Front-End Vision and Multi-Scale Image Analysis

Computational Imaging and Vision

Managing Editor

MAX A. VIERGEVER
Utrecht University, Utrecht, The Netherlands

Editorial Board
GUNILLA BORGEFORS, *Centre for Image Analysis, SLU, Uppsala, Sweden*
THOMAS S. HUANG, *University of Illinois, Urbana, USA*
SABURO TSUJI, *Wakayama University, Wakayama, Japan*

Volume 27

Front-End Vision and Multi-Scale Image Analysis

Multi-Scale Computer Vision Theory and Applications, written in Mathematica

by

Bart M. ter Haar Romeny

Eindhoven University of Technology,
Department of Biomedical Engineering,
Biomedical Imaging and Informatics,
Eindhoven, The Netherlands

Springer

Prof. Dr. Bart M. ter Haar Romeny
Eindhoven University of Technology
Fac. Biomedical Engineering
Dept. Image Analysis & Interpretation
5600 MB Eindhoven
Netherlands

This book is available in Mathematica, a mathematical programming language, at the following website:
http://bmia.bmt.tue.nl/Education/Courses/FEV/book/index.html

ISBN 978-1-4020-1503-8 (HB) e-ISBN 978-1-4020-8840-7 (e-book)
ISBN 978-1-4020-1507-6 (PB)

Library of Congress Control Number: 2008929603

Reprinted in 2008

© 2003 Springer Science + Business Media B.V.
No part of this work may be reproduced, stored in a retrieval system, or transmitted
in any form or by any means, electronic, mechanical, photocopying, microfilming, recording
or otherwise, without written permission from the Publisher, with the exception
of any material supplied specifically for the purpose of being entered
and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Table of contents

Front-End Vision and Multi-Scale Image Analysis xiii
The purpose of this book xiii
Scale-space theory is biologically motivated computer vision xiv
This book has been written in *Mathematica*. xvi
Acknowledgements xviii

1. Apertures and the notion of scale 1
1.1 Observations and the size of apertures 1
1.2 Mathematics, physics, and vision 2
1.3 We blur by looking 5
1.4 A critical view on observations 9
1.5 Summary of this chapter 12

2. Foundations of scale-space 13
2.1 Constraints for an uncommitted front-end 13
2.2 Axioms of a visual front-end 15
2.2.1 Dimensional analysis 15
2.2.2 The cooking of a turkey 16
2.2.3 Reynold's number 18
2.2.4 Rowing: more oarsmen, higher speed? 19
2.3 Axiomatic derivation of the Gaussian kernel 21
2.4 Scale-space from causality 23
2.5 Scale-space from entropy maximization 25
2.6 Derivatives of sampled, observed data 27
2.7 Scale-space stack 31
2.8 Sampling the scale-axis 32
2.9 Summary of this chapter 35

3. The Gaussian kernel 37
3.1 The Gaussian kernel 37
3.2 Normalization 38
3.3 Cascade property, selfsimilarity 39
3.4 The scale parameter 40
3.5 Relation to generalized functions 40
3.6 Separability 43
3.7 Relation to binomial coefficients 43
3.8 The Fourier transform of the Gaussian kernel 44
3.9 Central limit theorem 46
3.10 Anisotropy 48
3.11 The diffusion equation 49
3.12 Summary of this chapter 50

4. Gaussian derivatives ... 53
4.1 Introduction ... 53
4.2 Shape and algebraic structure ... 53
4.3 Gaussian derivatives in the Fourier domain ... 57
4.4 Zero crossings of Gaussian derivative functions ... 59
4.5 The correlation between Gaussian derivatives ... 60
4.6 Discrete Gaussian kernels ... 64
4.7 Other families of kernels ... 65
4.8 Higher dimensions and separability ... 67
4.9 Summary of this chapter ... 69

5. Multi-scale derivatives: implementations ... 71
5.1 Implementation in the spatial domain ... 71
5.2 Separable implementation ... 73
5.3 Some examples ... 74
5.4 N-dim Gaussian derivative operator implementation ... 78
5.5 Implementation in the Fourier domain ... 79
5.6 Boundaries ... 83
5.7 Advanced topic: speed concerns in *Mathematica* ... 85
5.8 Summary of this chapter ... 89

