . Thus exactly as the conservation law predicts.

On the right frame in fig.13.1, we present an example of singularity annihilation in scale space for the image defined by the function  $L(x, y, t) = x^2y + y^3 + 8yt$ . In this case there are four singular points at  $t < 0$ . Two of them have winding numbers  $-2\pi$  and the other two have winding numbers  $2\pi$ . Therefore, the annihilation is permitted by the current conservation law and indeed occurs at  $t = 0$ .

## 13.5 Summary and discussion

In this chapter we introduced a quantity characterizing isolated singular points in scalar images of any dimensions. In summary, the properties of this topological quantity are

- *Discrete*: It takes only a discrete set of values.
- *Localized*: It is zero in regular points and nonzero in the singular points.
- *Non-perturbative*: It can be operationally defined for singularity points of any (even infinite) order using hyper-surface integrals.
- *Conserved*: Under smooth image deformations, including Gaussian blurring, the topological density obeys a conservation law.
- *Invariant*: It is a topological quantity in the sense that smooth deformations of the surfaces on which this quantity is computed do not affect its value (unless other singularities are crossed).
- *Additive*: The sum of topological numbers in a given volume is a sum of these numbers in each sub-volume.

A direction for further research can be aimed at finding an adequate topological or geometrical quantity characterizing in a non-perturbative way the structure of the top-points in scale space. It may be interesting to analyze and compare the different diffusion schemes in terms of the characteristic quantity of their top-points.

# Chapter 14

# Multi-Scale Watershed Segmentation

Ole Fogh Olsen

*Department of Computer Science, University of Copenhagen  
Universitetsparken 1, DK-2100 Copenhagen, Denmark*

## 14.1 Introduction

On the path from acquisition to interpretation of an image one of the biggest hurdles is segmentation. That is, partitioning the image into basic structures corresponding to meaningful objects. One of the goals in this chapter is to outline how a segmentation tool can be built on top of linear Gaussian scale space. Another goal is to show the usefulness of catastrophe theory for analysing the deep structure of features in scale space.

A precise description of what is meaningful is needed in order to make them operationally defined in the sense of being expressible in terms of linear Gaussian scale space. Here we use a well known dissimilarity measure, the gradient magnitude, and define region boundaries to be the watersheds of this measure. Watersheds of the gradient magnitude have been reported (Griffen, 1995) to give good intuitive segmentations.

We do not assume any a priori knowledge of preferred scales for the image structure, therefore the segmentation tool probes the image over all scales. Accessing the structure at all scales is not enough; the correspondence between scales has to be established for a correct interpretation. The importance of the deep structure has been emphasised ever since the early literature on scale-space theory (Koenderink, 1984). In this article we will analyse the deep structure of segmentation using catastrophe theory (Saunders, 1980).

## 14.2 Morse functions and catastrophe theory

Before a model of objects can be established, the notion of an image also has to be defined. Morse functions have been suggested by many authors (Griffin and Colchester, 1995) and in chapter 11 as basis for describing images. The Morse functions form an open<sup>1</sup> and dense<sup>2</sup> subset in  $C^\infty$ . This implies that (1) an infinitesimal perturbation can transform any given function into a Morse function, (2) picking a scalar function at random yields a Morse function with probability one and (3) a perturbation of a Morse function yields a Morse function with probability one. The justification for using Morse functions in the analysis is that the difference between the real measurement and the corresponding Morse function is infinitesimal. The result of any analysis should vary continuously with the input, i.e. an infinitesimal change of input should give an infinitesimal change of the result. The Morse lemma states that the qualitative structure of a function can locally be described by its second order derivatives. The critical points of Morse functions are countable in number, isolated and non-degenerate. If the function depends on control parameters, however, its structure may change abruptly after a smooth variation of the parameters. The precise way in which this happens is the subject of catastrophe theory. The function can in general be split into a Morse and a non-Morse part (the catastrophe germ). The Morse part is not qualitatively influenced by a perturbation, but the non-Morse part generally is. This latter part may split up in several Morse parts, and the result is a qualitatively change of the function, the event of which is called a “Catastrophe”. Thom’s Classification Theorem (see Lindeberg (Lindeberg, 1994e) for a condensed version) lists the canonical form of the germs and the perturbation needed to cause the catastrophe for up to four control parameters. These events of sudden change in structure are generic and cannot be removed by a mere perturbation of the initial function. It is expected that on the traversal through function space one passes non-Morse functions albeit the non-Morse functions are countable in number. An frequently studied example for scale-space theory (one control parameter) is the generic event of an annihilation or creation of a extremum/saddle pair in the image (see chapters 11 and 10).

## 14.3 Watersheds and Catchment basins

The notion of *watersheds* and *catchment basins* arises when a function is viewed as a topographic relief with height identified with the image’s scalar value. The watersheds are the boundaries between areas that drain to different local minima. The area draining to one minimum is referred to as the catchment basin. The duality between regions and minima is fully exploited in the analysis of the deep structure. In the following we study functions  $f$  which are sufficiently regular to be differentiated ( $f \in C^2$  is sufficient). Similar to Najmans maximal lines (Najman and Schmitt, 1994), we define:

---

<sup>1</sup>A subset  $X$  is open if and only if  $\forall x \in X \exists \epsilon > 0 : N_\epsilon(x) \subset X$ , where  $N_\epsilon$  is the neighbourhood.

<sup>2</sup>A subset  $X$  is dense in  $Y$  if and only if  $\forall y \in Y \forall \epsilon > 0 : N_\epsilon(y) \cap X \neq \emptyset$

**Definition 14.3.1 (Slope line).** A path  $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$  is called a slope line if

$$\begin{aligned}\forall s \in \mathbb{R}, \gamma_s(s) &= \pm \nabla f(\gamma(s)) \neq 0 \\ \lim_{s \rightarrow -\infty} \gamma_s(s) &= \lim_{s \rightarrow +\infty} \gamma_s(s) = 0\end{aligned}$$

A slope line is descending if  $\forall s \in \mathbb{R}, \gamma_s(s) = -\nabla f(\gamma(s))$ , otherwise it is ascending.

The symbol  $\gamma_s(s)$  is the derivative of the path with respect to the path parameter. Note that critical points are connected with slope lines but not part of any slope line. In general slope lines are the well known steepest descent lines. In 2D domain, the common descending slope lines run from a maximum to minimum, the rare kind of slope lines, the so-called *separatrices*, starts or ends at a saddle. Figure 14.1 shows a mesh plot (a) and a contour plot (b) of a surface patch. The slope lines run perpendicular to the contours. Subfigure (c) shows schematically the critical points and the separatrices.

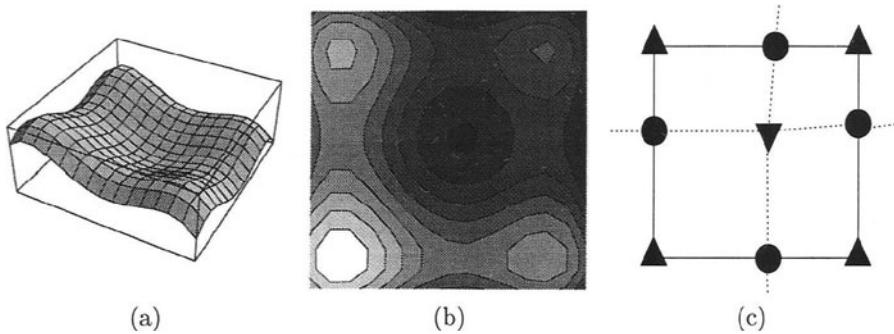


Figure 14.1: A mesh plot (a) and a contour plot (b) for a simple surface patch is illustrated. The positions of the critical points and the separatrices are projected to the domain in subfigure (c). Maxima, minimum and saddles are in (c) indicated with respectively upward triangles, downwards triangle and circles.

**Definition 14.3.2 (Partial ordering of critical points).** Let  $a$  and  $b$  be two critical points of  $f$ . We define a partial ordering of the critical points of  $f$  by saying that  $a$  is above  $b$  if there exists a descending slope line linking  $a$  to  $b$ . If there is a sequence of critical points  $a_i$  such that  $a_i$  is above  $a_{i+1}$  we say that  $a_0$  is above  $a_n$  for  $n > 0$ . In this way a partial ordering of the critical points is established.

Hence, in figure 14.1 all maxima are above two saddles and the one minimum, and all saddles are above the minimum.

**Definition 14.3.3 (Catchment basins).** A point belongs to a catchment basin for a minimum  $M$  if one of the following three conditions is fulfilled.

1. The point is  $M$ .
2. The point is on a slope line which is connected to  $M$ .

3. The point is a critical point which is above exactly one minimum, namely M.
4. The point is on a slope line which is connected to one of the critical points fulfilling condition 2.

**Definition 14.3.4 (Watersheds).** The watersheds form the boundaries between catchment basins. Let  $P(f)$  be the subset of the critical points which are above more than one local minimum of  $f$ . Let  $S(f)$  be all slope lines which connect two points from  $P(f)$ . The watersheds for  $f$  are the union of the points in  $P(f)$  and the points on a slope line in  $S(f)$ .

Hence in figure 14.1 a watershed runs from a maximum to a saddle to a maximum to a saddle etc (the solid line). The critical points outside  $P(f)$  consists of two groups: the local minima of  $f$  and the saddles and maxima which are only above one minimum. From definition 14.3.4 it is seen that the watersheds form a subset of the separatrices.

An important property of the watersheds is the representation of global structure with the implication that watersheds are not in general locally detectable. Formally stated from Najman (Najman and Schmitt, 1994);

**Proposition 14.3.5.** Let  $b$  be a point of the domain of  $f$  such that  $\nabla f(b) \neq 0$ . Let  $N_b$  be a neighbourhood of  $b$  which does not contain any critical point. Let  $\gamma$  be a path containing  $b$  and parallel to the gradient of  $f$  on  $N_b$ . Then there exists a function  $f_0$ , equal to  $f$  on  $N_b$  such that  $\gamma$  is in the watershed of  $f_0$ .

Hence, a gradient field patch without critical points in it does not contain enough information for determining whether or not a watershed is present. The watersheds are basically a limiting set of the slope lines. This probing of the global topology gives a stable detection.

Another important property especially interesting for segmentation is the fact that *watersheds form closed curves for Morse functions*. Hence the watersheds of a function give a full partitioning of the domain; there is no need for closing or connecting edges to get a partition. The closing of edges in a consistent way is one of the problems which has to be solved in segmentation methods based on edge detection (Ayache, 1995).

## Watersheds for the gradient magnitude

The idea of segmenting by watersheds of the gradient magnitude image is well known from the field of morphological segmentation (Griffen, 1995). Hence, there is a segment for each minimum in the gradient magnitude. The watersheds corresponds to the edges in the image. The global probing of structure is one of reasons that watersheds of the gradient magnitude handle junctions fairly gracefully: In local edge-detection with zero-crossings the edges so to speak fight for the right to the junction, with watershed on the contrary the edges are founded as a limit and simply merge together at junctions.

The singularities (and hence the minima) in the gradient magnitude come in two flavours formally stated by the following equations, respectively:

$$L_x = 0 \wedge L_y = 0 \quad (14.1)$$

$$\sqrt{L_x^2 + L_y^2} \neq 0 \wedge L_{ww} = 0 \wedge L_{wv} = 0. \quad (14.2)$$

In the latter equation, the so-called “gauge coordinates” (v,w) have been introduced, which are Cartesian coordinates relative to a frame in which the w-axis is aligned with the gradient (and thus the v-axis is tangential to the iso-greylevel curve). The latter conditions imply that the slope line curvature  $\mu = \frac{-L_{vw}}{L_w}$  and the relative variation of the scalar field  $L_w$  along slope lines  $\delta = \frac{-L_{ww}}{L_w}$  must equal zero (see (Florack et al., 1993a; Florack et al., 1994a) for elaboration on gauge coordinates and invariants). In other words there is either a singularity or the slope lines are straight and the isophotes are equally spaced along the slope lines.

When the image evolves according to the diffusion equation the squared gradient magnitude follows a certain evolution scheme. In tensor notation (Florack et al., 1993a) and Einstein summation convention we have:

$$\frac{\partial(L_i L_i)}{\partial t} = 2L_i L_{ikk} = \partial_{jj}(L_i L_i) - 2L_{ij} L_{ij} \quad (14.3)$$

The first term on the right hand side of equation 14.3 is the Laplacian of the squared gradient magnitude. From the last term it shows that the squared gradient magnitude does not evolve according to the diffusion equation so the principle of causality does not apply.

## 14.4 Catastrophe theory applied

In this section we analyse how the singularities of the magnitude of the gradient can split and merge with scale. Usually one has a family of functions over a multi dimensional domain with possibly degenerate points. The analysis then starts by extracting the degenerated directions and ignoring the non-degenerated, using the *splitting lemma* (Saunders, 1980). For simplicity we will take a different approach and only study the one dimensional case from the beginning, but note that adding Morse parts in perpendicular directions will bring the catastrophe into a higher dimensional space. The image L is expanded in the heat polynomials, see chapter 10 and 11, to fourth order:

$${}^4\tilde{L} = L + L_x x + \frac{1}{2!}L_{xx}(x^2 + 2t) + \frac{1}{3!}L_{xxx}(x^3 + 6xt) + \frac{1}{4!}L_{xxxx}(x^4 + 12x^2t + 24t^2) \quad (14.4)$$

Equation 14.4 is our local model of the image written in an arbitrary coordinate system. The coefficients are image derivatives evaluated at ( $x = 0, t = 0$ ). The derivatives and, in particular, the gradient magnitude can easily be calculated as:

$${}^4\tilde{L}_x = L_x + L_{xx}x + \frac{1}{2!}L_{xxx}(x^2 + 2t) + \frac{1}{3!}L_{xxxx}(x^3 + 6xt) \quad (14.5)$$

so that

$$\begin{aligned} {}^4\tilde{L}_x^2(x, t) = & L_x^2 + 2L_x L_{xxx}t + L_{xxx}^2t^2 + \\ & (2L_x L_{xx} + L_{xx} L_{xxx}t + 2L_x L_{xxxx}t + 2L_{xxx} L_{xxxx}t^2)x \\ & (L_{xx}^2 + L_x L_{xxx} + L_{xxx}^2t + 2L_{xx} L_{xxxx}t + L_{xxxx}^2t^2)x^2 \\ & (L_{xx} L_{xxx} + L_x L_{xxxx}/3 + 4L_{xxx} L_{xxxx}t/3)x^3 \\ & (L_{xxx}^2/4 + L_{xx} L_{xxxx}/3 + L_{xxxx}^2t/3)x^4 + O(x^5) \end{aligned} \quad (14.6)$$

Now we make a simple transformation  $x = z + u(t)$  to eliminate the first order term in equation 14.6. In the  $z$  coordinate system we move along with the singularity when varying  $t$ :

$${}^4\tilde{L}_x^2(z, t) = f_0(t) + f_2(t)z^2 + f_3(t)z^3 + f_4(t)z^4 + f_5(t)z^5 + f_6(t)z^6 \quad (14.7)$$

Equation 14.7 correspond to equation 14.6 after the transformation. The  $f_k(t)$  are specific combinations of the derivatives and  $t$ . The expression for  $u$  is determined by substituting  $x$  with  $z + u(t)$  in equation 14.6 and then determine  $u$  so the first order term in  $z$  vanish. Now the coefficient  $f_0(t)$  can be removed by a mere translation of the axis for  ${}^4\tilde{L}_x^2(z, t)$  and is not considered from here on. If the coefficient  $f_2(t)$  equals zero for a real value of  $t$  then a catastrophe will occur for this  $t$ -value. Hence, it is interesting to solve:

$$f_2(t) = 0 \quad (14.8)$$

Equation 14.8 has four solutions, two real valued ( $t_1$  and  $t_2$ ) and two complex valued ones. Equation 14.7 takes the following form for  $t = t_1$ , say:

$${}^4\tilde{L}_x^2(z, t_1) = f_3(t_1)z^3 + f_4(t_1)z^4 + f_5(t_1)z^5 + f_6(t_1)z^6 \quad (14.9)$$

The second order structure disappears leaving only third order structure. This is known as the *fold catastrophe* where a maximum/minimum pair is created or annihilated. Inserting the  $t_2$  solution in equation 14.7 yields a different result. The special structure of the gradient magnitude image gives this dependency. The equation  $f_3(t) = 0$  has three solutions which happen to coincide with the three of the four solutions to equation 14.8:

$${}^4\tilde{L}_x^2(z, t_2) = f_4(t_2)z^4 + f_5(t_2)z^5 + f_6(t_2)z^6 \quad (14.10)$$

This event is known as a *cusp catastrophe* where all terms up to fourth order disappear.

We have proved that the fold and cusp catastrophes occur generically. Now let us make the Taylor expansion in a point where a catastrophe takes place. Let us look at the general expansion in equation 14.6, so we have the catastrophe at  $(x, t) = (0, 0)$ . Since the point is critical and degenerate the first and second order derivatives must vanish:

$${}^4\tilde{L}_{xx}^2(0, 0) = 2L_x L_{xx} = 0 \quad (14.11)$$

$${}^4\tilde{L}_{xxx}^2(0, 0) = L_{xx}^2 + L_x L_{xxx} = 0 \quad (14.12)$$

We have two solutions  $L_{xx} = 0 \wedge L_x = 0$  and  $L_{xx} = 0 \wedge L_{xxx} = 0$  corresponding to a degeneracy of the singularities described in section 14.3, equation 14.1 and 14.2. For  $L_{xx} = 0 \wedge L_x = 0$  equation 14.6 reduces to this:

$$\begin{aligned} {}^4\tilde{L}_x^2|_{L_{xx}=0 \wedge L_x=0}(x, t) &= L_{xxx}^2 t^2 + 2L_{xxx} L_{xxxx} t^2 x \\ &\quad + (L_{xxx}^2 t + L_{xxxx}^2 t^2) x^2 + (4L_{xxx} L_{xxxx} t / 3) x^3 \\ &\quad + (L_{xxx}^2 / 4 + L_{xxxx}^2 t / 3) x^4 + O(x^5) \end{aligned} \quad (14.13)$$

The structure for this function in the neighbourhood of  $(x, t) = (0, 0)$  is determined by differentiating equation 14.13 and keeping significant terms:

$$({}^4\tilde{L}_x^2|_{L_{xx}=0 \wedge L_x=0})_x(x, t) = 2L_{xxx}^2 t x + 4L_{xxx} L_{xxxx} t x^2 + L_{xxx}^2 x^3 \quad (14.14)$$

The number of singularities can now be found by setting the l.h.s. of equation 14.14 equal zero and solve. For  $t \leq 0$  there are three real-valued solutions; they coincide for  $t = 0$ . For positive  $t$  values there is only one real-valued solution. Hence, for  $t = 0$  we pass a catastrophe, going from a minimum/maximum/minimum configuration into a single minimum. When an extra dimension is added the reverse event is also possible.

For  $L_{xx} = 0 \wedge L_{xxx} = 0$  in equation 14.6 we have:

$$\begin{aligned} {}^4\tilde{L}_x^2(x, t) &= L_x^2 + (2L_x L_{xxxx} t + L_{xxxx}^2 t^2) x + (L_{xxxx}^2 t^2) x^2 + (L_x L_{xxxx} / 3) x^3 \\ &\quad + (L_{xxxx}^2 t / 3) x^4 + O(x^5) \end{aligned} \quad (14.15)$$

The same kind of strategy as in the previous case can be applied. The conclusion is that for  $t$ -values less than zero the function has a maximum and a minimum. For  $t$ -values larger than zero the function is without critical points.

The analysis has shown two types of events:

1. A degenerate critical point in the image corresponds to the cusp catastrophe in the gradient magnitude
2. and vanishing second and third order structure in the same direction in the image corresponds to the fold catastrophe in the gradient magnitude.

## 14.5 The segmentation tool

In the previous sections we established the superficial and deep structure of the gradient magnitude. The duality between segments on the one hand and the minima of the gradient magnitude on the other hand suggests the linking scheme for the segments. The generic events are merging, splitting, annihilations and creations of segments. In the cases of annihilation and merging two minima and a saddle are reduced to one minimum, corresponding to the disappearing of a border between the two segments. The cases of splitting and creation are the reverse events where the emerging saddle corresponds to the appearing of a border

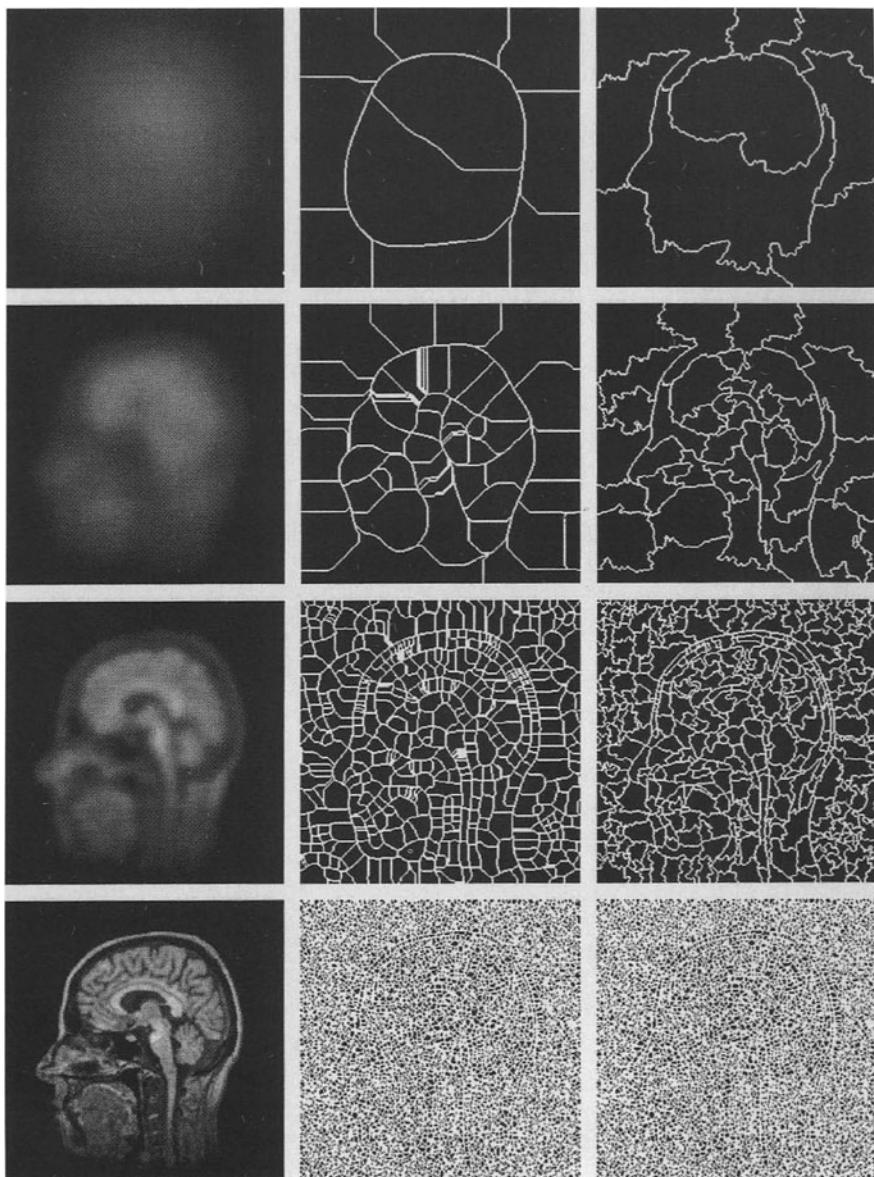


Figure 14.2: Each row corresponds to the same scale (bottom-up we have the  $\sigma$ -values 0.750, 4.59, 11.9, 30.9 pixels). The leftmost column displays slices of the scale space. The original image was  $256 \times 256$  pixels and is a slice of a 3D MR scan of a head. The middle column shows the watershed segmentation of the left column at detection scale. The right column shows the segments at localization scale caused by the segments at the detection scale.

between the segments (dual to the two minima). Hence, the linking is in all cases given by the saddle connecting the involved minima.

