Progress in Nonlinear Differential Equations and Their Applications
Volume 14

Jean-Michel Morel
Sergio Solimini

Variational Methods in Image Segmentation

with seven image processing experiments

Birkhäuser
Boston • Basel • Berlin

Jean Michel Morel
CEREMADE
Université Paris-Dauphine
Place de Lattre de Tassigny
75775 Paris Cedex 16, France

Sergio Solimini
Università degli Studi di Lecce
Srada provinciale Lecce-Arnesano
73100 Lecce Italy

Library of Congress Cataloging-in-Publication Data

Morel, Jean-Michel, 1953-
Variational methods for image segmentation / Jean-Michel Morel, Sergio Solimini.
p. cm. -- (progress in nonlinear differential equations and their applications ; v. 14)
Includes bibliographical references and index.
ISBN 0-8176-3720-6 (acid-free)
1. Image processing--Digital techniques--Mathematics.
2. Geometric measure theory. I. Solimini, Sergio, 1956-
II. Title. III. Series.
TA1637.M67 1994 94-36639
621.3'67--dc20 CIP

Printed on acid-free paper

Birkhäuser ®

Softcover reprint of the hardcover 1st edition 1995

ISBN-13: 978-1-4684-0569-9 e-ISBN-13: 978-1-4684-0567-5
DOI: 10.1007/978-1-4684-0567-5

Typeset and reformatted from the author's disk in $\mathcal{AMS}$-TEX

9 8 7 6 5 4 3 2 1

Contents

Preface

This book contains both a synthesis and mathematical analysis of a wide set of algorithms and theories whose aim is the automatic segmentation of digital images as well as the understanding of visual perception. A common formalism for these theories and algorithms is obtained in a variational form. Thank to this formalization, mathematical questions about the soundness of algorithms can be raised and answered.

Perception theory has to deal with the complex interaction between regions and "edges" (or boundaries) in an image: in the variational segmentation energies, "edge" terms compete with "region" terms in a way which is supposed to impose regularity on both regions and boundaries. This fact was an experimental guess in perception phenomenology and computer vision until it was proposed as a mathematical conjecture by Mumford and Shah.

The third part of the book presents a unified presentation of the evidences in favour of the conjecture. It is proved that the competition of one-dimensional and two-dimensional energy terms in a variational formulation cannot create fractal-like behaviour for the edges. The proof of regularity for the edges of a segmentation constantly involves concepts from geometric measure theory, which proves to be central in image processing theory. The second part of the book provides a fast and self-contained presentation of the classical theory of rectifiable sets (the "edges") and unrectifiable sets ("fractals"). This part contains a discussion of new uniform density properties of rectifiable sets, which prove extremely useful in image processing theory. Several image processing experiments and many figures illustrate algorithmic discussions and mathematical proofs.

This book will be accessible to graduate science students with some mathematical background. It will be of interest to mathematicians concerned with the interaction of analysis and geometry and to vision researchers.

Introduction

Natural, digital and perceptual images.

When one looks directly at scenes from the natural or the human world, or at any image (painting, photograph, drawing,...) representing such scenes, it is impossible to avoid seeing in them structures, which in many cases can be identified with real objects. These objects can be somehow concrete, as in photographs where we see trees, roads, windows, people, etc., or abstract perceptual structures, as the ones which appear in abstract paintings and can only be described in geometrical terms. However, we know that the "visual information" arriving at our retina, far from being structured, is purely local information for which a good model is given by digital images.

From the mathematical (and engineering) viewpoint, digital images simply are functions $g(x)$, where x is a point of the image domain Ω (the plane, a rectangle, ...) and $g(x)$ is a real number representing the "brightness" or "grey level" of the image at point x. This is the unstructured datum with which engineers have to deal in image analysis, robotics, etc.... And it also is somehow the basic datum which arrives at our retina. There is therefore not much in common between what we think we "immediately" see and what the physical information about light reflected by objects is. This striking difference between "images" in the engineering or the biological sense and "images" in the perceptual, artistic, semantic sense is well known, but the question is: How do we pass from the unstructured digital image to the structured perceptual one?

