25 Mumford and Shah Model and its Applications to Image Segmentation and Image Restoration

Leah Bar · Tony F. Chan · Ginmo Chung · Miyoun Jung ·
Nahum Kiryati · Rami Mohieddine · Nir Sochen ·
Luminita A. Vese

Abstract: We present in this chapter an overview of the Mumford and Shah model for image segmentation. We discuss its various formulations, some of its properties, the mathematical framework, and several approximations. We also present numerical algorithms and segmentation results using the Ambrosio–Tortorelli phase-field approximations on one hand, and using the level set formulations on the other hand. Several applications of the Mumford–Shah problem to image restoration are also presented.

25.1 Introduction: Description of the Mumford and Shah Model

An important problem in image analysis and computer vision is the segmentation one, that aims to partition a given image into its constituent objects, or to find boundaries of such objects. This chapter is devoted to the description, analysis, approximations, and applications of the classical Mumford and Shah functional proposed for image segmentation. In [62–64], David Mumford and Jayant Shah have formulated an energy minimization problem that allows to compute optimal piecewise-smooth or piecewise-constant approximations u of a given initial image g. Since then, their model has been analyzed and considered in depth by many authors, by studying properties of minimizers, approximations, and applications to image segmentation, image partition, image restoration, and more generally to image analysis and computer vision.

We denote by $\Omega \subset \mathbb{R}^d$ the image domain (an interval if $d = 1$, or a rectangle in the plane if $d = 2$). More generally, we assume that Ω is open, bounded, and connected. Let $g : \Omega \to \mathbb{R}$ be a given gray-scale image (a signal in one dimension, a planar image in two dimensions, or a volumetric image in three dimensions). It is natural and without losing any generality to assume that g is a bounded function in Ω, $g \in L^\infty(\Omega)$.

As formulated by Mumford and Shah [64], the *segmentation problem* in image analysis and computer vision consists in computing a decomposition

$$\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_n \cup K$$

of the domain of the image g such that

(a) The image g varies smoothly and/or slowly *within* each Ω_i.
(b) The image g varies discontinuously and/or rapidly across most of the boundary K between different Ω_i.

From the point of view of approximation theory, the segmentation problem may be restated as seeking ways to define and compute *optimal approximations* of a general function $g(x)$ by piecewise-smooth functions $u(x)$, i.e., functions u whose restrictions u_i to the pieces Ω_i of a decomposition of the domain Ω are continuous or differentiable.

In what follows, Ω_i will be disjoint connected open subsets of a domain Ω, each one with a piecewise-smooth boundary, and K will be a closed set, as the union of boundaries of Ω_i inside Ω, thus

$$\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \cup \Omega_n \cup K, \quad K = \Omega \cap (\partial\Omega_1 \cup \ldots \cup \partial\Omega_n).$$

The functional E to be minimized for image segmentation is defined by [62–64],

$$E(u, K) = \mu^2 \int_\Omega (u - g)^2 dx + \int_{\Omega \setminus K} |\nabla u|^2 dx + \nu|K|, \tag{25.1}$$

where $u : \Omega \to \mathbb{R}$ is continuous or even differentiable inside each Ω_i (or $u \in H^1(\Omega_i)$) and may be discontinuous across K. Here, $|K|$ stands for the total surface measure of the hypersurface K (the counting measure if $d = 1$, the length measure if $d = 2$, the area measure if $d = 3$). Later, we will define $|K|$ by $\mathcal{H}^{d-1}(K)$, the $d - 1$ dimensional Hausdorff measure in $\mathbb{R}^d$.

As explained by Mumford and Shah, dropping any of these three terms in (❯ 25.1), inf $E = 0$: without the first, take $u = 0$, $K = \emptyset$; without the second, take $u = g$, $K = \emptyset$; without the third, take for example, in the discrete case K to be the boundary of all pixels of the image g, each Ω_i be a pixel and u to be the average (value) of g over each pixel. The presence of all three terms leads to nontrivial solutions u, and an optimal pair (u, K) can be seen as a cartoon of the actual image g, providing a simplification of g.

An important particular case is obtained when we restrict E to piecewise-constant functions u, i.e., $u =$ constant c_i on each open set Ω_i. Multiplying E by μ^{-2}, we have

$$\mu^{-2} E(u, K) = \sum_i \int_{\Omega_i} (g - c_i)^2 dx + \nu_0|K|,$$

where $\nu_0 = \nu/\mu^2$. It is easy to verify that this is minimized in the variables c_i by setting

$$c_i = \text{mean}_{\Omega_i}(g) = \frac{\int_{\Omega_i} g(x) dx}{|\Omega_i|},$$

where $|\Omega_i|$ denotes here the Lebesgue measure of Ω_i (e.g., area if $d = 2$, volume if $d = 3$), so it is sufficient to minimize

$$E_0(K) = \sum_i \int_{\Omega_i} (g - \text{mean}_{\Omega_i} g)^2 dx + \nu_0|K|.$$

It is possible to interpret E_0 as the limit functional of E as $\mu \to 0$ [64].

Finally, the Mumford and Shah model can also be seen as a deterministic refinement of Geman and Geman's image restoration model [42].

25.2 Background: The First Variation

In order to better understand, analyze, and use the minimization problem (❯ 25.1), it is useful to compute its first variation with respect to each of the unknowns.

We first recall the definition of Sobolev functions $u \in W^{1,2}(U)$ [1], necessary to properly define a minimizer u when K is fixed.

Definition 1   Let $U \subset \mathbb{R}^d$ be an open set. We denote by $W^{1,2}(U)$ (or by $H^1(U)$) the set of functions $u \in L^2(\Omega)$, whose first-order distributional partial derivatives belong to $L^2(U)$. This means that there are functions $u_1, \ldots, u_d \in L^2(U)$ such that

$$\int_U u(x) \frac{\partial \varphi}{\partial x_i}(x) dx = - \int_U u_i(x) \varphi(x) dx$$

for $1 \le i \le d$ and for all functions $\varphi \in C_c^\infty(U)$.

We may denote by $\frac{\partial u}{\partial x_i}$ the distributional derivative u_i of u and by $\nabla u = \left(\frac{\partial u}{\partial x_1}, \ldots, \frac{\partial u}{\partial x_d} \right)$ its distributional gradient. In what follows, we denote by $|\nabla u|(x)$ the Euclidean norm of the gradient vector at x. $H^1(U) = W^{1,2}(U)$ becomes a Banach space endowed with the norm

$$\|u\|_{W^{1,2}(U)} = \left[\int_U u^2 dx + \sum_{i=1}^d \int_U \left(\frac{\partial u}{\partial x_i} \right)^2 dx \right]^{1/2}.$$

25.2.1   Minimizing in *u* with *K* Fixed

Let us assume first that K is fixed, as a closed subset of the open and bounded set $\Omega \subset \mathbb{R}^d$, and denote by

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

for $u \in W^{1,2}(\Omega \setminus K)$, where $\Omega \setminus K$ is open and bounded, and $g \in L^2(\Omega \setminus K)$. We have the following classical results obtained as a consequence of the standard method of calculus of variations.

Proposition 1   *There is a unique minimizer of the problem*

$$\inf_{u \in W^{1,2}(\Omega \setminus K)} E(u). \tag{25.2}$$

Proof   [39] First, we note that $0 \le \inf E < +\infty$, since we can choose $u_0 \equiv 0$ and $E(u_0) = \mu^2 \int_{\Omega \setminus K} g^2(x) dx < +\infty$. Thus, we can denote by $m = \inf_u E(u)$ and let $\{u_j\}_{j \ge 1} \in W^{1,2}(\Omega \setminus K)$ be a minimizing sequence such that $\lim_{j \to \infty} E(u_j) = m$.

Recall that for $u, v \in L^2$,

$$\left\| \frac{u+v}{2} \right\|_2^2 + \left\| \frac{u-v}{2} \right\|_2^2 = \frac{1}{2} \|u\|_2^2 + \frac{1}{2} \|v\|_2^2,$$

and so

$$\left\| \frac{u+v}{2} \right\|_2^2 = \frac{1}{2} \|u\|_2^2 + \frac{1}{2} \|v\|_2^2 - \left\| \frac{u-v}{2} \right\|_2^2. \tag{25.3}$$

Let $u, v \in W^{1,2}(\Omega \backslash K)$, thus $E(u), E(v) < \infty$, and apply (❯ 25.3) to $u - g$ and $v - g$, and then to ∇u and ∇v; we obtain

$$E\left(\frac{u+v}{2}\right) = \frac{1}{2}E(u) + \frac{1}{2}E(v) - \frac{\mu^2}{4}\int_{\Omega\backslash K}|u-v|^2 dx - \frac{1}{4}\int_{\Omega\backslash K}|\nabla(u-v)|^2 dx$$

$$= \frac{1}{2}E(u) + \frac{1}{2}E(v) -$$

$$\begin{cases} \frac{\mu^2}{4}\|u-v\|^2_{W^{1,2}(\Omega\backslash K)} + \left(1 - \frac{\mu^2}{4}\right)\|\nabla(u-v)\|^2_2 \text{ if } \frac{1}{4} \geq \frac{\mu^2}{4} \\ \frac{1}{4}\|u-v\|^2_{W^{1,2}(\Omega\backslash K)} + \left(\frac{\mu^2}{4} - 1\right)\|u-v\|^2_2 \text{ if } \frac{1}{4} \leq \frac{\mu^2}{4} \end{cases}. \quad (25.4)$$

If we choose $u, v \in W^{1,2}(\Omega\backslash K)$, such that $E(u), E(v) \leq m + \epsilon$, then

$$m \leq E\left(\frac{u+v}{2}\right) \leq m + \epsilon -$$

$$\begin{cases} \frac{\mu^2}{4}\|u-v\|^2_{W^{1,2}(\Omega\backslash K)} + \left(1 - \frac{\mu^2}{4}\right)\|\nabla(u-v)\|^2_2 \text{ if } \frac{1}{4} \geq \frac{\mu^2}{4} \\ \frac{1}{4}\|u-v\|^2_{W^{1,2}(\Omega\backslash K)} + \left(\frac{\mu^2}{4} - 1\right)\|u-v\|^2_2 \text{ if } \frac{1}{4} \leq \frac{\mu^2}{4} \end{cases}$$

thus,

$$\|u-v\|^2_{W^{1,2}(\Omega\backslash K)} \leq \begin{cases} \frac{4\epsilon}{\mu^2} \text{ if } \frac{1}{4} \geq \frac{\mu^2}{4} \\ 4\epsilon \text{ if } \frac{1}{4} \leq \frac{\mu^2}{4} \end{cases}. \quad (25.5)$$

We let $w_j = u_j - u_1$. From (❯ 25.5), $\{w_j\}$ is a Cauchy sequence in $W^{1,2}(\Omega\backslash K)$; let w denote its limit and set $u_0 = u_1 + w$. Then

$$E(u_0) = \mu^2\|u_0 - g\|^2_2 + \|\nabla u_0\|^2_2 = \mu^2\|(u_1 - g) + w\|^2_2 + \|\nabla u_1 + \nabla w\|^2_2$$

$$= \lim_{j\to+\infty}\left[\mu^2\|(u_1 - g) + w_j\|^2_2 + \|\nabla u_1 + \nabla w_j\|^2_2\right]$$

$$= \lim_{j\to+\infty} E(u_j) = m,$$

by the continuity of L^2-norms. This shows the existence of minimizers. The uniqueness follows from (❯ 25.5) by taking $\epsilon = 0$. ∎

Proposition 2 *The unique solution u of (❯ 25.2) is solution of the elliptic problem*

$$\int_{\Omega\backslash K}\nabla u(x) \cdot \nabla v(x)dx = -\mu^2\int_{\Omega\backslash K}[u(x) - g(x)]v(x)dx, \quad \forall v \in W^{1,2}(\Omega\backslash K), \quad (25.6)$$

or of

$$\triangle u = \mu^2(u - g)$$

in the sense of distributions in $\Omega\backslash K$, with associated boundary condition $\frac{\partial u}{\partial \vec{N}} = 0$ on $\partial(\Omega\backslash K)$, where $\vec{N}$ is the exterior unit normal to the boundary.

Proof   Indeed, let $\epsilon \mapsto A(\epsilon) = E(u + \epsilon v)$ for $\epsilon \in \mathbb{R}$ and arbitrary $v \in W^{1,2}(\Omega\backslash K)$. Then A is a quadratic function of ϵ, given by

$$A(\epsilon) = \mu^2 \int_{\Omega\backslash K} (u - g)^2 dx + \epsilon^2 \mu^2 \int_{\Omega\backslash K} v^2 dx + 2\epsilon\mu^2 \int_{\Omega\backslash K} (u - g)v dx$$
$$+ \int_{\Omega\backslash K} |\nabla u|^2 dx + \epsilon^2 \int_{\Omega\backslash K} |\nabla v|^2 dx + 2\epsilon \int_{\Omega\backslash K} \nabla u \cdot \nabla v dx,$$

and we have

$$A'(\epsilon) = 2\epsilon\mu^2 \int_{\Omega\backslash K} v^2 dx + 2\mu^2 \int_{\Omega\backslash K} (u - g)v dx + 2\epsilon \int_{\Omega\backslash K} |\nabla v|^2 dx$$
$$+ 2 \int_{\Omega\backslash K} \nabla u \cdot \nabla v dx,$$

and

$$A'(0) = 2\mu^2 \int_{\Omega\backslash K} (u - g)v dx + 2 \int_{\Omega\backslash K} \nabla u \cdot \nabla v dx.$$

Since we must have $E(u) = A(0) \leq A(\epsilon) = E(u + \epsilon v)$ for all $\epsilon \in \mathbb{R}$ and all $v \in W^{1,2}(\Omega\backslash K)$, we impose $A'(0) = 0$ for all such v, which yields the weak formulation (❯ 25.6).

If in addition u would be a strong classical solution of the problem, or if it would belong to $W^{2,2}(\Omega\backslash K)$, then integrating by parts in the last relation we obtain

$$A'(0) = 2\mu^2 \int_{\Omega\backslash K} (u - g)v dx - 2 \int_{\Omega\backslash K} (\triangle u)v dx + 2 \int_{\partial(\Omega\backslash K)} \nabla u \cdot \vec{N} v ds = 0.$$

Taking now $v \in C_0^1(\Omega\backslash K) \subset W^{1,2}(\Omega\backslash K)$, we obtain

$$\triangle u = \mu^2(u - g) \text{ in } \Omega\backslash K.$$

Using this and taking now $v \in C^1(\Omega\backslash K)$, we deduce the associated implicit boundary condition $\nabla u \cdot \vec{N} = \frac{\partial u}{\partial \vec{N}} = 0$ on the boundary of $\Omega\backslash K$ (in other words, on the boundary of Ω and of each Ω_i).   ∎

Assume now that $g \in L^\infty(\Omega\backslash K)$, which is not a restrictive assumption when g represents an image. We can deduce that the unique minimizer u of (❯ 25.2) satisfies $\|u\|_\infty \leq \|g\|_\infty$ (as expected, due to the smoothing properties of the energy). To prove this, we first state the following classical lemma (see e.g., ref.[39], Chapter A3).

Lemma 1   *If $\Omega\backslash K$ is open, and if $u \in W^{1,2}(\Omega\backslash K)$, then $u^+ = \max(u, 0)$ also lies in $W^{1,2}(\Omega\backslash K)$ and $|\nabla u^+(x)| \leq |\nabla u(x)|$ almost everywhere.*

Now let $u^*(x) = \max\{-\|g\|_\infty, \min(\|g\|_\infty, u(x))\}$ be the obvious truncation of u. Lemma 1 implies that $u^* \in W^{1,2}(\Omega\backslash K)$ and that $\int_{\Omega\backslash K} |\nabla u^*(x)|^2 dx \leq \int_{\Omega\backslash K} |\nabla u(x)|^2 dx$. We also obviously have $\int_{\Omega\backslash K} (u^* - g)^2 dx \leq \int_{\Omega\backslash K} (u - g)^2 dx$, and we deduce that $E(u^*) \leq E(u)$. But u is the unique minimizer of E, thus $u(x) = u^*(x)$ almost everywhere and we deduce $\|u\|_\infty \leq \|g\|_\infty$.

Remark 1 Several classical regularity results for a weak solution u of (❯ 25.2) can be stated:

- If $g \in L^\infty(\Omega \setminus K)$, then $u \in C^1_{loc}(\Omega \setminus K)$ (see e.g., ref.[39], Chapter A3).
- If $g \in L^2(\Omega \setminus K)$, then $u \in W^{2,2}_{loc}(\Omega \setminus K) = H^2_{loc}(\Omega \setminus K)$, which implies that u solves the PDE (see e.g., ref.[40], Chapter 6.3).

$$\triangle u = \mu^2(u - g) \text{ a.e. in } \Omega \setminus K.$$

25.2.2 Minimizing in K

We wish to formally compute here the first variation of $E(u, K)$ with respect to K. Let us assume that (u, K) is a minimizer of E from (❯ 25.1), and we vary K. Let us assume that locally, $K \cap U$ is the graph of a regular function ϕ, where U is a small neighborhood near a regular, simple point P of K. Without loss of generality, we can assume that $U = D \times I$ where I is an interval in $\mathbb{R}$ and $K \cap U = \{(x_1, x_2, \ldots, x_d) \in U = D \times I : x_d = \phi(x_1, \ldots, x_{d-1})\}$. Let u^+ denote the restriction of u to

$$U^+ = \{(x_1, x_2, \ldots, x_d) : x_d > \phi(x_1, \ldots, x_{d-1})\} \cap U,$$

and u^- the restriction of u to

$$U^- = \{(x_1, x_2, \ldots, x_d) : x_d < \phi(x_1, \ldots, x_{d-1})\} \cap U,$$

and choose H^1 extensions of u^+ from U^+ to U, and of u^- from U^- to U. For small ϵ, define a deformation K_ϵ of K inside U as the graph of

$$x_d = \phi(x_1, \ldots, x_{d-1}) + \epsilon \psi(x_1, \ldots, x_{d-1}),$$

such that ψ is regular and zero outside D, and $K_\epsilon = K$ outside U. Define

$$u_\epsilon(x) = \begin{cases} u(x) & \text{if} & x \notin U, \\ (\text{extension of } u^+)(x) & \text{if} & x \in U, \ x \text{ above } K_\epsilon \cap U \\ (\text{extension of } u^-)(x) & \text{if} & x \in U, \ x \text{ below } K_\epsilon \cap U. \end{cases}$$

Now, using $z = (x_1, \ldots, x_{d-1})$,

$$\begin{aligned} E(u_\epsilon, K_\epsilon) - E(u, K) &= \mu^2 \int_U \left[(u_\epsilon - g)^2 dx - (u - g)^2 \right] dx \\ &\quad + \int_{U \setminus K_\epsilon} |\nabla u_\epsilon|^2 dx - \int_{U \setminus K} |\nabla u|^2 dx + \nu \left[|K_\epsilon \cap U| - |K \cap U| \right] \\ &= \mu^2 \int_D \left(\int_{\phi(z)}^{\phi(z) + \epsilon \psi(z)} \left[(u^- - g)^2 - (u^+ - g)^2 \right] dx_d \right) dz \\ &\quad + \int_D \left(\int_{\phi(z)}^{\phi(z) + \epsilon \psi(z)} \left[|\nabla u^-|^2 - |\nabla u^+|^2 \right] dx_d \right) dz \\ &\quad + \nu \int_D \left[\sqrt{1 + |\nabla(\phi + \epsilon \psi)|^2} - \sqrt{1 + |\nabla \phi|^2} \right] dz. \end{aligned}$$

Thus,

$$\lim_{\epsilon \to 0} \frac{E(u_\epsilon, K_\epsilon) - E(u, K)}{\epsilon} = \mu^2 \int_D \left[(u^- - g)^2 - (u^+ - g)^2 \right] \Big|_{x_d = \phi(z)} \psi(z) dz$$

$$+ \int_D \left[|\nabla u^-|^2 - |\nabla u^+|^2 \right] \Big|_{x_d = \phi(z)} \psi(z) dz + \nu \int_D \frac{\nabla \phi \cdot \nabla \psi}{\sqrt{1 + |\nabla \phi|^2}} dz = 0$$

for all such ψ, since (u, K) is a minimizer. Integrating by parts, we formally obtain for all ψ:

$$\int_D \left\{ \left[(\mu^2 (u^- - g)^2 + |\nabla u^-|^2) - (\mu^2 (u^+ - g)^2 + |\nabla u^+|^2) \right] \Big|_{x_d = \phi(z)} \right.$$

$$\left. -\nu \mathrm{div} \left(\frac{\nabla \phi}{\sqrt{1 + |\nabla \phi|^2}} \right) \right\} \psi(z) dz = 0,$$

and we obtain the first variation with respect to K,

$$\left[\mu^2 (u^- - g)^2 + |\nabla u^-|^2 \right] - \left[\mu^2 (u^+ - g)^2 + |\nabla u^+|^2 \right] - \nu \mathrm{div} \left(\frac{\nabla \phi}{\sqrt{1 + |\nabla \phi|^2}} \right) = 0 \qquad (25.7)$$

on $K \cap U$. Noticing that the last term represents the curvature of $K \cap U$, and if we write the energy density as

$$e(u; x) = \mu^2 (u(x) - g(x))^2 + |\nabla u(x)|^2,$$

we finally obtain

$$e(u^+) - e(u^-) + \nu \mathrm{curv}(K) = 0 \text{ on } K$$

(at regular points of K, provided that the traces of u and of $|\nabla u|$ on each side of K are taken in the sense of Sobolev traces).