The segmentation tool is implemented in the following way. For each scale a segmentation is defined as the watersheds of the gradient magnitude image. The basic segments of the image are the watersheds of the gradient magnitude at the localization scale. A segmentation at a coarse scale (*detection scale*) is tracked to the *localization scale* using the linking of objects. Hence the user specifies two scale values and gets an automatic partitioning of the domain into structurally meaningful parts.

The segmentation can be probed for a number of different detection scales. An equidistant quantization in the natural scale parameter  $\tau$  defined by the reparametrisation:  $\sigma = \sigma_0 \exp(\tau)$ ,  $\tau \in (-\infty, \infty)$  is used to obtain scale invariance (Florack et al., 1992b). The implementation of the linking scheme uses the fact that image structure changes smoothly with scale, and therefore a spatial maximum correlation between segments at neighbouring scales can be used as linking criterion. A similar idea was used for linking of blobs by Lindeberg (Lindeberg, 1994e).

Figure 14.2 shows a sequence of blurred versions of an image for  $\sigma$ -values equal to respectively 0.750, 4.59, 11.9 and 30.9 pixels. The watersheds of the gradient magnitude of the image are shown in the middle column. In the rightmost column, the boundaries of the middle column have been tracked down to the localization scale using the linking. The width of the Gaussian is increased by 10 percent between each slice of the scale-space image.

In figure 14.3 is the segmentation at localization scale  $\sigma = 0.75$  for two detection scales  $\sigma = 4.59$  and  $11.9$  superimposed on the image at scale  $\sigma = 0.75$ . The huge over-segmentation at fine scale is reduced tremendously when the scale of interest is raised above noise level.

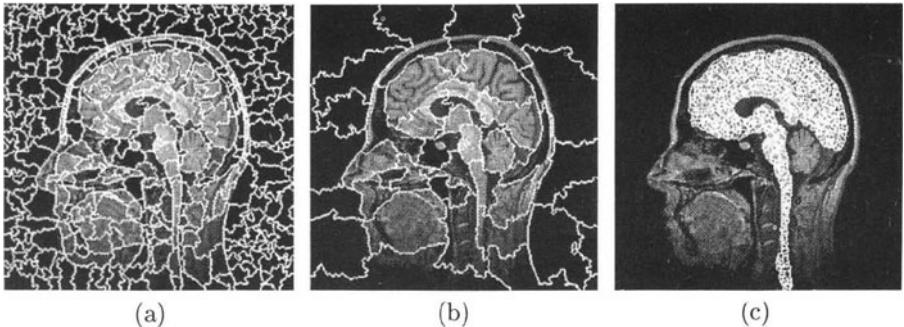


Figure 14.3: The image at scale  $\sigma = 0.75$  with the projected watershed segmentation from scale (a)  $\sigma = 4.59$  and (b)  $\sigma = 11.9$  superimposed. Selected segments (c) shown at localization scale  $\sigma = 0.75$  superimposed on the image.

In figure 14.3 (c) the segmentations from figure 14.3 (a,b) have been used for selecting regions of interest by clicking the mouse on the relevant area. This selection scheme is much faster than drawing the boundaries explicitly and the

required user interaction is much simpler than selecting from the overwhelming number of segments at localization scale. The MR segmentation is meant as an illustration of the multi-scale watershed technique. Better results can be obtain when doing the segmentation in fully 3D. The multi-scale watershed segmentation has been developed and implemented in full 3D by the author, see (Olsen, 1996).

## 14.6 Conclusions

A general framework for multi-scale segmentation has been presented. The idea of partitioning the domain by the watersheds of the gradient magnitude has been analysed and implemented in the case of linear Gaussian scale space. The generic events for the gradient magnitude is the annihilation and creation of saddle/extremum pairs and the transition of a minimum/saddle/minimum from/to a minimum corresponding to a merge or a split of two segments. The selection mechanism of interactively picking objects at an appropriate scale and combining the result at localization scale provides a fast way of doing semi-automatic segmentation.

The semantics of the segmentation can be changed by using another measure of dissimilarity instead of the gradient magnitude or a different diffusion scheme. This is possible within the same general framework although the linking structure will change. The tools for analysis from catastrophe theory are not confined to the presented example but are broadly applicable and useful when studying qualitative and quantitative structure of scalar functions on  $\mathbb{R}^n$ .

## **Part IV**

# **Non-linear Extensions**

# Chapter 15

## The Morphological Equivalent of Gaussian Scale-Space

Rein van den Boomgaard and Leo Dorst

*Faculty of Mathematics & Computer Science, University of Amsterdam  
Kruislaan 403, NL-1098 SJ Amsterdam, The Netherlands*

### 15.1 Introduction

An image is the result of a physical measurement, e.g. the luminance on the retina or the distance from the observer to the “depicted” objects. All physical measurements are the result of the interaction of a measurement probe of finite spatial and temporal size with the physical world. The size of the measurement probe determines at what scale the world is observed. This observation scale is often called the inner scale as it is proportional to the size of the smallest details that can be meaningful distinguished in the image.

Linear scale-space is based on a local *density* measure of the physical entity of interest, e.g. the luminance on a retina (be it either the retina in a human eye or the sensitive area in a CCD camera). Such a density measure, although scale-dependent, measures a local (point) property. As Florack(Florack, 1993) pointed out: linear scale-space theory gives a physical meaning to the mathematical notion of a point.

Linear scale-space also gives a physical meaning to the notion of the local differential geometry of the observations (as a function of position and scale). The mathematical tool of differentiation can be given a robust interpretation as a scale-dependent measurement. Linear scale-space theory, i.e. the combination

of scale-dependent density observations and the differential geometrical structure of the observations as measured through the N-jet, thus provides the computer vision scientist with robust tools to detect local image features (loosely defined as points in the image with a geometrical structure corresponding with a particular semantic meaning attached to it) and to select the appropriate scale to look for a particular feature (through the analysis of the multi-scale behavior of the feature).

With the immense success of linear scale-space theory in providing a unifying view, robust implementations and a common language to think about local image geometry, also came the use of the scale parameter not only to identify local features but to detect essentially multi-local features. In our opinion however, multi-scale is not equivalent to multi-local and thus enlarging the scale in local operators to capture multi-local structure is questionable.

In this chapter we look at a truly non-linear formalism: *mathematical morphology*. In contrast to most of the non-linear extensions of scale-space we still adhere to the notion of an uncommitted visual front-end. Mathematical morphology is based on the morphological superposition principle which is a truly non-linear one (see section 15.2). This however does not mean that there is no connection with physical measurements. In fact the image formation process of the scanning probe microscope is best described with a morphological dilation (Keller, 1991). In general, tactile sensing based on the contact of (non-deformable) surfaces is best described by morphological superposition. Contact probing and thus mathematical morphology is closely connected to multi-local geometry as morphological operators transport local geometry in images from one point to another.

In this paper the above observations serve as the motivation for a closer look at morphological scale-space. The Gaussian linear scale-space is well accepted as the unique *linear* multi-scale analysis of images that satisfies basic principles of symmetry. Several derivations of the uniqueness of the Gaussian kernel are known (see chapter 4). Less well known is that the same symmetry principles, but tied to a morphological (i.e. non-linear) superposition principle, uniquely lead to the morphological scale-space generated by erosions/dilations using parabolic structuring functions (the morphological equivalents of convolution kernels).

It is satisfying that the morphological scale-space bears great resemblance with the linear scale-space (from the algebraic point of view), also the proofs involved are very alike. The elegant way to derive the linear scale-space operator is in the frequency domain, whereas the simplest way to derive the morphological scale-space operator is in the *slope domain*, being the morphological equivalent of the frequency domain. Section 15.2 gives a short introduction to the morphological operators and section 15.4 introduces tangential dilation and the slope transform (which is the morphological equivalent of the Fourier transform). Section 15.5 then uses these results to derive the morphological scale-space operator from the basic principles.

In section 15.6 it will be shown that the morphological scale-space is generated by a partial differential equation similar to Burgers' equation (which describes the propagation of shock waves) in which the scale has the interpreted as the "distance" that points have traveled from their original position. The connection of morphological operators to the differential structure of images is only recently studied

extensively(Brockett and Maragos, 1994; Boomgaard and Smeulders, 1994). The work of Kimia(Kimia et al., 1990) also contains the morphological PDE's as he was the first to combine the linear diffusion with the morphological shock-wave propagation. The morphological infinitesimal generators do not only prove that morphological scale-space indeed can be generated ("grown") but they also provide us with a different view to well-known PDE's and image processing algorithms.

Having said all the above, the goal set forward in this chapter is rather modest. In the first three sections after this introduction a short introduction to mathematical morphology, including the slope transform, will be given. Those familiar with mathematical morphology, can jump to section 15.5 where the morphological scale-space operator is derived from basic principles. In section 15.6 the PDE governing morphological scale-space will be derived.

## 15.2 Morphological Operators

The fundamental difference between linear image processing and morphological image processing (from an algebraic point of view at least) is the superposition principle chosen. For a linear system the requirement is that the linear transform of a multiplicative weighted sum of images should be equal to the weighted sum of the individually transformed images. A morphological system is characterized by additive weighting and the use of the supremum (maximum) or infimum (minimum) superposition of images.

An image transform  $\Psi$  is said to be a *dilation* if and only if it distributes over the supremum (denoted with  $\vee$ ), i.e.  $\Psi$  is a dilation iff:

$$\Psi \left( \bigvee_i a_i + f_i \right) = \bigvee_i (a_i + \Psi(f_i)), \quad (15.1)$$

where  $f_i$  is some collection of images and the  $a_i$  are real scalars. Equivalently a transform is said to be an *erosion* if it distributes over the infimum of images.

One of the basic principles underlying both linear scale-space and morphological scale-space is spatial homogeneity (see section 15.5). This is satisfied in case the scale-space operator is translation invariant, i.e.  $\Psi(f_t) = (\Psi f)_t$  where the morphological convention is used that translation of functions is denoted by subscripting:  $f_t(x) = f(x - t)$ .

Any (finite valued) image  $f$  can be decomposed into impulses  $f = \bigvee_y f(y) + \Delta_y$  where  $\Delta$  is the *morphological pulse* (defined by  $\Delta(x) = 0$  if  $x = 0$  and  $-\infty$  elsewhere). Using this pulse representation of an image  $f$  and using the morphological superposition principle one obtains:

$$\Psi(f) = \Psi \left( \bigvee_y f(y) + \Delta_y \right) = \bigvee_y f(y) + \Psi(\Delta_y)$$

Translation invariance of  $\Psi$  implies that  $\Psi(\Delta_y) = (\Psi\Delta)_y$ . The function  $\psi = \Psi\Delta$  is called the *impulse response* (or kernel or structuring function) of the translation invariant morphological dilation  $\Psi$ . It fully characterizes  $\Psi$ .

The above defined interaction of the kernel  $\psi$  with the image under study  $f$  is called the dilation of  $f$  with structuring function  $\psi$  and is denoted as

$$(f \oplus \psi)(x) \equiv \bigvee_y f(y) + \psi(x - y)$$

Equivalently, the erosion is defined as:

$$(f \ominus \psi)(x) \equiv \bigwedge_y f(y) - \psi(x - y),$$

where  $\wedge$  denotes the infimum operator. The signs are chosen such that erosion and dilation form an adjunction in the sense that the operators are dual with respect to complementation of the images:  $f \oplus \psi = -(-f \ominus \check{\psi})$ , where  $\check{\psi}$  denotes the transpose (or mirroring) of  $\psi$ :  $\check{\psi}(x) = -\psi(-x)$ .

The erosion and dilation are *not* inverse operators, a dilation followed by an erosion is not the identity operator. Actually, it is precisely this fact that leads to many interesting and useful morphological operators. An erosion followed by a dilation is called a (structural) *opening* and is often denoted as  $f \circ \psi = (f \ominus \psi) \oplus \psi$ . The reverse, a dilation followed by an erosion, is called a (structural) *closing*  $f \bullet \psi = (f \oplus \psi) \ominus \psi$ . See the example application discussed in the introduction in which an image was closed to “fill” the narrow dark lines with the background color. The above definitions of dilation and erosion lie at the core of a vast amount of dedicated literature. For the interested reader we refer to the books of Serra(Serra, 1982; Serra, 1988) and Heijmans(Heijmans, 1994).

In this chapter we will not pursue the algebraic properties of the morphological operators any further. Instead we will look in more detail at the differential geometrical interpretation of the morphological operations. Whereas linear operators are often used to probe the local geometrical structure of the image data (through the N-jet calculation using Gaussian derivative kernels), morphological operators change the geometrical structure by transporting local geometry from one point to another. Mathematical morphology thus provides tools to probe the multi-local structure of images.

Many of the morphological operators indeed can be identified as detecting (and/or changing) the multi-local structure of images. Using a closing (a dilation followed by an erosion) for example to “wipe out” the thin black lines on a drawing indeed is a multi-local operator (see Fig. 15.1). First the edges of the lines are transported towards the middle of the line. The traveling distance is determined by the size of the dilation kernel (i.e. the size of the structuring element). In case this size is chosen appropriately, both inwards traveling edges “annihilate” thus destroying the multi-local geometry that was represented by the two opposite edges of the line. An erosion is then used to transport the remaining geometry back to their original position. In this process of transportation, annihilation and inverse transportation, the narrow black lines are destroyed. Note that the grey values within the line are irrelevant. Note in the above discussed application of detecting and localizing lines the explicit searching for the edges would not be necessary in case we are able to mark the points where the opposite edges of the line annihilate.

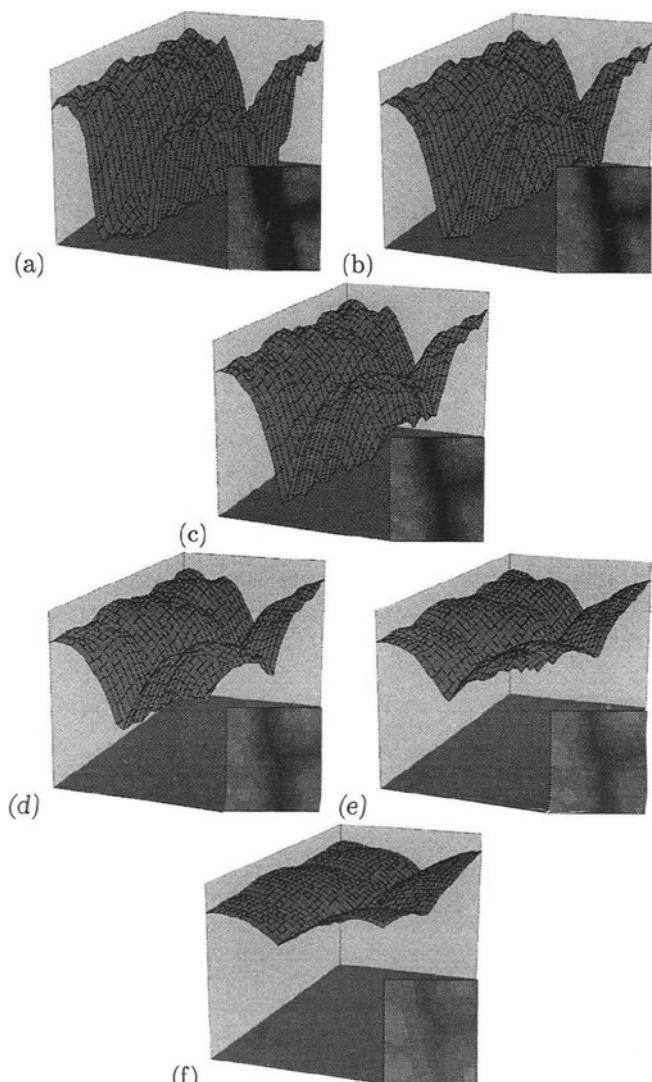


Figure 15.1: Morphological Dilation Scale Space. (a) A detail from a drawing, (b)-(f) Parabolic dilations of the image in (a) with parabolic structuring functions of increasing width (scale).

This of course is what the medial axis (or skeleton) algorithm does. Another example of a multi-local morphological operator is watershed segmentation(Vincent and Soille, 1991). The watershed algorithm can be used to find the ridges and ravines in images which are *not* truly local geometrical features.

### 15.3 The Geometrical Interpretation of Morphological Operators

Erosion and dilation form the basics of mathematical morphology. Whereas the convolution, being the basic operator in linear signal and image processing, can be interpreted in physical terms as a density probe, the morphological operators should be interpreted as contact probes. Consider the dilation of the function  $f$

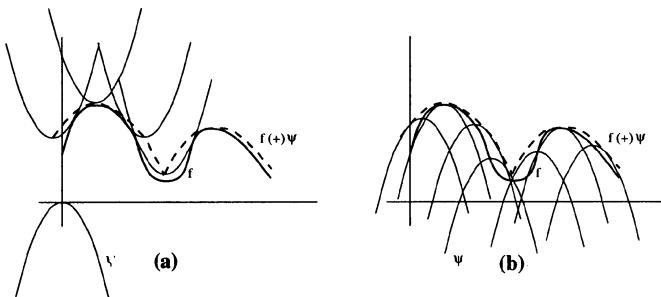


Figure 15.2: Dilation. (a) Hit-property interpretation and (b) Union-of-translates interpretation

depicted in Fig. 15.2.a with structuring function  $\psi$ . Geometrically this can be interpreted as follows. At position  $x$ , lower the graph of the mirrored structuring function  $\check{\psi}$  until it hits the graph of the function  $f$ . The origin of the graph of the structuring function  $\check{\psi}$  marks the point on the dilated function. The erosion is equally simple to interpret geometrically. Instead of lowering the structuring function from above, it has to raised from below (or just turn the paper upside down).

A second interpretation of the dilation is the *wave front propagation* interpretation, see Fig. 15.2.b. Any point  $(x, f(x))$  on the graph of the function  $f$  can be interpreted as the source of a wavefront as described by the structuring function  $\psi$  (translated to the point  $(x, f(x))$ ). Huygens' principle states that the initial wave front (the function  $f$ ) is moved to the curve formed by taking the envelope of all individual wave fronts. Within the functional description, taking the envelope is equivalent to the supremum of all translated versions of the structuring function (with translation vectors taken from the graph of  $f$ ).

These two interpretations are known in mathematical morphology as the *hit-property* interpretation and the *union-of-translates* interpretation, respectively. From a computer scientist's point of view these two interpretations can be termed

the *read* and *write* formalisms. In the first case, at each point in the image the information from all neighbouring points is collected (read) and combined into the new value for this point. In the write formalism each point in the image distributes its information over all points in its neighbourhood. The read/write formalism of course also applies to linear operators (but with no significant practical advantage of preferring one above the other). For morphological operators choosing either the read or write formalism will lead to quite different algorithms. Most often the write formalism allows one to carefully select a restricted number of neighbourhood points to write to and to process only those points in image which need to be processed. Most of the efficient algorithms in morphology indeed use some kind of wave-front propagation in such a way to ensure that every pixel is ‘visited’ as few times as possible. See Verwer and Van Vliet(Vliet and Verwer, 1988) and Vincent(Vincent and Soille, 1991) for some nice examples of morphological wave-front propagation algorithms.

## 15.4 Tangential Dilation and the Slope Transform

A translation invariant *linear* operator preserves the frequency of signals. Thus a sinusoidal function (a single frequency signal) is not changed by a linear filter, with the exception of a possible change in amplitude and phase. A translation invariant morphological operator preserves the slope of signals. Indeed a “single slope signal”—the planar function—is left unchanged by any morphological operator, with the possible exception of a translation of the planar function. Thus the planar functions are the *eigenfunctions* of mathematical morphology.

This observation led to the definition of the slope transform that is for morphological operators what the Fourier transform is for linear filters: the invertible decomposition of a signal into its constituent eigenfunctions. The slope transform is introduced in (Dorst and Boomgaard, 1994a). The slope transform is a generalization of the Legendre transform. Maragos (Maragos, 1995) independently introduced the A-transform which is a non-invertible simplification of the slope transform.

The invertibility of the slope transform comes at a price though. The slope transform provides a representation of signals in which the morphological spatial operators become pointwise operations (in the same way that the convolution becomes a pointwise multiplication in the frequency domain). It does so for the *tangential dilation* which is generalization of the classical dilation (which is contained in the tangential dilation).

Consider again the geometrical interpretation of the dilation where the structuring function  $\psi$  hits the function  $f$  from above (see figure 15.3). Simple geometrical analysis shows that the slope of the dilated function at  $x$  is equal to the slope of the original function in the *point-of-contact*  $y$ . Due to the touching contact, this slope can also be found in the structuring function at position  $x - y$ . This property is called the parallel slope transport(Boomgaard and Smeulders, 1994). Given the slope in a point on the dilated function, it is possible to calculate where that point was caused by contact on the original function. This tracing back

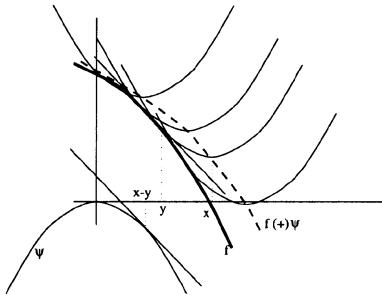


Figure 15.3: Tangential dilation

of points requires that there is a one to one relation between slope and position for the structuring function, i.e. the structuring function  $\psi$  should be concave (in the sense that a straight line segment connecting two points on the graph of the function is not below the graph except for the two given points).