This question has been addressed in a scientific framework by psychologists of the Gestalt school in the twenties (Koehler [Koeh]) and then, with the first computers allowing to display synthetic images, by psychophysicists who did measurements of how much was perceived in the very first milliseconds after the arrival of an image to the retina (Julesz [Ju1, Ju2]). At the end of the sixties, the same problem was addressed in a very practical framework by engineers who wished to extract automatically useful information from digital pictures. The aim of this extraction was to build perception-driven robots and to better understand the human and animal vision. Then arose the *Segmentation Problem*, which is the object of this book.

The segmentation problem and its algorithms.

Segmenting a digital image means finding (by a numerical algorithm) its *homogeneous regions* and its *edges*, or *boundaries*. Of course, the homogeneous regions are supposed to correspond to meaningful parts of objects in the real world, and the edges to their apparent contours. Gestaltists and psychophysicists agree that such a segmentation process is at work at the very first stages of the visual perception process. In addition they proved that these first stages are highly independent of any learning or *a priori* knowledge of the world. In other words, the Vision research assumes that these first stages are accomplished by a *Geometry machine*, and it is also assumed that computer algorithms, working on digital images, should be able to do the same (Marr [Marr3], Koenderink [Koen2]). It is not the aim of this book to directly discuss this thesis, but rather to classify the proposed algorithms and their mathematical properties. More than a thousand algorithms have been proposed for segmenting images or detecting "edges". It is of course impossible (and unnecessary) to review them all, but the first part of this book (Chapters 1 to 4) is devoted to a classification of these algorithms and their translation from a discrete into a continuous framework (more adapted to the mathematical analysis). The result of this discussion was unexpected to the authors of this book because they became aware that under the very large diversity of numerical tools, **there essentially was only one segmentation (or "edge detection") model.** Indeed, as we hope these chapters will convince the reader, most segmentation algorithms try to minimize, by several very different procedures, one and the same *Segmentation energy.* This energy measures how smooth the regions are, how faithful the "analyzed image" to the original image and the obtained "edges" to the image discontinuities are.

One can write, and we indeed write in Chapter 4, the "most general" segmentation energy underlying all the analysed computational models. It has six or seven partly redundant terms, however, and therefore does not fit elegant mathematical analysis. If we keep the three more meaningful terms of the functional, we obtain the Mumford-Shah energy. Thus the Mumford-Shah variational model, although initially proposed as one model more, happens to somehow be **the general model** of image segmentation, and all the other ones are variants, or algorithms tending to minimize these variants. The Mumford-Shah model defines the segmentation problem as a joint smoothing/edge detection problem: given an image $g(x)$, one seeks simultaneously a "piecewise smoothed image" $u(x)$ with a set K of abrupt discontinuities, the "edges" of g. Then the "best" segmentation of a given image is obtained by minimizing the functional

$$E(u, K) = \int_{\Omega \setminus K} (|\nabla u(x)|^2 + (u - g)^2)dx + \text{length}(K).$$

The first term imposes that u is smooth outside the edges, the second that the piecewise smooth image $u(x)$ indeed approximates $g(x)$, and the third that the discontinuity set K has minimal length (and therefore in particular is as smooth as possible). The model is minimal in the sense that removing one of the above three terms would yield a trivial solution. Needless to say, such a simple functional cannot give a good account of the geometric intricacy of most natural images, or of our perception of them. What is expected from algorithms minimizing such a functional is a sketchy, cartoon-like version of the image, and these algorithms will give perceptually good results when the processed images somehow match this *a priori model*: contrasted images with objects presenting piecewise smooth surfaces. We shall see some experimental examples in Chapters 2 to 5.

The Mumford-Shah conjecture.