We conclude this section by stating another important result from [64] regarding the type of singular points of K, when (u, K) is a minimizer of E from (❯ 25.1), in two dimensions, $d = 2$. For the rather technical proof of this result, we refer the reader to the instructive and inspiring constructions from [64].

Theorem 1 *Let $d = 2$. If (u, K) is a minimizer of $E(u, K)$ such that K is a union of simple $C^{1,1}$-curves K_i meeting $\partial \Omega$ and meeting each other only at their endpoints, then the only vertices of K are:*

(1) *Points P on the boundary $\partial \Omega$ where one K_i meets $\partial \Omega$ perpendicularly*
(2) *Triple points P where three K_i meet with angles $2\pi/3$*
(3) *Crack-tips where a K_i ends and meets nothing.*

In the later sections we will discuss cases when the minimizer u is restricted to a specific class of piecewise-constant or piecewise-smooth functions.

25.3 Mathematical Modeling and Analysis: The Weak Formulation of the Mumford and Shah Functional

To better study the mathematical properties of the Mumford and Shah functional (❯ 25.1), it is necessary to define the measure of K as its $d - 1$-dimensional Hausdorff measure $\mathcal{H}^{d-1}(K)$, which is the most natural way to extend the notion of length to nonsmooth sets. We recall the definition of the Hausdorff measure [4, 39, 41].

Definition 2 For $K \subset \mathbb{R}^d$ and $n > 0$, set

$$\mathcal{H}^n(K) = \sup_{\epsilon > 0} \mathcal{H}^n_\epsilon(K),$$

called the n-dimensional Hausdorff measure of the set K, where

$$\mathcal{H}^n_\epsilon(K) = c_n \inf \left\{ \sum_{i=1}^{\infty} (\mathrm{diam} A_i)^n \right\},$$

where the infimum is taken over all countable families $\{A_i\}_{i=1}^{\infty}$ of open sets A_i such that

$$K \subset \bigcup_{i=1}^{\infty} A_i \text{ and } \mathrm{diam} \, A_i \le \epsilon \text{ for all } i.$$

Here, the constant c_n is chosen so that $\mathcal{H}^n$ coincides with the Lebesgue measure on n-planes.

Remark 2 When n is an integer and K is contained in a C^1-surface of dimension n, $\mathcal{H}^n(K)$ coincides with its n-dimensional surface measure.

We consider a first variant of the functional,

$$E(u, K) = \mu^2 \int_{\Omega \setminus K} (u - g)^2 dx + \int_{\Omega \setminus K} |\nabla u|^2 dx + \nu \mathcal{H}^{d-1}(K). \tag{25.8}$$

In order to apply the direct method of calculus of variations for proving existence of minimizers, it is necessary to find a topology for which the functional is lower semi-continuous, while ensuring compactness of minimizing sequences. Unfortunately, the last functional $K \mapsto \mathcal{H}^{d-1}(K)$ is not lower semi-continuous with respect to any compact topology [4, 8, 39].

To overcome this difficulty, the set K is substituted by the jump set S_u of u, thus K is eliminated, and the problem, called the weak formulation, becomes, in its second variant,

$$\inf_u \left\{ F(u) = \mu^2 \int_{\Omega \setminus S_u} (u - g)^2 dx + \int_{\Omega \setminus S_u} |\nabla u|^2 dx + \nu \mathcal{H}^{d-1}(S_u) \right\}. \tag{25.9}$$

For illustration, we also give the weak formulation in one dimension, for signals. The problem of reconstructing and segmenting a signal u from a degraded input g deriving from a distorted transmission, can be modeled as finding the minimum

$$\inf_u \left\{ \mu^2 \int_a^b (u-g)^2 dt + \int_{(a,b)\backslash S_u} |u'|^2 dt + \nu \#(S_u) \right\},$$

where $\Omega = (a,b)$, S_u denotes the set of discontinuity points of u in the interval (a,b), and $\#(S_u) = \mathcal{H}^0(S_u)$ denotes the counting measure of S_u or its cardinal.

In order to show that (❯ 25.9) has a solution, the following notion of special functions of bounded variation and the following important lemma due to Ambrosio [3, 4] are necessary.

Definition 3 A function $u \in L^1(\Omega)$ is a special function of bounded variation on Ω if its distributional derivative can be written as

$$Du = \nabla u\, dx + (u^+ - u^-) \vec{N}_u \mathcal{H}^{d-1}|_{S_u}$$

such that $\nabla u \in L^1(\Omega)$, S_u is of finite Hausdorff measure, $(u^+ - u^-)\vec{N}_u \chi_{S_u} \in L^1(\Omega, \mathcal{H}^{d-1}|_{S_u}, \mathbb{R}^d)$, where u^+ and u^- are the traces of u on each side of the jump part S_u, and $\vec{N}_u$ is the unit normal to S_u. The space of special functions of bounded variation is denoted by $SBV(\Omega)$.

Lemma 2 Let $u_n \in SBV(\Omega)$ be a sequence of functions such that there exists a constant $C > 0$ with $|u_n(x)| \le C < \infty$ a.e. $x \in \Omega$ and $\int_\Omega |\nabla u_n|^2 dx + \mathcal{H}^{d-1}(S_{u_n}) \le C$. Then there exists a subsequence u_{n_k} converging a.e. to a function $u \in SBV(\Omega)$. Moreover, ∇u_{n_k} converges weakly in $L^2(\Omega)^d$ to ∇u, and

$$\mathcal{H}^{d-1}(S_u) \le \liminf_{n_k \to \infty} \mathcal{H}^{d-1}(S_{u_{n_k}}).$$

Theorem 2 Let $g \in L^\infty(\Omega)$, with $\Omega \subset \mathbb{R}^d$ open, bounded, and connected. There is a minimizer $u \in SBV(\Omega) \cap L^\infty(\Omega)$ of

$$F(u) = \mu^2 \int_{\Omega\backslash S_u} (u-g)^2 dx + \int_{\Omega\backslash S_u} |\nabla u|^2 dx + \nu \mathcal{H}^{d-1}(S_u).$$

Proof We notice that $0 \le \inf_{SBV(\Omega) \cap L^\infty(\Omega)} F < \infty$, because we can take $u_0 = 0 \in SBV(\Omega) \cap L^\infty(\Omega)$ and using the fact that $g \in L^\infty(\Omega) \subset L^2(\Omega)$, $F(u_0) < \infty$. Thus, there is a minimizing sequence $u_n \in SBV(\Omega) \cap L^\infty(\Omega)$ satisfying $\lim_{n\to\infty} F(u_n) = \inf F$. We also notice that, by the truncation argument from before, we can assume that $\|u_n\|_\infty \le \|g\|_\infty < \infty$. Since $F(u_n) \le C < \infty$ for all $n \ge 0$, and using $g \in L^\infty(\Omega) \subset L^2(\Omega)$, we deduce that $\|u_n\|_2 \le C$ and $\int_{\Omega\backslash S_{u_n}} |\nabla u_n|^2 dx + \mathcal{H}^{d-1}(S_{u_n}) < C$ for some positive real constant C. Using these and Ambrosio's compactness result, we deduce that there is a subsequence u_{n_k} of u_n, and $u \in SBV(\Omega)$, such that $u_{n_k} \rightharpoonup u$ in $L^2(\Omega)$, $\nabla u_{n_k} \rightharpoonup \nabla u$ in $L^2(\Omega)^d$. Therefore, $F(u) \le \liminf_{n_k \to \infty} F(u_{n_k}) = \inf F$, and we can also deduce that $\|u\|_\infty \le \|g\|_\infty$. ∎

For additional existence, regularity results and fine properties of minimizers, and for the connections between problems (❯ 25.8) and (❯ 25.9), we refer the reader to Dal Maso et al. [55, 56], the important monographs by Morel and Solimini [61], Chambolle [26], by Ambrosio et al. [4], by David [39], and by Braides [19]. Existence and regularity of minimizers for the piecewise-constant case can be found in [64], Congedo and Tamanini [53, 57, 80, 81], Larsen [52], among other works.

25.4   Numerical Methods: Approximations to the Mumford and Shah Functional

Since the original Mumford and Shah functional (❯ 25.1) (or its weak formulation (❯ 25.9)) is non-convex, it has an unknown set K of lower dimension, and it is not the lower-semicontinuous envelope of a more amenable functional, it is difficult to find smooth approximations and to solve the minimization in practice. Several approximations have been proposed, including: the weak membrane model and the graduate non-convexity of Blake and Zisserman [16] (which can be seen as a discrete version of the Mumford and Shah segmentation problem); discrete finite differences approximations starting with the work of Chambolle [23–25] (also proving the Γ-convergence of Blake-Zisserman approximations to the weak Mumford–Shah functional in one dimension); finite element approximations by Chambolle and Dal Maso [27] and by Chambolle and Bourdin [17, 18]; phase-field elliptic approximations due to Ambrosio and Tortorelli [5, 6] (with generalizations presented by [19] and extensions by Shah [78], and Alicandro et al. [2]); region growing and merging methods proposed by Koepfler et al. [49], by Morel and Solimini [61], by Dal Maso et al. [55, 56] and level set approximations proposed by Chan and Vese [28–31, 84], by Samson et al. [75], and by Tsai et al. [83]; approximations by nonlocal functionals by Braides and Dal Maso [20], among other approximations. We present in this section in many more details the phase-field elliptic approximations and the level set approximations together with their applications.

For proving the convergence of some of these approximations to the Mumford and Shah functional, the notion of Γ-convergence is used, which is briefly recalled below. We refer the interested reader to Dal Maso [38] for a comprehensive introduction to Γ-convergence.

We would like to refer the reader to the monographs and textbooks by Braides [19], by Morel and Solimini [61], and by Ambrosio et al. [4] on detailed presentations of approximations to the Mumford and Shah functional.

Definition 4   Let $X = (X, D)$ be a metric space. We say that a sequence $F_j : X \to [-\infty, +\infty]$ Γ-converges to $F : X \to [-\infty, +\infty]$ (as $j \to \infty$) if for all $u \in X$ we have
  (1) (liminf inequality) for every sequence $(u_j) \subset X$ converging to u,

$$F(u) \leq \liminf_j F_j(u_j) \tag{25.10}$$

(2) (existence of a recovery sequence) there exists a sequence $(u_j) \subset X$ converging to u such that

$$F(u) \geq \limsup_j F_j(u_j),$$

or, equivalently by (❷ 25.10),

$$F(u) = \lim_j F_j(u_j).$$

The function F is called the Γ-limit of (F_j) (with respect to D), and we write $F = \Gamma\text{-}\lim_j F_j$.

The following fundamental theorem is essential in the convergence of some of the approximations.

Theorem 3 *(Fundamental Theorem of Γ-convergence) Let us suppose that $F = \Gamma\text{-}\lim_j F_j$, and let a compact set $C \subset X$ exist such that $\inf_X F_j = \inf_C F_j$ for all j. Then there is minimum of F over X such that*

$$\min_X F = \liminf_j \inf_X F_j,$$

and if $(u_j) \subset X$ is a converging sequence such that $\lim_j F_j(u_j) = \lim_j \inf_X F_j$, then its limit is a minimum point of F.

25.4.1 Ambrosio and Tortorelli Phase-Field Elliptic Approximations

A specific strategy, closer to the initial formulation of the Mumford–Shah problem in terms of pairs $(u, K = S_u)$, is based on the approximation by functionals depending on two variables (u, v), the second one related to the set $K = S_u$.

25.4.1.1 Approximations of the Perimeter by Elliptic Functionals

The Modica–Mortola theorem [58, 59] enables the variational approximation of the perimeter functional $E \mapsto P(E, \Omega) = \int_\Omega |D\chi_E| < \infty$ of an open subset E of Ω by the quadratic, elliptic functionals

$$MM_\epsilon(v) = \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{W(v)}{\epsilon} \right) dx, \quad v \in W^{1,2}(\Omega),$$

where $W(t)$ is a "double-well" potential. For instance, choosing $W(t) = t^2(1 - t)^2$, assuming that Ω is bounded with Lipschitz boundary and setting $MM_\epsilon(v) = \infty$ if $v \in L^2(\Omega) \backslash W^{1,2}(\Omega)$, the functionals $MM_\epsilon(v)$ Γ-converge in $L^2(\Omega)$ to

$$F(v) = \begin{cases} \frac{1}{3} P(E, \Omega) & \text{if } v = \chi_E \text{ for some } E \in \mathcal{B}(\Omega), \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathcal{B}(\Omega)$ denotes the σ-algebra of Borel subsets of Ω.

Minimizing the functional $MM_\epsilon(v)$ with respect to v yields the associated Euler–Lagrange equation and boundary condition,

$$W'(v) = 2\epsilon^2 \triangle v \text{ in } \Omega, \quad \frac{\partial v}{\partial \vec{N}} = 0 \text{ on } \partial\Omega,$$

which can be easily solved in practice by finite differences.

25.4.1.2 Ambrosio-Tortorelli Approximations

In the Mumford and Shah functional the set $K = S_u$ is not necessarily the boundary of an open and bounded domain, but a construction similar to $MM_\epsilon(v)$ can still be used, with the potential $W(t) = \frac{1}{4}(1-t)^2$ instead. Ambrosio and Tortorelli proposed two elliptic approximations [5, 6] to the weak formulation of the Mumford and Shah problem. We present the second one [6], being simpler than the first one [5], and commonly used in practice.

Let $X = L^2(\Omega)^2$ and let us define

$$AT_\epsilon(u,v) = \int_\Omega (u-g)^2 dx + \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx \quad (25.11)$$

if $(u,v) \in W^{1,2}(\Omega)^2$, $0 \le v \le 1$, and $AT_\epsilon(u,v) = +\infty$ otherwise.

We also define the limiting Mumford–Shah functional,

$$F(u,v) = \begin{cases} \int_\Omega (u-g)^2 dx + \beta \int_\Omega |\nabla u|^2 + \alpha \mathcal{H}^{d-1}(S_u) & \text{if } u \in SBV(\Omega), \ v \equiv 1, \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 4 *AT_ϵ Γ-converges to F as $\epsilon \searrow 0$ in $L^2(\Omega)$. Moreover, AT_ϵ admits a minimizer (u_ϵ, v_ϵ) such that up to subsequences, u_ϵ converges to some $u \in SBV(\Omega)$ a minimizer of $F(u,1)$ and $\inf AT_\epsilon(u_\epsilon, v_\epsilon) \to F(u,1)$.*

Interesting generalizations of this result are given and proved by Braides in [19].

In practice, the Euler–Lagrange equations associated with the alternating minimization of AT_ϵ with respect to $u = u_\epsilon$ and $v = v_\epsilon$ are used and discretized by finite differences. These are

$$\frac{\partial AT_\epsilon(u,v)}{\partial u} = 2(u-g) - 2\beta \text{div}(v^2 \nabla u) = 0$$

$$\frac{\partial AT_\epsilon(u,v)}{\partial v} = 2\beta v |\nabla u|^2 - 2\alpha\epsilon \triangle v + \frac{\alpha}{2\epsilon}(v-1) = 0.$$

One of the finite differences approximations to compute u and v in two dimensions $x = (x_1, x_2)$ is as follows. We use a time-dependent scheme in $u = u(x_1, x_2, t)$ and a stationary semi-implicit fixed-point scheme in $v = v(x_1, x_2)$. Let $\triangle x_1 = \triangle x_2 = h$ be the step space, $\triangle t$ be the time step, and $g_{i,j}$, $u_{i,j}^n$, $v_{i,j}^n$ be the discrete versions of g, and of u and v at iteration $n \ge 0$, for $1 \le i \le M, 1 \le j \le N$. Initialize $u^0 = g$ and $v^0 = 0$.

For $n \geq 1$, compute and repeat to steady state, for $i = 2, \ldots, M - 1$ and $j = 2, \ldots, N - 1$ (combined with Neumann boundary conditions on $\partial\Omega$):

$$|\nabla u^n|^2_{i,j} = \left(\frac{u^n_{i+1,j} - u^n_{i,j}}{h}\right)^2 + \left(\frac{u^n_{i,j+1} - u^n_{i,j}}{h}\right)^2,$$

$$0 = 2\beta v^{n+1}_{i,j}|\nabla u^n|^2_{i,j} - 2\frac{\alpha\epsilon}{h^2}\left(v^n_{i+1,j} + v^n_{i-1,j} + v^n_{i,j+1} + v^n_{i,j-1}\right.$$

$$\left. -4v^{n+1}_{i,j}\right) + \frac{\alpha}{2\epsilon}\left(v^{n+1}_{i,j} - 1\right),$$

$$\frac{u^{n+1}_{i,j} - u^n_{i,j}}{\triangle t} = -\left(u^n_{i,j} - g_{i,j}\right) + \frac{\beta}{h^2}\left[\left(v^{n+1}_{i,j}\right)^2\left(u^n_{i+1,j} - u^n_{i,j}\right) + \left(v^{n+1}_{i,j}\right)^2\left(u^n_{i,j+1} - u^n_{i,j}\right)\right.$$

$$\left. -\left(v^{n+1}_{i-1,j}\right)^2\left(u^n_{i,j} - u^n_{i-1,j}\right) - \left(v^{n+1}_{i,j-1}\right)^2\left(u^n_{i,j} - u^n_{i,j-1}\right)\right]$$

which is equivalent with

$$|\nabla u^n|^2_{i,j} = \left(\frac{u^n_{i+1,j} - u^n_{i,j}}{h}\right)^2 + \left(\frac{u^n_{i,j+1} - u^n_{i,j}}{h}\right)^2,$$

$$v^{n+1}_{i,j} = \frac{\frac{\alpha}{2\epsilon} + \frac{2\alpha\epsilon}{h^2}\left(v^n_{i+1,j} + v^n_{i-1,j} + v^n_{i,j+1} + v^n_{i,j-1}\right)}{\frac{\alpha}{2\epsilon} + 2\beta|\nabla u^n|^2_{i,j} + \frac{8\alpha\epsilon}{h^2}},$$

$$u^{n+1}_{i,j} = u^n_{i,j} + \triangle t\left\{-\left(u^n_{i,j} - g_{i,j}\right) + \frac{\beta}{h^2}\left[\left(v^{n+1}_{i,j}\right)^2\left(u^n_{i+1,j} - u^n_{i,j}\right)\right.\right.$$

$$\left. + \left(v^{n+1}_{i,j}\right)^2\left(u^n_{i,j+1} - u^n_{i,j}\right) - \left(v^{n+1}_{i-1,j}\right)^2\left(u^n_{i,j} - u^n_{i-1,j}\right)\right.$$

$$\left.\left. - \left(v^{n+1}_{i,j-1}\right)^2\left(u^n_{i,j} - u^n_{i,j-1}\right)\right]\right\}.$$

We present experimental results obtained using the above Ambrosio–Tortorelli approximations applied to the well-known Barbara image shown in ❯ *Fig. 25-1* left. Segmented images u are shown in ❯ *Fig. 25-2* and the corresponding edge sets v are shown in ❯ *Fig. 25-3* for varying coefficients $\alpha, \beta \in \{1, 5, 10\}$. We notice that less regularization (decreasing both α and β) gives more edges in v, as expected, thus u is closer to g. Fixed α and increasing β gives smoother image u and fewer edges in v. Keeping fixed β but varying α does not produce much variation in the results. We also show in ❯ *Fig. 25-1* right the numerical energy versus iterations for the case $\alpha = \beta = 10$, $\epsilon = 0.0001$.

Applications of the Ambrosio–Tortorelli approximations to image restoration will be presented in details in ❯ Sect. 25.5.