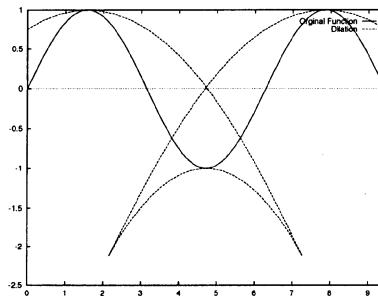


Figure 15.4: Example of Tangential Dilation

Obviously we can also take a point  $(y, f(y))$  on the original function and calculate where that point will move to under a dilation. Again a one-to-one relation between slope and position of the structuring function is needed. The tangential dilation is based on this forward reasoning. Tangential dilation differs from the classical definition of dilation. In classical dilation, not every point on the original function has its associated point in the dilation result. In the tangential dilation these points lead to a connected surface underneath the classical dilation result. An example is given in figure 15.4. The tangential dilation can be defined in a more functional notation as follows:

**Definition 15.4.1 (Tangential Dilation).** The tangential dilation of a function  $f$  with structuring function  $g$  is defined as:

$$(f \check{\oplus} g)(x) = \text{stat}_y [f(y) + g(x - y)]$$

where the stat operator is defined to return the set of stationary values of its argument:  $\text{stat}_uf(u) = \{f(\bar{u}) \mid (\nabla f)(\bar{u}) = 0\}$ .

The tangential dilation is a superset of the classical dilation, the latter of which is obtained by taking the supremum function value at every spatial coordinate  $x$ . The tangential dilation in general is multi-valued when represented as a “function”. Seen as a surface, tangential dilation transforms one differentiable surface into another (albeit a possibly self intersecting one). A more geometrically oriented description and analysis of the tangential dilation (and of the slope transform to be introduced shortly) can be found in (Dorst and Boomgaard, 1994b).

Since a tangential dilation does not change slopes locally, but just translates the point carrying that slope, it follows that it translates a function with a constant slope as a whole. Thus such functions are *morphological eigenfunctions*: they may change their location (which we will describe like amplitude and phase), but not their shape. Obviously such functions are the *planar functions*  $e_\omega(x) = \langle \omega, x \rangle$  (where  $\langle x, y \rangle$  denotes the inner product of  $x$  and  $y$ ).

**Proposition 15.4.2 (Dilating a planar function).** The planar functions  $e_\omega(x) = \langle \omega, x \rangle$  is translated by a dilation:

$$e_\omega \circledast \psi = e_\omega + S[\psi](\omega)$$

where  $S[\psi]$  is a function only dependent on  $\psi$ , which we may consider as an additive ‘phase factor’.

Each eigenfunction  $e_\omega$  thus gets translated by an amount that only depends on the structuring function  $\psi$  (and of course on  $\omega$ ). This amount is characteristic for  $\psi$ . If we can now invertibly decompose arbitrary functions into the planar eigenfunctions, then a dilation by  $\psi$  can be described as the composition of the shifted eigenfunctions. And if we then also decompose  $\psi$  into eigenfunctions, then the description of dilation involves the eigenfunctions only. This idea leads immediately to the *slope transform*.

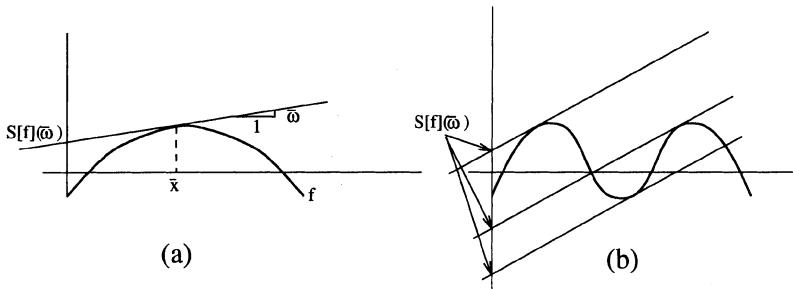


Figure 15.5: The definition of the slope transform for (a) concave and (b) non-concave functions

Fig. 15.5.a illustrates the geometric construction, known as the Legendre transform, on which a dual representation of functions by slopes can be based. At a

point  $(\bar{x}, f(\bar{x}))$  on the concave function  $f$ , the tangent plane is drawn. The slope of this tangent plane is  $\nabla f(\bar{x})$ ; denote it by  $\bar{\omega}$ . With  $\nabla f$  invertible, we thus have  $\bar{x} = (\nabla f)^{-1}(\bar{\omega})$ . Now note the intercept of the tangent plane with the functional axis. This real number equals  $f(\bar{x}) - \langle \bar{\omega}, \bar{x} \rangle$ . We call this the *slope transform* of  $f$  at  $\bar{\omega}$ , denoted by  $S[f](\bar{\omega})$ . Using the ‘stat’ operator, we may write:

**Definition 15.4.3 (Slope Transform).**

$$S[f](\omega) = \text{stat}_x [f(x) - \langle \omega, x \rangle].$$

Since stationarity of  $f(x) - \langle \omega, x \rangle$  at  $\bar{x}$  requires  $\nabla f(\bar{x}) - \omega = 0$ , and hence  $\omega = \nabla f(\bar{x})$ . This definition, motivated for the concave case, actually also serves well for the case where a given slope  $\omega$  is assumed at more than one value  $x$  (see Fig. 15.5b). Therefore we take it as the definition of the slope transform  $S[f]$  for arbitrary functions  $f$ . The slope transform thus defined extends some known transforms: the Legendre transformation, the Young-Fenchel conjugate(Doyen et al., 1995) and the  $\mathcal{A}$ -transform(Maragos, 1995).

Given the intercepts as a function of the slope, the original function can be reconstructed. Geometrically, this amounts to considering the function as the *caustic* of its tangent planar functions. The corresponding formula reconstructs  $f$  by determining the stationary points on the tangent planes when they vary with  $\omega$ . The *inverse slope transform* is given by:

**Definition 15.4.4 (Inverse Slope Transform).**

$$f(x) = \text{stat}_{\omega} [S[f](\omega) + \langle \omega, x \rangle].$$

The property of the slope transform that makes it the morphological equivalent of the Fourier transform is the fact that dilation in the spatial domain becomes addition in the slope domain:

**Proposition 15.4.5 (Dilation in Slope Domain).** Let  $f$  be a function and let  $g$  be a concave structuring function, then:

$$S[f \circledast g] = S[f] + S[g]$$

In the above proposition the addition of  $S[f]$  and  $S[g]$  should be interpreted as the Minkowski addition of the value set  $S[f](\omega)$  and  $S[g](\omega)$ . A geometrical proof is easily constructed in the one dimensional case (details can be found in (Dorst and Boomgaard, 1994a)).

In table 15.1 several properties of the slope transform are given. Especially the scaling and stretching properties will be needed in the derivation of the morphological scale-space operator in section 15.5.

## 15.5 Scale-Space from Basic Principles

In this section we return to the main objective of this chapter which is the derivation of the morphological scale-space operator from basic principles. These principles are the mathematical expression of the lack of priori knowledge about the

Table 15.1: Definition and Properties of the Slope Transform.

tangential dilation	$(f \oplus g)(x) = \text{stat}_u [f(x) + g(x - u)]$
eigenfunctions	$\langle \omega, x \rangle$
slope transform	$S[f](\omega) = \text{stat}_x [f(x) - \langle \omega, x \rangle]$
inverse slope transform	$S^{-1}[F](x) = \text{stat}_\omega [F(\omega) + \langle \omega, x \rangle]$
translation $f(x - a)$	$S[f](\omega) - \langle \omega, a \rangle$
stretching $f(ax)$	$S[f](\omega/a)$
scaling $a f(x)$	$a S[f](\omega/a)$
umbral scaling $a f(x/a)$	$a S[f](\omega)$
central theorem	$S[f \oplus g] = S[f] + S[g]$
band-limitation	dilation with $\ x\ $

physical observation method and about the symmetries in observation space. The principles stated here are the same as those commonly used to derive linear scale-space. An overview of several derivations of the linear scale-space can be found in chapter 4.

Let the image  $f$  be at the ground level in the scale-space. (Theoretically, such a zero-scale image does not exist as it cannot be observed. In practice the observed image is taken to be the zero scale image.) The images at larger scale  $\rho$  will collectively be called a scale-space and denoted by the function  $F$  such that  $F(x, \rho)$  is the grey value at position  $x$  measured at scale  $\rho$ . By definition we have  $F(\cdot, 0) = f$ . The remainder of this section is concerned with the derivation of the scale-space operator  $\Psi^\rho$  needed to construct the ‘observation’ at scale  $\rho$  from the zero scale image:  $F(\cdot, \rho) = \Psi^\rho f$ .

**Superposition principle.** In the introduction to this chapter it was already stated that linear superposition is only one of the superposition principles underlying physical measurement (albeit a very important one). In this chapter we look at morphological superposition (defined in eq. 15.1):

$$\Psi^\rho \left( \bigvee_i a_i + f^i \right) = \bigvee_i (a_i + \Psi^\rho(f^i)),$$

The superscript  $\rho$  in  $\Psi^\rho$  denotes the scale dependency of the scale-space operator.

**Spatial Homogeneity.** In the absence of any a priori preferred position in space the obvious requirement for a scale-space operator is translation invariance. Translation invariance combined with the morphological superposition principle (eq. 15.1) implies, as shown in section 15.2, that the scale-space operator has the algebraic form of a dilation with structuring function  $\psi^\rho = \Psi^\rho \Delta$ :

$$F(\cdot, \rho) = \Psi^\rho f = f \oplus \psi^\rho.$$

**Spatial Isotropy.** Not only all positions are considered equal, also at each point space is assumed isotropic. This requires the scale-space kernel to be rotational symmetric:

$$\psi^\rho(x) = \psi^\rho(\|x\|)$$

**Atlas Principle.** The notion of a zero scale image is a mathematical notion. Physically it has no real significance. In the mathematical model this should be reflected in the fact that the scale-space does not depend on the zero scale. If the image at scale  $\mu$  is taken as the new zero scale image and a new scale-space is constructed from this image (thus obtaining scale operators  $\Psi^\rho \Psi^\mu$ ), this scale space should be compatible with the orginal scale space ( $\Psi^\lambda$ ). This implies that it should make no difference if the scale is increased from zero to  $\rho + \mu$  in one step or that first the scale is increased to an intermediate scale value  $\mu$  and from there increased to  $\rho + \mu$ .

This *atlas principle* corresponds with the mathematical notion of an “additive semi group structure”  $\Psi^\rho \Psi^\mu = \Psi^{\rho+\mu}$ . The atlas principle and the semi-group property are often taken to be meaning the same thing. Note however that a semi-group structure comprises only a collection of objects (in our case the scale operators  $\Psi^\rho$ ) and an internal operator that obeys the associative law (in our case the operator composition  $\Psi^\rho \Psi^\mu$ ). Thus the collection of operators with associative law  $\Psi^\rho \Psi^\mu = \Psi^{\rho+\mu}$  is also a semi-group. However, it does not obey the atlas principle as it depends on the zero-scale. Going from scale  $\rho$  to scale  $\rho + \mu$  depends on the value  $\rho + \mu$  and not just on the difference between the two scales. This “supremum semi-group” corresponds with a morphological operator. The opening and closing using umbral scalings of structuring functions do possess this semi-group property and are indeed proposed in literature as morphological scale-space operators. In this chapter the atlas principle is taken as the basic principle and thus these ‘scale-spaces’ are not considered.

The atlas-principle  $\Psi^\rho \Psi^\mu = \Psi^{\rho+\mu}$  is a strong restriction on scale-space as it implies the existence of an infinitesimal generator that produces the image at scale  $\rho + d\rho$  given the image a scale  $\rho$ . The infinitesimal is independent of the scale. For the linear scale space the infinitesimal generator is known to be the Laplacian (the propagator of diffusion). Given the fact that dilation can be interpreted as geometry controlled translation of points it should come as no surprise that the infinitesimal generator of the morphological scale-space is the gradient. This will be treated in section 15.6

Using  $\Psi^\rho f = f \oplus \psi^\rho$  the semi group property of the operator can be reformulated as a semi group property of the kernel under dilation:

$$\psi^\rho \oplus \psi^\mu = \psi^{\rho+\mu}.$$

In the analysis of the above expression governing the scale dependency of the dilation kernel the classical dilation  $\oplus$  is replaced with tangential dilation  $\circledast$ . These dilations are equivalent only in case the functions involved are concave (in this case  $\psi^\rho$ ). This indeed is the case as will be shown shortly.

When the classical dilation  $\oplus$  is replaced with the tangential dilation  $\check{\oplus}$  and the slope transform of both sides of the above expression is taken, we obtain:

$$\mathcal{S}[\psi^\rho] + \mathcal{S}[\psi^\mu] = \mathcal{S}[\psi^{\rho+\mu}]$$

This is evidently true in case  $\mathcal{S}[\psi^\rho] = \rho \mathcal{S}[\psi]$ . According to one of the properties listed in table 15.1, scalar multiplication in the slope domain amounts to an umbral scaling in the spatial domain. Thus scale dependency of the morphological scale space kernel has the following form:

$$\psi^\rho = \rho \cdot \psi$$

where  $\rho \cdot \psi$  denotes the *umbral scaling* of the function  $\rho$  (defined as  $(\rho \cdot \psi)(x) = \rho\psi(x/\rho)$ ) and  $\psi$  is the scale-space kernel at unit scale. In umbral scaling not only the spatial dimensions are scaled with a factor  $\rho$  but also the grey value dimension is scaled. Essentially the graph of the function, interpreted as a geometrical object, is scaled with respect to the origin.

In the derivation of the scale dependency, classical dilation was replaced with tangential dilation. This is only allowed in case the kernel  $\psi$  is concave. Consider the case that  $\psi$  would be a non-concave function, i.e. has both convex and concave parts. It is a well-known result from mathematical morphology that repeatedly dilating a non-concave function with itself approaches the “concave hull” of the function. Therefore an infinitesimal generator can be taken to be concave; any non-concavities are smoothed out anyway after a finite scale step.

**Scale Invariance.** The scale-space formation should be obviously invariant under changes of scale. We consider two types of scaling: grey scaling  $f \mapsto \alpha f$  and umbral scaling  $f \mapsto \alpha \cdot f$ . The scale-space function  $F$  does change when  $f$  is replaced with  $\alpha f$  or  $\alpha \cdot f$ . It is required however that the change is equivalent with a reparameterization of the scale. The scale-space operator  $\Psi^\rho$  thus is scale invariant iff:

$$\Psi^\rho(\alpha f) = \alpha \Psi^{\Gamma(\rho)} f \tag{15.2}$$

$$\Psi^\rho(\alpha \cdot f) = \alpha \cdot \Psi^{\Lambda(\rho)} f \tag{15.3}$$

where  $\Gamma$  and  $\Lambda$  are arbitrary monotonically increasing functions.

For the linear scale-space grey scale invariance is trivial. It is the umbral scale invariance that determines the shape of the convolution kernel. In the morphological scale-space umbral scaling is trivial and the grey scale invariance determines the shape of the dilation kernel. Grey scale invariance demands that:

$$\Psi^\rho(\alpha f) = \alpha \Psi^{\Gamma(\rho)} f$$

For the morphological operator this is equivalent to:

$$\alpha f \check{\oplus} \psi^\rho = \alpha(f \check{\oplus} \psi^{\Gamma(\rho)}).$$

In the slope domain this gives:

$$\mathcal{S}[\alpha f] + \mathcal{S}[\psi^\rho] = \mathcal{S}[\alpha(f \check{\oplus} \psi^{\Gamma(\rho)})].$$

Using the stretching and scaling properties from table 15.1 and the scale dependency of the scale-space kernel the above can be rewritten as:

$$\alpha \cdot \mathcal{S}[f] + \rho \mathcal{S}[\psi] = \alpha \cdot \mathcal{S}[f] + \alpha \cdot (\Gamma(\rho) \mathcal{S}[\psi]).$$

Rearranging terms, the above leads to:

$$\frac{\alpha \mathcal{S}[\psi](\omega/\alpha)}{\mathcal{S}[\psi](\omega)} = \frac{\rho}{\Gamma(\rho)}.$$

This equality must be true for all values of  $\omega$  and thus both sides of the equality should be a constant. This is true in case  $\mathcal{S}[\psi]$  has the form:

$$\mathcal{S}[\psi](\omega) = a \|\omega\|^b.$$

(Note that the rotational symmetry of the kernel requires that  $\mathcal{S}[\psi]$  only depends on  $\|\omega\|$ ). The choice for the constant  $a$  only amounts to an umbral scaling in the spatial domain and thus to a trivial reparametrisation of scale. The classical choice for  $b = 2$  is a pragmatic one (within this derivation). This choice for  $b = 2$  facilitates efficient implementations as it is the unique value that makes the resultant rotational symmetric kernel separable.

Thus we arrive at the morphological scale-space kernel represented in the slope domain as:  $\mathcal{S}[\psi](\omega) = \|\omega\|^2$ . The inverse slope transform gives the spatial representation  $\psi(x) = -\frac{1}{4}\|x\|^2$ . Summarizing: morphological (dilation) scale-space is generated by parabolic dilations:

$$F(\cdot, \rho) = f \oplus \psi^\rho$$

with

$$\psi^\rho(x) = -\frac{1}{4\rho}\|x\|^2$$

Just like the linear Gaussian scale-space, the parabolic morphological scale-space has been “found” several times by several different authors. Verbeek(Verbeek and Verwer, 1989) was one of the first to look at the scaling properties of poweroids (including the parabola) used as structuring functions in mathematical morphology. Jackway(Jackway, 1992) and van den Boomgaard(Boomgaard and Smeulders, 1992) independently derived parabolic scale-space on the basis of those scaling properties. Separability of morphological structuring functions is discussed in (Boomgaard et al., 1996) where it was also shown that the parabolic structuring functions are indeed the only rotationally symmetric ones that can be separated. A property that the parabola’s have in common with the Gaussian functions in the linear domain.

## 15.6 The Infinitesimal Generator of Scale-Space

Deriving the scale-space operators from symmetry principles is only one way to do it. Several authors start with a causality principle stating that when going towards

larger scale, no new details can be introduced. Obviously the exact definition of what “details” are is the crucial step (zero crossings in the Laplacian for instance works in 1D but not in 2D, the number of extrema doesn’t work either in 2D<sup>1</sup>). Koenderink(Koenderink, 1984) interprets causality as the demand that grey values should be traceable towards smaller scales (a grey value cannot “appear” from nowhere). When doing so, it can be shown that the diffusion equation is the simplest *linear* partial differential equation linking an infinitesimal change in scale with the local differential geometry of the image. It should be noted though that continuity of the N-jet is implicitly assumed in this derivation. Morphological operators do not result in differentiable functions; all that can be said is that the function is Lipschitz(Boomgaard and Smeulders, 1994), i.e. the function is continuous and *almost everywhere* differentiable. Unfortunately the interesting points are the non differentiable ones of course (for instance the annihilation or medial axis points in the line detection application discussed earlier).

The derivation of the scale space operator from symmetry principles defines the macroscopic properties of the scale-space operator, whereas the causality principle defines its microscopic behaviour. Although the microscopic causality principle first introduced by Koenderink also did possess the required macroscopic symmetry properties, this requirement was soon to be abolished. By introducing non-linearities in the diffusion equation (see Perona and Malik(Perona and Malik, 1990)) the macroscopic notion of scale was lost. Scale became the “time” parameter in a non-linear PDE and choosing an appropriate time to stop the evolution became somewhat of an art. Evolutionary processes whose asymptotic behaviour is of more interest than its intermediate states were welcomed into the realm of scale-space theory.

The additive *atlas-principle*  $\Psi^\rho \Psi^\mu = \Psi^{\rho+\mu}$  is a very strong restriction on scale-space as it implies the existence of an infinitesimal generator independent of the scale (the operator that produces the image at scale  $\rho + d\rho$  given the image at scale  $\rho$ ).

It has only been realized recently that there is a close connection between a class of non-linear partial differential equations and the morphological operators (Brockett and Maragos, 1994; Boomgaard and Smeulders, 1994; Doyen et al., 1995). Consider the one-parameter family of images:

$$F(x, \rho) = (f \oplus \rho \cdot \psi)(x)$$

with  $\psi$  a concave function and  $\rho \cdot \psi$  the umbral scaling of  $\psi$  defined as  $(\rho \cdot \psi)(x) = \rho\psi(x/\rho)$ . In umbral scaling not only the spatial dimensions are scaled with a factor  $\rho$  but also the grey value dimension is scaled. Essentially the graph of the function, interpreted as a geometrical object, is scaled.

In this section it will be shown what happens if we change the scale from  $\rho$  to  $\rho + d\rho$ . Because the umbral scaling of any concave function forms a semi group

---

<sup>1</sup>It was shown by Paul Jackway(Jackway and Deriche, 1996) that in morphological scale-space the number of extrema is decreasing

under dilation (i.e.  $\lambda \cdot \psi \oplus \mu \cdot \psi = (\lambda + \mu) \cdot \psi$ ) we can write:

$$\begin{aligned} F(\cdot, \rho + d\rho) &= f \oplus (\rho + d\rho) \cdot \psi \\ &= (f \oplus \rho \cdot \psi) \oplus d\rho \cdot \psi \\ &= F(\cdot, \rho) \oplus d\rho \cdot \psi \end{aligned}$$

As we are interested in the case that  $d\rho$  is infinitesimal small, meaning that  $\rho \cdot \psi$  becomes very sharply pointed and looks like the morphological pulse, we may approximate the function  $F(\cdot, \rho)$  with its tangent line locally around the point  $x$ .

The dilation of a planar function is given in section 15.4 as  $e_\omega \oplus \psi = e_\omega + \mathcal{S}[\psi](\omega)$ . In the case of the tangent plane we obtain:

$$F(x, \rho + d\rho) = F(x, \rho) + \mathcal{S}[d\rho \cdot \psi](\nabla F(x, \rho))$$

Using one of the scaling properties from table 15.1 we have:

$$F(x, \rho + d\rho) = F(x, \rho) + d\rho \mathcal{S}[\psi](\nabla F(x, \rho))$$

Rearranging terms and taking the limit  $d\rho \rightarrow 0$  we obtain:

$$\frac{\partial F}{\partial \rho}(x, \rho) = \mathcal{S}[\psi](\nabla F)$$

This analysis shows that the stack of images generated by dilation with the umbral scaling of a concave structuring function is causal in the scale-parameter. The change in grey value going from scale  $\rho$  to scale  $\rho + d\rho$  is determined by the (first order) differential structure of the image at scale  $\rho$ . It can also be shown that the above PDE (with initial condition  $F(\cdot, 0) = f$ ) is solved by the dilation. A proof, by analogy with a classical proof by Lax (Lax, 1973), can be found in (Boomgaard, 1992a)).