Two main problems have to be solved once we have fixed a single universal segmentation model. The first one is practical: How can we define algorithms minimizing the Mumford-Shah energy on computers? Now, as we pointed out, the algorithms have in many cases **preexisted** the theory and many of them, though perfectible, yield perceptually reasonable results, as we shall see in Chapters 2 to 5. The second question is simply a mathematical one: Is the model consistent? That is, do segmentations exist which indeed minimize the Mumford-Shah energy, are they unique and are the boundaries thus obtained smooth?

Mumford and Shah [MumS1] conjectured the existence of minimal segmentations made of a finite set of C^1 curves. So far, this conjecture has not been fully proved and only partial but meaningful enough results are at hand. The problem has proved difficult for the present mathematical technique because of the subtle interaction of the two-dimensional term (in u) and the one-dimensional term length(K). The same difficulty arises when one wishes to define a computer program minimizing the energy and we hope that the mathematical analysis will somehow clarify the numerical debate.

Edge sets and rectifiable sets.

As a matter of fact, the first mathematical task is to correctly define the functional $E(u, K)$. Indeed, we cannot *a priori* impose that an edge set K minimizing E is made of a finite set of curves. This is precisely what has to be proved, and if we imposed this condition to all "candidates" to be minimizers, we would have, for the time being, no existence theorem at all.

This kind of situation is classical in mathematical analysis and is dealt with by enlarging the Search Space, that is, in our case, by looking for a solution in a wider class of sets with finite length than just finite sets of curves. This is done by defining the "length" of K as its one-dimensional Hausdorff measure, which is the most natural way of extending the concept of length to non-smooth sets (Carathéodory, 1915 [Cara], Hausdorff, 1919 [Haus]). So we shall have to work in the main part of this book with "1-sets", i.e., sets with finite, positive, one-dimensional Hausdorff measure. The theory of these sets was developed by Besicovitch between 1928 and 1944 [Besi1, 2, 3], and was completed by Federer (1947, [Fed1]), Marstrand (1961 [Marst]), Mattila (1975 [Matt]) and Preiss (1987 [Prei]). As the dates indicate, it has not been a straightforward theory.

Now, it is a extremely suggestive theory for image processing. Besicovitch conjectured a remarkable structural classification of sets with finite m-dimensional Hausdorff measure. This conjecture was proved by himself and the above-mentioned authors. We shall now explain the results of the theory in the case $m = 1$, which is particularly significant for image processing, but they are identical in any dimension.

Besicovitch calls *rectifiable* any 1-set which is essentially contained in a countable family of curves with finite length. Clearly, the "edge sets" sought for in image processing belong to this class. A 1-set is called *fully unrectifiable* if it has no rectifiable part. Of course, any 1-set can be divided into its rectifiable and its unrectifiable part, but the question arises: How can we separate them? Here comes the remarkable Besicovitch discovery: Besicovitch conjectured and thereafter proved that there is a simple *local density criterion* to decide whether a point x of K belongs to the rectifiable part of K or not. This criterion is, from the computational viewpoint, clear cut: If x is in the rectifiable part, then the length of $K \cap B(x,r)$ is equivalent to $2r$ when r tends to zero. Otherwise, one can find balls $B(x,r)$ with arbitrarily small radii such that the length of $K \cap B(x,r)$ is less than $\frac{3}{4}2r$. This is the *density criterion.* Another geometric criterion is the following. If x is in the rectifiable part, then K admits a tangent line at x, and if x is in the unrectifiable part, every cone with vertex x and arbitrarily small diameter contains parts of K with positive length. Here again, strong *local criteria* give an account of the (nonlocal) rectifiability property. Last but not least, the unrectifiable part has a "fractal" structure which, among other properties, implies that its projection on almost every line has zero length.