25.4.2   Level Set Formulations of the Mumford and Shah Functional

We review in this section the level set formulations for minimizing the Mumford and Shah functional, as proposed initially by Chan and Vese [28–31, 84], and by Tsai et al. [83]
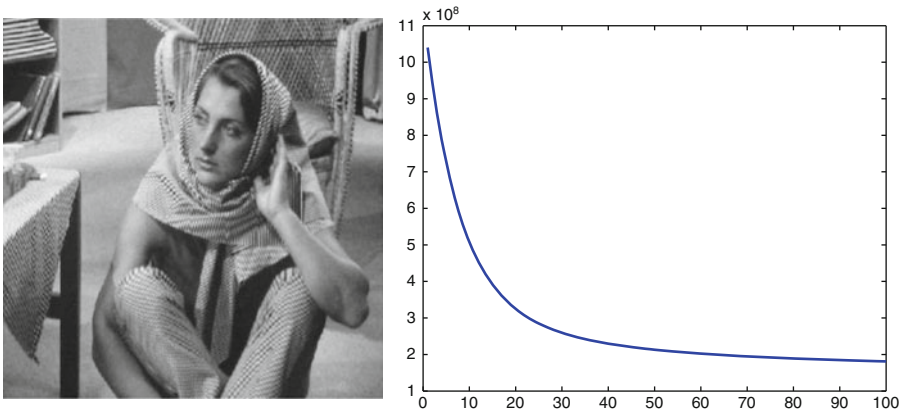
◼ Fig. 25-1
Left: original image *g*. *Right*: numerical energy versus iterations for the Ambrosio–Tortorelli approximations ($\alpha = \beta = 10$, $\epsilon = 0.0001$)

(see also the related work by Samson et al. [75] and Cohen et al. [36, 37]). These make the link between curve evolution, active contours, and Mumford–Shah segmentation. These models have been proposed by restricting the set of minimizers u to specific classes of functions: piecewise constant, piecewise smooth, with the edge set K represented by a union of curves or surfaces that are boundaries of open subsets of Ω. For example, if K is the boundary of an open-bounded subset of Ω, then it can be represented implicitly, as the zero-level line of a Lipschitz-continuous level set function. Thus the set K as an unknown is substituted by an unknown function, that defines it implicitly, and the Euler–Lagrange equations with respect to the unknowns can be easily computed and discretized.

Following the level set approach [69, 70, 76, 77], let $\phi : \Omega \to \mathbb{R}$ be a Lipschitz continuous function. We recall the variational level set terminology that will be useful to rewrite the Mumford and Shah functional in terms of (u, ϕ), instead of (u, K). We are inspired by the work of Zhao et al. [88] for a variational level set approach for motion of triple junctions in the plane.

We will use the one-dimensional (1D) Heaviside function H, defined by

$$H(z) = \begin{cases} 1 \text{ if } z \geq 0 \\ 0 \text{ if } z < 0 \end{cases},$$

and its distributional derivative $\delta = H'$ (in the weak sense). In practice, we may need to work with smooth approximations of the Heaviside and δ functions. Here, we will use the following C^∞ approximations as $\epsilon \to 0$ given by [28, 30],

$$H_\epsilon(z) = \frac{1}{2}\left[1 + \frac{2}{\pi}\arctan\left(\frac{z}{\epsilon}\right)\right], \quad \delta_\epsilon = H_\epsilon'.$$

(1,1)
SEGMENTED IMAGE u

(1,5)
SEGMENTED IMAGE u

(1,10)
SEGMENTED IMAGE u

(5,1)
SEGMENTED IMAGE u

(5,5)
SEGMENTED IMAGE u

(5,10)
SEGMENTED IMAGE u

(10,1)
SEGMENTED IMAGE u

(10,5)
SEGMENTED IMAGE u

(10,10)
SEGMENTED IMAGE u

◨ **Fig. 25-2**

Piecewise-smooth images *u* as minimizers of the Ambrosio–Tortorelli approximations for
$\epsilon = 0.0001$ **and various values of** (α, β)

The area (or the volume) of the region $\{x \in \Omega : \phi(x) > 0\}$ is

$$A\{x \in \Omega : \phi(x) > 0\} = \int_\Omega H(\phi(x))dx,$$

and for a level parameter $l \in \mathbb{R}$, the area (or volume) of the region $\{x \in \Omega : \phi(x) > l\}$ is

$$A\{x \in \Omega : \phi(x) > l\} = \int_\Omega H(\phi(x) - l)dx.$$
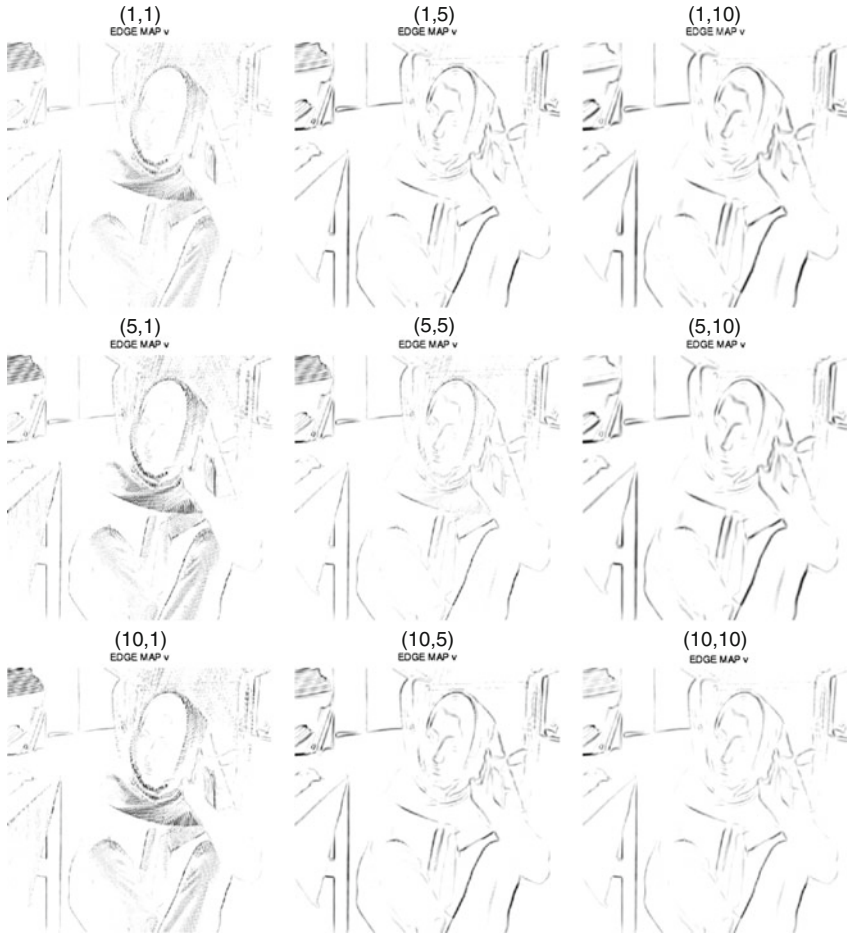
□ **Fig. 25-3**

Corresponding edge sets *v* as minimizers of the Ambrosio–Tortorelli approximations for $\epsilon = 0.0001$ and various values of (α, β)

The perimeter (or more generally the surface area) of the region $\{x \in \Omega : \phi(x) > 0\}$ is given by

$$L\{x \in \Omega : \phi(x) > 0\} = \int_\Omega |DH(\phi)|,$$

which is the total variation of $H(\phi)$ in Ω, and the perimeter (or surface area) of $\{x \in \Omega : \phi(x) > l\}$ is

$$L\{x \in \Omega : \phi(x) > l\} = \int_\Omega |DH(\phi - l)|.$$

Given the image data $g \in L^\infty(\Omega) \subset L^2(\Omega)$ to be segmented, the averages of g over the (nonempty) regions $\{x \in \Omega : \phi(x) > 0\}$ and $\{x \in \Omega : \phi(x) < 0\}$ respectively, are

$$\frac{\int_\Omega g(x)H(\phi(x))dx}{\int_\Omega H(\phi(x))dx} \quad \text{and} \quad \frac{\int_\Omega g(x)(1-H(\phi(x)))dx}{\int_\Omega(1-H(\phi(x)))dx} = \frac{\int_\Omega g(x)H(-\phi(x))dx}{\int_\Omega H(-\phi(x))dx}.$$

More generally, for a given level parameter $l \in \mathbb{R}$, the averages of g over the corresponding (nonempty) regions $\{x \in \Omega : \phi(x) > l\}$ and $\{x \in \Omega : \phi(x) < l\}$ respectively, are

$$\frac{\int_\Omega g(x)H(\phi(x)-l)dx}{\int_\Omega H(\phi(x)-l)dx} \quad \text{and} \quad \frac{\int_\Omega g(x)H(l-\phi(x))dx}{\int_\Omega H(l-\phi(x))dx}.$$

We prove next that if H and δ are substituted by the above C^∞ approximations H_ϵ, δ_ϵ as $\epsilon \to 0$, we obtain approximations of the area and length (perimeter) measures. We obviously have that $H_\epsilon(z) \to H(z)$ for all $z \in \mathbb{R}$, as $\epsilon \to 0$, and that the approximating area term $A_\epsilon(\phi) = \int_\Omega H_\epsilon(\phi(x))dx$ converges to $A(\phi) = \int_\Omega H(\phi(x))dx$.

Generalizing a result of Samson et al. [75], we can show [35] that our approximating functional $L_\epsilon(\phi) = \int_\Omega |DH_\epsilon(\phi)|dx = \int_\Omega \delta_\epsilon(\phi)|\nabla\phi|dx$ converges to the length $|K|$ of the zero-level line $K = \{x \in \Omega : \phi(x) = 0\}$, under the assumption that $\phi : \Omega \to \mathbb{R}$ is Lipschitz. The same result holds for the case of any l-level curve of ϕ and not only for the 0-level curve.

Theorem 5 *Let us define*

$$L_\epsilon(\phi) = \int_\Omega |\nabla H_\epsilon(\phi)|dx = \int_\Omega \delta_\epsilon(\phi)|\nabla\phi|dx.$$

Then we have

$$\lim_{\epsilon\to0} L_\epsilon(\phi) = \int_{\{\phi=0\}} ds = |K|,$$

where $K = \{x \in \Omega : \phi(x) = 0\}$.

Proof Using co-area formula [41], we have:

$$L_\epsilon(\phi) = \int_\mathbb{R}\left[\int_{\phi=\rho}\delta_\epsilon(\phi(x))ds\right]d\rho = \int_\mathbb{R}\left[\delta_\epsilon(\rho)\int_{\phi=\rho}ds\right]d\rho.$$

By setting $h(\rho) = \int_{\phi=\rho}ds$, we obtain

$$L_\epsilon(\phi) = \int_\mathbb{R}\delta_\epsilon(\rho)h(\rho)d\rho = \int_\mathbb{R}\frac{1}{\pi}\frac{\epsilon}{\epsilon^2+\rho^2}h(\rho)d\rho.$$

By the change of variable $\theta = \frac{\rho}{\epsilon}$, we obtain

$$\lim_{\epsilon\to0}L_\epsilon(\phi) = \lim_{\epsilon\to0}\int_R\frac{1}{\pi}\frac{\epsilon^2}{\epsilon^2+\epsilon^2\theta^2}h(\theta\epsilon)d\theta = \lim_{\epsilon\to0}\int_\mathbb{R}\frac{1}{\pi}\frac{1}{1+\theta^2}h(\theta\epsilon)d\theta$$

$$= h(0)\int_\mathbb{R}\frac{1}{\pi}\frac{1}{1+\theta^2}d\theta = h(0)\frac{1}{\pi}\arctan\theta|_{-\infty}^{+\infty} = h(0) = \int_{\phi=0}ds = |K|,$$

which concludes the proof. ∎

In general, this convergence result is valid for any approximations H_ϵ, δ_ϵ, under the assumptions

$$\lim_{\epsilon \to 0} H_\epsilon(z) = H(z) \text{ in } \mathbb{R}\backslash\{0\},$$

$\delta_\epsilon = H'_\epsilon$, $H_\epsilon \in C^1(\mathbb{R})$, $\int_{-\infty}^{+\infty} \delta_1(x)dx = 1$.

25.4.2.1 Piecewise-Constant Mumford and Shah Segmentation Using Level Sets

Our first formulation is for the case when the unknown set of edges K can be represented by $K = \{x \in \Omega : \phi(x) = 0\}$ for some (unknown) Lipschitz function $\phi : \Omega \to \mathbb{R}$. In this case we restrict the unknown minimizers u to functions taking two unknown values c_1, c_2, and the corresponding Mumford–Shah minimization problem can be expressed as [28, 30]

$$\inf_{c_1, c_2, \phi} E(c_1, c_2, \phi) = \int_\Omega (g(x) - c_1)^2 H(\phi)dx + \int_\Omega (g(x) - c_2)^2 H(-\phi)dx$$
$$+ \nu_0 \int_\Omega |DH(\phi)|. \tag{25.12}$$

The unknown minimizer u is expressed as

$$u(x) = c_1 H(\phi(x)) + c_2(1 - H(\phi(x))) = c_1 H(\phi(x)) + c_2 H(-\phi(x)).$$

We substitute H by its C^∞ approximation H_ϵ and we minimize instead

$$E_\epsilon(c_1, c_2, \phi) = \int_\Omega (g(x) - c_1)^2 H_\epsilon(\phi)dx + \int_\Omega (g(x) - c_2)^2 H_\epsilon(-\phi)dx$$
$$+ \nu_0 \int_\Omega |\nabla H_\epsilon(\phi)|dx. \tag{25.13}$$

The associated Euler–Lagrange equations with respect to c_1, c_2, and ϕ are

$$c_1(\phi) = \frac{\int_\Omega g(x) H_\epsilon(\phi(x))dx}{\int_\Omega H_\epsilon(\phi(x))dx}, \quad c_2(\phi) = \frac{\int_\Omega g(x) H_\epsilon(-\phi(x))dx}{\int_\Omega H_\epsilon(-\phi(x))dx},$$

and, after simplifications,

$$\delta_\epsilon(\phi)\left[(g(x) - c_1)^2 - (g(x) - c_2)^2 - \text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right] = 0 \text{ in } \Omega, \tag{25.14}$$

with boundary conditions $\nabla\phi \cdot \vec{N} = 0$ on $\partial\Omega$. Since $\delta_\epsilon > 0$ as defined, the factor $\delta_\epsilon(\phi)$ can be removed from (❷ 25.14), or substituted by $|\nabla\phi|$ to obtain a more geometric motion extended to all level lines of ϕ, as in the standard level set approach.

This approach has been generalized by Chung and Vese in [34, 35], where more than one level line of the same level set function ϕ can be used to represent the edge set K.

Using m distinct real levels $\{l_1 < l_2 < \ldots < l_m\}$, the function ϕ partitions the domain Ω into the following $m+1$ disjoint open regions, making up Ω, together with their boundaries:

$$\Omega_0 = \{x \in \Omega : -\infty < \phi(x) < l_1\},$$
$$\Omega_j = \{x \in \Omega : l_j < \phi(x) < l_{j+1}\}, \qquad 1 \le j \le m-1$$
$$\Omega_m = \{x \in \Omega : l_m < \phi(x) < +\infty\}.$$

The energy to minimize in this case, depending on $c_0, c_1, \ldots, c_m, \phi$, will be

$$E(c_0, c_1, \ldots, c_m, \phi) = \int_\Omega |g(x) - c_0|^2 H(l_1 - \phi(x))dx + \sum_{j=1}^{m-1} \int_\Omega |g(x)$$
$$- c_j|^2 H(\phi(x) - l_j) H(l_{j+1} - \phi(x))dx + \int_\Omega |g(x)$$
$$- c_m|^2 H(\phi(x) - l_m)dx + \nu_0 \sum_{j=1}^{m} \int_\Omega |DH(\phi - l_j)|. \quad (25.15)$$

The segmented image will be given by

$$u(x) = c_0 H(l_1 - \phi(x)) + \sum_{j=1}^{m-1} c_j H(\phi(x) - l_j) H(l_{j+1} - \phi(x)) + c_m H(\phi(x) - l_m).$$

As before, to minimize the above energy, we approximate and substitute the Heaviside function H by H_ϵ, as $\epsilon \to 0$. The Euler–Lagrange equations associated with the corresponding minimization

$$\inf_{c_0, c_1, \ldots, c_m, \phi} E_\epsilon(c_0, c_1, \ldots, c_m, \phi), \quad (25.16)$$

can be expressed as

$$\begin{cases} c_0(\phi) = \frac{\int_\Omega g(x) H_\epsilon(l_1 - \phi(t,x))dx}{\int_\Omega H_\epsilon(l_1 - \phi(t,x))dx}, \\ c_j(\phi) = \frac{\int_\Omega g(x) H_\epsilon(\phi(t,x) - l_j) H_\epsilon(l_{j+1} - \phi(t,x))dx}{\int_\Omega H_\epsilon(\phi(t,x) - l_j) H_\epsilon(l_{j+1} - \phi(t,x))dx}, \\ c_m(\phi) = \frac{\int_\Omega g(x) H_\epsilon(\phi(t,x) - l_m)dx}{\int_\Omega H_\epsilon(\phi(t,x) - l_m)dx}, \end{cases}$$

and

$$0 = |g - c_0|^2 \delta_\epsilon(l_1 - \phi) + \sum_{j=1}^{m-1} |g - c_j|^2 \left[\delta_\epsilon(l_{j+1} - \phi) H_\epsilon(\phi - l_j) - \delta_\epsilon(\phi - l_j) H_\epsilon(l_{j+1} - \phi) \right]$$

$$- |g - c_m|^2 \delta_\epsilon(\phi - l_m) + \nu_0 \sum_{j=1}^{m} \left[\delta_\epsilon(\phi - l_j) \operatorname{div}\left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right],$$

$$\left. \frac{\partial \phi}{\partial \vec{n}} \right|_{\partial \Omega} = 0, \quad (25.17)$$

where $\vec{N}$ is the exterior unit normal to the boundary $\partial \Omega$.

We give here the details of the numerical algorithm for solving (❯ 25.17) in two dimensions (x, y), using gradient descent, in the case of one function ϕ with two levels $l_1 = 0$,

$l_2 = l > 0$. Let $h = \triangle x = \triangle y$ be the space steps, $\triangle t$ be the time step, and $\epsilon = h$. Let (x_i, y_j) be the discrete points, for $1 \le i, j \le M$, and $g_{i,j} \approx g(x_i, y_j)$, $\phi_{i,j}^n \approx \phi(n \triangle t, x_i, y_j)$, with $n \ge 0$. Recall the usual finite differences formulas

$$\triangle_+^x \phi_{i,j} = \phi_{i+1,j} - \phi_{i,j}, \quad \triangle_-^x \phi_{i,j} = \phi_{i,j} - \phi_{i-1,j},$$

$$\triangle_+^y \phi_{i,j} = \phi_{i,j+1} - \phi_{i,j}, \quad \triangle_-^y \phi_{i,j} = \phi_{i,j} - \phi_{i,j-1}.$$

Set $n = 0$, and start with $\phi_{i,j}^0$ given (defining the initial set of curves). Then, for each $n > 0$ until steady state:

(1) compute averages $c_0(\phi^n)$, $c_1(\phi^n)$, and $c_2(\phi^n)$.
(2) compute $\phi_{i,j}^{n+1}$, derived from the finite differences scheme:

$$\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\triangle t} = \delta_\epsilon\left(\phi_{i,j}^n\right)\left[\frac{\nu_0}{h^2}\left(\triangle_-^x\left(\frac{\phi_{i+1,j}^n - \phi_{i,j}^{n+1}}{\left|\nabla\phi_{i,j}^n\right|}\right) + \triangle_-^y\left(\frac{\phi_{i,j+1}^n - \phi_{i,j}^{n+1}}{\left|\nabla\phi_{i,j}^n\right|}\right)\right) + |g_{i,j} - c_0|^2\right.$$

$$\left. - |g_{i,j} - c_1|^2 H_\epsilon\left(l - \phi_{i,j}^n\right)\right] + \delta_\epsilon\left(\phi_{i,j}^n - l\right)$$

$$\left[\frac{\nu_0}{h^2}\left(\triangle_-^x\left(\frac{\phi_{i+1,j}^n - \phi_{i,j}^{n+1}}{\left|\nabla\phi_{i,j}^n\right|}\right) + \triangle_-^y\left(\frac{\phi_{i,j+1}^n - \phi_{i,j}^{n+1}}{\left|\nabla\phi_{i,j}^n\right|}\right)\right)\right.$$

$$\left. - |g_{i,j} - c_2|^2 + |g_{i,j} - c_1|^2 H_\epsilon\left(\phi_{i,j}^n\right)\right],$$

where $\left|\nabla\phi_{i,j}^n\right| = \sqrt{\left(\frac{\phi_{i+1,j}^n - \phi_{i,j}^n}{h}\right)^2 + \left(\frac{\phi_{i,j+1}^n - \phi_{i,j}^n}{h}\right)^2}$. Let

$$C_1 = \frac{1}{\sqrt{\left(\frac{\phi_{i+1,j}^n - \phi_{i,j}^n}{h}\right)^2 + \left(\frac{\phi_{i,j+1}^n - \phi_{i,j}^n}{h}\right)^2}},$$

$$C_2 = \frac{1}{\sqrt{\left(\frac{\phi_{i,j}^n - \phi_{i-1,j}^n}{h}\right)^2 + \left(\frac{\phi_{i-1,j+1}^n - \phi_{i-1,j}^n}{h}\right)^2}},$$

$$C_3 = \frac{1}{\sqrt{\left(\frac{\phi_{i+1,j}^n - \phi_{i,j}^n}{h}\right)^2 + \left(\frac{\phi_{i,j+1}^n - \phi_{i,j}^n}{h}\right)^2}},$$

$$C_4 = \frac{1}{\sqrt{\left(\frac{\phi_{i+1,j-1}^n - \phi_{i,j-1}^n}{h}\right)^2 + \left(\frac{\phi_{i,j}^n - \phi_{i,j-1}^n}{h}\right)^2}}.$$

Let $m_1 = \frac{\triangle t}{h^2} \left(\delta_\epsilon \left(\phi^n_{i,j} \right) + \delta_\epsilon \left(\phi^n_{i,j} - l \right) \right) \nu_0$, $C = 1 + m_1 (C_1 + C_2 + C_3 + C_4)$. The main update equation for ϕ becomes

$$
\begin{aligned}
\phi^{n+1}_{i,j} = \frac{1}{C} \Big[&\phi^n_{i,j} + m_1 \left(C_1 \phi^n_{i+1,j} + C_2 \phi^n_{i-1,j} + C_3 \phi^n_{i,j+1} + C_4 \phi^n_{i,j-1} \right) \\
&+ \triangle t \delta_\epsilon \left(\phi^n_{i,j} \right) \left(-(g_{i,j} - c_1)^2 \left(1 - H_\epsilon \left(\phi^n_{i,j} - l \right) \right) \right. \\
&+ \left. (g_{i,j} - c_0)^2 \right) + \triangle t \delta_\epsilon \left(\phi^n_{i,j} - l \right) \left(-(g_{i,j} - c_2)^2 + (g_{i,j} - c_1)^2 H_\epsilon \left(\phi^n_{i,j} \right) \right) \Big],
\end{aligned}
$$

and repeat, until steady state is reached.