The close connection between this class of PDE's and the morphological operators serves to understand some of the intricate details in the numerical integration schemes needed for robust solution of these type of PDE's. As an example consider the PDE  $F_\rho = \|\nabla F\|$ . According to the above analysis we have to dilate the initial condition  $f(x)$  with the structuring function  $\rho \cdot \psi = \rho \cdot \mathcal{S}^{-1}[\|\cdot\|]$ . The inverse slope transform of the "cone function" is the disk with radius 1 (i.e. the function  $u(x) = 0$  for  $|x| \leq 1$  and  $-\infty$  elsewhere). Thus  $F(x, \rho)$  is obtained by taking the local maximum in a circular neighbourhood with radius  $\rho$  of  $x$ . The infinitesimal generator needed to go from  $\rho$  to  $\rho + d\rho$  is the dilation using a disk with radius  $d\rho$  as structuring function. From the morphological point of view it is thus no surprise that maxima and minima calculations are proposed in the more elaborate numerical solution schemes for these type of PDE's (see for example the discussion in (Weickert, 1996a)).

For the morphological scale-space generated with parabolic dilations:

$$F(\cdot, \rho) = f \oplus \psi^\rho$$

where

$$\psi^\rho(x) = -\frac{1}{4\rho} \|x\|^2$$

we have that:

$$\frac{\partial F}{\partial \rho} = \mathcal{S}[\psi](\nabla F) = \|\nabla F\|^2$$

The parabolic dilation thus is the morphological Green's function for the above partial differential equation.

## 15.7 Conclusions

This chapter expresses the view that morphological scale-space theory is not an alternative for linear scale-space theory in the sense that it provides a different methodology to look at the same image phenomena. Instead it provides a different, multi-local view on images. Despite this difference, the correspondence between the two scale-space approaches bear great resemblance on the algebraic level. Even the proofs of the derivation of the scale-space operator are similar, especially in case the proofs for the morphological scale-space are done in the slope-domain, being the morphological equivalent of the Fourier transform. The slope transform was briefly discussed in this paper as a description of images (i.e. the grey value landscape) in which morphological operations are simplified (like the convolution operation is simplified in the frequency domain).

The morphological scale-space is generated by dilations (erosions) with a parabolic structuring function of increasing width. Most often (see for instance Jackway (Jackway and Deriche, 1996)) the erosion and dilation scale-space are combined into one by extending the scale parameter to the entire real axis. Negative scales then are to be interpreted as the erosion scales (background) and positive scales as the dilation scales (objects). The possibility to treat objects and background in a different way is a great difference from linear theory. The convolution cannot make this difference as it is a self-dual operator.

As the morphological scale-space operator (the parabolic dilation) is based on translation of points on the basis of the local geometry, it does not provide a resampling of the image in the sense that the dilated image is a more smooth version of the zero scale image. After the dilation, the points then can be translated back to their original position, using a parabolic erosion, with the exception of course that the points that were annihilated in the dilation are “lost”. The closing indeed more closely resembles the original image, but it is not a smooth resampling of the image. A small change in scale can result in a non-local (and thus non-causal) change in the closing (or opening) result.

The partial differential equation that governs the morphological scale-space is given in section 15.6. This PDE is closely connected to Burgers' equation that describes the propagation of shock waves through a medium. Burgers' equation is known to have an infinite number of solutions although just one that satisfies the entropy condition (a physical constraint). It is surprising that (classical) mathematical morphology selects that unique entropy solution through the supremum or infimum operator.

The PDE  $F_\rho = \|\nabla F\|^2$  indicates that local maxima don't move (neither in spatial direction nor in grey value). They distribute their local geometry over the

neighbourhood, at the expense of the geometry of those neighbouring points. The local maxima stay at their place until they are annihilated by a local maximum of larger grey value. The number of local maxima thus decreases as scale increases. The local minima in the image do change their position under a dilation, although their number is also decreasing (see (Jackway and Deriche, 1996)).

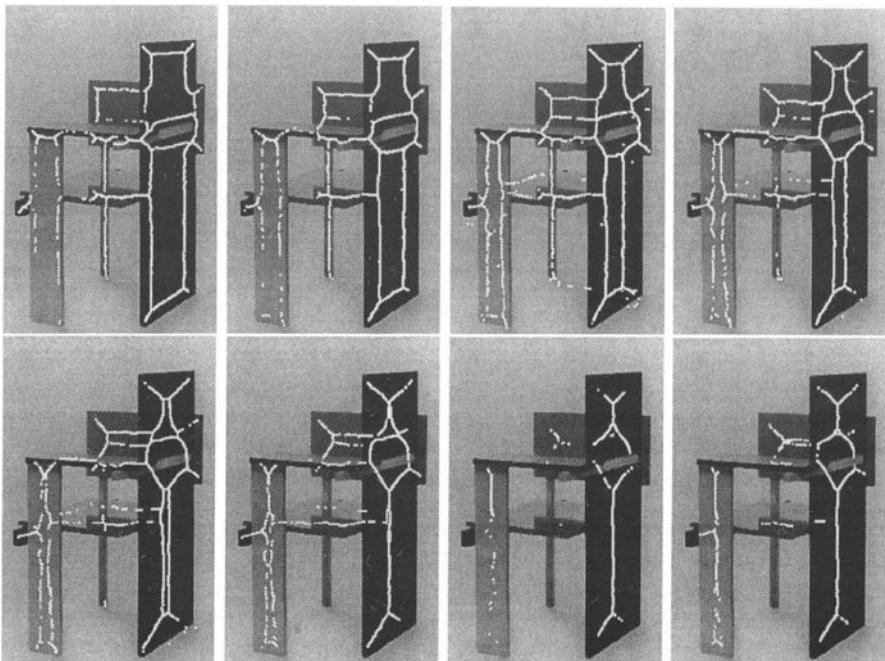


Figure 15.6: Scale-Space Fingerprint. From left to right, top to bottom the scale increases. The singularities at each scale level are superimposed on the original image (the singularities are drawn 3 pixels thick so that they become clearly visible). Note that the thin leg of the chair (designed by G.T. Rietveld) only shows up at lower scales, it completely disappears at higher scales.

In (Boomgaard and Smeulders, 1994) we propose another signature in morphological scale-space. There we conjecture that the collection of singularities (the non-differentiable points in a dilation result) traced over scale provides a complete representation of the original data. This conjecture generalizes classical results on the use of skeletons to describe the original set. In Fig. 15.6 an example is shown of the scale-space signature defined by the singularities.

We believe that mathematical morphology could provide a useful extension to classical scale-space theory in providing a framework for data driven multi-local analysis of images. This chapter has shown that it is possible to incorporate multi-local analysis this way without abolishing the idea of an uncommitted visual front-end.

# Chapter 16

## Nonlinear Diffusion Scale-Spaces

**Joachim Weickert**

*RWCP<sup>1</sup> Novel Function SNN<sup>2</sup> Laboratory,  
Imaging Center, Utrecht University & University Hospital Utrecht,  
E01.334, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

### 16.1 Introduction

Although Gaussian scale-space is the canonical way to define a *linear* scale evolution, people may regard some of its properties as less desirable under certain conditions:

- (a) Semantically useful information is eliminated in the same way as noise. Since the Gaussian scale-space is completely uncommitted, one cannot incorporate image-driven information in order to bias the scale-space evolution towards a desired task, for instance edge detection.
- (b) Linear diffusion filtering dislocates edges when moving from finer to coarser scales, see e.g. Witkin (Witkin, 1983). So structures which are identified at a coarse scale have to be traced back to the original image (Witkin, 1983; Bergholm, 1987). In practice, this correspondence problem can be difficult to handle and may give rise to instabilities.

Due to the uniqueness of the Gaussian scale-space within a linear framework we know that any modification in order to overcome these problems must violate some of the linear scale-space axioms. The goal of this chapter is to present solutions

---

<sup>1</sup>Real World Computing Partnership

<sup>2</sup>Dutch Foundation for Neural Networks

for (a) and (b) which renounce linearity. To this end, we are concerned with the following topics:

- (a) We study nonlinear diffusion equations with diffusivities which can be adapted to the local image structure. They may also enhance interesting structures such as edges.
- (b) Within this class one can find scale-space evolutions where edges remain better localized, so that they need not to be traced back.

In order to establish these filters as scale-space transformations, we describe their smoothing behaviour by means of Lyapunov functionals. Their existence guarantees that these filters are image-simplifying in a *global* sense, while – locally – they may act contrast-enhancing. Such a behaviour is highly desirable for restoration purposes. All results of this scale-space framework apply also to the Gaussian scale-space. Moreover, we shall see that many linear results can be generalized to the nonlinear setting, for instance the existence of a unique smooth solution, a maximum–minimum principle, conservation of the average grey value, and convergence to a constant steady state. These results do not only hold in the continuous case, but also carry over to the practically relevant semidiscrete and fully discrete settings. Simple criteria for choosing suitable numerical schemes can be found in these cases.

This chapter is intended to give an introduction to the basic ideas which helps the reader to understand the original papers (Catté et al., 1992; Cottet and Germain, 1993; Nitzberg and Shiota, 1992; Perona and Malik, 1990; Weickert, 1994; Weickert, 1995; Weickert, 1996c; Weickert, 1996a; Weickert, 1996b; Whitaker, 1993; Whitaker and Pizer, 1993) in a better way. For this reason, many discussions are based on one-dimensional examples, discretizations are analysed for the least complicated schemes, and relations to the Gaussian scale-space are emphasized.

The outline of the chapter is as follows: In Section 16.2 we discuss the first nonlinear diffusion filter due to Perona and Malik. Its theoretical shortcomings lead us to the study of regularizations, for which we shall review well-posedness and scale-space results in Section 16.3. The Sections 16.4 and 16.5 state prerequisites under which one can find similar results in a semidiscrete and discrete framework. Generalizations of these approaches are sketched in Section 16.6.

## 16.2 The Perona–Malik model

### 16.2.1 Basic idea

Let us consider a rectangular image domain  $\Omega = (0, a_1) \times (0, a_2)$  with boundary  $\partial\Omega$ , and a (grey-value) image which is given by a bounded mapping  $f : \Omega \rightarrow \mathbb{R}$ . In order to avoid the blurring and localization problems of linear diffusion filtering, Perona and Malik use a nonlinear diffusion method (Perona and Malik, 1990).

Their nonuniform process (which they name anisotropic<sup>3</sup>) reduces the diffusivity at those locations which have a larger likelihood to be edges, i.e. which have larger gradients. To obtain a filtered image  $u(x, t)$ , they solve the diffusion equation<sup>4</sup>

$$\partial_t u = \operatorname{div}(g(|\nabla u|^2) \nabla u) \quad \text{on} \quad \Omega \times (0, \infty), \quad (16.1)$$

with the original image as initial condition,

$$u(x, t) = f(x) \quad \text{on} \quad \Omega,$$

and reflecting boundary conditions ( $\partial_n$  denotes the derivative normal to the image boundary  $\partial\Omega$ ):

$$\partial_n u = 0 \quad \text{on} \quad \partial\Omega \times (0, \infty).$$

Among the diffusivities they propose is

$$g(|\nabla u|^2) = \frac{1}{1 + |\nabla u|^2 / \lambda^2} \quad (\lambda > 0). \quad (16.2)$$

The experiments of Perona and Malik were visually impressive: edges remained stable over a very long time. It was demonstrated (Perona and Malik, 1990) that edge detection based on this process clearly outperforms the linear Canny edge detector, even without applying non-maximum suppression and hysteresis thresholding. This is due to the fact that diffusion and edge detection interact in one single process instead of being treated as two independent processes which are to be applied subsequently. However, the Perona–Malik approach reveals some problems which we shall discuss next.

### 16.2.2 Ill-posedness

To study the theoretical behaviour of the Perona–Malik filter, let us for simplicity of notation restrict ourselves to the one-dimensional case.

For the diffusivity (16.2) it follows that the *flux function*  $\Phi(s) := sg(s^2)$  satisfies  $\Phi'(s) \geq 0$  for  $|s| \leq \lambda$ , and  $\Phi'(s) < 0$  for  $|s| > \lambda$ , see Figure 16.1. Since (16.1) can be rewritten as

$$\partial_t u = \partial_x(\Phi(\partial_x u)) = \Phi'(\partial_x u) \partial_{xx} u, \quad (16.3)$$

we observe that – in spite of its nonnegative diffusivity – the Perona–Malik model is of *forward parabolic type* for  $|\partial_x u| \leq \lambda$ , and of *backward parabolic type* for  $|\partial_x u| > \lambda$ . Hence,  $\lambda$  plays the role of a *contrast parameter* separating forward (low contrast) from backward (high contrast) diffusion areas. In the backward region the Perona–Malik equation behaves similar as the backward diffusion equation

$$\partial_t u = -\partial_{xx} u,$$

---

<sup>3</sup>It should be noted that – in our terminology – the Perona–Malik filter is regarded as an isotropic model, since it reveals a scalar-valued diffusivity and not a diffusion tensor.

<sup>4</sup>For smoothness reasons we write  $|\nabla u|^2$  instead of  $|\nabla u|$ .

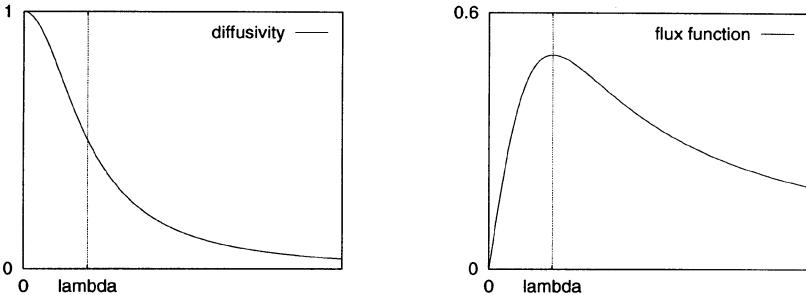


Figure 16.1: (a) LEFT: Diffusivity  $g(s^2) = \frac{1}{1+s^2/\lambda^2}$ . (b) RIGHT: Flux  $\Phi(s) = \frac{s}{1+s^2/\lambda^2}$ .

a classical example for an ill-posed equation, which can become immediately unstable, unless additional stabilizing constraints are introduced (*regularization*). In the same way as the forward diffusion smoothes contrasts, the backward diffusion enhances them. Thus, the Perona–Malik model may enhance gradients whose absolute value is larger than  $\lambda$ , see (Perona and Malik, 1990) for more details.

The forward–backward diffusion behaviour is explicitly intended in the Perona–Malik method, since it gives the desirable result of blurring small fluctuations and sharpening edges.

However, for general smooth nonmonotone flux function  $\Phi$  there is no mathematical theory available which guarantees the existence of the solution of the Perona–Malik equation. Höllig (Höllig, 1983) has proved that if  $\Phi$  is piecewise linear, decreasing in a bounded interval, and increasing elsewhere, there exist initial functions for which the corresponding initial value problem has infinitely many solutions.

Recently Kichenassamy (Kichenassamy, 1996) proposed a class of generalized solutions of the Perona–Malik process which are piecewise linear and contain jumps. Nevertheless, uniqueness and stability with respect to the initial image cannot be expected (Catté et al., 1992; Perona et al., 1994).

Interestingly, such forward–backward diffusion equations are not as unnatural as they look at first glance: besides the fact that they are closely related to other image processing methods (see the references in (Kichenassamy, 1996)), they have been proposed as models for explaining the evolution of stepwise constant temperature or salinity profiles in the ocean, cf. (Barenblatt et al., 1993). In this context, numerical experiments were carried out by Posmentier (Posmentier, 1977) in 1977. Starting from a smoothly increasing initial distribution he reported the creation of perturbations which led to a stepwise constant salinity profile after some time. He also observed instabilities, a first experimental hint to the ill-posedness of this equation. Instabilities were also reported later on by Dzhu Magaziewa (Magazieva, 1983).

In image processing, numerical simulations on the ill-posedness of the 1D Perona–Malik filter were performed by Nitzberg and Shiota (Nitzberg and Shio-

ta, 1992), Fröhlich and Weickert (Fröhlich and Weickert, 1994), and Benhamouda (Benhamouda, 1994). All results point in the same direction: the solution depends strongly on the regularizing effect of the discretization. Finer discretizations are less regularizing and reveal a larger danger of *staircasing effects*, where the number of inflection points of a smooth step edge increases.

Since the Perona–Malik equation is ill-posed, it makes no sense trying to find scale-space reasonings for it. Thus, let us turn our attention to regularizations which are well-posed and for which we can state scale-space results.

## 16.3 Regularized nonlinear diffusion filtering

### 16.3.1 Basic idea

The first regularization attempt is probably due to Posmentier who observed numerically the stabilizing effect of averaging the gradient within the diffusivity (Posmentier, 1977).

A mathematically sound formulation of this idea is given by Catté, Lions, Morel and Coll (Catté et al., 1992). They propose to regularize this gradient by convolving with a Gaussian  $K_\sigma$  with standard deviation  $\sigma > 0$ . By replacing the diffusivity  $g(|\nabla u|^2)$  of the Perona–Malik model by  $g(|\nabla u_\sigma|^2)$  with  $u_\sigma := K_\sigma * u$  they end up with

$$\partial_t u = \operatorname{div}(g(|\nabla u_\sigma|^2) \nabla u). \quad (16.4)$$

Due to its simplicity, this equation will serve as a prototype for all (regularized) nonlinear diffusion filters that we study below. Other regularizations of equations of Perona–Malik type have been proposed by Nitzberg and Shiota (Nitzberg and Shiota, 1992), Lions (cf. (Perona et al., 1994)), Barenblatt et al. (Barenblatt et al., 1993) and Weickert (Weickert, 1994; Weickert, 1996c).

The regularizing effect of the modification by Catté et al. is due to the fact that  $\nabla u_\sigma$  remains bounded. Moreover, the convolution with a Gaussian  $K_\sigma$  makes the filter insensitive to noise at scales smaller than  $\sigma$ . Therefore, when regarding (16.4) as an image restoration equation, it reveals – besides the contrast parameter  $\lambda$  – an additional *scale parameter*  $\sigma$ . This avoids a shortcoming of the genuine Perona–Malik process which misinterprets strong oscillations due to noise as edges which should be preserved or even enhanced.

Figure 16.2 depicts a Gaussian-like image and its filtered versions without and with regularization, respectively. We observe that the regularization stabilizes the ill-posed Perona–Malik process without renouncing the desired edge-enhancing properties (if  $\sigma$  is not too large).

### 16.3.2 General well-posedness and scale-space framework

In order to establish theoretical results for regularized nonlinear diffusion filters, let us consider the following filter structure:

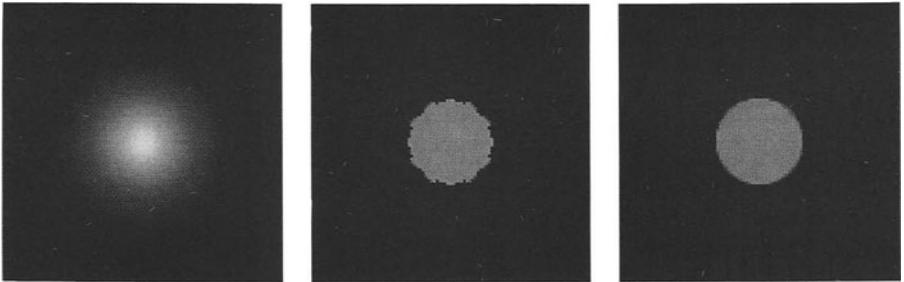


Figure 16.2: Nonlinear diffusion of a Gaussian-like image. (a) LEFT: Original image,  $\Omega = (0, 101)^2$ . (b) MIDDLE: Filtered without regularization,  $\lambda = 9$ ,  $\sigma = 0$ ,  $t = 250$ . (c) RIGHT: Filtered with regularization,  $\lambda = 9$ ,  $\sigma = 0.7$ ,  $t = 250$ .

Assume that  $\sigma > 0$ , and consider the problem

$$\left. \begin{array}{lll} \partial_t u = \operatorname{div}(g(|\nabla u_\sigma|^2) \nabla u) & \text{on} & \Omega \times (0, \infty), \\ u(x, 0) = f(x) & \text{on} & \Omega, \\ \partial_n u = 0 & \text{on} & \partial\Omega \times (0, \infty), \end{array} \right\} (P_c)$$

where the diffusivity  $g$  satisfies

$$\left. \begin{array}{ll} (\text{C1}) \text{ smoothness:} & g \in C^\infty[0, \infty), \\ (\text{C2}) \text{ positivity:} & g(s) > 0 \quad \forall s \in [0, \infty). \end{array} \right\}$$

We observe that  $(P_c)$  also comprises linear diffusion filtering, since the choice  $g(|\nabla u_\sigma|^2) := 1$  does not contradict (C1), (C2). So let us now establish some theoretical results for this class.<sup>5</sup>

### Theorem 16.3.1. (Properties of the continuous filter class)

For the continuous filter class  $(P_c)$  the following statements are valid:

- (a) (Well-posedness and smoothness results)

There exists a unique unique solution  $u(x, t)$  in the distributional sense which is in  $C^\infty(\bar{\Omega} \times (0, \infty))$  and depends continuously on  $f$  with respect to the  $L^2(\Omega)$  norm.

- (b) (Extremum principle)

Let  $a := \inf_{\Omega} f$  and  $b := \sup_{\Omega} f$ . Then,  $a \leq u(x, t) \leq b$  on  $\Omega \times [0, \infty)$ .

- (c) (Average grey level invariance)

The average grey level  $\mu := \frac{1}{|\Omega|} \int_{\Omega} f(x) dx$  is not affected by nonlinear diffusion filtering:  $\frac{1}{|\Omega|} \int_{\Omega} u(x, t) dx = \mu$  for all  $t > 0$ .

---

<sup>5</sup>The spaces  $L^p(\Omega)$ ,  $1 \leq p < \infty$  consist of the functions  $w$ , for which the (Lebesgue) integral  $\int_{\Omega} |w(x)|^p dx$  exists. They are supplemented with the norm  $\|w\|_{L^p(\Omega)} := (\int_{\Omega} |w(x)|^p dx)^{1/p}$ .

(d) (Lyapunov functionals)

$V(t) := \int_{\Omega} r(u(x,t)) dx$  is a Lyapunov function for all convex  $r \in C^2[a,b]$ :  
 $V(t)$  is decreasing and bounded from below by  $\int_{\Omega} r(\mu) dx$ .

(e) (Convergence to a constant steady state)

$$\lim_{t \rightarrow \infty} u(x,t) = \mu \text{ in } L^p(\Omega), 1 \leq p < \infty.$$

The existence, uniqueness and regularity proof of Theorem 16.3.1 is due to (Catté et al., 1992), all other results are proved in (Weickert, 1996a).