6. Differential structure of images ... 91
6.1 The differential structure of images ... 91
6.2 Isophotes and flowlines ... 92
6.3 Coordinate systems and transformations ... 96
6.4 Directional derivatives ... 102
6.5 First order gauge coordinates ... 103
6.6 Gauge coordinate invariants: examples ... 108
6.6.1 Ridge detection ... 108
6.6.2 Isophote and flowline curvature in gauge coord ... 110
6.6.3 Affine invariant corner detection ... 113
6.7 A curvature illusion ... 115
6.8 Second order structure ... 117
6.8.1 The Hessian matrix and principal curvatures ... 119
6.8.2 The shape index ... 120
6.8.3 Principal directions ... 122
6.8.4 Gaussian and mean curvature ... 123
6.8.5 Minimal and zero Gaussian curvature surfaces ... 126
6.9 Third order image structure: T-junction detection ... 127
6.10 Fourth order image structure: junction detection ... 131
6.11 Scale invariance and natural coordinates ... 132
6.12 Irreducible invariants ... 134
Intermezzo: Tensor notation ... 135
6.13 Summary of this chapter ... 136

7. Natural limits on observations 137
7.1 Limits on differentiation: scale, accuracy and order 137
7.2 Summary of this chapter 141

8. Differentiation and regularization 143
8.1 Regularization 143
8.2 Regular tempered distributions and test functions 144
8.3 An example of regularization 147
8.4 Relation regularization $\Longleftrightarrow$ Gaussian scale-space 148
8.5 Summary of this chapter 152

9. The front-end visual system - the retina 153
9.1 Introduction 153
9.2 Studies of vision 154
9.3 The eye 156
9.4 The retina 157
9.1 Retinal receptive fields 160
9.6 Sensitivity profile measurement of a receptive field 162
9.7 Summary of this chapter 165

10. A scale-space model for the retinal sampling 167
10.1 The size and spatial distribution of receptive fields 167
10.2 A scale-space model for the retinal receptive fields 172
10.3 Summary of this chapter 177

11. The front-end visual system - LGN and cortex 179
11.1 The thalamus 179
11.2 The lateral geniculate nucleus (LGN) 181
11.3 Corticofugal connections to the LGN 183
11.4 The primary visual cortex 185
11.4.1 Simple cells 187
11.4.2 Complex cells 188
11.4.3 Directional selectivity 189
11.5 Intermezzo: Measurement of neural activity in the brain 191
Electro-Encephalography (EEG) 191
Magneto-Encephalography (MEG) 192
Functional MRI (fMRI) 193
Optical imaging with voltage sensitive dyes 194
Positron Emission Tomography (PET) 194
11.6 Summary of this chapter 195

12. The front-end visual system - cortical columns 197
12.1 Hypercolumns and orientation structure 197
12.2 Stabilized retinal images 200
12.3 The concept of local sign 202
12.4 Gaussian derivatives and Eigen-images 204

12.5 Plasticity and self-organization ... 208
12.6 Higher cortical visual areas ... 210
12.7 Summary of this chapter ... 211
12.8 Vision dictionary ... 211
12.8.1 Further reading on the web: ... 212

13 Deep structure I. watershed segmentation ... 215
13.1 Multi-scale measurements ... 215
13.2 Scale selection ... 216
13.3 Normalized feature detection ... 218
13.4 Automatic scale selection ... 219
13.4.1 λ-Normalized scale selection ... 220
13.4.2 Is this really deep structure? ... 220
13.5 Edge focusing ... 221
13.5.1 Simplification followed by focusing ... 221
13.5.2 Linking in 1D ... 222
13.6 Follicle detection in 3D ultrasound ... 225
13.6.1 Fitting spherical harmonics to 3D points ... 229
13.7 Multi-scale segmentation ... 231
13.7.1 Dissimilarity measure in scale-space ... 231
13.7.2 Watershed segmentation ... 232
13.7.3 Linking of regions ... 234
13.7.4 The multi-scale watershed segmentation ... 237
13.8 Deep structure and nonlinear diffusion ... 239
13.8.1 Non-linear diffusion watershed segmentation ... 239

