Nonlinear Scale Space Analysis in Image Processing

by

Ilya Pollak

B.S. and M.Eng, Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1995

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

August, 1999

Signature of Author: ______________________________
Dept. of Electrical Engineering and Computer Science
July 20, 2000

Certified by: ______________________________
Alan S. Willsky
Professor of EECS
Thesis Supervisor

Certified by: ______________________________
Hamid Krim
Professor of ECE, North Carolina State University,
Thesis Supervisor

Accepted by: ______________________________
Arthur C. Smith
Professor of EECS
Chair, Committee for Graduate Students

Nonlinear Scale Space Analysis in Image Processing

by Ilya Pollak

ipollak@alum.mit.edu

Submitted to the Department of Electrical Engineering
and Computer Science on July 20, 2000
in Partial Fulfillment of the Requirements for the Degree
of Doctor of Philosophy in Electrical Engineering and Computer Science

Abstract

The objective of this work is to develop and analyze robust and fast image segmentation algorithms. They must be robust to pervasive, large-amplitude noise, which cannot be well characterized in terms of probabilistic distributions. This is because the applications of interest include synthetic aperture radar (SAR) segmentation in which speckle noise is a well-known problem that has defeated many algorithms. The methods must also be robust to blur, because many imaging techniques result in smoothed images. For example, SAR image formation has a natural blur associated with it, due to the finite aperture used in forming the image. We introduce a family of first-order multi-dimensional ordinary differential equations with discontinuous right-hand sides and demonstrate their applicability to segmenting both scalar-valued and vector-valued images, as well as images taking values on a circle. An equation belonging to this family is an inverse diffusion everywhere except at local extrema, where some stabilization is introduced. For this reason, we call these equations "stabilized inverse diffusion equations" ("SIDEs"). Existence and uniqueness of solutions, as well as stability, are proven for SIDEs. A SIDE in one spatial dimension may be interpreted as a limiting case of a semi-discretized Perona-Malik equation [49,50], which, in turn, was proposed in order to overcome certain shortcomings of Gaussian scale spaces [72]. These existing techniques are reviewed in a background chapter. SIDEs are then described and experimentally shown to suppress noise while sharpening edges present in the input image. Their application to the detection of abrupt changes in 1-D signals is also demonstrated. It is shown that a version of the SIDEs optimally solves certain change detection problems. Its relations to the Mumford-Shah functional [44] and to linear programming are discussed. Theoretical performance analysis is carried out, and a fast implementation of the algorithm is described.

Thesis Supervisor: Alan S. Willsky
Title: Professor of Electrical Engineering and Computer Science

Thesis Supervisor: Hamid Krim
Title: Assistant Professor of Electrical and Computer Engineering, North Carolina State University

This thesis is dedicated to the memory of my mother, Tamara B. Skomorovskaya (1935-1997).

Acknowledgments

After spending eight years at MIT, all in the same group and in the same office, I have the dubious distinction of having been in the Stochastic Systems Group (SSG) longer than any current (and, I suspect, past) member except our leader Alan Willsky[1]. I have seen many people go through the group: dozens of graduate students, six secretaries, numerous post-docs, research scientists, and visitors—sufficient to staff an Electrical Engineering department of a medium-size college, and still have enough administrative assistants left over to sustain "Murphy Brown" for at least a couple of seasons.

I have seen two people engaged in a rubber-band fight; they are now a quantitative analyst at a large financial company and an assistant professor at a top-ranked university. I remember an attempted coup d'etat in the group when two other people (also now assistant professors), using the fact that Alan was traveling, decided to take a shot at the sacred tradition of naming the computer equipment after fine wines, by calling our new printer "Gallo". Luckily, Alan's speedy and decisive actions upon his return restored the order: the printer was re-named "Chenas". Yet another member of the group (also an assistant professor now) once surprised me with his masterful use of a few Russian choice words; it turned out that he learned those especially for me, using his five-language dictionary of obscenities. Even more entertaining is the ongoing feud between two current group members: actions taken have included sealing the opponent inside a room, putting something interesting onto his desk (e.g., *another desk*), inventing a nickname for him and cleverly working it—in a disguised form—into a seminar on one's research, etc. In short, I have seen a lot. Therefore, despite the fact that my acknowledgment section is unusually large, I cannot possibly name all the people who have contributed to making my experience at MIT an enjoyable and entertaining one.