The fact that the solution of the nonlinear diffusion filtering is infinitely times differentiable for  $t > 0$  shows that it reveals the same smoothness properties as the Gaussian scale-space. Continuous dependence of the solution on the initial image is of significant practical importance, since it guarantees stability under perturbations. This is of importance when considering stereo images, image sequences or slices from medical CT or MR sequences, since we know that similar images remain similar after filtering.

Many smoothing scale-space properties are closely related to extremum principles: Hummel (Hummel, 1986) for instance shows that under certain conditions the maximum principle for parabolic operators is equivalent to the property that the corresponding scale-space never creates additional level-crossings for  $t > 0$ .

Average grey level invariance is a property which distinguishes diffusion filters from morphological scale-spaces. The latter ones cannot be written in divergence form and, thus, they are not conservative. Moreover, average grey level invariance is required in scale-space based segmentation algorithms such as the hyperstack (Niessen et al., 1996c). In addition to this invariance it is evident that  $(P_c)$  satisfies classical scale-space invariances like grey level shift invariance, reverse contrast invariance, translation invariance and isometry invariance (see (Alvarez et al., 1993) for detailed explanations and applications to other nonlinear scale-spaces). Usual architectural properties of scale-spaces (e.g. the semi-group property) are satisfied as well.

The Lyapunov functionals introduced in (d) show that the considered evolution equation is a simplifying, information-reducing transform with respect to many aspects: Indeed, the special choices  $r(s) := |s|^p$ ,  $r(s) := (s-\mu)^{2n}$  and  $r(s) := s \ln s$ , respectively, imply that all  $L^p(\Omega)$  norms with  $2 \leq p \leq \infty$  are decreasing (e.g. the energy  $\|u(t)\|_{L^2(\Omega)}^2$ ), all even central moments are decreasing (e.g. the variance), and the entropy  $S[u(t)] := - \int_{\Omega} u(x,t) \ln(u(x,t)) dx$ , a measure of uncertainty and missing information, is increasing with respect to  $t$  (Weickert, 1996a; Weickert, 1996c).

The result (e) tells us that, for  $t \rightarrow \infty$ , this simplifying scale-space representation tends to the most global image representation that is possible: a constant image with the same average grey level as  $f$ .

Figure 16.3 gives an impression of the properties of nonlinear diffusion filtering. It depicts the temporal evolution of an MR image. Edges remain well-localized so that they do not need to be traced back, and structures which can be distinguished from their neighbourhood by a sufficiently large grey value difference reveal an extended lifetime across the scales. At the chin the effect of contrast enhancement becomes visible. The filtered results are very segmentation-like.

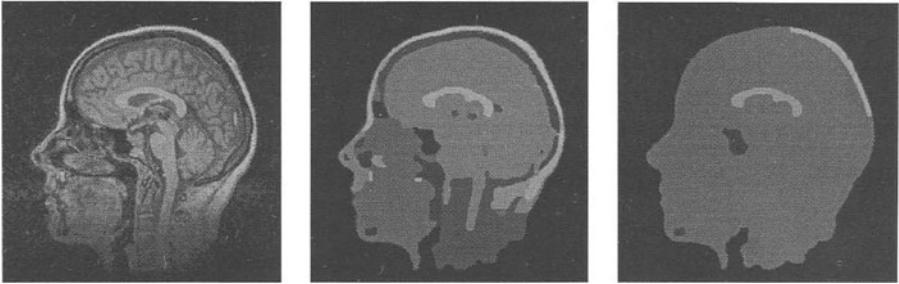


Figure 16.3: Nonlinear diffusion of an MR image ( $\Omega = (0, 256)^2$ ,  $\lambda = 3$ ,  $\sigma = 1$ ). (a) LEFT: Original MR image. (b) MIDDLE:  $t = 2.5 \cdot 10^4$ . (c) RIGHT:  $t = 5 \cdot 10^5$ .

Figure 16.4 shows an example from computer aided quality control. For furniture production it is of importance to classify the quality of wood surfaces. If one aims to automatize this evaluation, one has to process the image in such a way that quality relevant features become better visible and unimportant structures disappear. Fig. 16.4(a) depicts a wood surface possessing one defect. To visualize this defect, equation (16.4) can be successfully applied (Fig. 16.4(b)). In (Weickert, 1994) it is demonstrated how a modified nonlinear diffusion process yields even more accurate results with less roundings at the corners.

## 16.4 Semidiscrete nonlinear diffusion scale-spaces

Although the continuous setting presents a nice theoretical framework, real images are discretized on a fixed pixel grid. Therefore, we need a *semidiscrete* scale-space framework which is discrete in space and continuous in scale. While discrete analogs of linear scale-space can be found in (Lindeberg, 1994e), the goal of this section is to study how nonlinear results carry over under spatial discretization.

### 16.4.1 An example

To illustrate the ideas of spatial discretization, let us restrict ourselves to the one-dimensional equation

$$\partial_t u = \partial_x (g(|\partial_x u_\sigma|^2) \partial_x u). \quad (16.5)$$

A discrete image can be regarded as a vector  $f \in \mathbb{R}^N$ , whose components  $f_i$ ,  $i \in J := \{1, \dots, N\}$  display the grey values at the pixels. A pixel  $i$  represents the location  $x_i = (i - \frac{1}{2}) h$ , where  $h$  is the grid size, and the function  $u(t) : [0, \infty) \rightarrow \mathbb{R}^N$  describes the scale-space evolution of the image vector  $f$ .

A spatial approximation of (16.5) is given by

$$\frac{du_i}{dt} = \frac{1}{h} \left( g_{i+1/2} \frac{u_{i+1} - u_i}{h} - g_{i-1/2} \frac{u_i - u_{i-1}}{h} \right), \quad i \in J. \quad (16.6)$$

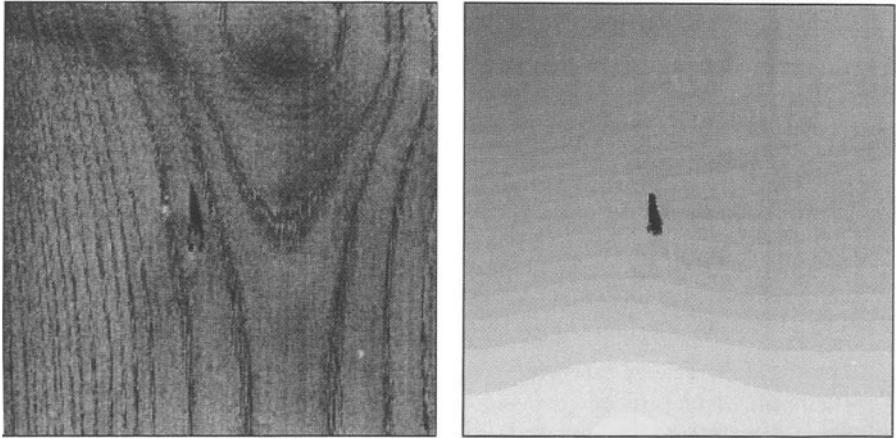


Figure 16.4: Defect detection in wood. (a) LEFT: Wood surface,  $\Omega = (0, 256)^2$ . (b) RIGHT: Nonlinear diffusion,  $\lambda = 4$ ,  $\sigma = 2$ ,  $t = 2000$ .

This scheme remains valid at the boundary points  $i = 1$  and  $i = N$  by using the dummy values  $u_0 := u_1$  and  $u_{N+1} := u_N$ , which can be regarded as discrete reflecting boundary conditions. The expressions  $g_{i \pm 1/2}(u)$  denote approximations of  $g(|\partial_x u_\sigma|^2)$  at the points  $x_{i \pm 1/2}$ , respectively. They can be obtained as follows:

In the spatially discrete case the convolution  $u_\sigma = K_\sigma * u$  comes down to a multiplication of  $u \in \mathbb{R}^N$  with a suitable matrix  $H \in \mathbb{R}^{N \times N}$ , and the choice

$$g_{i \pm 1/2}(u) := \begin{cases} g\left(\left|\frac{(Hu)_{i+1} - (Hu)_i}{h}\right|^2\right) & (i \in \{1, \dots, N-1\}) \\ 0 & (\text{else}), \end{cases} \quad (16.7)$$

takes into account the reflecting boundary conditions.

We can write (16.6) in vector notation as

$$\frac{du}{dt} = A(u)u \quad (16.8)$$

with  $A(u) = (a_{ij}(u))$  and

$$a_{ij}(u) := \begin{cases} \frac{g_{i \pm 1/2}(u)}{h^2} & (j = i \pm 1), \\ -\frac{g_{i+1/2}(u) + g_{i-1/2}(u)}{h^2} & (j = i), \\ 0 & (\text{else}). \end{cases} \quad (16.9)$$

This is a coupled system of nonlinear ordinary differential equations (ODEs). For the Gaussian scale-space we have  $g_{i+1/2} = 1$  for  $i = 1, \dots, N-1$ , which leads to a linear system of ODEs.

### 16.4.2 General well-posedness and scale-space framework

Let us now consider a more general framework for these kinds of semidiscrete diffusion filters. To this end, we recall the definition of irreducibility,<sup>6</sup> which expresses connectivity among the pixels.

**Definition 16.4.1. (Irreducibility).** A matrix  $A = (a_{ij}) \in \mathbb{R}^{N \times N}$  is called irreducible if for any  $i, j \in J$  there exist  $k_0, \dots, k_r \in J$  with  $k_0 = i$  and  $k_r = j$  such that  $a_{k_p k_{p+1}} \neq 0$  for  $p = 0, \dots, r - 1$ .

The problem class  $(P_s)$  we are concerned with is defined as follows:

Let  $f \in \mathbb{R}^N$ . Find a function  $u \in C^1([0, \infty), \mathbb{R}^N)$  which satisfies the initial value problem

$$\left. \begin{aligned} \frac{du}{dt} &= A(u) u, \\ u(0) &= f, \end{aligned} \right\} (P_s)$$

where  $A = (a_{ij})$  has the following properties:

- (S1) smoothness:  $A \in C^1(\mathbb{R}^N, \mathbb{R}^{N \times N})$ ,
- (S2) symmetry:  $a_{ij}(u) = a_{ji}(u) \quad \forall i, j \in J, \forall u \in \mathbb{R}^N$ ,
- (S3) vanishing row sums:  $\sum_{j \in J} a_{ij}(u) = 0 \quad \forall i \in J, \forall u \in \mathbb{R}^N$ ,
- (S4) nonnegative off-diagonals:  $a_{ij}(u) \geq 0 \quad \forall i \neq j, \forall u \in \mathbb{R}^N$ ,
- (S5) irreducibility for all  $u \in \mathbb{R}^N$ .

Under these prerequisites we obtain the subsequent well-posedness and scale-space results (Weickert, 1996b; Weickert, 1996a):

#### Theorem 16.4.2. (Properties of the semidiscrete filter class)

For the semidiscrete filter class  $(P_s)$  the following statements are valid:

(a) (Well-posedness)

The problem  $(P_s)$  has a unique solution  $u(t) \in C^1([0, \infty), \mathbb{R}^N)$ . This solution depends continuously on the initial value and the right-hand side of the ODE system.

(b) (Extremum principle)

Let  $a := \min_{j \in J} f_j$  and  $b := \max_{j \in J} f_j$ . Then,  $a \leq u_i(t) \leq b$  for all  $i \in J$  and  $t \in [0, \infty)$ .

(c) (Average grey level invariance)

The average grey level  $\mu := \frac{1}{N} \sum_{j \in J} f_j$  is not affected by the semidiscrete diffusion filter:  $\frac{1}{N} \sum_{j \in J} u_j(t) = \mu$  for all  $t > 0$ .

---

<sup>6</sup>This is also the correct mathematical concept to avoid undesirable effects for semidiscrete Gaussian scale-spaces such as the creation of two independent subgrids in (Lindeberg, 1994e, p. 112) for  $\gamma = 1$ .

(d) (Lyapunov functionals)

$V(t) := \sum_{i \in J} r(u_i(t))$  is a Lyapunov function for all convex  $r \in C^1[a, b]$ :  
 $V(t)$  is decreasing and bounded from below by  $\sum_{i \in J} r(\mu)$ .

(e) (Convergence to a constant steady state)

$$\lim_{t \rightarrow \infty} u_i(t) = \mu \text{ for all } i \in J.$$

These results allow the same interpretation as their continuous counterparts.

### 16.4.3 Application to the example

Let us now check whether our example (16.8) satisfies the requirements (S1)–(S5).

Since  $H \in \mathbb{R}^{N \times N}$  and  $g \in C^\infty[0, \infty)$ , we have  $A \in C^\infty(\mathbb{R}^N, \mathbb{R}^{N \times N})$ . This proves (S1).

The symmetry of  $A = (a_{ij})$  follows directly from (16.9): For instance, if  $j = i+1$ , we obtain

$$a_{i,i+1}(u) = \frac{g_{i+1/2}(u)}{h^2} = \frac{g_{(i+1)-1/2}(u)}{h^2} = a_{i+1,i}(u). \quad (16.10)$$

By the construction of  $A$  it is also evident that all row sums vanish, i.e. (S3) is satisfied. Moreover, since  $g$  is positive, it follows that  $a_{ij} \geq 0$  for all  $i, j \in J$  and, thus, (S4) holds.

In order to show that  $A$  is irreducible, let us consider two arbitrary pixels  $s_1$  and  $s_2$ . Then we have to find  $k_0, \dots, k_r \in J$  with  $k_0 = s_1$  and  $k_r = s_2$  such that  $a_{k_q, k_{q+1}} \neq 0$  for  $q = 0, \dots, r-1$ . To this end, let  $r := |s_2 - s_1|$  and

$$k_q := s_1 + q \cdot \text{sign}(s_2 - s_1) \quad (q = 0, \dots, |s_2 - s_1|), \quad (16.11)$$

Then, using (16.9) and the positivity of  $g$  we obtain  $k_0 = s_1$ ,  $k_r = s_2$ , and  $a_{k_q, k_{q+1}} > 0$  for  $q = 0, \dots, r-1$ . This establishes (S5).

Since (S1)–(S5) are satisfied, Theorem 16.4.2 guarantees that this scheme creates an evolution with all required well-posedness and scale-space results.

## 16.5 Fully discrete nonlinear diffusion scale-spaces

In general the analytical solution for the nonlinear ODE system of the semidiscrete framework is not known. Hence, we have to apply further numerical techniques which result in a discretization in  $t$  direction. This *fully discrete* case shall be treated next.

### 16.5.1 An example

The simplest and most widely-used way to discretize the semidiscrete model equation (16.8) in  $t$  direction is to use the scheme

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} = A(u^{(k)}) u^{(k)},$$

where the upper index denotes the time level and  $\tau > 0$  is the time step size. This comes down to the iteration scheme

$$u^{(k+1)} = (I + \tau A(u^{(k)})) u^{(k)}, \quad (16.12)$$

where  $I \in \mathbb{R}^N$  is the unit matrix. This scheme is called *explicit*, since  $u^{(k+1)}$  can be directly calculated from  $u^{(k)}$  without solving a system of equations.

### 16.5.2 General well-posedness and scale-space framework

In order to interpret the scheme (16.12) within a general framework, we study a filter class satisfying the subsequent requirements:

Let  $f \in \mathbb{R}^N$ . Calculate a sequence  $(u^{(k)})_{k \in \mathbb{N}_0}$  of processed versions of  $f$  by means of

$$\begin{aligned} u^{(0)} &= f, \\ u^{(k+1)} &= Q(u^{(k)}) u^{(k)}, \quad \forall k \in \mathbb{N}_0, \end{aligned}$$

where  $Q = (q_{ij})$  has the following properties:

- (D1) continuity in its argument:  $Q \in C(\mathbb{R}^N, \mathbb{R}^{N \times N})$ ,
- (D2) symmetry:  $q_{ij}(v) = q_{ji}(v) \quad \forall i, j \in J, \forall v \in \mathbb{R}^N$ ,
- (D3) unit row sum:  $\sum_{j \in J} q_{ij}(v) = 1 \quad \forall i \in J, \forall v \in \mathbb{R}^N$ ,
- (D4) nonnegativity:  $q_{ij}(v) \geq 0 \quad \forall i, j \in J, \forall v \in \mathbb{R}^N$ ,
- (D5) irreducibility for all  $v \in \mathbb{R}^N$ ,
- (D6) positive diagonal:  $q_{ii}(v) > 0 \quad \forall i \in J, \forall v \in \mathbb{R}^N$ .

$(P_d)$

Under these prerequisites we get the following theorem (Weickert, 1996b; Weickert, 1996a):

#### Theorem 16.5.1. (Properties of the discrete filter class)

For the discrete filter class  $(P_d)$  the following statements are valid:

- (a) (Continuous dependence on initial image)

For every  $k > 0$  the unique solution  $u^{(k)}$  of  $(P_d)$  depends continuously on the initial image  $f$ .

- (b) (Extremum principle)

Let  $a := \min_{j \in J} f_j$  and  $b := \max_{j \in J} f_j$ . Then,  $a \leq u_i^{(k)} \leq b$  for all  $i \in J$  and  $k \in \mathbb{N}_0$ .

(c) (Average grey level invariance)

The average grey level  $\mu := \frac{1}{N} \sum_{j \in J} f_j$  is not affected by the discrete diffusion filter:  $\frac{1}{N} \sum_{j \in J} u_j^{(k)} = \mu$  for all  $k \in \mathbb{N}_0$ .

(d) (Lyapunov sequences)

$V^{(k)} := \sum_{i \in J} r(u_i^{(k)})$  is a Lyapunov sequence for all convex  $r \in C[a, b]$ :  $V^{(k)}$  is decreasing and bounded from below by  $\sum_{i \in J} r(\mu)$ .

(e) (Convergence to a constant steady state)

$$\lim_{k \rightarrow \infty} u_i^{(k)} = \mu \text{ for all } i \in J.$$

### 16.5.3 Application to the example

Let us now investigate if the explicit scheme (16.12) belongs to the discrete filter class  $(P_d)$  by checking the properties (D1)–(D6). Let

$$Q(u^{(k)}) := (q_{ij}(u^{(k)})) := I + \tau A(u^{(k)}). \quad (16.13)$$

Then (D1)–(D3) follow immediately from (S1)–(S3).

Thus, let us investigate the nonnegativity. From  $a_{ij} \geq 0$  for  $i \neq j$ , we also have  $q_{ij} \geq 0$  for  $i \neq j$ . However, the diagonal entries yield

$$q_{ii} = 1 + \tau a_{ii}.$$

Using (S3)–(S5) we obtain

$$a_{ii} \stackrel{(S3)}{=} - \sum_{\substack{j \in J \\ j \neq i}} a_{ij} \stackrel{(S4), (S5)}{<} 0 \quad \forall i \in J.$$

Hence,  $Q(u^{(k)})$  is nonnegative if

$$\tau \leq \frac{1}{\max_{i \in J} |a_{ii}(u^{(k)})|}.$$

Now we want to prove (D5). For  $i \neq j$  we have  $q_{ij}(u^{(k)}) = a_{ij}(u^{(k)})$ . Now for

$$\tau < \frac{1}{\max_{i \in J} |a_{ii}(u^{(k)})|} \quad (16.14)$$

it follows that  $q_{ii}(u^{(k)}) > 0$  for all  $i \in J$  and, thus, the irreducibility of  $A(u^{(k)})$  carries over to  $Q(u^{(k)})$ .

Finally, the time step size restriction for ensuring irreducibility implies that all diagonal elements of  $Q(u^{(k)})$  are positive. This establishes (D6).

From these considerations we conclude that the explicit discretization of (16.8) reveals all well-posedness and scale-space properties provided that the time step size satisfies the restriction (16.14). By virtue of (16.2) and (16.9) we observe that this is fulfilled for  $\tau < h^2/2$ . This criterion is strict and cannot be relaxed when considering only the Gaussian scale-space.

### 16.5.4 A more efficient discretization

It is worthwhile to have a look at a slightly more complicated discretization of (16.5), namely

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} = A(u^{(k)}) u^{(k+1)}, \quad (16.15)$$

which leads to the scheme

$$(I - \tau A(u^{(k)})) u^{(k+1)} = u^{(k)}. \quad (16.16)$$

We observe that this scheme does not give the solution  $u^{(k+1)}$  directly (explicitly): It requires to solve a linear system first. For this reason it is called a *linear-implicit (semi-implicit)* scheme. Nevertheless, in our one-dimensional model example, the resulting system matrix is tridiagonal and diagonally dominant, such that the whole system can be resolved in linear (!) effort by the Thomas algorithm (cf. (Morton and Mayers, 1994, pp. 23–25)), requiring only  $5N$  multiplications/divisions and  $3N$  additions.

It can be shown (Weickert, 1996a) that for all  $\tau > 0$ , the matrix  $Q(u^{(k)}) = (I - \tau A(u^{(k)}))^{-1}$  exists and (16.16) satisfies (D1)–(D6). Thus, the scheme is unconditionally stable and gives the desired scale-space results for arbitrarily large time step sizes. Such schemes can be much more efficient than their explicit counterparts. Moreover, for this scheme one can prove (Weickert, 1996a) that  $q_{ij}(u^{(k)}) > 0$  for all  $i, j \in J$  and for all  $k \in \mathbb{N}_0$ . This means that every point at level  $k$  influences every point at level  $k + 1$ . Thus, unlike in explicit schemes, we have an infinite spread of information, which reflects the typical behaviour of many continuous parabolic processes, and – in the linear case – of the infinite support of the Gaussian.

## 16.6 Generalizations

The results that have been presented so far constitute only the simplest class of nonlinear diffusion scale-spaces. Indeed, the Theorems 1–3 are valid for signals in arbitrary dimensions (see (Gerig et al., 1992; Rambaux and Garçon, 1994) for 3D implementations of nonlinear diffusion filters), and they can be extended to the more flexible anisotropic filters using a diffusion tensor instead of a scalar-valued diffusivity (Nitzberg and Shiota, 1992; Cottet and Germain, 1993; Weickert, 1994; Weickert, 1995; Weickert, 1996c; Weickert, 1996a; Weickert, 1996b). In order to enhance corners or flow-like structures they may utilize more refined structure descriptors than the Gaussian-smoothed gradient (El-Fallah and Ford, 1994; Nitzberg and Shiota, 1992; Weickert, 1995), and they may also be extended to vector-valued images (Gerig et al., 1992; Whitaker, 1993).