We conclude this section with several experimental results obtained using the models presented here, that act as denoising, segmentation, and active contours. In ❯ *Fig. 25-4* we show an experimental result taken from [30] obtained using the binary piecewise-constant model (❯ 25.12); we notice how interior contours can be automatically detected. In ❯ *Fig. 25-5*, we show an experimental result using the multilayer model (❯ 25.15), with $m = 2$ and two levels l_1, l_2, applied to the segmentation of a brain image.

The work in [35, 84] also shows how the previous Mumford–Shah level set approaches can be extended to piecewise-constant segmentation of images with triple junctions, several non-nested regions, or with other complex topologies, by using two or more level set functions that form a perfect partition of the domain Ω.
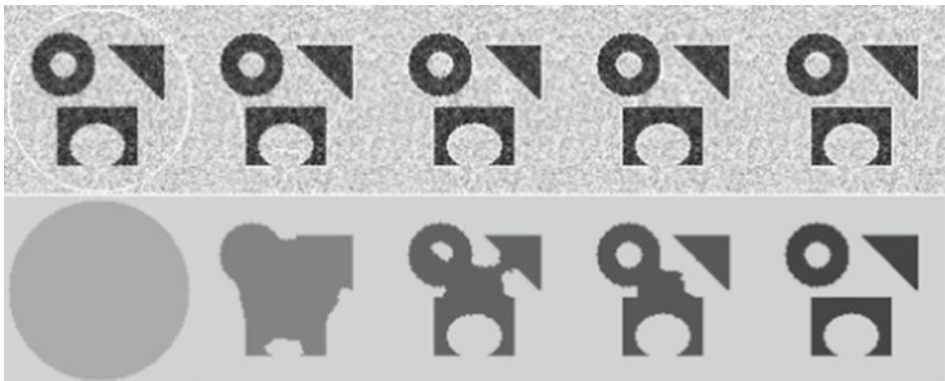
◨ **Fig. 25-4**

Detection of different objects in a noisy image, with various convexities and with an interior contour which is automatically detected, using only one initial curve. After a short time, an interior contour appears inside the torus, and then it expands. *Top*: g and the evolving contours. *Bottom*: the piecewise-constant approximations u of g over time, given by $u = c_1 H(\phi) + c_2 (1 - H(\phi))$
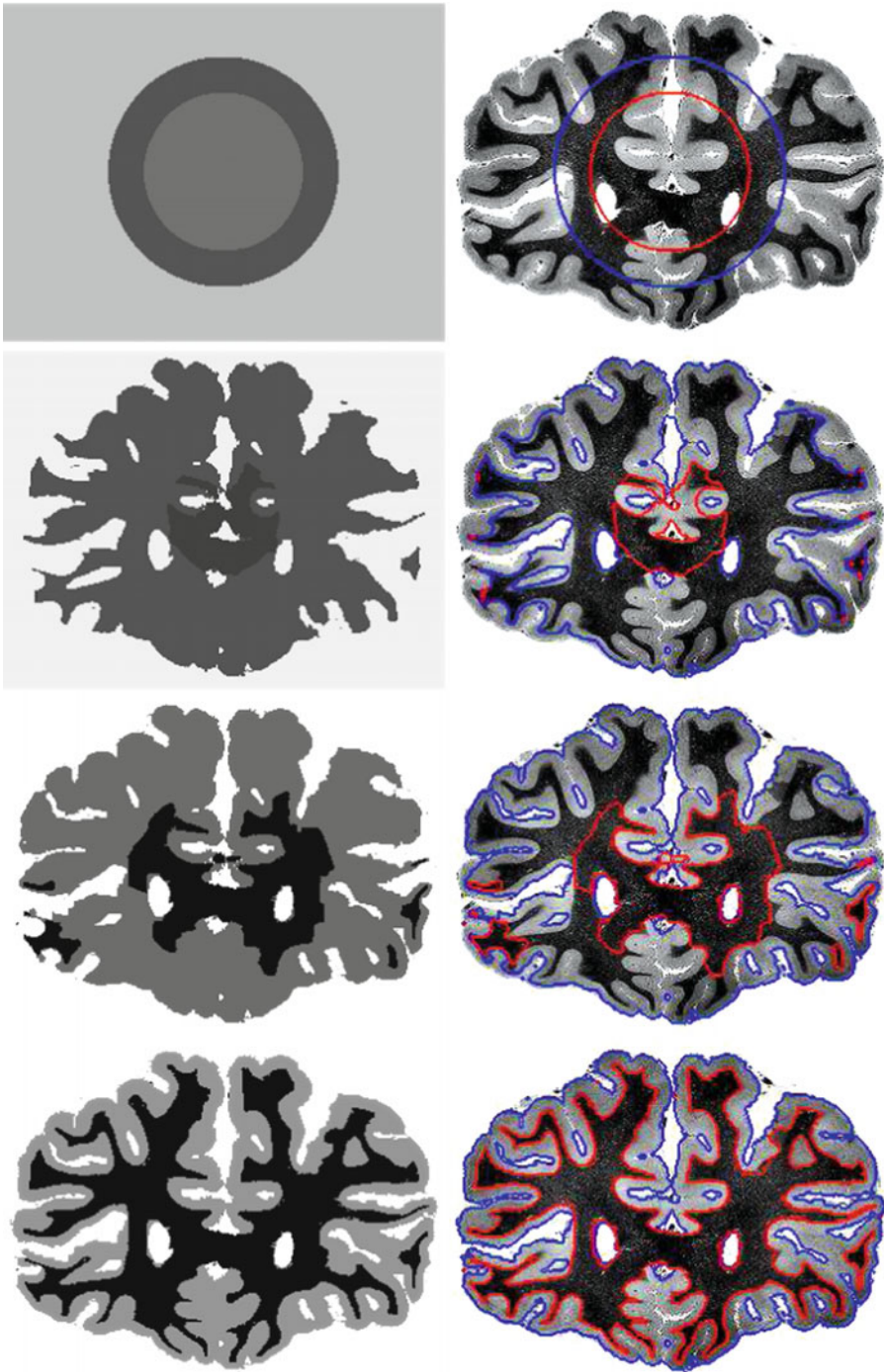
◼ **Fig. 25-5**
Segmentation of a brain image using one level set function with two levels. Parameters:
$l_1 = 0, l_2 = 25, \triangle t = 0.1, \nu_0 = 0.1 \cdot 255^2$, **1,500 iterations**

25.4.2.2   Piecewise-Smooth Mumford and Shah Segmentation Using Level Sets

We first consider the corresponding two-dimensional case under the assumption that the edges denoted by K in the image can be represented by one level set function ϕ, i.e., $K = \{x \in \Omega | \phi(x) = 0\}$, and we follow the approaches developed in parallel by Chan and Vese [31, 84] and by Tsai et al. [83], in order to minimize the general Mumford and Shah model. As in [84], the link between the unknowns u and ϕ can be expressed by introducing two functions u^+ and u^- (see ❯ *Fig. 25-6*) such that

$$u(x) = \begin{cases} u^+(x) \text{ if } \phi(x) \geq 0, \\ u^-(x) \text{ if } \phi(x) \leq 0. \end{cases}$$

We assume that u^+ and u^- are H^1 functions on $\phi \geq 0$ and on $\phi \leq 0$, respectively (with Sobolev traces up to all boundary points, i.e., up to the boundary $\{\phi = 0\}$). We can write the following minimization problem

$$\inf_{u^+, u^-, \phi} E(u^+, u^-, \phi),$$

where

$$E(u^+, u^-, \phi) = \mu^2 \int_\Omega |u^+ - g|^2 H(\phi) dx + \mu^2 \int_\Omega |u^- - g|^2 (1 - H(\phi)) dx$$
$$+ \int_\Omega |\nabla u^+|^2 H(\phi)) dx + \int_\Omega |\nabla u^-|^2 (1 - H(\phi)) dx + \nu \int_\Omega |DH(\phi)|$$

is the Mumford–Shah functional restricted to $u(x) = u^+(x)H(\phi(x)) + u^-(x)(1 - H(\phi(x))$.

Minimizing $E(u^+, u^-, \phi)$ with respect to u^+, u^-, and ϕ, we obtain the following Euler–Lagrange equations (embedded in a time-dependent dynamical scheme for ϕ):
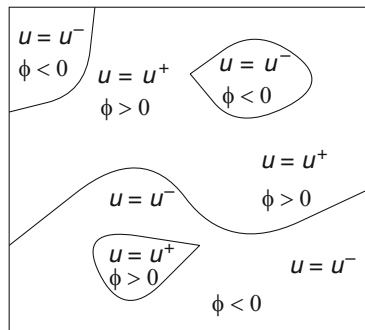

❒ **Fig. 25-6**
The functions u^+, u^- and the zero level lines of the level set function ϕ for piecewise-smooth image partition
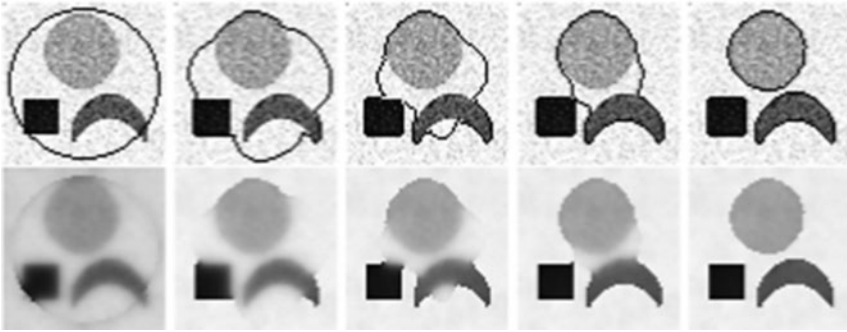
$$\mu^2(u^+ - g) = \triangle u^+ \text{ in } \{x : \phi(t, x) > 0\}, \quad \frac{\partial u^+}{\partial \vec{n}} = 0 \text{ on } \{x : \phi(t, x) = 0\} \cup \partial\Omega, \quad (25.18)$$

$$\mu^2(u^- - g) = \triangle u^- \text{ in } \{x : \phi(t, x) < 0\}, \quad \frac{\partial u^-}{\partial \vec{n}} = 0 \text{ on } \{x : \phi(t, x) = 0\} \cup \partial\Omega, \quad (25.19)$$

$$\frac{\partial \phi}{\partial t} = \delta_\varepsilon(\phi)\left[\nu\nabla\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \mu^2|u^+ - g|^2 - |\nabla u^+|^2 + \mu^2|u^- - g|^2 + |\nabla u^-|^2\right], \quad (25.20)$$

where $\partial/\partial\vec{n}$ denotes the partial derivative in the normal direction $\vec{n}$ at the corresponding boundary. We also associate the boundary condition $\frac{\partial\phi}{\partial\vec{n}} = 0$ on $\partial\Omega$ to ❷ Eq. (25.20).

We show in ❷ *Figs. 25-7* and ❷ *25-8* experimental results taken from [84] obtained with the piecewise-smooth two-phase model.

◨ **Fig. 25-7**
Results on a noisy image, using the level set algorithm for the piecewise-smooth Mumford–Shah model with one level set function. The algorithm performs as active contours, denoising, and edge detection
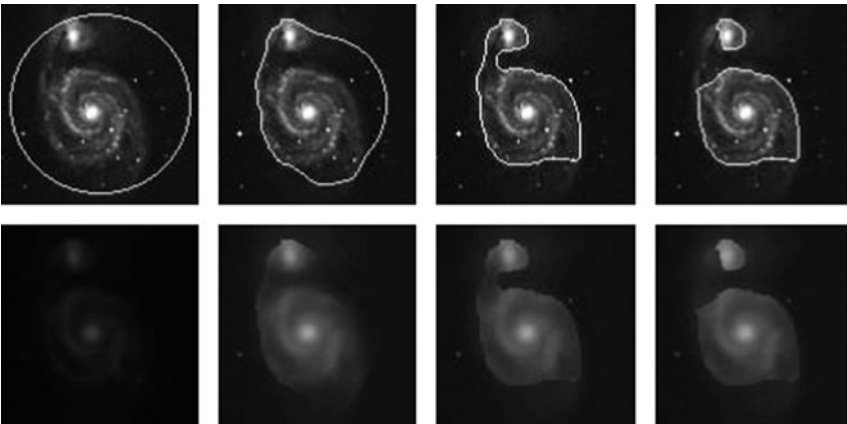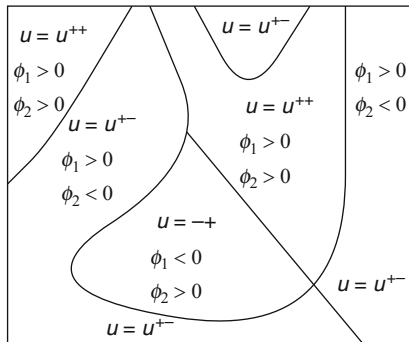

◨ **Fig. 25-8**
Numerical result using the piecewise-smooth Mumford–Shah level set algorithm with one level set function, on a piecewise-smooth real galaxy image

There are cases when the boundaries K of regions forming a partition of the image could not be represented by the boundary of an open domain. To overcome this, several solutions have been proposed in this framework and we mention two of them: (1) in the work by Tsai et al. [83], the minimization of $E(u^+, u^-, \phi)$ is repeated inside each of the two regions previously computed and (2) in the work of Chan and Vese [84], two or more level set functions are used. For example, in two dimensions, the problem can be solved using only two level set functions, and we do not have to know a priori how many gray levels the image has (or how many segments). The idea is based on the Four-Color Theorem. Based on this observation, we can "color" all the regions in a partition using only four "colors," such that any two adjacent regions have different "colors." Therefore, using two level set functions, we can identify the four "colors" by the following (disjoint) sets: $\{\phi_1 > 0, \phi_2 > 0\}$, $\{\phi_1 < 0, \phi_2 < 0\}$, $\{\phi_1 < 0, \phi_2 > 0\}$, $\{\phi_1 < 0, \phi_2 < 0\}$. The boundaries of the regions forming the partition will be given by $\{\phi_1 = 0\} \cup \{\phi_2 = 0\}$, and this will be the set of curves K. Note that, in this particular multiphase formulation of the problem, we do not have the problems of "overlapping" or "vacuum" (i.e., the phases are disjoint, and their union is the entire domain Ω, by definition).

As before, the link between the function u and the four regions can be made by introducing four functions $u^{++}, u^{+-}, u^{-+}, u^{--}$, which are in fact the restrictions of u to each of the four phases, as follows (see ❯ *Fig. 25-9*):

$$u(x) = \begin{cases} u^{++}(x), & \text{if } \phi_1(x) > 0 \text{ and } \phi_2(x) > 0, \\ u^{+-}(x), & \text{if } \phi_1(x) > 0 \text{ and } \phi_2(x) < 0, \\ u^{-+}(x), & \text{if } \phi_1(x) < 0 \text{ and } \phi_2(x) > 0, \\ u^{--}(x), & \text{if } \phi_1(x) < 0 \text{ and } \phi_2(x) < 0. \end{cases}$$

◨ Fig. 25-9
The functions $u^{++}, u^{+-}, u^{-+}, u^{--}$, and the zero level lines of the level set functions ϕ_1, ϕ_2 for piecewise-smooth image partition

Again, using the Heaviside function, the relation between u, the four functions u^{++}, u^{+-}, u^{-+}, u^{--}, and the level set functions ϕ_1 and ϕ_2 can be expressed by:

$$u = u^{++}H(\phi_1)H(\phi_2) + u^{+-}H(\phi_1)(1 - H(\phi_2)) + u^{-+}(1 - H(\phi_1))H(\phi_2) + u^{--}$$
$$(1 - H(\phi_1))(1 - H(\phi_2)).$$

We then introduce an energy in level set formulation based on the Mumford–Shah functional:

$$
\begin{aligned}
E(u, \phi_1, \phi_2) = {} & \mu^2 \int_\Omega |u^{++} - g|^2 H(\phi_1)H(\phi_2)dx \\
& + \int_\Omega |\nabla u^{++}|^2 H(\phi_1)H(\phi_2)dx \\
& + \mu^2 \int_\Omega |u^{+-} - g|^2 H(\phi_1)(1 - H(\phi_2))dx \\
& + \int_\Omega |\nabla u^{+-}|^2 H(\phi_1)(1 - H(\phi_2))dx \\
& + \mu^2 \int_\Omega |u^{-+} - g|^2 (1 - H(\phi_1))H(\phi_2)dx \\
& + \int_\Omega |\nabla u^{-+}|^2 (1 - H(\phi_1))H(\phi_2)dx \\
& + \mu^2 \int_\Omega |u^{--} - g|^2 (1 - H(\phi_1))(1 - H(\phi_2))dx \\
& + \int_\Omega |\nabla u^{--}|^2 (1 - H(\phi_1))(1 - H(\phi_2))dx \\
& + \nu \int_\Omega |DH(\phi_1)| + \nu \int_\Omega |DH(\phi_2)|.
\end{aligned}
$$

Note that the expression $\int_\Omega |DH(\phi_1)| + \int_\Omega |DH(\phi_2)|$ is not exactly the length term of $K = \{x \in \Omega : \phi_1(x) = 0 \text{ and } \phi_2(x) = 0\}$, it is just an approximation and simplification. In practice, satisfactory results using the above formula are obtained, and the associated Euler–Lagrange equations are simplified.

We obtain the associated Euler–Lagrange equations as in the previous cases, embedded in a dynamic scheme, assuming $(t, x, y) \mapsto \phi_i(t, x, y)$: minimizing the energy with respect to the functions u^{++}, u^{+-}, u^{-+}, u^{--}, we have, for each fixed t:

$$\mu^2(u^{++} - g) = \triangle u^{++} \text{ in } \{\phi_1 > 0, \phi_2 > 0\}, \quad \frac{\partial u^{++}}{\partial \vec{n}} = 0 \text{ on } \{\phi_1 = 0, \phi_2 \geq 0\}, \{\phi_1 \geq 0, \phi_2 = 0\};$$

$$\mu^2(u^{+-} - g) = \triangle u^{+-} \text{ in } \{\phi_1 > 0, \phi_2 < 0\}, \quad \frac{\partial u^{+-}}{\partial \vec{n}} = 0 \text{ on } \{\phi_1 = 0, \phi_2 \leq 0\}, \{\phi_1 \geq 0, \phi_2 = 0\};$$

$$\mu^2(u^{-+} - g) = \triangle u^{-+} \text{ in } \{\phi_1 < 0, \phi_2 > 0\}, \quad \frac{\partial u^{-+}}{\partial \vec{n}} = 0 \text{ on } \{\phi_1 = 0, \phi_2 \geq 0\}, \{\phi_1 \leq 0, \phi_2 = 0\};$$

$$\mu^2(u^{--} - g) = \triangle u^{--} \text{ in } \{\phi_1 < 0, \phi_2 < 0\}, \quad \frac{\partial u^{--}}{\partial \vec{n}} = 0 \text{ on } \{\phi_1 = 0, \phi_2 \leq 0\}, \{\phi_1 \leq 0, \phi_2 = 0\}.$$

The Euler–Lagrange equations evolving ϕ_1 and ϕ_2, embedded in a dynamic scheme, are formally:

$$\frac{\partial \phi_1}{\partial t} = \delta_\varepsilon(\phi_1) \left[\nu \nabla \left(\frac{\nabla \phi_1}{|\nabla \phi_1|} \right) - \mu^2 |u^{++} - g|^2 H(\phi_2) - |\nabla u^{++}|^2 H(\phi_2) \right.$$
$$- \mu^2 |u^{+-} - g|^2 (1 - H(\phi_2)) - |\nabla u^{+-}|^2 (1 - H(\phi_2)) + \mu^2 |u^{-+}$$
$$\left. - g|^2 H(\phi_2) + |\nabla u^{-+}|^2 H(\phi_2) + \mu^2 |u^{--} - g|^2 (1 - H(\phi_2)) + |\nabla u^{--}|^2 (1 - H(\phi_2)) \right] = 0,$$

$$\frac{\partial \phi_2}{\partial t} = \delta_\varepsilon(\phi_2) \left[\nu \nabla \left(\frac{\nabla \phi_2}{|\nabla \phi_2|} \right) - \mu^2 |u^{++} - g|^2 H(\phi_1) - |\nabla u^{++}|^2 H(\phi_1) \right.$$
$$+ \mu^2 |u^{+-} - g|^2 H(\phi_1) + |\nabla u^{+-}|^2 H(\phi_1) - \mu^2 |u^{-+} - g|^2 (1 - H(\phi_1)) - |\nabla u^{-+}|^2 (1 - H(\phi_1))$$
$$\left. + \mu^2 |u^{--} - g|^2 (1 - H(\phi_1)) + |\nabla u^{--}|^2 (1 - H(\phi_1)) \right].$$

We can show, by standard techniques of the calculus of variations on the space $SBV(\Omega)$ (special functions of bounded variations), and a compactness result due to Ambrosio [3], that the proposed minimization problems from this section, in the level set formulation, have a minimizer. Finally, because there is no uniqueness of minimizers, and because the problems are nonconvex, the numerical results may depend on the initial choice of the curves, and we may compute a local minimum only. We think that, using the seed initialization (see [84]) the algorithms have the tendency of computing a global minimum, most of the times. Additional experimental results are shown in [84].