14. Deep structure II. catastrophe theory ... 241
14.1 Catastrophes and singularities ... 241
14.2 Evolution of image singularities in scale-space ... 242
14.3 Catastrophe theory basics ... 243
14.3.1 Functions ... 243
14.3.2 Characterization of points ... 243
14.3.3 Structural equivalence ... 244
14.3.4 Local characterization of functions ... 244
14.3.5 Thom's theorem ... 245
14.3.6 Generic property ... 246
14.3.7 Dimensionality ... 246
14.3.8 Illustration of the concepts ... 247
14.4 Catastrophe theory in scale-space ... 250
14.4.1 Generic events for differential operators ... 251
14.4.2 Generic events for other differential operators ... 254
14.4.3 Annihilations and creations ... 255
14.5 Summary of this chapter ... 256

15. Deep structure III. topological numbers 257
15.1 Topological numbers 257
15.1.1 Topological numbers in scale-space 258
15.1.2 Topological number for a signal 259
15.1.3 Topological number for an image 259
15.1.4 The winding number on 2D images 260
15.2 Topological numbers and catastrophes 263
15.3 The deep structure toolbox 265
15.3.1 Detection of singularities 265
15.3.2 Linking of singularities 265
15.3.3 Linking of contours 268
15.3.4 Detection of catastrophes 268
15.3.5 General discrete geometry approach 269
15.4 From deep structure to global structure 271
15.4.1 Image representations 271
15.4.2 Hierarchical pre-segmentation 272
15.4.3 Perceptual grouping 273
15.4.4 Matching and registration 274
15.4.5 Image databases 274
15.4.6 Image understanding 275
15.5 Summary of this chapter 275

16. Deblurring Gaussian blur 277
16.1 Deblurring 277
16.2 Deblurring with a scale-space approach 277
16.3 Less accurate representation, noise and holes 281
16.4 Summary of this chapter 284

17. Multi-scale optic flow 285
17.1 Introduction 285
17.2 Motion detection with pairs of receptive fields 286
17.3 Image deformation by a discrete vectorfield 289
17.4 The optic flow constraint equation 290
17.5 Scalar and density images 292
17.6 Derivation of multi-scale optic flow constraint equation 292
17.6.1 Scalar images, normal flow. 296
17.6.2 Density images, normal flow. 301
17.7 Testing the optic flow constraint equations 303
17.8 Cleaning up the vector field 305
17.9 Scale selection 307
17.10 Discussion 309
17.11 Summary of this chapter 310

18. Color differential structure 311
18.1 Introduction 311
18.2 Color image formation and color invariants 311
18.3 Koenderink's Gaussian derivative color model 314
18.4 Implementation 320
18.5 Combination with spatial constraints 325
18.6 Summary of this chapter 327

19. Steerable kernels 329
19.1 Introduction 329
19.2 Multi-scale orientation 330
19.3 Orientation analysis with Gaussian derivatives 331
19.4 Steering with self-similar functions 332
19.5 Steering with Cartesian partial derivatives 336
19.6 Detection of stellate tumors 338
19.7 Classical papers and student tasks 342
19.8 Summary of this chapter 343

20. Scale-time 345
20.1 Introduction 345
20.2 Analysis of prerecorded time-sequences 346
20.3 Causal time-scale is logarithmic 349
20.4 Other derivations of logarithmic scale-time 351
20.5 Real-time receptive fields 353
20.6 A scale-space model for time-causal receptive fields 354
20.7 Conclusion 359
20.8 Summary of this chapter 360