**Acknowledgement.** The author thanks Antonio Lopez Peña for valuable comments on the manuscript for this chapter.

# Bibliography

- Abramowitz, M. and Stegun, I. A., editors (1964). *Handbook of Mathematical Functions*. Applied Mathematics Series. National Bureau of Standards, 55 edition.
- Adler, A. (1985). *The Geometry of Random Fields*. Wiley, New York.
- Almansa, A. (1994). Ridge enhancement of fingerprint images. Technical Report 2, CeCal, Engineering School, University of Uruguay, Montevideo, Uruguay.
- Almeida, E., Caro, R., Curbelo, R., and Ulivi, G. (1993). Clasificación de huellas dactilares. Technical Report 2, CeCal, Engineering School, University of Uruguay, Montevideo, Uruguay. (Taller V).
- Alvarez, L., Guichard, F., Lions, P.-L., and Morel, J.-M. (1993). Axioms and fundamental equations of image processing. *Archive Rational Mechanics and Analysis*, 123(3):199–257.
- Amini, A. A. (1994). A scalar function formulation for optical flow. In (Eklundh, 1994), pages 125–131.
- Aoki, Y. and Iijima, T. (1996). Pluralizing method of simple similarity. *IEICE Trans. Inf. Syst.*, E79-D:485–490.
- Arnspang, J. (1988a). Notes on local determination of smooth optic flow and the translational property of first order optic flow. Technical Report 88/1, Institute of Computer Science, University of Copenhagen.
- Arnspang, J. (1988b). Optic acceleration. Local determination of absolute depth and velocity, time to contact and geometry of an accelerating surface. Technical Report 88/2, Institute of Computer Science, University of Copenhagen.
- Arnspang, J. (1991). *Motion Constraint Equations in Vision Calculus*. Dissertation for the Danish Dr. scient. degree, University of Utrecht, Utrecht, The Netherlands.
- Arnspang, J. (1993). Motion constraint equations based on constant image irradiance. *Image and Vision Computing*, 11(9):577–587.

- Asai, K., Hoshino, Y., Yamashita, N., and Hratsuka, S. (1986). Fingerprint identification system. In *Second USA-Japan Computer Conference*, page 30.
- Åström, K. (1996). *Invariancy Methods for Points, Curves and Surfaces in Computational Vision*. PhD thesis, Dept of Mathematics, Lund University, Sweden.
- Åström, K. and Heyden, A. (1996a). Stochastic analysis of scale-space smoothing. In (ICPR96, 1996).
- Åström, K. and Heyden, A. (1996b). Stochastic analysis of sub-pixel edge detection. In (ICPR96, 1996).
- Ayache, N. (1995). Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, 13(4).
- Babaud, J., Witkin, A. P., Baudin, M., and Duda, R. O. (1986). Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):26–33.
- Barenblatt, G. I., Bertsch, M., Passo, R. D., and Ughi, M. (1993). A degenerate pseudoparabolic regularization of a nonlinear forward–backward heat equation arising in the theory of heat and mass exchange in stably stratified turbulent shear flow. *SIAM J. Math. Anal.*, 24:1414–1439.
- Barron, J. L., Fleet, D. J., and Beauchemin, S. S. (1994). Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77.
- Benhamouda, B. (1994). Parameter adaptation for nonlinear diffusion in image processing. Master's thesis, Dept. of Mathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany.
- Bergengruen, O. (1994). Preprocessing of poor quality fingerprint images. In *XIV International Conference of the Chilean Computer Science Society*.
- Bergholm, F. (1987). Edge focusing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9:726–741.
- Bigün, J., Granlund, G. H., and Wiklund, J. (1991). Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790.
- Blom, J. (1992). *Topological and Geometrical Aspects of Image Structure*. PhD thesis, University of Utrecht, Department of Medical and Physiological Physics, Utrecht, The Netherlands.
- Boomgaard, R. v. d. (1992a). *Mathematical Morphology: Extensions towards Computer Vision*. PhD thesis, University of Amsterdam.
- Boomgaard, R. v. d. (1992b). The morphological equivalent of the Gauss convolution. *Nieuw Archief voor Wiskunde*, 10(3):219–236.

- Boomgaard, R. v. d., Dorst, L., Schavemaker, J., and Smeulders, A. W. M. (1996). Quadratic structuring functions in mathematical morphology. In *Mathematical Morphology and its applications to signal processing 3*.
- Boomgaard, R. v. d. and Smeulders, A. W. M. (1992). The morphological structure of images. In *Proceedings 11th IAPR International Conference on Pattern Recognition*, pages 268–270.
- Boomgaard, R. v. d. and Smeulders, A. W. M. (1994). The morphological structure of images: The differential equations of morphological scale-space. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(11):1101–1113.
- Boothby, W. M. (1975). *An Introduction to Differential Geometry and Riemannian Geometry*. Academic Press, New York, San Francisco, London.
- Boscovitz, R. J. (1763). *A Theory of Natural Philosophy*. The M.I.T. Press, Cambridge Mass. 1st Venetian edition 1763, English edition 1966.
- Bowyer, K., Shapiro, L., and Tanimoto, S., editors (1994). *Conference on Computer Vision and Pattern Recognition*, Seattle, WA. IEEE Computer Society Press.
- Brockett, R. R. W. and Maragos, P. (1994). Evolution equations for continuous scale morphological filtering. *IEEE transactions on Signal Processing*, 42(12):3377–3386.
- Brown, L. G. and Shvaytser, H. (1990). Surface orientation from projective foreshortening of isotropic texture autocorrelation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(6):584–588.
- Bruckstein, A. M. and Shaked, D. (1993). On projective invariant smoothing and evolution of planar curves and polygons. Technical Report CIS Report No. 9328, Computer Science Dept., Technion, Haifa 32000, Israel.
- Burt, P. J. (1981). Fast filter transforms for image processing. *Computer Vision, Graphics, and Image Processing*, 16:20–51.
- Buxton, B. and Cipolla, R., editors (1996). *Proceedings of the 4th European Conference on Computer Vision*, number 1064–1065 in Lecture Notes in Computer Science, Cambridge, UK. Springer-Verlag.
- Calderón (1853). The constant prince. In *The dramas of Calderón*, volume I, pages 12–13. tr. Denis Florence McCarty, London.
- Candela, G. T., Grother, P. J., Watson, C. I., Wilkinson, R. A., and Wilson, C. L. (1995). PCASYS - A pattern-level classification automation system for fingerprints. U. S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Visual Image Processing Group, Gaithersburg, MD 20899.

- Canny, F. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):676–698.
- Casamayú, L., Curbelo, R., Gómez, A. G. G., Oliveira, R., and Steffen, H. (1992). Convenio facultad de ingeniería - suprema corte de justicia. Facultad de Ingeniería, Universidad de la República.
- Castan, S., Zhao, J., and Shen, J. (1990). Optimal filter for edge detection methods and results. In *Proceedings of the 1st European Conference on Computer Vision*, pages 13–17.
- Catté, F., Lions, P.-L., Morel, J.-M., and Coll, T. (1992). Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.*, 29:182–193.
- Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine computation of complex Fourier series. *Mathematics of Computation*, 19:297–301.
- Cottet, G.-H. and Germain, L. (1993). Image processing through reaction combined with nonlinear diffusion. *Math. Comp.*, 61:659–673.
- Cramér, H. and Leadbetter, M. R. (1967). *Stationary and Related Stochastic Processes*. Wiley, New York.
- Cressie, N. A. C. (1991). *Statistics for Spatial Data*. Wiley, New York.
- Crowley, J. L. (1981). *A Representation for Visual Information*. PhD thesis, Carnegie-Mellon University, Robotics Institute, Pittsburgh, Pennsylvania.
- CVPR96 (1996). *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, California. IEEE Computer Society Press.
- Damon, J. (1984). The unfolding and determinacy theorems for subgroups of  $\mathcal{A}$  and  $\mathcal{K}$ . *Memoirs Amer. Math. Soc.*, 50(306).
- Damon, J. (1995a). Local morse theory for solutions to the heat equation and gaussian blurring. *Journal of Differential Equations*, 115(2):368–401.
- Damon, J. (1995b). Properties of ridges and cores for two dimensional images. preliminary version.
- Damon, J. (1996). Generic properties of solutions to partial differential equations. to appear in Arch. Rat. Mech. Anal.
- Danielsson, P.-E. and Seger, O. (1990). Rotation invariance in gradient and higher order derivative detectors. *Computer Vision, Graphics, and Image Processing*, 49:198–221.
- Daugman, J. G. (1990). An information-theoretic view of analog representation in striate cortex. *Comput. Neurosci.*, 31:403–423.

- Deriche, R. (1987). Using Canny's criteria to derive an optimal edge detector recursively implemented. *International Journal of Computer Vision*, 1:167–187.
- Deriche, R. and Faugeras, O. (1996). Les EDP en traitement des images et vision par ordinateur. *Traitement du Signal*, in press.
- Devernay, F. and Faugeras, O. D. (1994). Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. In (Bowyer et al., 1994), pages 208–213.
- Devlaminck, V. and Dubus, J. (1996). Estimation of compressible or incompressible deformable motions for density images. In Delogne, P., editor, *Proceedings of the International Conference on Image Processing 1996*, volume 1, pages 125–128. Lausanne, Switzerland. IEEE Computer Society Press.
- D'Haeyer, J. (1986). Determining motion of image curves from local pattern changes. *Computer Vision, Graphics, and Image Processing*, 34:166–188.
- Dibos, F. (1995). Analyse multiéchelle invariante par transformations projectives. *C. R. Acad. Sci. Paris*, t. 320(Série I):917–921.
- Dorst, L. and Boomgaard, R. v. d. (1994a). Morphological signal processing and the slope transform. *Signal processing*, 38(1):79–98.
- Dorst, L. and Boomgaard, R. v. d. (1994b). Two dual representations of mathematical morphology based on the parallel normal transport property. In Serra, P. S. J., editor, *Mathematical morphology and its applications to signal processing 2*, pages 161–170. Kluwer Academic Publishers.
- Doyen, L., Najman, L., and Mattioli, J. (1995). Mutational equations of the morphological dilation tubes. *Journal of mathematical imaging and vision*, 5(3):219–230.
- Eberly, D., Gardner, R., Morse, B., Pizer, S., and Scharlach, C. (1994). Ridges for image analysis. *Journal of Mathematical Imaging and Vision*, 4:351–371.
- Eguchi, T., Gilkey, P., and Hanson, J. (1980). Gravitation, gauge theories and differential geometry. *Physics Reports*, 66(2):213–393.
- Eklundh, J.-O., editor (1994). *Proceedings of the 3rd European Conference on Computer Vision*, number 800–801 in Lecture Notes in Computer Science, Stockholm, Sweden. Springer-Verlag.
- El-Fallah, A. I. and Ford, G. E. (1994). The evolution of mean curvature in image filtering. In *Proceedings of the 1st International Conference on Image Processing*, volume 1, pages 298–302. IEEE Computer Society Press.
- et al, D. E. (1994). Ridge flow models for image segmentation. In *SPIE Proc. Med. Imag.*, volume 2167:7, pages 54–64.

- Faugeras, O., editor (1990). *Proceedings of the 1st European Conference on Computer Vision*, volume 427 of *Lecture Notes in Computer Science*, Antibes, France. Springer-Verlag.
- Faugeras, O. (1993). Sur l'évolution de courbes simples du plan projectif réel. *C. R. Acad. Sci. Paris*, t. 317(Série I):565–570.
- Faugeras, O. (1994). *Three-Dimensional Computer Vision*. MIT Press.
- Fennema, C. L. and Thompson, W. B. (1979). Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9:301–315.
- Fidrich, M. and Thirion, J.-P. (1995). Multiscale representation and analysis of features from medical images. In Ayache, N., editor, *International Conference on Computer Vision, Virtual Reality and Robotics in Medicine*, volume 905 of *LNCS*, pages 358–364, Nice.
- Florack, L. M. J. (1993). *The Syntactical Structure of Scalar Images*. PhD thesis, Utrecht University, Utrecht, The Netherlands.
- Florack, L. M. J. (1995). Grey-scale images. Technical Report ERCIMI-09/95-R039, INESC Aveiro, Portugal. [http://www-ercim.inria.fr/publication/technical\\_reports](http://www-ercim.inria.fr/publication/technical_reports).
- Florack, L. M. J. (1996). Data, models, and images. In (ICIP96, 1996), pages 469–472.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1992a). Families of tuned scale-space kernels. In Sandini, G., editor, *Proceedings of the 2nd European Conference on Computer Vision*, volume 588 of *Lecture Notes in Computer Science*, pages 19–23, Santa Margherita Ligure, Italy. Springer-Verlag.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1992b). Scale and the differential structure of images. *Image and Vision Computing*, 10(6):376–388.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1993a). Cartesian differential invariants in scale-space. *Journal of Mathematical Imaging and Vision*, 3(4):327–348.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1993b). Non-linear diffusion by metrical affinity. In preparation.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1994a). General intensity transformations and differential invariants. *Journal of Mathematical Imaging and Vision*, 4(2):171–187.

- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1994b). Images: Regular tempered distributions. In O, Y.-L., Toet, A., Heijmans, H. J. A. M., Foster, D. H., and Meer, P., editors, *Proc. of the NATO Advanced Research Workshop Shape in Picture - Mathematical description of shape in greylevel images*, volume 126 of *NATO ASI Series F*, pages 651–660. Springer Verlag, Berlin.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1994c). Linear scale-space. *Journal of Mathematical Imaging and Vision*, 4(4):325–351.
- Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1996a). The multiscale local jet. *International Journal of Computer Vision*.
- Florack, L. M. J. and Nielsen, M. (1994). The intrinsic structure of the optic flow field. Technical Report ERCIM-07/94-R033 or INRIA-RR-2350, INRIA Sophia-Antipolis, France. [http://www-ercim.inria.fr/publication/technical\\_reports](http://www-ercim.inria.fr/publication/technical_reports).
- Florack, L. M. J., Niessen, W. J., and Nielsen, M. (1996b). The intrinsic structure of optic flow incorporating measurement duality. Technical Report 96-15, Department of Computer Science, University of Copenhagen.
- Folland, G. B. (1976). Introduction to partial differential equations. Princeton Math. Notes, Princeton Univ. Press.
- Förstner, M. A. and Gülich, E. (1987). A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *ISPRS Intercommission Workshop*.
- Fourier, J. (1822). *Théorie analytique de la chaleur*. Paris. English translation: *The analytical theory of heat*, Dover, 1955.
- Friedman, M. (1983). *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton University Press, Princeton.
- Fröhlich, J. and Weickert, J. (1994). Image processing using a wavelet algorithm for nonlinear diffusion. Technical Report 104, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany.
- Fua, P. and Leclerc, Y. G. (1995). Object-centered surface reconstruction: Combining multi-image stereo and shading. *International Journal of Computer Vision*, 16:35–56.
- Fulton, W. (1989). *Algebraic Curves*. Addison-Wesley Publishing Co.
- Gabor, D. (1965). Information theory in electron microscopy. *Laboratory Investigation*, 14:801–807.
- Gerig, G., Kübler, O., Kikinis, R., and Jolesz, F. A. (1992). Nonlinear anisotropic filtering of MRI data. *IEEE Trans. Medical Imaging*, 11:221–232.

- Gibson, J. J. (1950). *The Perception of the Visual World.* Houghton-Mifflin, Boston.
- Gilmore, R. (1981). *Catastrophe theory for scientists and engineers.* Dover.
- Golubitsky, M. and Guillemin, V. (1974). *Stable Mappings and their Singularities.* Springer Graduate Texts in Mathematics, Springer-Verlag.
- Gordan, P. (1871). Ueber die Bildung der Resultante zweier Gleichungen. *Mathematische Annalen*, 50:355–414.
- Gray, A. (1993). *Modern Differential Geometry of Curves and Surfaces.* CRC Press, Boca Raton.
- Greenwood, D. D. (1964). *Mapping.* The University of Chicago Press, Chicago.
- Griffen, L. D. (1995). *Descriptions of Image Structure.* PhD thesis, University of London.
- Griffin, L. D. and Colchester, A. C. F. (1995). Superficial and deep structure in linear diffusion scale space: isophotes, critical points and separatrices. *Image and Vision Computing*, 13(7):543–557.
- Grimson, W. E. L. (1981). *From Images to Surfaces.* MIT Press, Cambridge, Massachusetts.
- Gårding, J. and Lindeberg, T. (1994). Direct estimation of local surface shape in a fixating binocular vision system. In (Eklundh, 1994), pages 365–376.
- Gårding, J. and Lindeberg, T. (1996). Direct computation of shape cues based on scale-adapted spatial derivative operators. *International Journal of Computer Vision*, 17(2):163–192.
- Haar Romeny, B. M. t., editor (1994). *Geometry-Driven Diffusion in Computer Vision.* Number 1 in the series Computational Imaging and Vision. Kluwer Academic Publishers, Dordrecht, Netherlands.
- Haar Romeny, B. M. t. (1996). Scale-space research at Utrecht University. In *Proc. 12th International Conference on Analysis and Optimization of Systems: Images, Wavelets and PDE's*, pages 15–30, CEREMADE / INRIA, Paris, France.
- Haar Romeny, B. M. t., Bastin, F. H., Steenbeek, J., Zuiderveld, K. J., and Viergever, M. A. (1994a). Three dimensional CT subtraction angiography. In Robb, R. A., editor, *SPIE Proc. Visualization in Biomedical Computing*, pages 262–271, Bellingham, WA. SPIE Press.
- Haar Romeny, B. M. t. and Binkhuysen, F. H. B. (1992). PACS as an infrastructural essential element in an advanced digital imaging department. In de Valk, J. P. J., editor, *Integrated Diagnostic Imaging: Digital PACS in Medicine*, pages 320–386. Elsevier Science Publishers B.V., Amsterdam.

- Haar Romeny, B. M. t., Florack, L. M. J., de Swart, M., Wilting, J., and Viergever, M. A. (1994b). Deblurring Gaussian blur. In *Proc. Mathematical Methods in Medical Imaging II*, volume 2299, pages 139–148, San Diego, CA. SPIE.
- Haar Romeny, B. M. t., Florack, L. M. J., Salden, A. H., and Viergever, M. A. (1994c). Higher order differential structure of images. *Image and Vision Computing*, 12(6):317–325.
- Haar Romeny, B. M. t., Niessen, W. J., Wilting, J., and Florack, L. M. J. (1994d). Differential structure of images: Accuracy of representation. In *Proc. First IEEE Internat. Conf. on Image Processing*, pages 21–25, Austin, TX. IEEE.
- Haring, S., Viergever, M. A., and Kok, J. N. (1994). Kohonen networks for multiscale image segmentation. *Image and Vision Computing*, 12(6):339–344.
- Hecht, E. (1987). *Optics*. Addison-Wesley, Reading, Mass.
- Heijmans, H. J. A. M. (1994). *Morphological Image Operators*. Academic Press.
- Hilbert, D. (1893). Ueber die vollen Invariantsystemen. *Mathematische Annalen*, 42:313–373.
- Hildreth, E. and Marr, D. (1980). Theory of edge detection. *Proceedings of Royal Society of London*, 207:187–217.
- Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial Intelligence*, 23:309–354.
- Hille, E. and Phillips, R. S. (1957). *Functional Analysis and Semi-Groups*, volume XXXI. American Mathematical Society Colloquium Publications.
- Hirschmann, I. I. and Widder, D. V. (1955). *The Convolution Transform*. Princeton University Press, Princeton, New Jersey.
- Höllig, K. (1983). Existence of infinitely many solutions for a forward–backward heat equation. *Trans. Amer. Math. Soc.*, 278:299–316.
- Horn, B. and Schunck, B. (1981). Determining optical flow. *Artificial Intelligence*, 23:185–203.
- Horn, B. K. P. (1986). *Robot Vision*. MIT Press, Cambridge.
- Horn, B. K. P. and Schunck, B. G. (1993). “Determining optical flow”: a retrospective. *Artificial Intelligence*, 59:81–87.
- Hummel, R. A. (1986). Representations based on zero-crossings in scale space. In *Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '86, Miami Beach, June 22–26, 1986)*, pages 204–209. IEEE Computer Society Press, Washington.

- Hummel, R. A. (1987). The scale-space formulation of pyramid data structures. In Uhr, L., editor, *Parallel Computer Vision*, pages 187–223. Academic Press, New York.
- ICCV95 (1995). *Proceedings of the 5th International Conference on Computer Vision*, Cambridge, MA. IEEE Computer Society Press.
- ICIP96 (1996). *Proceedings of the 2nd International Conference on Image Processing*, Lausanne, Switzerland. IEEE Computer Society Press.
- ICPR84 (1984). *Proceedings of the 7th International Conference on Pattern Recognition*, Montreal, Canada.
- ICPR96 (1996). *Proceedings of the 13th International Conference on Pattern Recognition*, Vienna, Austria. IEEE Computer Society Press.
- Iijima, T. (1962). Basic theory on normalization of a pattern (in case of typical one-dimensional pattern). *Bulletin of Electrical Laboratory*, 26:368–388. (in Japanese).
- Iijima, T. (1971a). Basic equation of figure and observational transformation. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(7):37–38. English Abstracts.
- Iijima, T. (1971b). Basic theory on feature extraction of figures. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(12):35–36. English Abstracts.
- Iijima, T. (1971c). Basic theory on normalization of figure. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(11):24–25. English Abstracts.
- Iijima, T. (1971d). Basic theory on the construction of figure space. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(8):35–36. English Abstracts.
- Iijima, T. (1971e). A suppression kernel of a figure and its mathematical characteristics. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(9):30–31. English Abstracts.
- Iijima, T. (1971f). A system of fundamental functions in an abstract figure space. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 54-C(11):26–27. English Abstracts.
- Iijima, T. (1972a). Basic theory on the structural recognition of a figure. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 55-D(8):28–29. English Abstracts.
- Iijima, T. (1972b). Theoretical studies on the figure identification by pattern matching. *Transactions of the Institute of Electronics and Communication Engineers of Japan*, 55-D(8):29–30. English Abstracts.