First and foremost, I would like to thank my thesis supervisor Alan Willsky, who was the ideal advisor: not only a great scientist, but also a very enthusiastic and supportive mentor, providing broad guidance and posing interesting and important problems, yet allowing me virtually unrestricted freedom. His encouragement when research progressed well and his patience when it did not, were extremely helpful. Equally important was the fact that Alan was a fantastic collaborator: he was always up to date on every minor technical detail (in spite of his large number of students), which is why I never felt isolated in my research. His many suggestions on how to deal with difficult technical issues, as well as his vast knowledge of the literature, were invaluable. The speed with which he read drafts of papers and thesis chapters was

[1] Despite this fact, I probably hold another dubious record—namely, having the shortest dissertation. I think there have been people in the group whose theses had sentences longer than this whole document.

sometimes frightening; by following his comments and suggestions, one would typically improve the quality of a paper by orders of magnitude. His help and advice—and, I am sure, his reputation—were also instrumental to the success of my job-hunting campaign.

My co-supervisor Hamid Krim was equally supportive; I am very thankful to him for having an "open door" and for the endless hours we spent discussing research, politics, soccer, religion, and life in general. I am grateful to Hamid for inviting me to spend several days in North Carolina, and for his immense help during my job search.

The suggestions and advice of the two remaining committee members—Olivier Faugeras and Sanjoy Mitter—also greatly contributed both to my thesis and to finding a good academic position. I am indebted to Olivier for inviting me to visit his group at INRIA. I thoroughly enjoyed the month I spent in Antibes and Sophia-Antipolis, for which I would like to thank the whole ROBOTVIS group: Didier Bondyfalat, Sylvain Bougnoux, Rachid (Nour-Eddine) Deriche, Cyrille Gauclin, José Gomez, Pierre Kornprobst, Marie-Cécile Lafont, Diane Lingrand, Théo Papadopoulo, Nikos Paragios, Luc Robert, Robert Stahr, Thierry Viéville, and Imad Zoghlami.

My special thanks go to Stéphane Mallat for inviting me to give a seminar at École Polytechnique, and for his help in my job search. He has greatly influenced my professional development, both through a class on wavelets which he taught very engagingly and enthusiastically at MIT in 1994, and through his research, some of which was the basis for my Master's thesis.

I would like to thank Stuart Geman for suggesting the dynamic programming solution of Chapter 4, as well as for a number of other interesting and useful insights.

I am thankful for the stimulating discussions with Michele Basseville, Charlie Bouman, Yoram Bresler, Patrick Combettes, David Donoho, Al Hero, Mohamed Khodja, Jitendra Malik, Jean-Michel Morel, David Mumford, Igor Nikiforov, Pietro Perona, Jean-Christophe Pesquet, Guillermo Sapiro, Eero Simoncelli, Gil Strang, Allen Tannenbaum, Vadim Utkin, and Song-Chun Zhu, all of whom have contributed to improving the quality of my thesis and to broadening my research horizon.

In addition to the people I acknowledged above, Eric Miller, Peyman Milanfar, and Andy Singer generously shared with me many intricacies of the academic job search.