25.4.2.3 Extension to Level Set Based Mumford–Shah Segmentation with Open Edge Set *K*

We have mentioned in ❯ Sect. 25.2.2, Theorem 1 that in two dimensions, the Mumford–Shah functional E from (❯ 25.1) allows for minimizers (u, K) such that the set K could include open curves or crack tips where a curve K_i of K ends and meets nothing. On the other hand, the level set formulations presented in the previous sections allow only for closed curves as pieces of K, an inherent property due to the implicit representation of boundaries. In this section, we show how we can modify the level set representation of the Mumford–Shah functional, so that images with edges made of open curves could also be segmented. For more details we refer the reader to Mohieddine–Vese [60].

The main idea is to use the open curve representation using level sets due to Smereka [79]. In [79], by adding a "dual" level set function, a level set formulation for open curves extending the standard methods is proposed. Given a level set function $\phi : \Omega \rightarrow \mathbb{R}$ and a "dual" level set function $\psi : \Omega \rightarrow \mathbb{R}$ (as Lipschitz continuous functions), an open curve K can be defined as $K = \{x \in \Omega : \phi(x) = 0, \psi(x) > 0\}$. This is illustrated in ❯ *Fig. 25-10*.
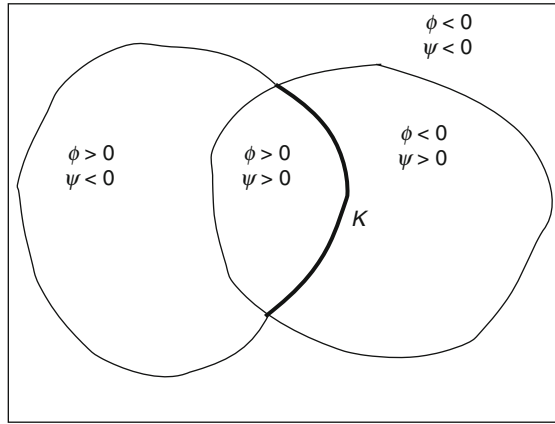
▫ Fig. 25-10
Representation of an open curve $K = \{\phi = 0\} \cap \{\psi > 0\}$

The method in [79] was applied to a curvature equation which models the dynamics of spiral crystal growth, with velocity

$$\mathbf{v}(t) = (1 - \lambda \kappa)\mathbf{n}, \qquad (25.21)$$

where κ is the curvature and $\mathbf{n}$ is the unit normal of the open spiral curve, and λ is a constant. After reformulating the equations with open level sets yields [79]:

$$\frac{\partial \phi}{\partial t} + sign(\psi)[1 - \lambda sign(\psi)\kappa(\phi)]|\nabla \phi| = 0,$$

$$\frac{\partial \psi}{\partial t} + sign(\phi)[1 - \lambda sign(\phi)\kappa(\psi)]|\nabla \psi| = 0.$$

In general, this method will give a symmetric system of the form

$$\frac{\partial \phi}{\partial t} + F(\phi, \psi) = 0, \qquad (25.22)$$

$$\frac{\partial \psi}{\partial t} + F(\psi, \phi) = 0 \qquad (25.23)$$

from where it is clear why these level set functions are called "dual" to each other. In the general form, ❷ Eqs. (25.22) and (❷ 25.23) may or may not be derived from functional minimization.

Here, we will use the idea of Smereka in the minimization of the Mumford–Shah model for segmentation with open edge curves. We first define the following characteristic functions over Ω: $\chi_1 = H(\phi)$, $\chi_2 = H(\psi)(1-H(\phi))$, $\chi_3 = H(\psi)$, and $\chi_0 = (1-H(\psi))(1-H(\phi))$.

Then we propose the following open curve formulation of Mumford–Shah, in a particular case: minimize

$$E(u_1, u_2, \phi, \psi) = \int_\Omega \left[|u_1 - g|^2 \chi_1 + |u_2 - g|^2 \chi_2 + \left| \frac{u_1 + u_2}{2} - g \right|^2 \chi_0 \right] dx$$

$$+ \mu \int_\Omega \left[|\nabla u_1|^2 \left(\chi_1 + \frac{\chi_0}{4} \right) + \mu |\nabla u_2|^2 \left(\chi_2 + \frac{\chi_0}{4} \right) + \frac{\mu}{2} \nabla u_1 \cdot \nabla u_2 \chi_0 \right]$$

$$dx + \lambda \int_\Omega \chi_3 |\nabla \chi_1|.$$

The segmented image will be $u = u_1 \chi_1 + u_2 \chi_2 + \frac{u_1 + u_2}{2} \chi_0$, the set $K = \{\phi = 0\} \cap \{\psi > 0\}$ models the open jump set, and the length of K is $|K| = \int_\Omega |\nabla H(\phi)| H(\psi) = \int_\Omega \chi_3 |\nabla \chi_1|$. In the above energy, the first term corresponds to the data fidelity, the second term corresponds to the regularization in u, while the third term is the length penalty. Thus, the functional imposes that $u \approx g$ over χ_1, χ_2, χ_0 (thus over Ω), and that u is of class H^1 over the regions whose characteristic functions are $\chi_1 + \chi_0$ and $\chi_2 + \chi_0$.

As in the previous sections, we first substitute the Heaviside function H by smooth approximations H_ϵ. Also, as in Theorem 5, it is possible to show that the approximating term $\int_\Omega H_\epsilon(\psi) |\nabla H_\epsilon(\phi)|$ converges, as $\epsilon \to 0$, to the length of the open set $K = \{x \in \Omega : \phi(x) = 0, \psi(x) > 0\}$.

The Euler–Lagrange equations associated with the minimization, expressed using the L^2 gradient descent, formally are

$$\frac{\partial u_1}{\partial t} = \mu \text{div} \left[2\chi_1 \nabla u_1 + \frac{\chi_0}{2} \nabla(u_1 + u_2) \right] - 2(u_1 - g)\chi_1 - \left(\frac{u_1 + u_2}{2} - g \right) \chi_0$$

$$\left[\chi_1 \nabla u_1 + \frac{\chi_0}{4} \nabla(u_1 + u_2) \right] \cdot \mathbf{n} = 0 \ on \ \partial\Omega$$

$$u_1(0, x) = u_{1-\text{initial}}(x),$$

$$\frac{\partial u_2}{\partial t} = \mu \text{div} \left[2\chi_2 \nabla u_2 + \frac{\chi_0}{2} \nabla(u_1 + u_2) \right] - 2(u_2 - g)\chi_2 - \left(\frac{u_1 + u_2}{2} - g \right) \chi_0$$

$$\left[\chi_2 \nabla u_2 + \frac{\chi_0}{4} \nabla(u_1 + u_2) \right] \cdot \mathbf{n} = 0 \ on \ \partial\Omega$$

$$u_2(0, x) = u_{2-\text{initial}}(x),$$

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\lambda \text{div} \left(\chi_3 \frac{\nabla \phi}{|\nabla \phi|} \right) - |u_1 - g|^2 + \chi_3 |u_2 - g|^2 + (1 - \chi_3) \right.$$

$$\left| \frac{u_1 + u_2}{2} - g \right|^2 - \mu |\nabla u_1|^2 \left(\frac{3}{4} + \frac{\chi_3}{4} \right)$$

$$\left. + \mu |\nabla u_2|^2 \left(\frac{1}{4} + \frac{3\chi_3}{4} \right) + \frac{1}{2} \mu \nabla u_1 \cdot \nabla u_2 (1 - \chi_3) \right]$$

$$\left(\chi_3 \frac{\nabla \phi}{|\nabla \phi|} \right) \cdot \mathbf{n} = 0 \ on \ \partial\Omega$$

$$\phi(0, x) = \phi_{\text{initial}}(x),$$

$$\frac{d\psi}{dt} = -\delta(\psi)\left[\lambda|\nabla\chi_1| + (1-\chi_1)\left(|u_2 - g|^2 - \left|\frac{u_1 + u_2}{2} - g\right|^2 - \frac{\mu}{4}|\nabla u_1|^2\right.\right.$$

$$\left.\left.+\frac{3\mu}{4}|\nabla u_2|^2 - \frac{\mu}{2}\nabla u_1 \cdot \nabla u_2\right)\right]$$

$$\psi(0, x) = \psi_{\text{initial}}(x).$$

Since we have doubled the amount of functions used to define one curve, we have also increased the computational cost. Moreover, from a theoretical point of view, the system of equations derived from the standard L^2 gradient decent may be ill posed. For illustration, following Neuberger [65] and Renka [71], assume that we have the energy functional with a potential F, i.e., $E(\phi) = \int_\Omega F(D\phi)$, to be minimized over $H^1(\Omega)$, where here $D : H^1(\Omega) \rightarrow H^1(\Omega) \times L^2(\Omega)$ is the operator $D\phi = (\phi, \nabla\phi)^T$. We assume that $\phi \in H^1(\Omega)$ and for any $h \in H_0^1(\Omega)$ we have the directional derivative:

$$(E'(\phi), h) = \int_\Omega F'(D\phi)Dh = \langle\nabla F(D\phi), Dh\rangle_{L^2} = \langle D^*\nabla F(D\phi), h\rangle_{L^2},$$

where D^* is the adjoint of D. We will call the first variation the L^2 gradient, $\nabla_{L^2}E(\phi) = D^*\nabla F(D\phi)$ and it defines the usual gradient descent method, $\frac{\partial\phi}{\partial t} = -\nabla_{L^2}E(\phi)$. In the semi-discrete case, we construct the sequence ϕ^n by $\phi^{n+1} = \phi^n - \triangle t\nabla_{L^2}E(\phi^n)$, with $\phi^0 \in H^1(\Omega)$, $\triangle t > 0$, such that $E(\phi^{n+1}) < E(\Phi^n)$. In order to have $\phi^{n+1} \in H^1(\Omega) \subset L^2(\Omega)$, this would require that $\nabla_{L^2}E(\phi^n) \in H^1(\Omega) \subset L^2(\Omega)$, in other words we would assume too strong regularity for the solution ϕ, which may not hold. This is one of the reasons for the small time steps necessary for stability when using L^2 gradient decent. Thus, the combination of small time steps and increased amount of functions to represent open curves can become problematic in practice. To avoid these issues, we derive an alternative decent direction which is better posed. The next simplest direction to the L^2 gradient is the Sobolev H^1 gradient direction. Denote the H^1 gradient as $\nabla_{H^1}E(\phi)$ and as before we will look at the directional derivative. Equating the directional derivative with the H^1 inner product yields the Sobolev gradient as follows:

$$(E'(\phi), h) = \langle\nabla_{L^2}E(\phi), h\rangle_{L^2} = \langle\nabla_{H^1}E(\phi)), h\rangle_{H^1}.$$

So we have

$$\langle\nabla_{H^1}E(\phi), h\rangle_{H^1} = \langle D(\nabla_{H^1}E(\phi)), Dh\rangle_{L^2} = \langle D^*D(\nabla_{H^1}E(\phi)), h\rangle_{L^2}$$

and therefore $\nabla_{H^1}E(\phi) = (D^*D)^{-1}(\nabla_{L^2}E(\phi)) = (I - \Delta)^{-1}(\nabla_{L^2}E(\phi))$. One way to look at this is applying gradient decent with respect to a different inner product. Numerically, it can be viewed as a preconditioning of the regular gradient decent method [72]. This will also have numerical benefits. For more details on the theory of Sobolev gradients, see [65]. Here, the Sobolev H^1 gradient is used for all four equations in u_1, u_2, ϕ, ψ.

We present a few experimental results for the segmentation of a simple synthetic image with noise. In ❯ *Fig. 25-11* we show a synthetic noisy image, the evolution of the unknown open curve K over iterations, and the denoised image u over iterations. ❯ *Figure 25-12* shows the surface plot of the unknown u during the iterative procedure. The numerical
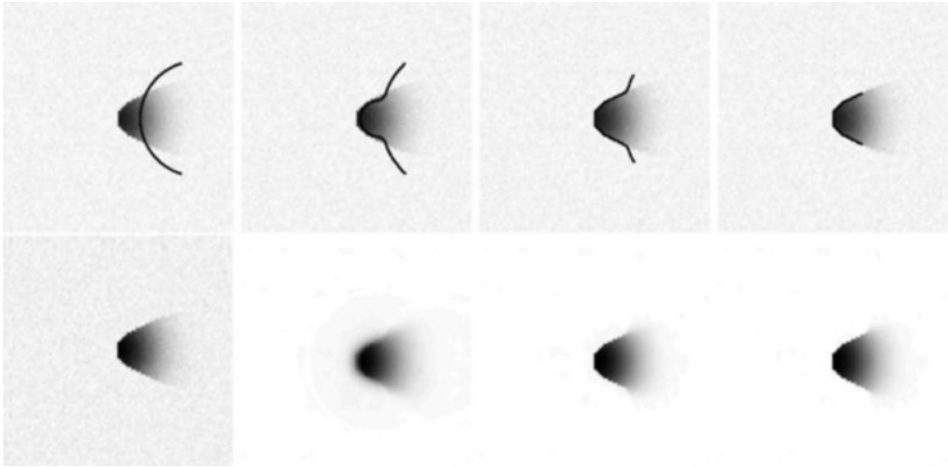
◘ **Fig. 25-11**

Segmentation of a synthetic noisy image with open curve discontinuity. *Top*, **from** *left* **to** *right*: **evolution of the unknown open curve** *K* **with iterations, superimposed over the noisy data** *g*. *Bottom*, **from** *left* **to** *right*: **the initial noisy image** *g* **and the restored image** *u* **over iterations**
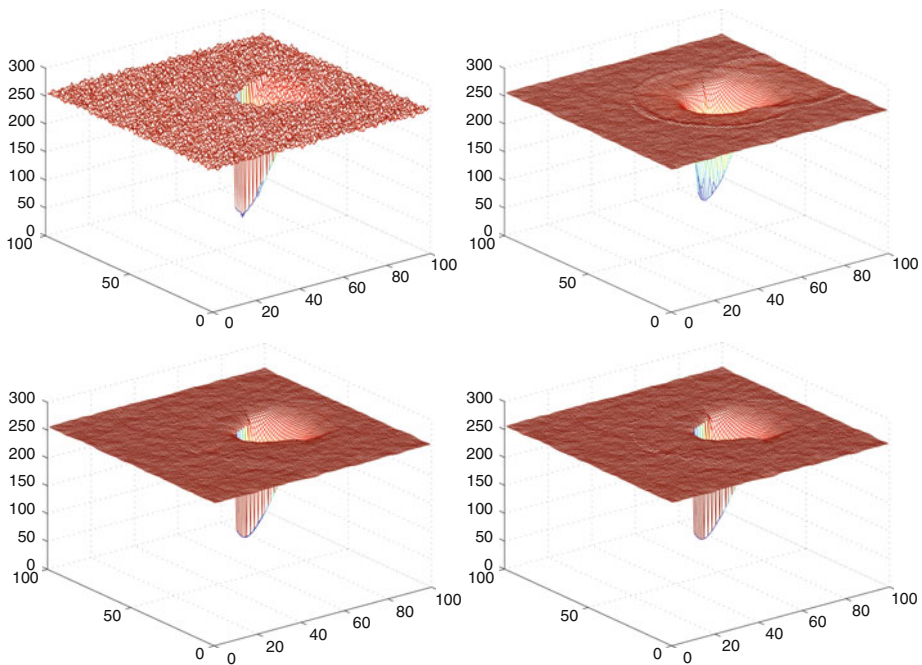
◘ **Fig. 25-12**

The surface plot of the image *u* **in** ◗ *Fig. 25-11* **over iterations**
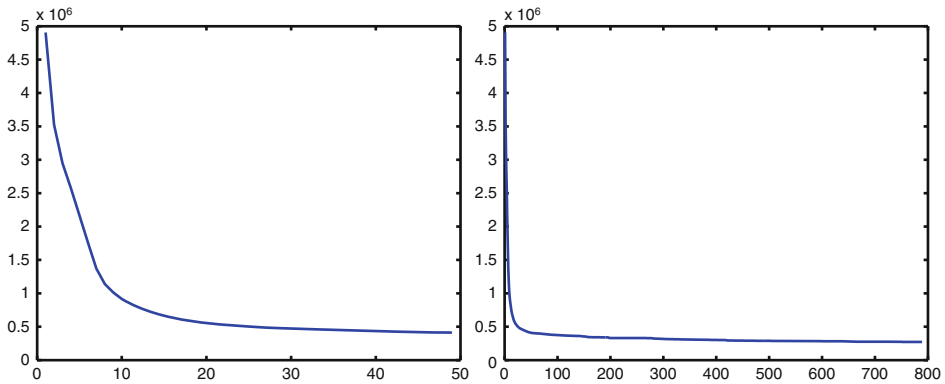
□ **Fig. 25-13**

Numerical energy versus iterations (*left*, first 50 iterations; *right*, first 800 iterations)

energy versus iterations is presented in ❯ *Fig. 25-13*, showing that the proposed numerical algorithm [60] is stable in practice. The boundary conditions for u_1 and u_2 can be simplified.

25.5   Case Examples: Variational Image Restoration with Segmentation-Based Regularization

This section focuses on the challenging task of edge-preserving variational image restoration. In this context, restoration is referred to as image deblurring and denoising, where we deal with Gaussian and impulsive noise models. Terms from the Mumford–Shah segmentation functional are used as regularizers, reflecting the model of piecewise-constant or piecewise-smooth images.

In the standard model of degradation the underlying assumptions are the linearity and shift invariance of the blur process and the additivity and normal distribution of the noise. Formally, let Ω be an open-bounded subset of $\mathbb{R}^n$. The observed image $g : \Omega \to \mathbb{R}^N \in L^\infty$ is given by

$$g = h * u + n, \tag{25.24}$$

where g is normalized to the hypercube $[0,1]^N$, h is the blur kernel such that $h(x) > 0$ and $\int h(x)dx = 1$, $u : \Omega \to \mathbb{R}^N$ is the ("ideal") original image, $n \sim N(0, \sigma^2)$ stands for a white Gaussian noise, and $*$ denotes the convolution operator. The restoration problem is the recovery of the original image u given ❯ Eq. (25.24). Non-blind image restoration is the problem whenever the blur kernel is known, while blind restoration refers to the case of unknown kernel [50, 51]. The recovery process in the non-blind case is a typical inverse problem where the image u is the minimizer of an objective functional of the form

$$\mathcal{F}(u) = \Phi(g - h * u) + \mathcal{J}(\nabla u). \tag{25.25}$$

The functional consists of fidelity term and a regularizer. The fidelity term Φ forces the smoothed image $h * u$ to be close to the observed image g. The commonly used model of a white Gaussian noise $n \sim N(0, \sigma^2)$ leads by the maximum likelihood estimation to the minimization of the L^2 norm of the noise

$$\Phi_{L^2} = \|g - h * u\|_{L^2(\Omega)}^2. \tag{25.26}$$

However, in the case of impulsive noise, some amount of pixels do not obey the Gaussian noise model. Minimization of outlier effects can be accomplished by replacing the quadratic form (❯ 25.26) with a robust ρ-function [45], e.g.,

$$\Phi_{L^1} = \|g - h * u\|_{L^1(\Omega)}. \tag{25.27}$$

The minimization of (❯ 25.26) or (❯ 25.27) alone with respect to u is an inverse problem which is known to be ill posed: small perturbations in the data g may produce unbounded variations in the solution. To alleviate this problem, a regularization term can be added. The Tikhonov L^2 stabilizer [82]

$$\mathcal{J}_{L^2} = \int_\Omega |\nabla u|^2 dx,$$

leads to over smoothing and loss of important edge information. A better edge preservation regularizer, the Total Variation (TV) term, was introduced by Rudin et al. [73, 74], where the L^2 norm was replaced by the L^1 norm of the image gradients

$$\mathcal{J}_{L^1} = \int_\Omega |\nabla u| dx.$$

Still, although the Total Variation regularization outperforms the L^2 norm, the image features – the edges, are not explicitly extracted. The edges are implicitly preserved only by the image gradients.