- Iijima, T. (1973). *Pattern recognition*. Corona-sha. (in Japanese).
- Iijima, T., Genchi, H., and Mori, K. (1972). A theoretical study of pattern identification by matching method. In *Proc. First USA-Japan Computer Conference (Tokyo, Oct. 3-5, 1972)*, pages 42-48.
- Jackway, P. T. (1992). Morphological scale-space. In *Proceedings 11th IAPR International Conference on Pattern Recognition*, pages 252-255. IEEE Computer Society Press.
- Jackway, P. T. and Deriche, M. (1996). Scale-space properties of the multi-scale morphological dilation erosion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):38-51.
- Jacobs, T. S. (1988). *Light for the artist*. Watson-Guptill Publications, New York.
- Johansen, P. (1994). On the classification of top points in scale space. *Journal of Mathematical Imaging and Vision*, 4:57-67.
- Johansen, P., editor (1996). *Proceedings of the Copenhagen Workshop on Gaussian Scale-Space Theory*. DIKU Tech. Rep. Nr. 96/19.
- Johansen, P., Skelboe, S., Grue, K., and Andersen, J. D. (1986). Representing signals by their top points in scale space. In Haton, J.-P., editor, *Proceedings of the 8th International Conference on Pattern Recognition*, Paris, France. IEEE Computer Society Press.
- Jones, J. and Palmer, L. (1987). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J. Neurophysiol.*, 58(6):1187-1211.
- Kalitzin, S. N., Haar Romeny, B. M. t., Salden, A. H., and Nacken, P. F. (1996). Topological numbers and singularities in scalar images. scale space evolution properties. Submitted.
- Kameswara, R. C. V. (1977). Pattern recognition techniques applied to fingerprints. Linköping Studies in Science and Technology. Dissertations No. 19, Linköping.
- Kanatani, K. (1984). Detection of surface orientation and motion from texture by stereological technique. *Journal of Artificial Intelligence*, 23:213-237.
- Karlin, S. (1968). *Total Positivity*. Stanford Univ. Press.
- Karssemeijer, N. (1995). Detection of stellate distortions in mammograms using scale-space operators. In Barillot, R. D. Y. B. C., editor, *Proceedings of the Information Processing in Medical Imaging 1995*, pages 335-346, Ile de Berder, France. Kluwer Academic Publishers.

- Kass, M., Witkin, A., and Terzopoulos, D. (1987). Snakes: Active contour models. In *Proceedings of the 1st International Conference on Computer Vision*, pages 679–698, London.
- Kaymaz, E. and Mitra, S. (1992). Analysis and matching of degraded and noisy fingerprints. In *Applications of Digital Image Processing XV*, volume 1771 of *SPIE*, pages 498–508.
- Keller, D. (1991). Reconstruction of STM and AFM images distorted by finite-size tips. *Surface Science*, 253:353–364.
- Kichenassamy, S. (1996). Nonlinear diffusions and hyperbolic smoothing for edge enhancement. In Berger, M.-O., Deriche, R., Herlin, I., Jaffré, J., and Morel, J.-M., editors, *ICAOS '96: Images, Wavelets and PDEs*, volume 219 of *Lecture Notes in Control and Informations Sciences*, pages 119–124. Springer, London.
- Kimia, B. B., Tannenbaum, A., and Zucker, S. W. (1990). Toward a computational theory of shape: An overview. In (Faugeras, 1990), pages 402–407.
- Klein, F. (1893). Vergleichende Betrachtungen über neuere geometrische Forschungen (Erlanger Programm). *Mathematische Annalen*, 43:63–100.
- Kobayashi, T., Honma, K., Nakajima, T., and Hanada, K. (1993). Masticatory function in patients with mandibular prognathism before and after orthognathic surgery. *J Oral Maxillofac Surg*, 51(9):997–1001.
- Koenderink, J. J. (1977). Current models in contrast processing. In Spekreijse, H. and Tweel, L. H. v. d., editors, *Spatial contrast, Report of a workshop*. (Koninklijke Nederlandse Akademie van Wetenschappen, Amsterdam 1976).
- Koenderink, J. J. (1978). Visual detection of spatial contrast: Influence of location in the visual field, target extent and illuminance level. *Biological Cybernetics*, 30:157–167.
- Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics*, 50:363–370.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26(1):161–180.
- Koenderink, J. J. (1988). Scale-time. *Biological Cybernetics*, 58:159–162.
- Koenderink, J. J. (1990). *Solid Shape*. MIT Press, Cambridge.
- Koenderink, J. J. and Doorn, A. J. v. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791.
- Koenderink, J. J. and Doorn, A. J. v. (1980). Photometric invariants related to solid shape. *Optica Acta*, 27:981–996.

- Koenderink, J. J. and Doorn, A. J. v. (1986a). Dynamic shape. *Biological Cybernetics*, 53:383–396.
- Koenderink, J. J. and Doorn, A. J. v. (1986b). Logical stratification of organic intelligence. In Trappi, R., editor, *Cybernetics and Systems*, pages 871–878. D. Reidel Publishing Company.
- Koenderink, J. J. and Doorn, A. J. v. (1987). Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375.
- Koenderink, J. J. and Doorn, A. J. v. (1990). Receptive field families. *Biological Cybernetics*, 63:291–298.
- Koenderink, J. J. and Doorn, A. J. v. (1992a). Generic neighborhood operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(6):597–605.
- Koenderink, J. J. and Doorn, A. J. v. (1992b). Second-order optic flow. *Journal of the Optical Society of America-A*, 9(4):530–538.
- Koster, A. S. E., Vincken, K. L., Graaf, C. N. D., Zander, O. C., and Viergever, M. A. (1996). Heuristic linking models in multi-scale image segmentation. *Computer Vision, Graphics, and Image Processing*. In press.
- Lax, P. D. (1973). *Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves*. SIAM.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 5:437–457.
- Lee, D. N. (1980). The optic flow field: the foundation of vision. *Phil. Trans. R. Soc. Lond. B*, 290:169–179.
- Lifshitz, L. M. (1987). *Image Segmentation using Global Knowledge and A Priori Information*. PhD thesis, Univ. of North Carolina. Technical Report 87-012.
- Lifshitz, L. M. and Pizer, S. M. (1990). A multiresolution hierarchical approach to image segmentation based on intensity extrema. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(6):529–541.
- Lindeberg, T. (1990). Scale-space for discrete signals. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3):234–254.
- Lindeberg, T. (1991). *Discrete Scale-Space Theory and the Scale-Space Primal Sketch*. PhD thesis, Dept. of Numerical Analysis and Computing Science, KTH. ISRN KTH/NA/P-91/08-SE. An extended and revised version published as book Scale-Space Theory in Computer Vision in The Kluwer International Series in Engineering and Computer Science.
- Lindeberg, T. (1992). Scale-space behaviour of local extrema and blobs. *Journal of Mathematical Imaging and Vision*, 1(1):65–99.

- Lindeberg, T. (1993a). Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, 11(3):283–318.
- Lindeberg, T. (1993b). Discrete derivative approximations with scale-space properties: A basis for low-level feature extraction. *Journal of Mathematical Imaging and Vision*, 3(4):349–376.
- Lindeberg, T. (1993c). On scale selection for differential operators. In Heia, K., Høgdra, K. A., and Braathen, B., editors, *Proceedings of the 8th Scandinavian Conference on Image Analysis*, pages 857–866, Tromsø, Norway. Norwegian Society for Image Processing and Pattern Recognition.
- Lindeberg, T. (1994a). On the axiomatic foundations of linear scale-space: Combining semi-group structure with causality vs. scale invariance. Technical Report ISRN KTH/NA/P-94/20-SE, Dept. of Numerical Analysis and Computing Science, KTH. (Submitted).
- Lindeberg, T. (1994b). Scale selection for differential operators. Technical Report ISRN KTH/NA/P-94/03-SE, Dept. of Numerical Analysis and Computing Science, KTH. Extended version to appear in Int. Jour. on Computer Vision.
- Lindeberg, T. (1994c). Scale-space for N-dimensional discrete signals. In O, Y.-L., Toet, A., Heijmans, H. J. A. M., Foster, D. H., and Meer, P., editors, *Proceedings of the NATO Advanced Research Workshop Shape in Picture - Mathematical Description of Shape in Greylevel Images*, volume 126 of *NATO ASI Series F: Computer and Systems Sciences*, pages 571–590, Berlin. Springer-Verlag. (Also available in Tech. Rep. ISRN KTH/NA/P-92/26-SE from Royal Inst of Technology).
- Lindeberg, T. (1994d). Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):225–270. Supplement *Advances in Applied Statistics: Statistics and Images: 2*.
- Lindeberg, T. (1994e). *Scale-Space Theory in Computer Vision*. The Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers.
- Lindeberg, T. (1995). Direct estimation of affine image deformations using visual front-end operations with automatic scale selection. In (ICCV95, 1995), pages 134–141.
- Lindeberg, T. (1996a). Edge detection and ridge detection with automatic scale selection. In (CVPR96, 1996), pages 465–470.
- Lindeberg, T. (1996b). Feature detection with automatic scale selection. Technical Report ISRN KTH/NA/P-96/18-SE, Dept. of Numerical Analysis and Computing Science, KTH.

- Lindeberg, T. (1996c). Linear spatio-temporal scale-space. in preparation.
- Lindeberg, T. (1996d). A scale selection principle for estimating image deformations. Technical Report ISRN KTH/NA/P-96/16--SE, Dept. of Numerical Analysis and Computing Science, KTH. To appear in *Image and Vision Computing*.
- Lindeberg, T. (1996e). Scale-space theory: A framework for handling image structures at multiple scales. In *Proc. CERN School of Computing*, Egmond aan Zee, The Netherlands.
- Lindeberg, T. and Fagerström, D. (1996). Scale-space with causal time direction. In (Buxton and Cipolla, 1996), pages 229–240.
- Lindeberg, T. and Gårding, J. (1993). Shape from texture from a multi-scale perspective. Technical Report ISRN KTH/NA/P-93/03-SE, Dept. of Numerical Analysis and Computing Science. Shortened version in *Proc. 4th Int. Conf. on Computer Vision*, Berlin, Germany, May 1993, pp. 683–691.
- Lindeberg, T. and Gårding, J. (1994). Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D structure. In (Eklundh, 1994), pages 389–400. To appear in *Image and Vision Computing*.
- Lindeberg, T. and Haar Romeny, B. M. t. (1994). Linear scale-space I: Basic theory. In (Haar Romeny, 1994), pages 1–41.
- Lindenbaum, M., Fischer, M., and Bruckstein, A. (1994). On Gabor's contribution to image enhancement. *Pattern Recognition*, 27:1–8.
- Lopez, C. and Morel, J. M. (1993). Formalisation mathematique et equations fondamentales de la pyramide visuelle. Technical report, CEREMADE. Université Paris Dauphine.
- Lucretius (1951). *On the Nature of the Universe*. Penguin Books Ltd, Harmondsworth, Middlesex, England. (Lucretius ca. 100–ca. 55 B.C.).
- Maas, R., Nielsen, M., Niessen, W. J., Haar Romeny, B. M. t., Florack, L. M. J., and Viergever, M. A. (1996). Local disparity measurements using scalable operators. In (Johansen, 1996), pages 80–87. DIKU Tech. Rep. Nr. 96/19.
- Magazieva, S. K. D. (1983). Numerical study of a partial differential equation. *U.S.S.R. Comput. Maths. Math. Phys.*, 23(4):45–49.
- Maintz, J. B. A., Elsen, P. A. v. d., and Viergever, M. A. (1994). *Medical imaging, analysis of multimodality 2D/3D images*, volume 19 of *Studies in Health, Technology and Informatics*, chapter Using geometrical features to match CT and MR brain images, pages 43–52. IOS Press, Amsterdam. Beolchi, L. and Kuhn, M., Eds.
- Maragos, P. (1995). Slope transforms: Theory and applications to nonlinear signal processing. *IEEE Transactions on Signal Processing*, 43(4):864–877.

- Marr, D. C. (1982). *Vision*. Freeman.
- Mason, B. J. (1962). *Clouds, rain and rainmaking*. Cambridge University Press, London.
- Mather, J. (1968). Stability of  $C^\infty$ -mappings: III. Finitely determined map germs. *Inst. Hautes Etudes Sci. Publ. Math.*, 36:127–156.
- Mather, J. (1969). Stability of  $C^\infty$ -mappings: II. Infinitesimal stability implies stability. *Ann. of Math.*, 89(2):254–291.
- Mather, J. (1970). Stability of  $C^\infty$ -mappings: V. Transversality. *Advances in Math.*, 37:301–336.
- McCabe, R., Willson, C., and Grubb, D. (1992). *Research considerations regarding FBI-IAFIS tasks & requirements*. NISTIR 4892, National Institute of Standards and Technology.
- Michelli, E. D., Caprile, B., Ottonello, P., and Torre, V. (1989). Localization and noise in edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:1106–1117.
- Milnor, J. (1963). Morse theory, volume 51 of *Annals of Mathematics Studies*. Princeton University Press.
- Misner, C. W., Thorne, K. S., and Wheeler, J. A. (1973). *Gravitation*. Freeman, San Francisco.
- Monga, O., Ayache, N., and Sander, P. T. (1992). From voxel to intrinsic surface features. *Image and Vision Computing*, 10(6):403–417.
- Morrison, P. and Morrison, P. (1982). *Powers of Ten, A book about the relative size of things in the universe and the effect of adding another zero*. Scientific American Library, New York.
- Morse, M. and Cairns, S. S. (1969). *Critical Point Theory in Global Analysis and Differential Topology*. Academic Press, New York and London.
- Morton, K. W. and Mayers, D. F. (1994). *Numerical solution of partial differential equations*. Cambridge University Press.
- Mundy, J. and Zisserman, A., editors (1992). *Applications of Invariance in Vision*. MIT Press.
- Nagel, H.-H. (1987). On the estimation of optical flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324.
- Nagel, H.-H. (1992). Direct estimation of optical flow and its derivatives. In Orban, G. A. and Nagel, H.-H., editors, *Artificial and Biological Vision Systems*, Basic Research Series, pages 193–224. Springer Verlag, Berlin.

- Najman, L. and Schmitt, M. (1994). Watershed of a continuos function. *Signal Processing*, 38(6):99–112.
- Nakahara, M. (1989). *Geometry, Topology and Physics*. Adam Hilger, Bristol and New York.
- Nakamura, O., Nagaoka, Y., and Minami, T. (1986). A restoration algorithm of fingerprint images. In *Systems and Computers in Japan*, volume 17:6, page 31.
- Nalwa, V. S. and Binford, T. O. (1986). On detecting edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:699–714.
- Nielsen, M., Florack, L., and Deriche, R. (1994). Regularization and scale-space. Technical Report RR-2352, INRIA, France.
- Nielsen, M., Florack, L., and Deriche, R. (1996a). Regularization, scale-space, and edge detection filters. In (Buxton and Cipolla, 1996), pages 70–81.
- Nielsen, M., Florack, L., and Deriche, R. (1996b). Regularization, scale-space, and edge detection filters. *Journal of Mathematical Imaging and Vision*. To appear.
- Niessen, W. J., Duncan, J. S., Florack, L. M. J., Haar Romeny, B. M. t., and Viergever, M. A. (1995). Spatiotemporal operators and optic flow. In Huang, S. T. and Metaxas, D. N., editors, *Physics-Based Modeling in Computer Vision*, pages 78–84. IEEE Computer Society Press.
- Niessen, W. J., Duncan, J. S., Nielsen, M., Florack, L. M. J., Haar Romeny, B. M. t., and Viergever, M. A. (In press). A multi-scale approach to image sequence analysis.
- Niessen, W. J., Haar Romeny, B. M. t., Florack, L. M. J., Salden, A. H., and Viergever, M. A. (1994a). Nonlinear diffusion of scalar images using well-posed differential operators. In (Bowyer et al., 1994), pages 92–97.
- Niessen, W. J., Haar Romeny, B. M. t., Florack, L. M. J., and Viergever, M. A. (1996a). A general framework for geometry-driven evolution equations. *International Journal of Computer Vision*. In press.
- Niessen, W. J., Haar Romeny, B. M. t., and Viergever, M. A. (1994b). Numerical analysis of geometry-driven diffusion equations. In (Haar Romeny, 1994), pages 393–410.
- Niessen, W. J., Nielsen, M., Florack, L. M. J., Maas, R., Haar Romeny, B. M. t., and Viergever, M. A. (1996b). Multiscale optic flow using physical constraints. DIKU Tech. Rep. Nr. 96/19.
- Niessen, W. J., Vincken, K. L., Koster, A. S. E., and Viergever, M. A. (1996c). A comparison of multiscale image representations for image segmentation. In Amini, A. A. and Bookstein, F. L., editors, *Proc. IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 263–272.

- Nitzberg, M. and Shiota, T. (1992). Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):826–833.
- Nordström, N. (1990). Biased anisotropic diffusion: A unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, 8:318–327.
- Nyquist, H. (1928). Certain topics in telegraph transmission theory. *AIEE Trans.*, 47.
- Okutomi, M. and Kanade, T. (1992). A locally adaptive window for signal matching. *International Journal of Computer Vision*, 7(2):143–162.
- Olsen, O. F. (1996). Multi-scale segmentation of grey-scale images. Technical report, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2200 Copenhagen East, Denmark. DIKU-rapport 96/30.
- Osher, S. and Sethian, S. (1988). Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formalism. *J. Computational Physics*, 79:12–49.
- Otsu, N. (1981). *Mathematical studies on feature extraction in pattern recognition*. PhD thesis, Researches of the Electrotechnical Laboratory, Ibaraki, Japan. (in Japanese).
- Otte, M. and Nagel, H.-H. (1994). Optical flow estimation: Advances and comparisons. In Eklundh, J.-O., editor, *Proceedings of the European Conference on Computer Vision*, volume 800 of *Lecture Notes in Computer Science*, pages 51–60, Stockholm, Sweden.
- Pauwels, E. J., Fiddelaers, P., Moons, T., and Gool, L. J. v. (1995). An extended class of scale-invariant and recursive scale-space filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(7):691–701.
- Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639.
- Perona, P., Shiota, T., and Malik, J. (1994). Anisotropic diffusion. In (Haar Romeny, 1994), pages 72–92.
- Piles, R. d. (Written 1668–1708). *Course de Peinture par Principes*. Éditions Jacqueline Chambon, Nîmes, 1990.
- Pizer, S., Gauch, J., and Lifshitz, L. M. (1988). Interactive 2D and 3D object definition in medical images based on multiresolution image descriptions. In *SPIE Proceedings*, volume 914:B, 438–449.
- Poggio, H. V. T. and Yuille, A. (1985). A regularized solution to edge detection. Technical Report A. I. Memo 883, M.I.T.

- Posmentier, E. S. (1977). The generation of salinity finestructure by vertical diffusion. *J. Phys. Oceanogr.*, 7:298–300.
- Pratt, B. (1978). *Digital Image Processing*. Wiley-Interscience.
- Rambaux, I. and Garçon, P. (1994). Nonlinear anisotropic diffusion filtering of 3D images. Technical report, Département Génie Mathématique, INSA de Rouen and Laboratory of Technomathematics, University of Kaiserslautern.
- Rao, A. R. and Schunk, B. G. (1991). Computing oriented texture fields. *Computer Vision, Graphics, and Image Processing : Graphical Models and Image Processing*, 53(2):157–185.
- Reichenbach, H. (1960). *The Theory of Relativity and A Priori Knowledge*. University of California Press, Berkeley. Translated by M. Reichenbach; original German edition published in 1920.
- Rieger, J. (1995). Generic evolution of edges on diffused greyvalue surfaces. *Journal of Mathematical Imaging and Vision*, 5:207–217.
- Rudin, W. (1981). *Real and Complex analysis*. McGraw-Hill, New York.
- Ruskin (1860). *Modern painters*, volume 5. (See: Library Edition of the Collected Works of John Ruskin, Eds. E.T.Cook and A.Wedderburn, 39 Vols, London, 1903–12), London.
- Ruskin, J. (1857). *The elements of drawing*. Smith, Elder & Co., London.
- Salden, A. H. (1996). *Dynamic Scale-Space Paradigms*. PhD thesis, Utrecht University, The Netherlands.
- Salden, A. H., Florack, L. M. J., Haar Romeny, B. M. t., Koenderink, J. J., and Viergever, M. A. (1992a). Multi-scale analysis and description of image structure. *Nieuw Archief voor Wiskunde*, 10(3):309–326.
- Salden, A. H., Haar Romeny, B. M. t., Florack, L. M. J., Koenderink, J. J., and Viergever, M. A. (1992b). A complete and irreducible set of local orthogonally invariant features of 2-dimensional images. In Young, I. T., editor, *Proceedings 11th IAPR Internat. Conf. on Pattern Recognition*, pages 180–184, The Hague, the Netherlands. IEEE Computer Society Press, Los Alamitos.
- Salden, A. H., Haar Romeny, B. M. t., and Viergever, M. A. (1994). Local and multilocal scale-space description. In O, Y.-L., Toet, A., Heijmans, H. J. A. M., Foster, D. H., and Meer, P., editors, *Proc. of the NATO Advanced Research Workshop Shape in Picture - Mathematical description of shape in greylevel images*, volume 126 of *NATO ASI Series F*, pages 661–670. Springer Verlag, Berlin.
- Salden, A. H., Haar Romeny, B. M. t., and Viergever, M. A. (1995a). Affine and projective scale space theories. In *Proc. Conf. on Differential Geometry and Computer Vision: From Pure over Applicable to Applied Differential Geometry*, Nordfjordeid, Norway.

- Salden, A. H., Haar Romeny, B. M. t., and Viergever, M. A. (1995b). Classical scale space theory from physical principles. Submitted.
- Salden, A. H., Haar Romeny, B. M. t., and Viergever, M. A. (1996). Algebraic invariants of linear scale spaces. Submitted.
- Sapiro, G. and Tannenbaum, A. (1993). Affine invariant scale-space. *International Journal of Computer Vision*, 11(1):25–44.
- Sapiro, G., Tannenbaum, A., and Olver, P. J. (1994). Classification and uniqueness of invariant geometric flows. *C. R. Acad. Sci. Paris*, t. 319(Série I):339–344.
- Sasco, G. and Almeida, A. (1995). Reconocimiento de huellas dactilares. Master's thesis, CeCal, Engineering School, University of Uruguay, Montevideo, Uruguay. Taller V.
- Saunders, P. T. (1980). *An introduction to Catastrophe Theory*. Cambridge University Press.
- Schmid, C. and Mohr, R. (1996). Combining greyvalue invariants with local constraints for object recognition. In (CVPR96, 1996).
- Schoenberg, I. J. (1950). On Pólya frequency functions. ii. Variation-diminishing integral operators of the convolution type. *Acta Sci. Math. (Szeged)*, 12:97–106.
- Schoenberg, I. J. (1953). On smoothing operations and their generating functions. *Bull. Amer. Math. Soc.*, 59:199–230.
- Schunck, B. G. (1984). The motion constraint equation for optical flow. In (ICPR84, 1984), pages 20–22.
- Schwartz, L. (1950–1951). *Théorie des Distributions*, volume I, II of *Actualités scientifiques et industrielles; 1091,1122*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, Paris.
- Schwartz, L. (1966). *Théorie des Distributions*. Hermann, Paris, 2nd edition edition.
- Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press.
- Serra, J. (1988). *Image Analysis and Mathematical Morphology, Vol 2*. Academic Press.
- Spivak, M. (1965). *Calculus on Manifolds*. W. A. Benjamin, New York.
- Spivak, M. (1975). *Differential Geometry*, volume 1–5. Publish or Perish, Berkeley.
- Stokking, R., Zuiderveld, K. J., Hulshoff Pol, H. E., and Viergever, M. A. (1994). Integrated visualization of SPECT and MR images for frontal lobe damaged regions. In Robb, R. A., editor, *Visualization in Biomedical Computing 1994*, pages 282–290. Proc. SPIE Vol. 2359, SPIE Press Bellingham WA, Rochester.