I thank the National Science Foundation and Alan Willsky for providing the financial support, through a Fellowship and a Research Assistantship, respectively. I thank Clem Karl, Paul Fieguth and Andrew Kim for all their help with computers; Ben Halpern for teaching me repeatedly and patiently how to use the fax machine, and for helping me to tame our color printer; Mike Daniel and Austin Frakt for sharing the LaTeXstyle file which was used to format this thesis; John Fisher for organizing a very informative reading group on learning theory; Asuman Koksal for making our office a little cozier; Andy Tsai for being a great officemate; Dewey Tucker for his help in generating my PowerPoint job talk; Jun Zhang and Tony Yezzi both for technical interactions and for in-depth discussions of the opera; Mike Schneider for sharing his encyclopedic knowledge (and peculiar tastes) of the art in general, and for educating me on a number of other topics, ranging from estimation and linear programming to botany and ethnography, as well as for volunteering to be my personal guide and chauffeur in Phoenix, AZ; Ron Dror for our numerous discussions of biological and computer vision; and Martin Wainwright for conveying some interesting facts about probability and natural image statistics, and for having even more peculiar tastes in arts than Mike. I also enjoyed the sometimes heated arguments with Khalid Daoudi over soccer, but I still think he was wrong about the penalty in that Brazil vs. Norway game! The unfortunate lack of a resident ballet expert in SSG was rectified by my summer employment at Alphatech, Inc. (in Burlington, MA), where I had a chance to exchange opinions and ballet tapes with Bob Tenney.

I enjoyed interacting with all other SSG group members and visitors, past and present. The ones whose names are not mentioned above are Lori Belcastro, François Bergeaud, Mickey Bhatia, Terrence Ho, Taylore Kelly, Junmo Kim, Alex Ihler, Bill Irving, Seema Jaggi, Rachel Learned, Cedric Logan, Mark Luettgen, Mike Perrott, Paula Place, John Richards, and Gerald Spica.

I thank my brother Nick for his leadership by example in showing me that an excellent PhD thesis *can* be finished even under very unfavorable circumstances, such as having two small children and a full-time job.

I am most grateful to my parents who gladly endured many sacrifices that accompanied our emigration, so that my brother and I could have a better life in the United States. Without their strength and courage I would not have had a chance to study in the best university in the world; without their tremendous support, I would not have been able to complete my studies. As I know how important my getting this degree was to my mother, I dedicate this thesis to her memory.

Contents

List of Figures

Chapter 1

Introduction

IN this chapter, we introduce the problem of segmentation and change detection addressed in this thesis, and describe the organization of the thesis.

1.1 Problem Description and Motivation

To segment a 1-D or 2-D signal means, roughly speaking, to partition the domain of its definition into several regions in such a way that the signal is homogeneous within each region and changes abruptly between regions. The exact meaning of the word "homogeneous" depends on the application: most often it means smoothly varying or constant intensity (this case is addressed in Chapters 3 and 4), or uniform texture (Chapters 4 and 5).

The objective of this thesis is to develop and analyze robust and fast image segmentation algorithms. They must be robust to pervasive, large-amplitude noise, which cannot be well characterized in terms of probabilistic distributions. This is because the applications of interest are exemplified by synthetic aperture radar (SAR) segmentation in which speckle noise is a well-known problem that has defeated many algorithms. (A prototypical SAR log-magnitude image of two textural regions–forest and grass–is shown in Figure 1.1.) The methods must also be robust to blur, because many imaging techniques result in smoothed images. For example, SAR image formation has a natural blur associated with it, due to the finite aperture used in forming the image.

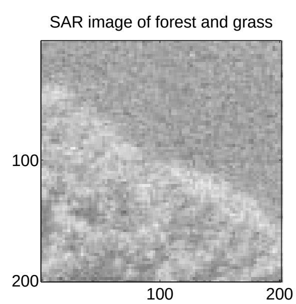

Figure 1.1. SAR image of trees and grass.

Image segmentation is closely related to restoration, that is, the problem of estimating an image based on its degraded observation. Indeed, the solution to one of these problems makes the other simpler: estimation is easier if the boundaries of homogeneous image regions are known, and vice versa, segmentation is easier once a good estimate of the image has been computed. It is therefore natural that many segmentation algorithms are related to restoration techniques, and in fact some methods combine the two, producing estimates of both the edge locations and image intensity [36, 44], as we will see in Chapter 2.