21. Geometry-driven diffusion 361
21.1 Adaptive Smoothing and Image Evolution 361
21.2 Nonlinear Diffusion Equations 362
21.3 The Perona & Malik Equation 364
21.4 Scale-space implementation of the P&M equation 366
21.5 The P&M equation is ill-posed. 370
21.6 Von Neumann stability of numerical PDE's 372
21.7 Stability of Gaussian linear diffusion 373
21.8 A practical example of numerical stability 376
21.9 Euclidean shortening flow 378
21.10 Grayscale invariance 379
21.11 Numerical examples shortening flow 379
21.12 Curve Evolution 382
21.13 Duality between PDE- and curve evolution 383
21.14 Mathematical Morphology 386
21.15 Mathematical morphology on grayvalued images 389
21.16 Mathematical morphology versus scale-space 390
21.17 Summary of this chapter 390

22. Epilog 393

A. Introduction to *Mathematica* 395
A.1 Quick overview of using Mathematica 395
A.2 Quick overview of the most useful commands 397
A.3 Pure functions 401
A.4 Pattern matching 401
A.5 Some special plot forms 404
A.6 A faster way to read binary 3D data 405
A.7 What often goes wrong 407
A.8 Suggested reading 410
A9. Web resources 412

B. The concept of convolution 413
B.1 Convolution 413
B.2 Convolution is a product in the Fourier domain 416

C. Installing the book and packages 419
C.1 Content 419
C.2 Installation for all systems 420
C.3 Viewing the book in the Help Browser 420
C.4 Sources of additional applications 421

D. First Start with *Mathematica*: Tips & Tricks 423
1. Evaluation 423
2. Images 423
3. Programming 424
4. 3D 424

References 425

Index 455

Front-End Vision and Multi-Scale Image Analysis

Bart M. ter Haar Romeny, PhD

Department of Biomedical Engineering
Eindhoven University of Technology
The Netherlands

Tell me, and I will forget. Show me, and I will remember.
Involve me, and I will understand.
Old Chinese proverb

The purpose of this book

Scale is not an important parameter in computer vision research. It is an *essential* parameter. It is an immediate consequence of the process of observation, of *measurements*. This book is about scale, and its fundamental notion in computer vision, as well as human vision.

Scale-space theory is the theory of apertures, through which we and machines observe the world. The apertures come in an astounding variety. They can be exploited to model the first stages of human vision, and they appear in all aspects of computer vision, such as the extraction of features, the measurement of optic flow and stereo disparity, to do orientation analysis, segmentation, image enhancement etc. They have an essential role in the fundamental processes of differentiation and regularization.

Scale-space theory is named after the space that is formed by looking at an image at many different scales simultaneously.

When stacked, we get one dimension extra, i.e. the scale dimension. The scale-space is the space of the spatial and scale dimensions, see figure 1.

This book is a tutorial course. The level of the book is undergraduate and first level graduate. Its main purpose is to be used as a coursebook in computer vision and front-end vision entry courses. It may also be useful as an entry point for research in biological vision.

Although there are excellent texts appearing on the notion of scale space, most of them are not easy reading for people just entering this field or lacking a solid mathematical background. This book is intended partly to fill this gap, to act as an entry point for the

growing literature on scale-space theory. Throughout the book we will work steadily through the necessary mathematics.

The book discusses the many classical papers published over the last two decades, when scale-space theory became mature. The different approaches and findings are put into context. First, linear scale-space theory is derived from first principles, giving it a sound mathematical basis.

The notion that a multi-scale approach is a natural consequence of the process of observation is interwoven in every chapter of this book. E.g. Horn and Schunck's famous optic flow equation gets a new meaning when we 'look at the data'. The concept of a point and local point operators like the derivative operator diffuse into versions with a Gaussian extent, making the process of differentiation well posed. It immediately makes large and mature fields like differential geometry, invariant theory, tensor analysis and singularity theory available for analysis on discrete data, such as images.

We develop ready-to-use applications of differential invariants of second, third and fourth order. The relation between accuracy, differential order and scale of the operator is developed, and an example of very high order derivatives is worked out in the analytical deblurring of Gaussian blur.

Practical examples are also developed in the chapters on multi-scale optic flow and multi-scale differential structure of color images. Again, the physics of the observation process forces the analytical solution to be multi-scale. Several examples of ways to come to proper scale-selection are treated underway.