An alternative regularizer is the one used in the Mumford–Shah functional [62, 64]. We recall that this is accomplished by searching for a pair (u, K) where $K \subset \Omega$ denotes the set of discontinuities of u, the unknown image, such that $u \in H^1(\Omega \backslash K)$, $K \subset \Omega$ closed in Ω, and

$$G(u, K) = \beta \int_{\Omega \backslash K} |\nabla u|^2 \, dx + \alpha \mathcal{H}^{n-1}(K) < \infty. \tag{25.28}$$

In our study, the regularizer to the restoration problem (❯ 25.25) is given by

$$\mathcal{J}_{MS} = G(u, K),$$

its L^1 variant [2, 78], and elliptic or level set approximations of these, as presented next. This enables the explicit extraction and preservation of the image edges in the course of the restoration process. We show the advantages of this regularizer in several applications and noise models (Gaussian and impulsive).

As we have mentioned, Ambrosio and Tortorelli [6] introduced an elliptic approximation $G_\epsilon(u, v)$ to $G(u, K)$, as $\epsilon \to 0^+$, that we recall here,

$$G_\epsilon(u, v) = \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 \ + \ \frac{(v-1)^2}{4\epsilon} \right) dx. \tag{25.29}$$

Replacing the Mumford–Shah regularization term (❯ 25.28) by $G_\epsilon(u, v)$ yields the proposed restoration model

$$\mathcal{F}_\epsilon(u, v) = \Phi(g - h * u) + \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx. \quad (25.30)$$

The functional (❯ 25.30) can also be understood from a generalized robust statistics viewpoint. This is beyond the scope of this chapter and the interested reader can find the details in [12].

In the rest of the chapter we consider the non-blind restoration problem presented in [13] and its generalizations to several more realistic situations. We consider the problem of (semi) blind deconvolution, the case of impulsive noise, the color restoration problem and the case of space-variant blur. We also consider the problem of restoration of piecewise-constant images from noisy-blurry data using the level set form of the Mumford–Shah regularizer and image restoration using nonlocal Mumford–Shah–Ambrosio–Tortorelli regularizers.

25.5.1 Non-blind Restoration

We first address the restoration problem with a known blur kernel h and additive Gaussian noise [10, 13]. In this case the fidelity term is the L^2 norm of the noise (❯ 25.26), and the regularizer $\mathcal{J}_{MS} = G_\epsilon(u, v)$ (❯ 25.29). The objective functional is therefore

$$\mathcal{F}_\epsilon(u, v) = \frac{1}{2} \int_\Omega (g - h * u)^2 dx + \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx.$$

$$(25.31)$$

The functional (❯ 25.31) is strictly convex, bounded from below and coercive with respect to the functions u and v if the other one is fixed. Following [33], the alternate minimization (AM) approach is applied: in each step of the iterative procedure we minimize with respect to one function and keep the other one fixed. The minimization is carried out using the Euler–Lagrange (E-L) equations with Neumann boundary conditions where u is initialized as the blurred image g and v is initialized to 1.

$$\frac{\delta \mathcal{F}_\epsilon}{\delta v} = 2\beta v |\nabla u|^2 + \alpha \frac{v-1}{2\epsilon} - 2\epsilon\alpha \triangle v = 0 \quad (25.32)$$

$$\frac{\delta \mathcal{F}_\epsilon}{\delta u} = (h * u - g) * h(-x, -y) - 2\beta \nabla \cdot (v^2 \nabla u) = 0 \quad (25.33)$$

❯ Equation (25.32) is linear with respect to v and can be easily solved after discretization by the Minimal Residual algorithm [86]. The integro-differential equation (❯ 25.33) can be solved by the conjugate-gradients method [13]. The iterative process is stopped whenever some convergence criterion is satisfied (e.g., $\|u^{n+1} - u^n\| < \varepsilon \|u^n\|$). ❯ *Figure 25-14* demonstrates the outcome of the algorithm. The top-left image is the blurred image g. The kernel corresponds to horizontal motion blur. The top-right image is the reconstruction obtained
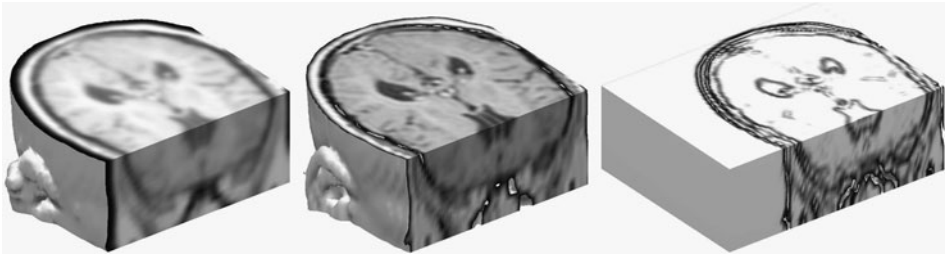
◨ Fig. 25-14
The case of a known (nine-pixel horizontal motion) blur kernel. *Top-left*: **corrupted image.**
Top-right: **restoration using the TV method** [74, 85]. *Bottom-left*: **restoration using the MS**
method. *Bottom-right*: **Edge map produced by the MS method**

using Total Variation (TV) regularization [74, 85]. The bottom-left image is the outcome
of the MS regularizer, with a known blur kernel. The bottom-right image shows the associ-
ated edge map v determined by the algorithm. Acceptable restoration is obtained with both
methods. Nevertheless, the MS method yields a sharper result and is almost free of "ghosts"
(white replications of notes) that can be seen in the top-right image (e.g., between the C
notes in the right part of the top stave). The algorithm can be also applied to 3D images as
shown in ❯ *Fig. 25-15*. In this example the blur kernel was anisotropic 3D Gaussian kernel.

25.5.2   Semi-Blind Restoration

Blind restoration refers to the case when the blur kernel h is not known in advance. In
addition to being ill posed with respect to the image, the blind restoration problem is ill
posed in the kernel as well. Blind image restoration with joint recovery of the image and

◾ Fig. 25-15
3D restoration of MRI image. *Left*: **blurred (**σ_x **= 1.0,** σ_y **= 1.0,** σ_z **= 0.2) image.** *Middle*: **recovered image.** *Right*: **edge map**

the kernel, and regularization of both, was presented by You and Kaveh [87], followed by Chan and Wong [33]. Chan and Wong suggested to minimize a functional consisting of a fidelity term and Total Variation (L^1 norm) regularization for both the image and the kernel:

$$\mathcal{F}(u,h) \ = \ \frac{1}{2}\|h * u - g\|^2_{L^2(\Omega)} + \alpha_1 \int_\Omega |\nabla u| dx + \alpha_2 \int_\Omega |\nabla h| dx. \tag{25.34}$$

By this approach the recovered kernel is highly dependent on the image characteristics. It allows the distribution of edge directions in the image to have an influence on the shape of the recovered kernel which may lead to inaccurate restoration [13]. Facing the ill-posedness of blind restoration with a general kernel, two approaches can be taken. One is to add relevant data; the other is to constrain the solution. In many practical situations, the blurring kernel can be modeled by the physics/optics of the imaging device and the setup. The blurring kernel can then be constrained and described as a member in a class of parametric functions. The blind restoration problem is then reduced to a semi-blind one. Let us consider the case of isotropic Gaussian blur parameterized by the width σ,

$$h_\sigma(x) = \frac{1}{2\pi\sigma^2} \, e^{-\frac{x^2}{2\sigma^2}}, \quad x \in \mathbb{R}^n.$$

The semi-blind objective functional then takes the form [13]

$$\mathcal{F}_\epsilon(v,u,\sigma) = \frac{1}{2} \int_\Omega (h_\sigma * u - g)^2 dx \ + \ G_\epsilon(u,v) + \gamma \int_\Omega |\nabla h_\sigma|^2 dx. \tag{25.35}$$

The last term in ❷ Eq. (25.35) stands for the regularization of the kernel, necessary to resolve the fundamental ambiguity in the division of the apparent blur between the recovered image and the blur kernel. This means that we prefer to reject the hypothesis that the blur originates from u, and assume that it is due to the convolution with the blur kernel. From the range of possible kernels, we thus select a wide one. This preference is represented by the kernel smoothness term: the width of the Gaussian corresponds to its smoothness, measured by the L^2 norm of its gradient. The optimization is carried out by using the alternate minimization approach. The recovered image u is initialized with g, the edge indicator

function v is initialized with 1, and σ with a small number ε which reflects a delta function kernel. The Euler–Lagrange equations with respect to v and u are given by (❯ 25.32) and (❯ 25.33) respectively. The parameter σ is the solution of

$$\frac{\partial \mathcal{F}_\epsilon}{\partial \sigma} = \int_\Omega \left[(h_\sigma * u - g)\left(\frac{\partial h_\sigma}{\partial \sigma} * u\right) + \gamma \frac{\partial}{\partial \sigma}|\nabla h_\sigma|^2 \right] dx = 0, \qquad (25.36)$$

which can be calculated by the bisection method. The functional (❯ 25.35) is not generally convex. Nevertheless, in practical numerical simulations the algorithm converges to visually appealing restoration results as can be seen in the second row of ❯ *Fig. 25-16*.

◨ **Fig. 25-16**
Semi-blind restoration. *Top row*: blurred images. *Second row*: restoration using the semi-blind method. *Third row*: original images. *Bottom row*: edge maps produced by the semi-blind method

25.5.3   Image Restoration with Impulsive Noise

Consider an image that has been blurred with a known blur kernel h and contaminated by impulsive noise. Salt and pepper noise, for instance, is a common model for the effects of bit errors in transmission, malfunctioning pixels, and faulty memory locations. Image deblurring algorithms that were designed for Gaussian noise produce inadequate results with impulsive noise.

The left image of ❷ *Fig. 25-17* is a blurred image contaminated by salt-and-pepper noise, and the right image is the outcome of the Total Variation restoration method [85]. A straight forward sequential approach is to first denoise the image, then to deblur it. This two-stage method is however prone to failure, especially at high noise density. Image denoising using median-type filtering creates distortion that depends on the neighborhood size, this error can be strongly amplified by the deblurring process. This is illustrated in ❷ *Fig. 25-18*. The top-left and top-right images are the blurred and blurred-noisy images, respectively. The outcome of 3×3 median filtering followed by Total Variation deblurring [85] is shown bottom left. At this noise level, the 3×3 neighborhood size of the median filter is insufficient, the noise is not entirely removed, and the residual noise is greatly amplified by the deblurring process. If the neighborhood size of the median filter increases to 5×5, the noise is fully removed, but the distortion leads to inadequate deblurring (bottom right).

In a unified variational framework, the "ideal" image u can be approximated as the minimizer of the objective functional [12, 14]
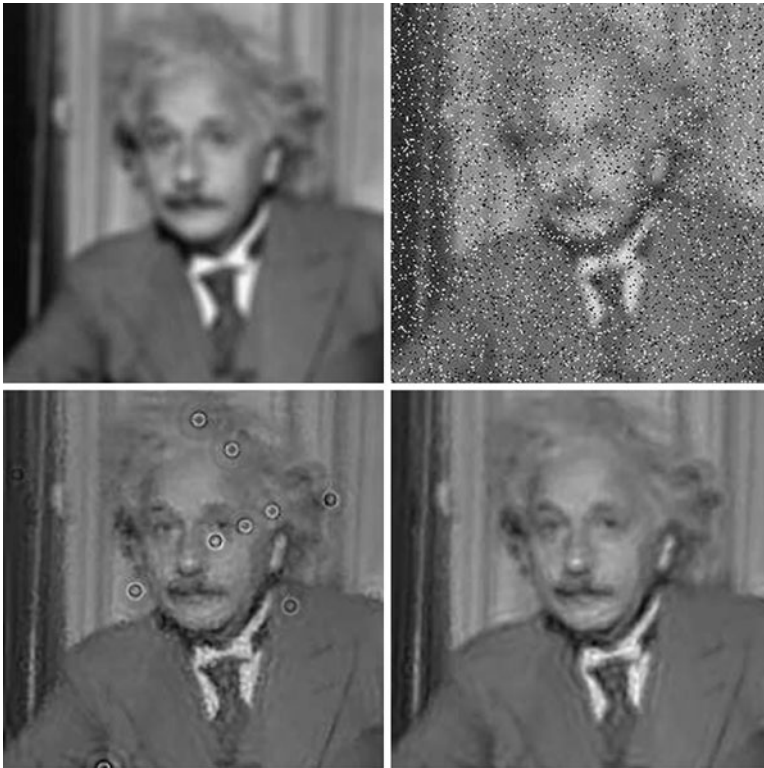
$$\mathcal{F}_\epsilon(u, v) = \int_\Omega \sqrt{(h * u - g)^2 + \eta} \, dx \; + \; G_\epsilon(u, v). \tag{25.37}$$

◼ **Fig. 25-17**
Current image deblurring algorithms fail in the presence of salt and pepper noise. *Left*: **blurred image with salt-and-pepper noise.** *Right*: **restoration using the TV method [85]**

▣ Fig. 25-18

The failure of the two-stage approach to salt-and-pepper noise removal and image deblurring. *Top-left*: **blurred image.** *Top-right*: **blurred image contaminated by salt-and-pepper noise.** *Bottom-left*: **the outcome of 3 × 3 median filtering, followed by deblurring.** *Bottom-right*: **the outcome of 5 × 5 median filtering, followed by deblurring**

The quadratic data-fidelity term is now replaced by the modified L^1 norm [66] which is robust to outliers, i.e., to impulse noise. The parameter $\eta \ll 1$ enforces the differentiability of (❱ 25.37) with respect to u. Optimization of the functional is carried out using the Euler–Lagrange equations subject to Neumann boundary conditions:

$$\frac{\delta \mathcal{F}_\epsilon}{\delta v} = 2\beta v |\nabla u|^2 + \alpha \left(\frac{v-1}{2\epsilon} \right) - 2\epsilon\alpha \triangle v = 0, \qquad (25.38)$$

$$\frac{\delta \mathcal{F}_\epsilon}{\delta u} = \frac{(h * u - g)}{\sqrt{(h * u - g)^2 + \eta}} * h(-x, -y) - 2\beta \nabla \cdot (v^2 \nabla u) = 0. \qquad (25.39)$$

The alternate minimization technique can be applied here as well since the functional (❱ 25.37) is convex, bounded from below and coercive with respect to either function u or v if the other one is fixed. Equation (❱ 25.38) is obviously linear with

respect to v. In contrast, (❯ 25.39) is a nonlinear integro-differential equation. Linearization of this equation is carried out using the fixed point iteration scheme as in [33, 85]. In this method, additional iteration index l serves as intermediate stage calculating u^{n+1}. We set $u = u^l$ in the denominator, and $u = u^{l+1}$ elsewhere, where l is the current iteration number. Equation (25.39) can thus be rewritten as

$$\mathcal{H}(v, u^l) u^{l+1} = G(u^l), \quad l = 0, 1, \ldots. \tag{25.40}$$

where $\mathcal{H}$ is the linear integro-differential operator

$$\mathcal{H}(v, u^l) u^{l+1} = \frac{h * u^{l+1}}{\sqrt{(h * u^l - g)^2 + \eta}} * h(-x, -y) - 2\beta \nabla \cdot (v^2 \nabla u^{l+1})$$

and

$$G(u^l) = \frac{g}{\sqrt{(h * u^l - g)^2 + \eta}} * h(-x, -y). \tag{25.41}$$

Note that (❯ 25.40) is now a *linear* integro-differential equation in u^{l+1}.

The discretization of ❯ Eqs. (25.38) and (❯ 25.40) yields two systems of linear algebraic equations. These systems are solved in alternation, leading to the following iterative algorithm [12]:

Initialization: $u^0 = g$, $v^0 = 1$.

1. Solve for v^{n+1}

$$\left(2\beta |\nabla u^n|^2 + \frac{\alpha}{2\epsilon} - 2\alpha \epsilon \triangle \right) v^{n+1} = \frac{\alpha}{2\epsilon}. \tag{25.42}$$

2. Set $u^{n+1,0} = u^n$ and solve for u^{n+1} (iterating on l)

$$\mathcal{H}(v^{n+1}, u^{n+1,l}) u^{n+1,l+1} = G(u^{n+1,l}). \tag{25.43}$$

3. If $\left(\|u^{n+1} - u^n\|_{L_2} < \varepsilon_1 \|u^n\|_{L_2} \right)$ stop.

The convergence of the algorithm was proved in [14]. ❯ *Figure 25-19* demonstrates the performance of the algorithm. The top row shows the blurred images with increasing salt-and-pepper noise level. The outcome of the restoration algorithm is shown in the bottom row.

A variant of the Mumford–Shah functional in its Γ-convergence approximation was suggested by Shah [78]. In this version the L^2 norm of $|\nabla u|$ in (❯ 25.29) was replaced by its L^1 norm in the first term of G_ϵ

$$\mathcal{J}_{MSTV}(u, v) = \beta \int_\Omega v^2 |\nabla u| dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx.$$

Alicandro et al. [2] proved the Γ-convergence of this functional to

$$\mathcal{J}_{MSTV}(u) = \beta \int_{\Omega \backslash K} |\nabla u| \, dx + \alpha \int_K \frac{|u^+ - u^-|}{1 + |u^+ - u^-|} d\mathcal{H}^1 + |D^c u|(\Omega),$$

■ **Fig. 25-19**
Top row: the *Lena* image blurred with a pill-box kernel of radius 3 and contaminated by salt-and-pepper noise. The noise density is (*left* to *right*) 0.01, 0.1 and 0.3. *Bottom row*: the corresponding recovered images

where u^+ and u^- denote the image values on two sides of the edge set K, $\mathcal{H}^1$ is the one-dimensional Hausdorff measure and $D^c u$ is the Cantor part of the measure-valued derivative Du. The Mumford–Shah and Shah regularizers are compared in ❯ *Fig. 25-20*. The blurred and noisy images are shown in the left column. The results of the restoration using the Mumford–Shah stabilizer (MS) are presented in the middle column and the images recovered using the Shah regularizer (MSTV) are shown in the right column.

The recovery using both methods is satisfactory, but it can be clearly seen that while the Mumford–Shah restoration performs better in the high-frequency image content (see the shades for instance), the Shah restoration attracts the image toward the piecewise constant or cartoon limit which yields images much closer to the "ideal." This can be explained by the fact that the Shah regularizer is more robust to image gradients and hence eliminates high-frequency contributions.

The special case of pure impulse denoising (no blur) is demonstrated in ❯ *Fig. 25-21*. The image on the left shows the outcome of the algorithm of [67] with L^1 norm for both the fidelity and regularization, while the recovery using the L^1 fidelity and MS regularizer is shown on the right. It can be observed that the better robustness of the MS regularizer leads to better performance in the presence of salt and pepper noise.

◼ Fig. 25-20
Left column: the *Window* image blurred with a pill-box kernel of radius 3 and contaminated by salt-and-pepper noise. The noise density is (*top* to *bottom*) 0.01, and 0.1. *Middle column*: the corresponding recovered images with Mumford–Shah (MS) regularization. *Right column*: the corresponding recovered images with Shah (MSTV) regularization

◼ Fig. 25-21
Pure impulse denoising. *Left column*: restoration using the L^1 regularization [67]. *Right column*: restoration using the MS regularizer

25.5.4   Color Image Restoration

We now extend the restoration problem to vector-valued images [9]. In the case of color images, the image intensity is defined as $u : \Omega \rightarrow [0,1]^3$. Here g^ν denotes the observed image at channel $\nu \in \{r, g, b\}$ such that $g^\nu = h * u^\nu + n^\nu$. The underlying assumption here

is that the blur kernel h is common to all of the channels. If the noise is randomly located in a random color channel, the fidelity term can be modeled as

$$\Phi_{L^2} = \int_\Omega \sum_\nu (h * u^\nu - g^\nu)^2 dx$$

in the case of Gaussian noise, and

$$\Phi_{L^1} = \int_\Omega \sum_\nu \sqrt{(h * u^\nu - g^\nu)^2 + \eta}\, dx, \quad \eta \ll 1, \tag{25.44}$$

in the case of impulsive noise. The TV regularization can be generalized to

$$\mathcal{J}_{TV}(u) = \int_\Omega \|\nabla u\|\, dx, \tag{25.45}$$

where

$$\|\nabla u\| = \sqrt{\sum_{\nu \in \{r,g,b\}} |\nabla u^\nu|^2 + \mu}, \quad \mu \ll 1. \tag{25.46}$$

The color MS regularizer thus takes the form

$$\mathcal{J}_{MS}(u, v) = \beta \int_\Omega v^2 \|\nabla u\|^2\, dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx. \tag{25.47}$$

Note that in this regularizer the edge map v is common for the three channels and provides the necessary coupling between colors. In the same fashion the color MSTV regularizer is given by

$$\mathcal{J}_{MSTV}(u, v) = \beta \int_\Omega v^2 \|\nabla u\|\, dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx. \tag{25.48}$$

Once again, the optimization technique is alternate minimization with respect to u^ν and v [9]. ❯ *Figure 25-22* demonstrates the outcome of the different regularizers for an image blurred by Gaussian kernel and corrupted by both Gaussian and salt-and-pepper noise. The fidelity term in all cases was selected as Φ_{L^1} (❯ 25.44).