In describing any restoration or segmentation technique, the notion of *scale* is very important. Any such technique incorporates a scale parameter—either directly in the computation procedure, or implicitly as a part of the image model—which controls the smoothness of the estimate and/or sizes of the segmented regions. The precise definitions of scale, in several contexts, are given in Chapters 2 and 3; intuitively, changing the parameter from zero to infinity will produce a so-called *scale space*, i.e. a set of increasingly coarse versions of the input image. There are two approaches to generating a scale space: one starts with a probabilistic model, the other starts with a set of "common-sense" heuristics. The difference between the two is conceptual: they both may lead to the same algorithm [39, 44], producing the same scale space. In the former case, one would build a probabilistic model of images of interest [6, 32, 39] and proceed to derive an algorithm for computing the solution which is, in some probabilistic sense, optimal. For example, one could model images as piecewise constant functions with additive white Gaussian noise, and the edges (i.e. the boundaries separating the constant pieces of the function) as continuous curves whose total length is a random variable with a known distribution. Assigning larger probabilities to the occurrence of edges would correspond to larger scales in such a model, which will be illustrated in Chapter 4. Given a realization of this random field, the objective could be to compute the maximum likelihood estimates [66] of the edge locations. The main shortcoming of this approach is that a good model is unavailable in many applications, and that usually any realistic model yields a complicated objective functional to be optimized. Obtaining the optimal solution is therefore not computationally feasible, and one typically settles for a local maximum [6, 63]. An alternative to such probabilistic methods of generating scale spaces is to devise an algorithm using a heuristic description of images of interest. Stabilized Inverse Diffusion Equations (SIDEs), which are the main topic of this thesis, belong to this latter category.

SIDEs are motivated by the great recent interest in using evolutions specified by partial differential equations (PDE's) as image processing procedures for tasks such as restoration and segmentation, among others [1, 12, 35, 46, 49, 50, 55–57, 72]. The basic paradigm behind SIDEs, borrowed from [1, 35, 49, 72], is to treat the input image as the initial data for a diffusion-like differential equation. The unknown in this equation is usually a function of three variables: two spatial variables (one for each image dimension) and the scale—which is also called *time* because of the similarity of such equations to evolution equations encountered in physics. In fact, one of the starting points of this line of investigation was the observation [72] that smoothing an image with Gaussians of varying width is equivalent to solving the linear heat diffusion equation with the image as the initial condition. Specifically, the solution to the heat equation at time t is the convolution of its initial condition with a Gaussian of variance $2t$. Gaussian filtering has been used both to remove noise and as a pre-processor for edge detection procedures [9]. It has serious drawbacks, however: it displaces and removes important image features, such as edges, corners, and T-junctions. (An example of this behavior will be given in Chapter 2 (Figure 2.1).) The interpretation of Gaussian filtering as a linear diffusion led to the design of other, nonlinear, evolution equations, which better preserve these features [1, 46, 49, 55–57]. For example, one motivation for the work of Perona and Malik in [49, 50] is achieving both noise removal and edge enhancement through the use of an equation which in essence acts as an unstable inverse diffusion near edges and as a stable linear-heat-equation-like diffusion in homogeneous regions without edges.

The point of departure for the development of our SIDEs are the anisotropic diffusions introduced by Perona and Malik and described in the next chapter. In a sense that we will make both precise and conceptually clear, the evolutions that we introduce may be viewed as a conceptually limiting case of the Perona-Malik diffusions. These evolutions have discontinuous right-hand sides and act as inverse diffusions "almost everywhere" with stabilization resulting from the presence of the discontinuities in the right-hand side. As we will see, the scale space of such an equation is a family of segmentations of the original image, with larger values of the scale parameter t corresponding to segmentations at coarser resolutions.

Since "segmentation" may have different meanings in different contexts, we close this section by further clarifying which segmentation problems are addressed in this thesis and which are not. This also gives us an opportunity to mention four important application areas.

Automatic Target Recognition. Segmentation problems arise in many aspects of image processing for automatic target recognition. Some examples are the identification of tree-grass boundaries in synthetic aperture radar images [18] (an example given in the beginning of this section will be further treated in Chapter 3), the localization of hot targets in infrared radar [15], and the separation of foreground and background in laser radar (see [23] and references therein). Due to high rates of data acquisition, it is desirable for some of these problems that the algorithm be fast, in addition to being robust. As shown in Chapters 3 and 4, the algorithms presented in this thesis outperform existing segmentation algorithms in speed and/or robustness.