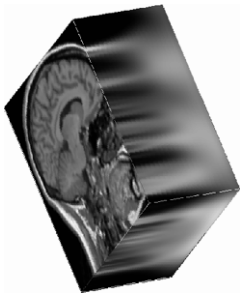

Figure 1. A scale-space of a sagittal MR image. The image is blurred with a Gaussian kernel with variable width σ, which is the third dimension in this image.

Scale-space theory is biologically motivated computer vision

We consider it very important to have many cross-links between multi-scale computer vision theory, and findings in the human (mammalian) visual system. We hope the reader will appreciate the mutual cross-fertilization between these fields. For that reason we elaborate the current state of the art in neurophysiological and psychophysical findings of the first stages of the visual system.

The chapters on time-scale and multi-scale orientation analysis are directly inspired by findings from biological vision. The grouping of local properties into meaningful larger subgroups (perceptual grouping) is treated both on the level of establishing neighborhood relationships through all measured properties of the points, and through

the study of the deep structure of images, where topology comes in as a mathematical toolkit. The natural hierarchical ordering is exploited in a practical application, where we discuss multi-scale watershed segmentation.

This book is meant to be a practical and interactive book. It is written as a series of notebooks in *Mathematica* 4, a modern computer algebra language/system.

For every technique discussed the complete code is presented. The reader can run and adapt all experiments himself, and learn by example and prototyping. The most effective way to master the content is to go through the notebooks on a computer running *Mathematica* and play with variations.

This book is a tribute to Jan Koenderink, professor at Utrecht University in the Netherlands, chairman of the Physics Department 'Physics of Man'. He can be considered the 'godfather' of modern scale-space theory. A brilliant physicist, combining broad knowledge on human visual perception with deep insight in the mathematics and physics of the problems.

Figure 2. Prof. dr. hon.dr. Jan Koenderink.

This book is just a humble introduction to his monumental oeuvre and the offspin of it. Many papers he wrote together with his wife, Ans van Doorn. They published on virtually every aspect of front-end vision *and* computer vision with a strong perceptually based inspiration, and the physical modeling of it.

This book is written for both the computer vision scientist with an interest in multi-scale approaches to image processing, and for the neuroscientist with an appeal for mathematical modeling of the early stages of the visual system. One of the purposes of this book is to bridge the gap between both worlds. To accommodate a wide readership, both from physics and biology, sometimes mathematical rigor is lacking (but can be found in the indicated references) in favor of clarity of the exposition.

Figure 3. Attendants of the first international Scale-Space conference, Summer 1997 in Utrecht, the Netherlands, chaired by the author (standing fourth from right).

Figure 4. Attendants of the second international Scale-Space conference, Summer 1999 in Corfu, Greece, chaired by Mads Nielsen, PhD (IT-Univ. Copenhagen, foreground fifth from right). See for the conference series: www.scalespace.org.

This book has been written in *Mathematica*

This book is written as a series of *Mathematica* notebooks. *Mathematica* is a high level interactive mathematical programming language, developed and marketed by Stephen Wolfram (www.wolfram.com). Notebooks are interactive scientific documents, containing both the text as the code.

Mathematica consists of a two separate programs, the kernel (the computing engine) and a front-end which handles all information for and from the user. The structure of 'cells' in the front-end enables the efficient mix of explaining text, computer code and graphics in an intuitive way. The reasons to write this book in *Mathematica* are plentiful:

- We can now do *mathematical* prototyping with computer vision principles/techniques on images. The integration of both symbolic and fast numerical capabilities, and the powerful pattern matching techniques make up for a new and efficient approach to apply and teach computer vision, more suitable for human mathematical reasoning. For, computer vision *is* mathematics on images. It is now easy to do *rapid prototyping*.

- Reading a scientific method now has something extra: the code of every method discussed is available, ready for testing, with modifications and applications to the reader's images. The

gap between the appreciation of a method in a theoretical paper and one's own working software that applies the method is now closing. David Donoho writes about his WaveLab package: "An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures. (http://www-stat.stanford.edu/~wavelab/).