The methods based on Mumford–Shah regularizer are superior to the TV stabilizers, where MSTV provides a result slightly closer to the "ideal" with little loss of details.

25.5.5 Space-Variant Restoration

The assumption of space-invariant blur kernel is sometimes inaccurate in real photographic images. For example, when multiple objects move at different velocities and in different directions in a scene, one gets space-variant motion blur. Likewise, when a camera lens is focused on one specific object, other objects nearer or farther away from the lens are not as sharp. In such situations, different blur kernels degrade different areas of the image. In some cases it can be assumed that the blur kernel is a piecewise space-variant function. This means that every sub-domain in the image is blurred by a different kernel. In the full

a   Blurred and noisy   b   TV

c   MS   d   MSTV

⬛ **Fig. 25-22**

Recovery of the *Lena* image blurred by 7 × 7 out-of-focus kernel contaminated by mixture of Gaussian and salt-and-pepper noise

blind restoration, several operations have to be simultaneously applied: (1) segmentation of the subregions, (2) estimation of the blur kernels, and (3) recovery of the "ideal" image. Here we present the simplest case where we assume that the subregions and blur kernels are known in advance. The segmentation procedure in a semi-blind restoration problem can be found in [15]. The non-blind space-variant restoration approach relies on the use of a global regularizer, which eliminates the requirement of dealing with region boundaries. As a result, the continuity of the gray levels in the recovered image is inherent. This method does not limit the number of subregions, their geometrical shape, and the kernel support size.

Let the open nonoverlapping subsets $w_i \subset \Omega$ denote regions that are blurred by kernels h_i, respectively. In addition, $\Omega \setminus \cup \overline{w_i}$, denotes the background region blurred by the background kernel h_b, and $\overline{w_i}$ stands for the closure of w_i. The region boundaries are denoted by ∂w_i. The recovered image u is the minimizer of the objective functional

$$\mathcal{F}(u, v) = \frac{1}{2} \sum_i \eta_i \int_{w_i} (h_i * u - g)^2 dx + \frac{\eta_b}{2} \int_{\Omega \setminus (\cup \overline{w_i})} (h_b * u - g)^2 dx + \mathcal{J}_{MS}(u, v), \quad (25.49)$$

where η_i and η_b are positive scalars and $\mathcal{J}_{MS}(u, v)$ is the Mumford–Shah regularizer (❷ 25.29). Following the formulation of Chan and Vese [30], the domains w_i can be replaced by the Heaviside function $H(\phi_i)$, where

$$H(\phi_i) = \begin{cases} 1, & \phi_i > 0, \\ 0, & \phi_i \leq 0, \end{cases} \qquad (25.50)$$

and $\phi_i : \Omega \to \mathbb{R}$ is a level set function such that

$$\partial w_i = \{x \in \Omega : \phi_i(x) = 0\}.$$

The functional then takes the form

$$\mathcal{F}(u,v) = \frac{1}{2} \sum_i \eta_i \int_\Omega (h_i * u - g)^2 H(\phi_i) dx +$$
$$\frac{\eta_b}{2} \int_\Omega (h_b * u - g)^2 \left(1 - \sum_i H(\phi_i)\right) dx + \mathcal{J}_{MS}(u,v). \qquad (25.51)$$

❱ *Figure 25-23* demonstrates the performance of the suggested algorithm. The two images in the left column were synthetically blurred by different blur kernels within the marked shapes. The corresponding recovered images are shown in the right column. Special handling of the region boundaries was not necessary because the MS

◩ **Fig. 25-23**
Non-blind space-variant restoration. *Left column*: spatially variant motion blurred images.
***Right column*: the corresponding recovered images using the suggested method**

regularizer was applied globally to the whole image, enforcing the piecewise smoothness constraint. This means that the boundaries of the blurred regions were smoothed within the restoration process while edges were preserved.

25.5.6 Level Set Formulations for Joint Restoration and Segmentation

We present here other joint formulations for denoising, deblurring, and piecewise-constant segmentation introduced in [46] that can be seen as applications and modifications of the piecewise-constant Mumford–Shah model in level set formulation presented in ❯ Sect. 25.4.2.1. For related work we refer the reader to [11–13, 48, 54]. We use a minimization approach and we consider the gradient descent method. Let $g = h * u + n$ be a given blurred noisy image, where h is a known blurring kernel (such as the Gaussian kernel) and n represents Gaussian additive noise of zero mean. We assume that the contours or jumps in the image u can be represented by the m distinct levels $\{-\infty = l_0 < l_1 < l_2 < \cdots < l_m < l_{m+1} = \infty\}$ of the same implicit (Lipschitz continuous) function $\phi : \Omega \to \mathbb{R}$ partitioning Ω into $m + 1$ disjoint open regions $R_j = \{x \in \Omega : l_{j-1} < \phi(x) < l_j\}, 1 \le j \le m + 1$. Thus, we can recover the denoised-deblurred image $u = c_1 H(\phi - l_m) + \sum_{j=2}^{m} c_j H(\phi - l_{m-j+1})H(l_{m-j+2} - \phi) + c_{m+1}H(l_1 - \phi)$ by minimizing the following energy functional ($\nu_0 > 0$):

$$E(c_1, c_2, \ldots, c_{m+1}, \phi) = \int_\Omega \left| g - h * \left(c_1 H(\phi - l_m) + \sum_{j=2}^{m} c_j H(\phi - l_{m-j+1})H(l_{m-j+2} - \phi) \right. \right.$$
$$\left. \left. + c_{m+1}H(l_1 - \phi) \right) \right|^2 dx + \nu_0 \sum_{j=1}^{m} \int_\Omega |\nabla H(\phi - l_j)| dx.$$

In the binary case (one level $m = 1, l_1 = 0$), we assume the degradation model $g = h * (c_1 H(\phi) + c_2(1 - H(\phi))) + n$, and we wish to recover $u = c_1 H(\phi) + c_2(1 - H(\phi))$ in Ω together with a segmentation of g. The modified binary segmentation model incorporating the blur becomes:

$$\inf_{c_1, c_2, \phi} \left\{ E(c_1, c_2, \phi) = \int_\Omega |g - h * (c_1 H(\phi) + c_2(1 - H(\phi)))|^2 dx \right.$$
$$\left. + \nu_0 \int_\Omega |\nabla H(\phi)| dx \right\}. \tag{25.52}$$

We compute the Euler–Lagrange equations minimizing this energy with respect to c_1, c_2, and ϕ. Using alternating minimization, keeping first ϕ fixed and minimizing the energy

with respect to the unknown constants c_1 and c_2, we obtain the following linear system of equations:

$$c_1 \int_\Omega h_1^2 dx + c_2 \int_\Omega h_1 h_2 dx = \int g h_1 dx,$$

$$c_1 \int_\Omega h_1 h_2 dx + c_2 \int_\Omega h_2^2 dx = \int g h_2 dx$$

with the notations $h_1 = h * H(\phi)$ and $h_2 = h * (1 - H(\phi))$. Note that the linear system has a unique solution because the determinant of the coefficient matrix is not zero due to the Cauchy–Schwartz inequality $\left(\int_\Omega h_1 h_2 dx\right)^2 \leq \int_\Omega h_1^2 dx \int_\Omega h_2^2 dx$, where the equality holds if and only if $h_1 = h_2$ for a.e. $x \in \Omega$. But clearly, $h_1 = h * H(\phi)$ and $h_2 = h * (1 - H(\phi))$ are distinct, thus we have strict inequality.

Keeping now the constants c_1 and c_2 fixed and minimizing the energy with respect to ϕ, we obtain the evolution equation by introducing an artificial time for the gradient descent in $\phi(t, x)$, $t > 0$, $x \in \Omega$

$$\frac{\partial \phi}{\partial t}(t, x) = \delta(\phi)\left[\left(\tilde{h} * g - c_1 \tilde{h} * (h * H(\phi)) - c_2 \tilde{h} * (h * (1 - H(\phi)))\right)\right.$$
$$\left. (c_1 - c_2) + \nu_0 \mathrm{div}\left(\frac{\nabla \phi}{|\nabla \phi|}\right)\right],$$

where $\tilde{h}(x) = h(-x)$.

We show in ❷ Fig. 25-24 a numerical result for joint denoising, deblurring and segmentation of a synthetic image, in a binary level set approach.

In the case of two distinct levels $l_1 < l_2$ of the level set function ϕ ($m = 2$), we wish to recover a piecewise-constant image of the form $u = c_1 H(\phi - l_2) + c_2 H(l_2 - \phi) H(\phi - l_1) + c_3 H(l_1 - \phi)$ and a segmentation of g, assuming the degradation model $g = h * (c_1 H(\phi - l_2) + c_2 H(l_2 - \phi) H(\phi - l_1) + c_3 H(l_1 - \phi)) + n$, by minimizing

$$\inf_{c_1, c_2, c_3, \phi} E(c_1, c_2, c_3, \phi) = \int_\Omega |g - h * (c_1 \ H(\phi - l_2) + c_2 H(l_2 - \phi) H(\phi - l_1)$$
$$+ \ c_3 H(l_1 - \phi))|^2 \, dx + \nu_0 \sum_{j=1}^{2} \int_\Omega |\nabla H(\phi - l_j)| dx. \quad (25.53)$$

Similar to the previous binary model with blur, for fixed ϕ, the unknown constants are computed by solving the linear system of three equations:

$$c_1 \int h_1^2 dx + c_2 \int h_1 h_2 dx + c_3 \int h_1 h_3 dx = \int g h_1 dx$$
$$c_1 \int h_1 h_2 dx + c_2 \int h_2^2 dx + c_3 \int h_2 h_3 dx = \int g h_2 dx$$
$$c_1 \int h_1 h_3 dx + c_2 \int h_2 h_3 dx + c_3 \int h_3^2 dx = \int g h_3 dx$$

where $h_1 = h * H(\phi - l_2)$, $h_2 = h * (H(l_2 - \phi) H(\phi - l_1))$, and $h_3 = h * H(l_1 - \phi)$.
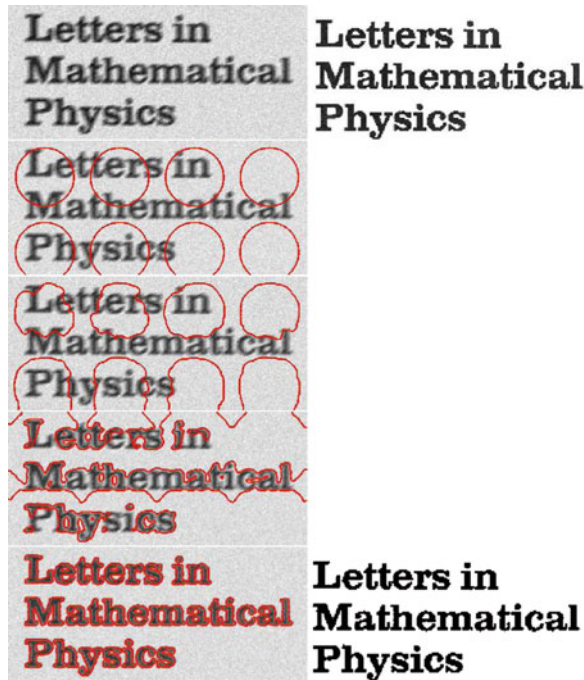
◪ **Fig. 25-24**
**Joint segmentation, denoising, and deblurring using the binary level set model.** *Top row*:
(from *left* to *right*) degraded image *g* (blurred with motion blur kernel of length 10, oriented
at an angle $\theta = 25°$ w.r.t. the horizon and contaminated by Gaussian noise with $\sigma_n = 10$),
original image. Rows 2–5: initial curves, curve evolution using (❷ 25.52) at iterations 50, 100,
300 with $\nu_0 = 5 \cdot 255^2$, and the restored image *u* (SNR = 28.1827). (c_1, c_2): original image
$\approx (62.7525, 259.8939)$, restored *u*, $(61.9194, 262.7795)$

For fixed c_1, c_2, and c_3, by minimizing the functional E with respect to ϕ, we obtain the
gradient descent for $\phi(t, x)$, $t > 0$, $x \in \Omega$:

$$\frac{\partial \phi}{\partial t}(t, x) = \tilde{h} * \big(g - h * (c_1 H(\phi - l_2) + c_2 H(l_2 - \phi) H(\phi - l_2)$$
$$+ c_3 H(l_1 - \phi))(c_1 \delta(\phi - l_2)$$
$$+ c_2 H(l_2 - \phi)\delta(\phi - l_1) - c_2 H(\phi - l_1)\delta(l_2 - \phi) - c_3 \delta(l_1 - \phi)))$$
$$+ \nu_0 \text{div}\left(\frac{\nabla \phi}{|\nabla \phi|}\right)(\delta(\phi - l_1) + \delta(\phi - l_2)). \tag{25.54}$$

We show in ❷ *Figs. 25-25* and ❷ *25-26* a numerical result for joint denoising,
deblurring, and segmentation of the brain image in a multilayer level set approach.
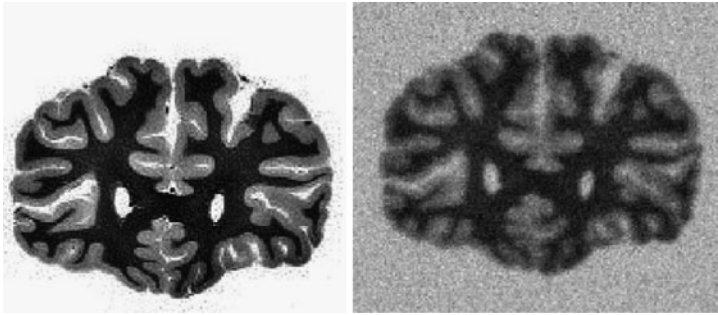
◘ Fig. 25-25

Original image (*left*) and its noisy, blurry version (*right*) blurred with Gaussian kernel with
$\sigma_b = 1$ **and contaminated by Gaussian noise** $\sigma_n = 20$

25.5.7   Image Restoration by Nonlocal Mumford–Shah Regularizers

The traditional regularization terms discussed in the previous sections (depending on the image gradient) are based on local image operators, which denoise and preserve edges very well, but may induce loss of fine structures like texture during the restoration process. Recently, Buades et al. [22] introduced the nonlocal means filter, which produces excellent denoising results. Gilboa and Osher [43, 44] formulated the variational framework of NL-means by proposing nonlocal regularizing functionals and the nonlocal operators such as the nonlocal gradient and divergence. Following Jung et al. [47], we present here nonlocal versions of the Mumford–Shah–Ambrosio–Tortorelli regularizing functionals, called NL/MSH[1] and NL/MSTV, by applying the nonlocal operators proposed by Gilboa–Osher to MSH[1] and MSTV respectively, for image restoration in the presence of blur and Gaussian or impulse noise. In addition, for the impulse noise model, we propose to use a preprocessed image to compute the weights w (the weights w defined in the NL-means filter are more appropriate for the additive Gaussian noise case).
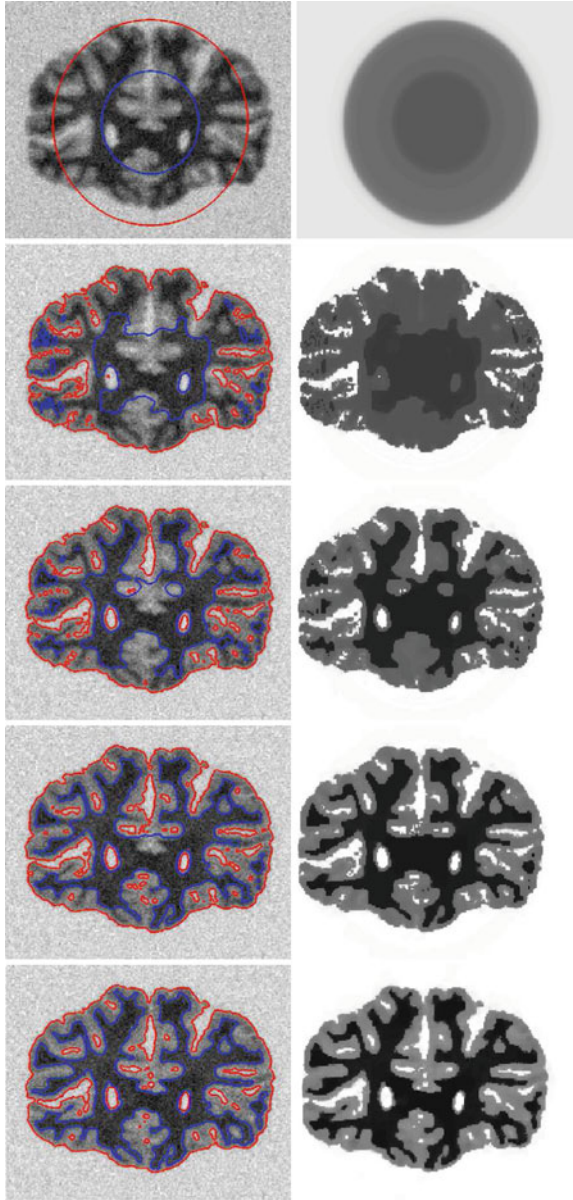
We first recall the Ambrosio–Tortorelli regularizer,

$$\Psi_\epsilon^{MSH^1}(u,v) = \beta \int_\Omega v^2 |\nabla u|^2 dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx,$$

where $0 \le v(x) \le 1$ represents the edges: $v(x) \approx 0$ if $x \in K$ and $v(x) \approx 1$ otherwise, ϵ is a small positive constant, α, β are positive weights.

Shah [78] suggested a modified version of the approximation (❷ 25.55) to the MS functional by replacing the norm square of $|\nabla u|$ by the norm in the first term:

$$\Psi_\epsilon^{MSTV}(u,v) = \beta \int_\Omega v^2 |\nabla u| dx + \alpha \int_\Omega \left(\epsilon |\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right) dx.$$

◨ **Fig. 25-26**

Curve evolution and restored *u* using (❯ 25.54), $\nu_0 = 0.02 \cdot 255^2$, (c_1, c_2, c_3): original image ≈ (12.7501, 125.3610, 255.6453), restored *u* ≈ (22.4797, 136.9884, 255.0074)

This functional Γ converges to the other Ψ^{MSTV} functional [2]:

$$\Psi^{MSTV}(u) = \beta \int_{\Omega\setminus K} |\nabla u| dx + \alpha \int_K \frac{|u^+ - u^-|}{1 + |u^+ - u^-|} d\mathcal{H}^1 + |D_c u|(\Omega),$$

where u^+ and u^- denote the image values on two sides of the jump set $K = J_u$ of u, and $D_c u$ is the Cantor part of the measure-valued derivative Du.

Nonlocal methods in image processing have been explored in many papers because they are well adapted to texture denoising, while the standard denoising models working with local image information seem to consider texture as noise, which results in losing texture. Nonlocal methods are generalized from the neighborhood filters and patch-based methods. The idea of neighborhood filter is to restore a pixel by averaging the values of neighboring pixels with a similar gray level value.

Buades et al. [22] generalized this idea by applying the patch-based methods, proposing a famous neighborhood filter called nonlocal-means (or NL-means):

$$NLu(x) = \frac{1}{C(x)} \int_\Omega e^{-\frac{d_a(u(x),u(y))}{h^2}} u(y) dy$$

$$d_a(u(x), u(y)) = \int_{\mathbb{R}^2} G_a(t) |u(x+t) - u(y+t)|^2 dt$$

where d_a is the patch distance, G_a is the Gaussian kernel with standard deviation a determining the patch size, $C(x) = \int_\Omega e^{-\frac{d_a(u(x),u(y))}{h^2}} dy$ is the normalization factor, and h is the filtering parameter which corresponds to the noise level; usually we set it to be the standard deviation of the noise. The NL-means not only compares the gray level at a single point but the geometrical configuration in a whole neighborhood (patch). Thus, to denoise a pixel, it is better to average the nearby pixels with similar structures rather than just with similar intensities.

