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.
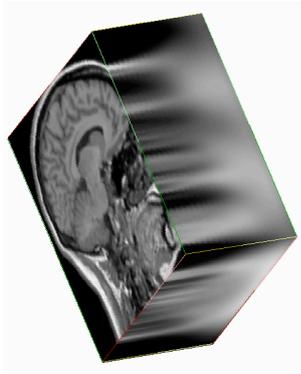

Figure 1. A scale-space of a sagittal MR image. The image is blurred with a Gaussian kernel with variable width σ, which is the third dimension in this image.

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 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