When a computer scientist has applied some *edge detector* to an image, he can see on the screen a set of points, and most edge detection or *segmentation devices* deal with so-called *connectivity algorithms.* What do these algorithms aim at? Clearly, a separation of the *rectifiable* part from the unrectifiable one. Indeed, "edges" are always assumed to be contained in curves or in finite unions of curves because they are assumed to correspond to the apparent contours of physical objects. Even if these objects may have "fractal" boundaries, like trees or clouds, the perception and the computer programs tend to define their boundary as a set of smooth lines. *The Besicovitch theory is a first answer as to the theoretical possibility of separating by local criteria the rectifiable part of a set from its unrectifiable part.* To our knowledge, no one of the Besicovitch criteria has been tried for "cleaning" edge sets from their spurious, unrectifiable part. The reason may be that the situation, in the numerical framework, is less contrasted: what really is aimed at is *a separation from a more rectifiable part from a less rectifiable part.*

We shall see that variational algorithms in image processing implicitly do this task, since minima of the Mumford-Shah functional are *uniformly rectifiable* in a sense which will be developed in the third part of this book. In any case, the second part of this book is devoted to a complete exposition of the Besicovitch theory. Our presentation does not follow the original Besicovitch presentation, which only works for dimension 1. (An excellent such presentation is done in Falconer [Fal1]). We tried to take into account all the new information about the geometric structure of rectifiable and unrectifiable sets given by the above mentioned Marstrand and Mattila contributions, as well as several new compactness results about sequences of m-sets. All of these results are proved because they are used in the third part of this book: We wished the exposition to be as complete and self-contained as possible.

What will be proved about the Mumford-Shah functional.

Although the Mumford and Shah conjecture is not yet proved, adequately weakened versions have been, and we shall list a series of questions which have been and will in this book be positively answered.

• Do minimal segmentations exist, and are they unique for each image? The answer is yes for existence and no for uniqueness, which matches the experimental intuition.

• Is the set of minimal segmentations small? Somehow, yes, because it is compact.

We now are concerned with the **structure** of the edge set.

• Are the edge sets obtained by the Mumford-Shah theory *rectifiable* in the Besicovitch sense? The answer is yes, and this is a first confirmation that the variational method is sound.

• Is the "edge set" made of a finite set of (not necessarily smooth) curves? The answer is unknown, but we shall see that the edge set is almost equal to a finite set of curves and can even be enclosed in a single curve with finite length, which is a remarkable result!
At this stage of the theory, one can deduce that the Mumford-Shah energy and conjecture are sound and that all the segmentation algorithms discussed in the first part of the book find, beyond their own experimental justification, a mathematical label of correctedness.

To summarize, this book is divided in three parts:

• Modelization (Chapters 1 to 5),
• Geometric measure theory and the structure of sets with finite Hausdorff measure (Chapters 6 to 12), and
• Existence and structural properties of the minimal segmentations for the Mumford-Shah model (Chapters 13 to 16).

We tried to make this book as self-contained as possible. The first part only asks that the reader have some familiarity with the formalism of partial differential operators; no further mathematical knowledge is required. The second part is fully self-contained. It will, however, seem easier to readers having some knowledge of the Lebesgue measure and integration theory. The last part asks the reader to know the elementary distribution theory, but not more than the mere definitions of derivatives in the distributional sense. The few results about elliptic equations and variational problems used therein are either proved or explicitly admitted.
The authors wish to thank Haim Brezis who had the idea that the present book could be written, and Ennio de Giorgi and Yves Meyer, who were a steady source of good suggestions and encouragement. The remarkable applications of the segmentation algorithm discussed in Chapter 5, developed at Cognitec-Inc by Lenny Rudin and Stanley Osher, brought surprise and excitement. We thank Isabeau Birindelli and Guy David for many useful comments. The first author thanks heartily his collaborators Luis Alvarez, Coloma Ballester, Vicent Caselles, Antonin Chambolle, Francine Catté, Tomeu Coll, Françoise Dibos, Manolo Gonzalez, Georges Koepfler, Frédéric Guichard and Christian Lopez, who kindly put to his disposition experimental results. Every successful image processing experiment condenses a lot of mathematical and algorithmic skill.