Segmentation of medical images. Given an image or a set of images resulting from a medical imaging procedure—such as ultrasound [13], magnetic resonance imaging [2, 64], tomography [6, 27], dermatoscopy [19, 60]—it is necessary to extract certain objects of interest, e. g. an internal organ, a tumor, or a boundary between the gray matter and white matter. The main challenges of the segmentation problem depend on the object and on the imaging modality. For example, ultrasound imaging introduces both significant blurring and speckle noise [8, 22, 70], and so the corresponding segmentation algorithms must be robust to such degradations. Robustness of the algorithm introduced in this thesis is experimentally demonstrated in Chapters 3 and 4; Chapter 4 also contains its theoretical analysis.

Detection of abrupt changes in 1-D signals [3]. Application areas include analysis of electrocardiograms and seismic signals, vibration monitoring in mechanical structures, and quality control. Several synthetic 1-D examples are considered in Chapters 3 and 4.

Computer vision. In the computer vision literature, the term "segmentation" often refers to finding the contours of objects in natural images—i.e., photographic pictures of scenes which we are likely to see in our everyday life [16]. This is an important problem in low-level vision, because it has been universally accepted since [71] that segmenting the perceived scene plays an important role in human vision. However, noise and other types of degradation are usually not as significant here as in the medical and radar images; the main challenge is the variety of objects, shapes, and textures in a typical picture. This problem is therefore not directly addressed in the present thesis; applying the algorithms developed here to this problem is a topic for future research.

■ 1.2 Summary of Contributions and Thesis Organization

Chapter 2 is a review of several methods of image restoration and segmentation, each of which has important connections to the main contribution of this thesis, namely Stabilized Inverse Diffusion Equations (SIDEs). We mentioned in the previous section that SIDE evolutions produce nonlinear scale spaces, similarly to the equations in [1,49]. What sets SIDEs apart from other frameworks is the form of their right-hand sides which results in improved performance. This is shown in Chapter 3, both through experiments confirming the SIDEs' speed and robustness, and through proving a number of their useful properties, such as stability with respect to small changes in the input data. It is also explained how the SIDEs are related to several existing methods, such as Perona-Malik diffusions [49], shock filters [46], total variation minimization [6, 46, 58], the Geman-Reynolds functional [21], and region merging [43].

The main drawback of the methods which are based on heuristic descriptions of images of interest (rather than on probabilistic models) is that it is usually unclear how they perform on noisy images; experimentation is the most popular way of demonstrating robustness [5, 49, 56, 63]. This has been the case with diffusion-based methods: they are complicated nonlinear procedures, and therefore theoretical analysis of what happens when a random field is subjected to such a procedure is usually impossible. Unlike their predecessors, SIDEs turn out to be amenable to some probabilistic analysis, which is carried out in Chapter 4. It is shown that a specific SIDE finds in $N \log N$ time the maximum likelihood solutions to certain binary classification problems. The likelihood function for one of these problems is essentially a 1-D version of the Mumford-Shah functional [44]. Thus, an interesting link is established between diffusion equations, Mumford and Shah's variational formulation, and probabilistic models. The robustness of the SIDE is explained by showing that, in a certain special case, it is optimal with respect to an H_∞-like criterion—which, roughly speaking, means that the SIDE achieves the minimum worst-case error. The performance is also analyzed by computing bounds on the probabilities of errors in edge location estimates. To summarize, the main contribution of Chapter 4 is establishing a connection between diffusion-based methods and maximum likelihood edge detection, as well as extensive performance analysis.

Chapter 5 extends SIDEs to vector-valued images and images taking values on a circle. We argue that most of the properties derived in Chapter 3 carry over. These results are applicable to color segmentation, where the image value at every pixel is a three-vector of red, green, and blue values. We also apply our algorithm to texture segmentation, in which the vector image to be processed is formed by extracting features from the raw texture image, as well as to segmenting orientation images.

Possible directions of future research are proposed in Chapter 6.