- Super, B. J. and Bovik, A. C. (1992). Shape-from-texture by wavelet-based measurement of local spectral moments. In *Conference on Computer Vision and Pattern Recognition*, pages 296–301, Champaign, IL.
- Sze, M. (1959). *The Way of Chinese Painting, its ideas and technique*. Vintage books, New York.
- Thirion, J.-P. and Gourdon, A. (1995). Computing the differential characteristics of isodensity surfaces. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 61(2):109–202.
- Thom, R. (1972). Stabilité structurelle et morphogénèse. Paris.
- Tikhonov, A. N. and Arsenin, V. Y. (1977). *Solution of ill-posed problems*. V.H Winston & Sons, John Wiley.
- Tistarelli, M. (1994). Multiple constraints for optical flow. In (Eklundh, 1994), pages 61–70.
- Torre, V. and Poggio, A. (1986). On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:147–163.
- Tretiak, O. and Pastor, L. (1984). Velocity estimation from image sequences with second order differential operators. In (ICPR84, 1984), pages 16–19.
- U. S. Department of Justice (1984). The science of fingerprints. Washington, D.C.
- Unser, M., Aldroubi, A., and Eden, M. (1991). Recursive regularization filters: Design, properties and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3).
- Verbeek, P. and Verwer, B. J. H. (1989). 2D adaptive smoothing by 3D distance transforms. *Pattern Recognition Letters*, 9:53–65.
- Verri, A., Girosi, F., and Torre, V. (1990). Differential techniques for optical flow. *Journal of the Optical Society of America-A*, 7(5):912–922.
- Vincent, L. and Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598.
- Vincken, K. L., de Graaf, C. N., Koster, A. S. E., Viergever, M. A., Appelman, F. J. R., and Timmens, G. R. (1990). Multiresolution segmentation of 3D images by the hyperstack. In *Proc. First Conference on Visualization in Biomedical Computing*, pages 115–122. IEEE Computer Society Press, Los Alamitos, CA.
- Vincken, K. L., Koster, A. S. E., and Viergever, M. A. (1994). Probabilistic segmentation of partial volume voxels. *Pattern Recognition Letters*, 15(5):477–484.

- Vincken, K. L., Niessen, W. J., Koster, A. S. E., and Viergever, M. A. (1996). Blurring strategies for image segmentation using a multiscale linking model. In (CVPR96, 1996).
- Vliet, L. v. and Verwer, B. J. H. (1988). A contour processing method for fast binary neighbourhood operations. *Pattern Recognition Letters*, 7(1):27–36.
- Waerden, B. L. v. d. (1940). *Moderne Algebra*, volume I-II. Springer-Verlag, Berlin.
- Webster, A. G. (1955). *Partial differential equations of mathematical physics*. Dover, New York. (First edition 1927).
- Weickert, J. (1994). Anisotropic diffusion filters for image processing based quality control. In Fasano, A. and Primicerio, M., editors, *Proc. Seventh European Conf. on Mathematics in Industry*, pages 355–362, Teubner, Stuttgart.
- Weickert, J. (1995). Multiscale texture enhancement. In Hlaváč, V. and Šára, R., editors, *Computer Analysis of Images and Patterns*, volume 970 of *Lecture Notes in Comp. Science*, pages 230–237. Springer, Berlin. (Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns, Prague, Czech Republic, September 1995).
- Weickert, J. (1996a). *Anisotropic diffusion in image processing*. PhD thesis, Dept. of Mathematics, University of Kaiserslautern, Germany. Revised and extended version to be published by Teubner Verlag, Stuttgart, Germany, 1997.
- Weickert, J. (1996b). Nonlinear diffusion scale-spaces: From the continuous to the discrete setting. In Berger, M. O., Deriche, R., Herlin, I., Jaffré, J., and Morel, J.-M., editors, *ICAOS '96: Images, wavelets and PDEs*, volume 219 of *Lecture Notes in Control and Information Sciences*, pages 111–118. Springer, London.
- Weickert, J. (1996c). Theoretical foundations of anisotropic diffusion in image processing. In Kropatsch, W., Klette, R., and Solina, F., editors, *Theoretical foundations of computer vision*, Computing Supplement 11, pages 221–236. Springer, Wien.
- Weickert, J. A., Haar Romeny, B. M. t., and Viergever, M. A. (1996). Conservative image transformations with restoration and scale-space properties. In (ICIP96, 1996), pages 465–468.
- Weitzenböck, R. (1923). *Invariantentheorie*. P. Noordhoff, Groningen.
- Werkhoven, P. (1990). *Visual Perception of Successive Order*. PhD thesis, University of Utrecht, Department of Medical Physics, Utrecht, the Netherlands.

- Whitaker, R. T. (1993). Characterizing first and second-order patches using geometry-limited diffusion. In Barrett, H. H. and Gmitro, A. F., editors, *Information processing in medical imaging*, volume 687 of *Lecture Notes in Comp. Science*, pages 149–167. Springer, Berlin.
- Whitaker, R. T. and Pizer, S. M. (1993). A multi-scale approach to nonuniform diffusion. *CVGIP*, 57(1):99–110.
- Widder, D. V. (1975). *The Heat Equation*. Academic Press.
- Wilson, C. I. W. C. L. (1992). NIST special database 4, fingerprint database. U. S. Department of Commerce, Technology Administration, National Institute of Standards and Technology, Advanced Systems Division, Image Recognition Group, Gaithersburg, MD 20899.
- Witkin, A. P. (1983). Scale-space filtering. In *Proc. 8th Int. Joint Conf. on Artificial Intelligence (IJCAI '83)*, volume 2, pages 1019–1022, Karlsruhe, Germany.
- Witkin, A. P. (1984). Scale space filtering: A new approach to multi-scale description. In Ullman, S. and Richards, W., editors, *Image understanding 1984*. Norwood Ablex, NJ.
- Wolferen, W. A. v., Boschma, T. S., Hofstee-Hoofstman, T. W. A., Matthijssen, R. A. H., Doorn, J. J., Polman, E. B. H. M., Zuiderveld, K., Zonneveld, F. W., den Boer, J. A., and Maat, G. J. R. (1993). Paranasal sinuses & anterior skull base. volume 1-I. Elsevier, Amsterdam, the Netherlands.
- Yuille, A. L. (1988). The creation of structure in dynamic shape. In *Proceedings of the 2nd International Conference on Computer Vision*, pages 685–689, Tarpon Springs, Florida.
- Yuille, A. L. and Poggio, T. A. (1985). Fingerprint theorems for zero crossings. *J. of the Optical Society of America (A)*, 2:683–692.
- Yuille, A. L. and Poggio, T. A. (1986). Scaling theorems for zero crossings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:15–25.
- Zuiderveld, K. J. (1995). *Visualization of Multimodality Medical Volume Data using Object-Oriented Methods*. PhD thesis, Utrecht University.
- Zuiderveld, K. J. and Viergever, M. A. (1992). Visualization of volumetric medical image data. In Dewilde, P. and Vandewalle, J., editors, *Computer Systems and Software Engineering*, pages 363–385. Kluwer Academic Publishers, Dordrecht.
- Zuiderveld, K. J. and Viergever, M. A. (1994). Multi-modal volume visualization using object-oriented methods. In *1994 Symposium on Volume Visualization*, pages 59–66, New York. ACM SIGGRAPH.

# Index

- $\mathcal{A}$ -equivalence, 161  
 $\mathcal{A}$ -transform, 212  
affine  
    corner detection, 4  
    equivalence problem, 122  
group, 117  
    invariant evolution, 19  
scale-space, 23, 50n, 77, 92ff, 165ff  
    scale-space foliation, 166  
anisotropic diffusion, 13, 83  
anisotropy measure, 24  
annihilate, 139, 149, 159, 188, 206  
    four critical points, 179  
aperture problem, 32, 34, 42, 66  
atlas principle, 214, 217  
autoconvolution algebra, 64  
average grey level invariance, 52, 56,  
    226, 230, 233  
axiomatics, 53, 76, 86, 91, 99, 212  
Baire category theorem, 151  
Baire space, 151  
band limited signal, 92, 130  
basic equation of figure, 50  
basic principles, *see* axiomatics  
Bessel functions, 81  
binary form, 8, 119  
binocular stereo, 35ff  
binomial smoothing, 79  
boundary conditions, 229  
brightness constraint equation, 32  
Burgers vectors, 125  
Canny-Deriche filter, 100  
carry-along, 65  
Cartesian notation, 4  
Castan filter, 100  
catastrophe, 168, 170, 187  
germ, 192  
points, 182  
theory, 149  
catchment basins, 192, 193  
Cauchy problem, 116  
causality, 55ff, 77ff, 86, 195  
Cayley's omega process, 119  
Cayley-form, 120  
central moments, 227  
 $C^\infty$ -topology, 151  
 $C^k$  functions, 151  
closing, 206  
codimension, 139ff, 158, 162  
    of a germ, 158  
completely transverse, 161ff  
computational molecule, 95  
concatenation property, 167  
concave hull, 215  
conductivity, 96  
contact probes, 208  
continuity requirement, 87  
continuous scale parameter, 86  
continuous signals, 91  
contrast enhancement, 227  
contrast parameter, 223  
convergence, 227ff  
convolution, 53  
convolution algebra, 101  
corners, 5  
correlation, 135  
correspondence problem, 221  
covariance, 117  
covariance function, 134, 135  
covariance matrix, 93  
covectors, 63  
creation, 139, 149, 159  
critical curve, 140ff

- critical point, 140ff, 144, 148ff, 168, 169, 192
- CT, 11
- CT/MR matching, 11
- curve evolution, 13
- cusp catastrophe, 169ff, 196
- cusp catastrophe curve, 175
- deblurring, 9, 58
- deep structure, 12, 93
- deformation of a germ, 152
- degenerate topoints, 141, 146
- denoising, 13
- density, 65, 69
  - filter current, 69
  - probe, 208
- derivative
  - covariant, 184
  - Lie, 69–71
  - partial, 68
- detection scale, 199
- detector, 61
- device space, 63
- diffeomorphic, 168
- differentiability axiom, 88
- diffusion tensor, 234
- diffusivity, 223ff
- dilation, 46, 205
- dimensional analysis, 83
- discrete, 80
  - affine scale-space, 94ff
  - diffusion, 231
  - scale-space, 130ff
  - signal, 75ff, 92, 129ff
- discretisation, 129ff, 136, 225, 228, 231, 233
- discretisation grid, 92, 135, 228
- discriminant, 8
- dislocates edges, 221
- disparity, 31, 35, 92
- duality, 63, 93n
  - morphological, 63
  - principle, 62
  - topological, 53, 63
- edge
- detection, 135, 221ff
- enhancement, 224ff
- preserving smoothing, 13ff
- Einstein convention, 4
- elliptic
  - aperture, 166
  - umbilic catastrophe, 170, 177ff
- entropy, 227
- epipolar constraint, 35
- epsilon tensor, 4
- equivalence, 148, 162
  - problem, Euclidean, 118, 127
  - problem, similarity, 117
- Erlangen Programm, 121
- erosion, 46, 205
- event, *see* catastrophe
- evolution
  - equation, 54, 88, 96
  - of level curves, 46
  - scheme, 95
- explicit scheme, 232ff
- extremum, 139ff, 168ff, 181ff, 192ff
  - curve, 140
  - cusp, 174
  - cusp curve, 177
  - principle, 226, 230ff
- Feature detection, 4
- filter, *see* detector
- Fingerprint Identification Systems, 21
- fingerprints, 78
- finite determinacy theorem, 160ff
- first principles, *see* axiomatics
- flux function, 223
- fold catastrophe, 169ff, 172, 196
- fold surfaces, 169ff
- foliated, 166
- forward–backward diffusion, 224
- Fourier transform of derivative, 9
- Frank vectors, 125
- Frobenius norm, 38
- function space topology, 150
- gauge
  - condition, 67
  - condition, canonical, 67

- condition, temporal, 69
- coordinate, 4ff, 195
- field, 66
- theory, 66
- Gaussian
  - curvature, 6
  - family, 64
- Gaussian scale-space axiomatics, *see* axiomatics
- generic event, 140ff, 147ff
  - concave, 143
  - convex, 143
  - non-linear PDE, 158
- geometric diffusion, 13
- geometric reasoning, 6
- germ, 149ff, 157
- Green's function, 100
- group
  - affine, *see* affine group
  - homotopy, *see* homotopy
  - similarity, *see* similarity group
  - symmetry, *see* symmetry group
- H**
  - equivalence, 149ff, 157ff
  - equivalence, example, 150, 158
  - generic, 153
  - stable, 153, 157
- $\mathcal{H}(U)$ , 151
- handkerchief surface, 170
- harmonic polynomials, 155ff
- heat operator, 154
- heat polynomial
  - decomposition, 154
  - weighted decomposition, 154
- Hessian, 4, 140ff, 148, 161
- Hilbert space, 152
- hit-property, 208
- homogeneity, 57, 78, 213
- homogeneous polynomials, 154
- homotopy, 66, 183
- Hopf bifurcation, 168
- Horn & Schunck equation, 34
- Huygens' principle, 208
- hyperbolic umbilic catastrophe, 170, 179ff
- hyperstack, 12
- ideal interpolation, 135
- ideal sampling, 130
- ill-posedness, 223ff
- image
  - acquisition, 129ff
  - concept, 66
  - enhancement, 58
  - processing paradigm, 64
  - raw, 61
- implicit scheme, 234
- index notation, 4
- infinite codimension, 158
- infinitesimal generator, 54, 81, 88, 91, 95
- initial functions, 152
- inner scale, 203
- integrated visualization, 11
- integration scale, 23
- interactive anatomy atlas, 19
- interocular line, 36
- interpolation, 129, 133
  - ideal, 132–134
  - random field, 134
- intrinsic invariants, 115
- invariance
  - projective, 46
  - rotational, 51, 56, 86
  - scale, 48, 57ff, 77, 83ff, 91ff, 104, 215
  - theory for forms, 121
  - topological, 183
  - translational, 47ff, 86
- invariant, 77, 117
- inverse slope transform, 212
- inversion of the scale-space transformation, 53
- irreducibility, 8, 230, 232, 233
- IS
  - equivalence, 150ff
  - generic, 153
  - stable, 153ff
- isometry invariance, 56ff
- isophote curvature, 5
- isotropy, 57, 78, 167, 214

- jet extension map, 161
- jet space, 161, 204
- junction detection, 9
- Kronecker tensor, 4
- $L^2$ -metric, 152
- Lambertian reflection, 36
- Legendre transformation, 212
- level
  - crossings, 227
  - curve, 96
  - curve evolution, 94
  - surface, 78
- Levi-Civita tensor, 4
- linear forms, 63
- linearity
  - assumption, 46ff, 86
  - integral operator, 51, 53
- linking, 12
- local
  - density measure, 203
  - diffeomorphism, 149
  - equivalence, 149
  - germ, 169
  - image features, 204
  - infinitesimal generator, 96
  - normal forms, 158
  - operator, 90
  - sample, 62
  - sample, cross-section of, 64
- locality, 54, 82ff, 91, 94
- localization problems, 222
- localization scale, 199
- Lyapunov functionals, 222, 227, 231
- Lyapunov sequences, 233
- manifest invariant notation, 4
- mathematical morphology, 204
- maximum principle, 55
- maximum–minimum principle, 222
- mean curvature, 6
- measurement probe, 203
- medial axis algorithm, 208
- merging
  - regions, 160, 197
  - singularities, 139, 188
- minutiae, 22
- mirror symmetry, 86
- monkey saddle, 170, 177, 186
- monotone intensity transformations, 96
- monotony, 55
- morphological
  - eigenfunctions, 209, 211
  - impulse response, 205
  - operators, 205ff
  - PDE interpretation, 218
  - pulse, 218
  - scale-spaces, 227
  - superposition principle, 204
- Morse
  - critical points, 149
  - functions, 168, 192
  - lemma, 150
  - theory, 148ff, 159
- MR, 11
- multi index, 151
- multi-camera, 31
- multi-frame, 31
- multimodality matching, 11
- N-jet, *see* jet space
- natural coordinates, 34
- noise, 129
- non-creation of extrema, 78, 81, 85
- non-degenerate critical points, 148, 152
- non-degenerate saddle, 186
- non-enhancement of local extrema, 82ff, 88ff, 95
- non-generic, *see* generic event
- non-linear
  - diffusion, 13, 46, 92, 95, 158, 222
  - scale-space, 46, 77, 96
- non-negativity, 52, 55
- non-uniform diffusion, 75
- non-uniform scale-space, 92
- normal flow, 34, 66
- normal form, 148ff
- normalization, 52, 56ff, 86
- numerical implementation, 9, 231
- object velocities, 31

- opening, 206
- optic flow, 34ff, 71ff, 94
  - intrinsic, 35
  - semantical, 67
  - syntactical, 71
- parabolic differential equation, 91
- parallel slope transport, 209
- Perona–Malik model, 222
- PET, 11
- point operator, 64
- Poisson bracket, 119
- polar of binary forms, 120
- positivity, 48ff, 86, 226, 230ff
- poweroids, 216
- pre-scale-space, 87
- principal curvature, 6
- projective invariances, 46
- propagator, 66
- pull back, 65
- push forward, 65
- pyramid, 76, 80
- quad-trees, 76
- quadratic structuring element, 64
- quantisation, 130
- quartic hump, 174
- quartic saddle, 174
- random field, 133ff
  - interpolation of, 134
- recursive filtering, 79, 111ff
- regularity, 54
- regularization, 99, 224ff
- residual, 151
- resultant, 8, 120
- ridge, 11
- ridge enhancement, 25ff
- ridge operator, 11
- root labeling, 12
- rotation invariance, *see* invariance, rotational
- rotational symmetry, *see* invariance, rotational
- saddle
  - extremum pair, 139
- annihilation, 172
- creation, 172
- cusp, 174ff
- cusp curve, 177
- point, 139ff, 168ff
- sampling, 129ff
- scale invariance, *see* invariance, scale
- scale selection, 6, 26, 37ff, 95
- scale-invariant filtering, 101
- scale-recursive filters, 101
- scale-space
  - axioms, *see* axiomatics
  - first citation, 58n
  - generator, 52, 99
  - infinitesimal generator, *see* infinitesimal generator
  - operator, 99
- Schwartz space, 63
- second moment matrix, 23, 95
- segmentation, 12, 191
- semi-discrete diffusion, 228ff
- semi-group property, 48, 52ff, 57, 77, 80ff, 105, 227
- sensorium, *see* device space
- separability, 52, 57, 85
- separating
  - points, 145, 187
  - regions, 160, 197
- separatrices, 193
- shape adaptation, 28, 95
- shape distortions, 92
- shift-invariant operators, 76
- shoe surface, 170
- similarity group, 116
- similarity jet, 118
- singularity, 5, *see also* critical point
  - annihilation, *see* annihilate
  - creation, *see* creation
  - curve, 140
  - Morse, 185
  - non-degenerate, 185
  - of gradient magnitude, 194
  - order, 181
  - scattering, 188
  - strings, 187
  - topological, 185

- skeleton algorithm, 208
- slope
  - domain, 204
  - inverse transform, 212
  - line, **193**, 195
  - line curvature, 195
  - points, 168
  - transform, 209ff, **212**
- smoothness property, 54ff, 63, 226ff
- Sobolev space, 100
- spacetime, 61, 64ff
- spatio-temporal operators, 32
- spatio-temporal scale-space, 64, 77, 95
- SPECT, 11
- spectral density, 134
- splitting, *see* separating
- splitting lemma, 195
- stability, 148
  - under deformations, 153
- stabilizing constraints, 224
- staircasing effect, 225
- state space, 62
- stationary
  - random field, 130ff
  - strictly, 134
  - wide sense, 134
- steepest descent lines, 193
- stereo matching, 31ff, 94
- stochastic analysis, 129ff
- Stoke's theorem, 184
- strong continuity, 88, 91
- structuring element, 63, 206
- subset
  - dense, 151
  - open, 151
- superposition principle, 205, **213**
- swallowtail catastrophe, 170, 177ff
- Sylvester determinant, 120
- symmetry group, 146
- system states, 118
- tangent, 141
- tangential dilation, **210**
- tangential flow, 32, 66
- tempered distributions, 63
- temporal scale-space, 81
- tensorial notation, 4
- Thom's classification theorem, 168ff, 192
- Thom's transversality theorem, 161
- Tikhonov regularization, 56, 100
- time-causality, 81
- topological
  - current, 187
  - duality, 53, 63
  - invariants, 183
  - numbers, 181ff
  - singularity, 185
- toppoints, 140ff
  - image representation by, 146
- translational invariance, 47ff, 86
- transvectants, 8, 122
- transvection, 119
- transversality, 148, 160ff
- umbral scaling, 215
- unconditional stability, 234
- unfolding, 158ff, 169, 174
  - theorem, 162
  - versal, 153, 158, 162
- unimodality, 86
- union-of-translates, 208
- variance, 227
- velocity estimation, 31
- velocity field, 33
- visual front-end, 86
- visual system, 92
- watershed segmentation, 208
- watersheds, 192, **194**
- wave front propagation, 208
- wavelets, 76
- weighted homogeneity, 158
- well-posedness, 225ff
- white noise, 135
- winding number, 182ff
- Young-Fenchel conjugate, 212

## Computational Imaging and Vision

---

1. B.M. ter Haar Romeny (ed.): *Geometry-Driven Diffusion in Computer Vision*. 1994 ISBN 0-7923-3087-0
2. J. Serra and P. Soille (eds.): *Mathematical Morphology and Its Applications to Image Processing*. 1994 ISBN 0-7923-3093-5
3. Y. Bizais, C. Barillot, and R. Di Paola (eds.): *Information Processing in Medical Imaging*. 1995 ISBN 0-7923-3593-7
4. P. Grangeat and J.-L. Amans (eds.): *Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*. 1996 ISBN 0-7923-4129-5
5. P. Maragos, R.W. Schafer and M.A. Butt (eds.): *Mathematical Morphology and Its Applications to Image and Signal Processing*. 1996 ISBN 0-7923-9733-9
6. G. Xu and Z. Zhang: *Epipolar Geometry in Stereo, Motion and Object Recognition*. A Unified Approach. 1996 ISBN 0-7923-4199-6
7. D. Eberly: *Ridges in Image and Data Analysis*. 1996 ISBN 0-7923-4268-2
8. J. Sporring, M. Nielsen, L. Florack and P. Johansen (eds.): *Gaussian Scale-Space Theory*. 1997 ISBN 0-7923-4561-4