- *Mathematica* stays close to the traditional notation. The mathematical notation of e.g. symbols, operators and rules are virtually identical to traditional math

The notebooks are WYSIWYG. Notebooks can easily be saved as LaTeX or HTML/MathML documents. Notebooks are portable ASCII documents, they appear virtually identical on a wide range of computer platforms.

- All programs become compact. In this book no example exceeds 20 lines of code. There are no **`for`**, **`while`** or **`do`** loops. Most commands are **`Listable`**, i.e. operate on any member of the operand list. The language contains hardly abbreviations, and is so intuitive that learning the language may be mastered during reading the book. In the appendix a list of tutorial books on *Mathematica* is given, and a summary of the command structure of the most popular commands used in this book.

- Wolfram Research Inc. indicates that over 2 million licenses are sold. It may serve as a (WWW-based) starting set in exchangeable *Mathematica* computer vision routines.

- *Mathematica* is complete. Over 2500 highly optimized mathematical routines are on board, which relieves the computer vision programmer from searching for routines in Numerical Recipes, IMSL etc. It has graphical capabilities for 1D to 4D (animations). It is now integrated with Java (**`JLink`**), which is available anywhere and ideally suited for further development of the GUI and real-time manipulation with the data. *Mathematica* code can be compiled for further speed increase.

The parallel version of *Mathematica* now enables the easy distribution over a range of kernels on different computers on a network.

- Last but not least: the present version of *Mathematica* is fast. From release 4 it has reached a point where, from being an ideal rapid prototyping tool, it is now turning into an all-round prototyping *and* application tool. The run-time of most experiments described in this book is within fractions of seconds to tens of seconds on a typical 1.7 GHz 256 MB Pentium IV system under Windows.

It is platform independent, and is available on any type of computer.

Acknowledgements

The contents of this book is based on work of: Jan J. Koenderink, Luc M. J. Florack, Tony Lindeberg, Wiro J. Niessen, Alfons H. Salden, Mads Nielsen, Jon Sporring, Ole F. Olsen, Jan-Mark Geuzebroek, Erik Dam, Peter Johansen, Avan Suinesiaputra, Hans J. Blom, Bram van Ginneken, Stiliyan Kalitzin, Joes Staal, Robert Maas, Max Viergever, Antonio López Peña, Nico Karssemeijer, Stephen Pizer, Dave Eberley, Jean-Philippe Thirion, Stephen Zucker, Wim van de Grind, Ans Koenderink-van Doorn, Alex Frangi, Arjan Kuijper, Remco Duits, Bram Platel, Frans Kanters, Taizo Iijima, Ted Adelson, Pietro Perona, Jitendra Malik, David Mumford, Luis Alvarez, Markus van Almsick, William Freeman, Izumi Ohzawa, Ralph Freeman, Russell DeValois, David Hubel, Torsten Wiesel, Semir Zeki, Erik Kandel, Amiram Grinvald, Brian Wandell, Robert Rodieck and many more, as well asmany of the 3D Computer Vision / Image Sciences Institute master's students.

The careful proofreading and remarks by Michel Bister were very helpful. Markus van Almsick gave many *Mathematica* hints, insights and valuable help with editing. I thank my wife Hetty, for the patience and loving support during the endeavor to write this book.

This book originated from coursenotes of the annual graduate course on "Multiscale Image Analysis and Font-End Vision" at the Image Sciences Institute of Utrecht University 1996-2000 and at Faculty of Biomedical Engineering at Eindhoven University of Technology in the Netherlands since 2001. This course has also been given at a summerschool at CIMAT in Guanajuato, Mexico, at IMPA in Rio de Janeiro, Brazil, at the Multimedia University in Cyberjaya, Malaysia and as a tutorial at several international conferences (VBC 98, CVPR 1999, MICCAI 2001). The author was kindly hosted for two months at the new IT University at Copenhagen, Denmark in the summer of 2000 to get a lot of the writing done. The highly competent and kind atmosphere of the collaborating labs there considerably contributed to the pleasure it was to write and to program this interactive book. I learned a lot by *doing* multi-scale computer vision, and I hope the reader will do the same.

Eindhoven - Utrecht - Copenhagen, February 2003

Email: B.M.terHaarRomeny@tue.nl