In practice, we use the search window $\Omega_w = \{y \in \Omega : |y - x| \le r\}$ instead of Ω (semi-local) and the weight function at $(x, y) \in \Omega \times \Omega$ depending on a function $u : \Omega \to \mathbb{R}$

$$w(x, y) = exp\left(-\frac{d_a(u(x), u(y))}{h^2}\right).$$

The weight function $w(x, y)$ gives the similarity of image features between two pixels x and y, which is normally computed based on the blurry noisy image g.

Based on the gradient and divergence definitions on graphs in the context of machine learning, Gilboa and Osher [44] derived the nonlocal operators. Let $u : \Omega \to \mathbb{R}$ be a function, and $w : \Omega \times \Omega \to \mathbb{R}$ is a weight function assumed to be nonnegative and symmetric. The nonlocal gradient $\nabla_w u : \Omega \times \Omega \to \mathbb{R}$ is defined as the vector $(\nabla_w u)(x, y) :=$ $(u(y) - u(x))\sqrt{w(x, y)}$. Hence, the norm of the nonlocal gradient of u at $x \in \Omega$ is defined as

$$|\nabla_w u|(x) = \sqrt{\int_\Omega (u(y) - u(x))^2 w(x, y) dy}.$$

The nonlocal divergence $div_w \vec{v} : \Omega \to \mathbb{R}$ of the vector $\vec{v} : \Omega \times \Omega \to \mathbb{R}$ is defined as the adjoint of the nonlocal gradient

$$(div_w \vec{v})(x) := \int_\Omega (v(x,y) - v(y,x))\sqrt{w(x,y)}dy.$$

Based on these nonlocal operators, they introduced nonlocal regularizing functionals of the general form

$$\Psi(u) = \int_\Omega \phi(|\nabla_w u|^2)dx,$$

where $\phi(s)$ is a positive function, convex in $\sqrt{s}$ with $\phi(0) = 0$. Inspired by these ideas, we present nonlocal versions of Ambrosio–Tortorelli and Shah approximations to the MS regularizer for image denoising-deblurring. This is also continuation of work by Bar et al. [11–13], as presented in the first part of this section.

We propose the following nonlocal approximated Mumford–Shah and Ambrosio–Tortorelli regularizing functionals (NL/MS) by applying the nonlocal operators to the approximations of the MS regularizer,

$$\Psi^{NL/MS}(u,v) = \beta \int_\Omega v^2 \phi(|\nabla_w u|^2)dx + \alpha \int_\Omega \left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right)dx,$$

where $\phi(s) = s$ and $\phi(s) = \sqrt{s}$ correspond to the nonlocal version of MSH[1] and MSTV regularizers, called here NL/MSH[1] and NL/MSTV, respectively:

$$\Psi^{NL/MSH^1}(u,v) = \beta \int_\Omega v^2|\nabla_w u|^2 dx + \alpha \int_\Omega \left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right)dx$$

$$\Psi^{NL/MSTV}(u,v) = \beta \int_\Omega v^2|\nabla_w u|dx + \alpha \int_\Omega \left(\epsilon|\nabla v|^2 + \frac{(v-1)^2}{4\epsilon} \right)dx.$$

In addition, we use these nonlocal regularizers to deblur images in the presence of Gaussian or impulse noise. Thus, by incorporating the proper fidelity term depending on the noise model, we design two types of total energies as

Gaussian noise model:

$$E^G(u,v) = \int_\Omega (g - h * u)^2 dx + \Psi^{NL/MS}(u,v),$$

Impulse noise model:

$$E^{Im}(u,v) = \int_\Omega |g - h * u|dx + \Psi^{NL/MS}(u,v).$$

Minimizing these functionals in u and v, we obtain the Euler–Lagrange equations:

Gaussian noise model:

$$\frac{\partial E^G}{\partial v} = 2\beta v\phi(|\nabla_w u|^2) - 2\epsilon\alpha \triangle v + \alpha \left(\frac{v-1}{2\epsilon} \right) = 0,$$

$$\frac{\partial E^G}{\partial u} = h^* * (h * u - g) + L^{NL/MS}u = 0.$$

Impulse noise model:

$$\frac{\partial E^{Im}}{\partial v} = 2\beta v \phi(|\nabla_w u|^2) - 2\epsilon\alpha \,\triangle\, v + \alpha\left(\frac{v-1}{2\epsilon}\right) = 0,$$

$$\frac{\partial E^{Im}}{\partial u} = h^* * \text{sign}(h * u - g) + L^{NL/MS} u = 0,$$

where $h^*(x) = h(-x)$ and

$$L^{NL/MS} u = -2 \int_\Omega (u(y) - u(x)) w(x, y)$$
$$\left[(v^2(y)\phi'(|\nabla_w(u)|^2(y))\right.$$
$$\left. + v^2(x)\phi'(|\nabla_w(u)|^2(x))\right] dy.$$

More specifically, the NL/MSH[1] and NL/MSTV regularizers give

$$L^{NL/MSH^1} u = -2\nabla_w \cdot \left(v^2(x)\nabla_w u(x)\right)$$
$$= -2 \int_\Omega (u(y) - u(x)) w(x, y)$$
$$\left[v^2(y) + v^2(x)\right] dy,$$

$$L^{NL/MSTV} u = -\nabla_w \cdot \left(v^2(x)\frac{\nabla_w u(x)}{|\nabla_w u(x)|}\right)$$
$$= -\int_\Omega (u(y) - u(x)) w(x, y)\left[\frac{v^2(y)}{|\nabla_w u|(y)} + \frac{v^2(x)}{|\nabla_w u|(x)}\right] dy.$$

The energy functionals $E^G(u, v)$ and $E^{Im}(u, v)$ are convex in each variable and bounded from below. Therefore, to solve two Euler–Lagrange equations simultaneously, the alternate minimization (AM) approach is applied: in each step of the iterative procedure, we minimize with respect to one function while keeping the other one fixed. Due to its simplicity, we use the explicit scheme for u based on the gradient descent method and the Gauss–Seidel scheme for v. Note that since both energy functionals are not convex in the joint variable, we may compute only a local minimizer. However, this is not a drawback in practice, since the initial guess for u in our algorithm is the data g.

◘ **Fig. 25-27**
Original and noisy blurry images (noisy blurry image using the pill-box kernel of radius 2 and Gaussian noise with $\sigma_n = 5$)

Furthermore, to extend the nonlocal methods to the impulse noise case, we need a preprocessing step for the weight function $w(x, y)$ since we cannot directly use the data g to compute w. In other words, in the presence of impulse noise, the noisy pixels tend to have larger weights than the other neighboring points, so it is likely to keep the noise value at such pixel. Thus, we propose a simple algorithm to obtain first a preprocessed image f, which removes the impulse noise (outliers) as well as preserves the textures as much
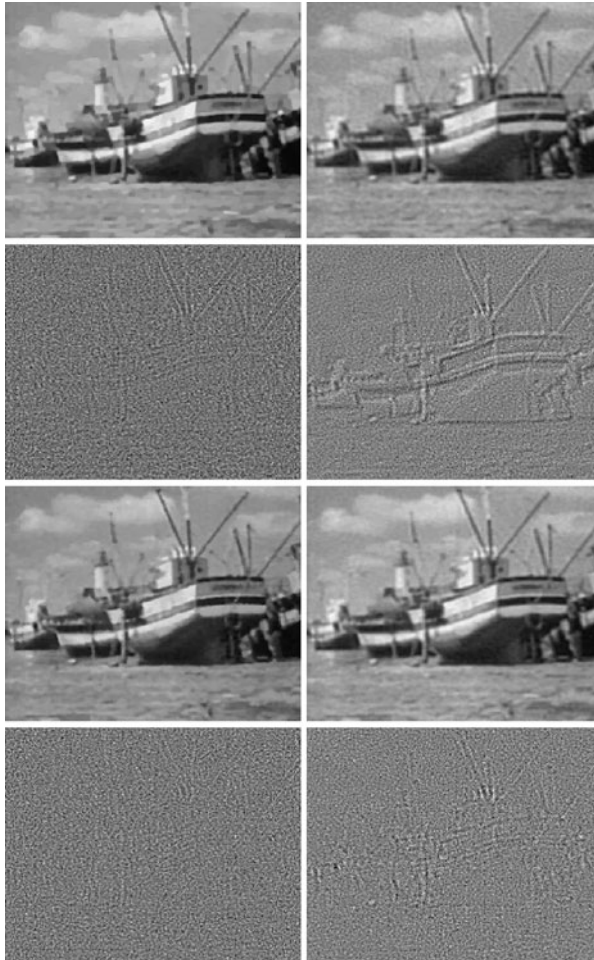
◾ **Fig. 25-28**

Recovery of noisy blurry image from ❯ *Fig. 25-27*. ***Top row*: recovered image u using MSTV (SNR = 25.1968), MSH[1] (SNR = 23.1324).** ***Third row*: recovered image u using NL/MSTV (SNR = 26.4696), NL/MSH[1] (SNR = 24.7164). Second**, *bottom rows*: **corresponding residuals $g - h * u$.** β = **0.0045 (MSTV), 0.001 (NL/MSTV), 0.06 (MSH[1]), 0.006 (NL/MSH[1]), α = 0.00000001,** ϵ = **0.00002**

as possible. Basically, we use the median filter, well known for removing impulse noise. However, if we apply one step of the median filter, then the output may be too smoothed out. In order to preserve the fine structures as well as to remove the noise properly, we use the idea of Bregman iteration [21, 68], and we propose the following algorithm to obtain a preprocessed image f that will be used only in the computation of the weight function:

Initialize : $r_0 = 0, f_0 = 0$.
do (iterate $n = 0, 1, 2, \ldots$)
$\qquad f_{n+1} = median(g + r_n, [a \; a])$
$\qquad r_{n+1} = r_n + g - h * f_{n+1}$
while $\|g - h * f_n\|_1 > \|g - h * f_{n+1}\|_1$
[Optional] $f_m = median(f_m, [b \; b])$

■ **Fig. 25-29**

Recovery of noisy blurry image with Gaussian kernel with $\sigma = 1$ **and salt-and-pepper noise with** $d = 0.3$. *Top row*: **original image, blurry image, noisy-blurry image.** *Middle row*: **recovered images using MSTV (SNR = 27.8336), MSH[1] (SNR = 23.2052).** *Bottom row*: **recovered images using NL/MSTV (SNR = 29.3503), NL/MSH[1] (SNR = 27.1477).** *Parameters*: $\beta = 0.25$ **(MSTV), 0.1 (NL/MSTV),** $\alpha = 0.01, \epsilon = 0.002$. *Parameters*: $\beta = 2$ **(MSH[1]), 0.55 (NL/MSH[1]),** $\alpha = 0.001, \epsilon = 0.0001$

where g is the given noisy blurry data, $median(u, [a\ a])$ is the median filter of size $a \times a$ with input u; the optional step is needed in the case when the final f_m still has some salt-and-pepper-like noise. This algorithm is simple and requires a few iterations only, so it takes less than 1 s for a 256×256 size image. The preprocessed image f will be used only in the computation of the weights w, while keeping g in the data fidelity term, thus artifacts are not introduced by the median filter.

We show in ❯ *Figs. 25-27* and ❯ *25-28* an experimental result for image restoration of a boat image degraded by the pill-box kernel blur of radius 2 and additive Gaussian noise. The nonlocal methods give better reconstruction.

We show in ❯ *Figs. 25-29* and ❯ *25-30* an experimental result for image restoration of a woman image degraded by Gaussian kernel blur and salt-and-pepper noise. ❯ *Figure 25-30* shows the edge set v for the four results. The nonlocal methods give better reconstruction.

We show in ❯ *Fig. 25-31* an experimental result for restoration of the Einstein image degraded by motion kernel blur and random-valued impulse noise. The nonlocal methods give better reconstruction.

◧ **Fig. 25-30**

Edge map *v* using the MS regularizers in the recovery of the Lena image blurred with Gaussian blur kernel with σ_b = 1 and contaminated by salt-and-pepper noise with density d = 0.3. *Top*: (*left*) MSTV, (*right*) NL/MSTV. *Bottom*: (*left*) MSH[1], (*right*) NL/MSH[1]
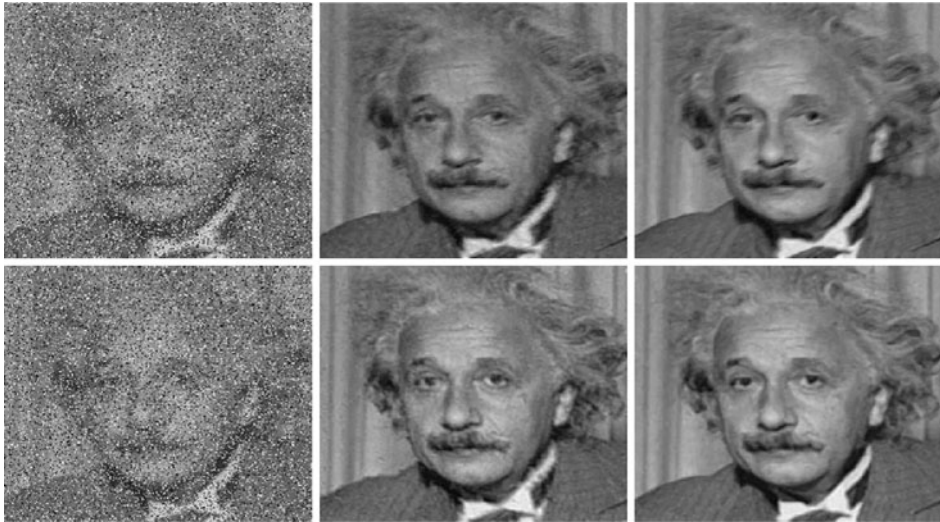
▣ Fig. 25-31

Comparison between MSH[1] and NL/MSH[1] with the image blurred and contaminated by high density (d = 0.4) of random-valued impulse noise. *Top*: **noisy blurry image blurred with the motion blur in recovered images using MSH[1] (*left*, SNR = 17.9608) and NL/MSH[1] (*right*, SNR = 20.7563).** *Bottom*: **noisy blurry image blurred with the Gaussian blur in recovered images using MSH[1] (*left*, SNR = 16.6960) and NL/MSH[1] (*right*, SNR = 24.2500).** *Top*: β = **1.5 (MSH[1]), 0.5 (NL/MSH[1]),** α = **0.0001,** ϵ = **0.002.** *Bottom*: β = **2.5 (MSH[1]), 0.65 (NL/MSH[1]),** α = **0.000001,** ϵ = **0.002**

25.6   Conclusion

We conclude this chapter by first summarizing its main results. The Mumford–Shah model for image segmentation has been presented, together with its main properties. Several approximations to the Mumford and Shah energy have been discussed, with an emphasis on phase-field approximations and level set approximations. Several numerical results for image segmentation by these methods have been presented. In the last section of the chapter, several restoration problems were addressed in a variational framework. The fidelity term was formulated according to the noise model (Gaussian, impulse, multi-channel impulse). First, the a priori piecewise-smooth image model was mathematically integrated into the functional as an approximation of the Mumford–Shah segmentation elements by the Γ-convergence formulation. Comparative experimental results show the superiority of this regularizer with respect to modern state-of-the-art restoration techniques. Also, the piecewise-constant level set formulations of the Mumford–Shah energy have been applied to image restoration (related to relevant work by Kim et al. [48]), joint with segmentation. Finally, in the last section, the Ambrosio–Tortorelli approximations and Bar et al. restoration models have been extended to nonlocal regularizers, inspired

by the work of Gilboa et al. These models produce much improved restoration results for images with texture and fine details.

25.7   Recommended Reading

Many more topics on the Mumford–Shah model and its applications have been explored in image processing, computer vision, and more generally in inverse problems. This chapter contains only a small sample of results and methods. As mentioned before, we recommend detailed monographs on the Mumford–Shah problem and related theoretical and applications topics by Blake and Zisserman [16], by Morel and Solimini [61], by Chambolle [26], by Ambrosio et al. [4], by David [39], and by Braides [19]. Also, the monographs by Aubert and Kornprobst [8] and by Chan and Shen [32] contain chapters presenting the Mumford and Shah problem and its main properties.

We would like to mention the work by Cohen et al. [36, 37] on using curve evolution approach and the Mumford–Shah functional for detecting the boundary of a lake. The work by Aubert et al. [7] also proposes an interesting approximation of the Mumford–Shah energy by a family of discrete edge-preserving functionals, with Γ-convergence result.

References and Further Reading

1. Adams RA (1975) Sobolev spaces. Academic, New York

2. Alicandro R, Braides A, Shah J (1999) Free-discontinuity problems via functionals involving the L1-norm of the gradient and their approximation. Interfaces Free Bound 1:17–37

3. Ambrosio L (1989) A compactness theorem for a special class of functions of bounded variation. Boll Un Mat Ital 3(B):857–881

4. Ambrosio L, Fusco N, Pallara D (2000) Functions of bounded variation and free discontinuity problems. Oxford University Press, New York

5. Ambrosio L, Tortorelli VM (1990) Approximation of functionals depending on jumps by elliptic functionals via Γ-convergence. Comm Pure Appl Math 43(8):999–1036

6. Ambrosio L, Tortorelli VM (1992) On the approximation of free discontinuity problems. Boll Un Mat Ital B7(6):105–123

7. Aubert G, Blanc-Féraud L, March R (2006) An approximation of the Mumford-Shah energy by a family of discrete edge-preserving functionals. Nonlinear Anal 64(9):1908–1930

8. Aubert G, Kornprobst P (2006) Mathematical problems in image processing. Springer, New York

9. Bar L, Brook A, Sochen N, Kiryati N (2007) Deblurring of color images corrupted by impulsive noise. IEEE Trans Image Process 16(4):1101–1111

10. Bar L, Sochen N, Kiryati N (2004) Variational pairing of image segmentation and blind restoration. In Proceedings of 8th European conference on computer vision, vol 3022 of LNCS, pp 166–177

11. Bar L, Sochen N, Kiryati N (2005) Image deblurring in the presence of salt-and-pepper noise. In Proceedings of 5th international conference on scale space and PDE methods in computer vision, vol 3459 of LNCS, pp 107–118

12. Bar L, Sochen N, Kiryati N (2006) Image deblurring in the presence of impulsive noise. Int J Comput Vis 70:279–298

13. Bar L, Sochen N, Kiryati N (2006) Semi-blind image restoration via Mumford-Shah regularization. IEEE Trans Image Process 15(2):483–493

14. Bar L, Sochen N, Kiryati N (2007) Convergence of an iterative method for variational deconvolution and impulsive noise removal. SIAM J Multiscale Model Simulat 6:983–994

15. Bar L, Sochen N, Kiryati N (2007) Restoration of images with piecewise space-variant blur. In Proceedings of 1st international conference on scale space and variational methods in computer vision, pp 533–544

16. Blake A, Zisserman A (1987) Visual reconstruction. MIT Press, Cambridge

17. Bourdin B (1999) Image segmentation with a finite element method. M2AN Math Model Numer Anal 33(2):229–244

18. Bourdin B, Chambolle A (2000) Implementation of an adaptive finite-element approximation of the Mumford-Shah functional. Numer Math 85(4):609–646

19. Braides A (1998) Approximation of free-discontinuity problems, vol 1694 of Lecture notes in mathematics. Springer, Berlin

20. Braides A, Dal Maso G (1997) Nonlocal approximation of the Mumford-Shah functional. Calc Var 5:293–322

21. Bregman LM (1967) The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming. USSR Comp Math Phys 7:200–217

22. Buades A, Coll B, Morel JM (2005) A review of image denoising algorithms, with a new one. SIAM MMS 4(2):490–530

23. Chambolle A (1992) Un théorème de γ-convergence pour la segmentation des signaux. C R Acad Sci Paris Sér. I Math 314(3):191–196

24. Chambolle A (1995) Image segmentation by variational methods: Mumford and Shah functional, and the discrete approximation. SIAM J Appl Math 55:827–863

25. Chambolle A (1999) Finite-differences discretizations of the Mumford-Shah functional. M2AN Math Model Numer Anal 33(2):261–288

26. Chambolle A (2000) Inverse problems in image processing and image segmentation: some mathematical and numerical aspects. In: Chidume CE (ed) ICTP Lecture notes series, vol 2. ICTP

27. Chambolle A, Dal Maso G (1999) Discrete approximation of the Mumford-Shah functional in dimension two. M2AN Math Model Numer Anal 33(4):651–672

28. Chan T, Vese L (1999) An active contour model without edges. Lecture Notes Comput Sci 1682:141–151

29. Chan T, Vese L (2000) An efficient variational multiphase motion for the Mumford-Shah segmentation model. In 34th Asilomar conference on signals, systems, and computers, vol 1, pp 490–494

30. Chan T, Vese L (2001) Active contours without edges. IEEE Trans Image Process 10:266–277

31. Chan T, Vese L (2001) A level set algorithm for minimizing the Mumford-Shah functional in image processing. In IEEE/Computer Society proceedings of the 1st IEEE workshop on variational and level set methods in computer vision, pp 161–168

32. Chan TF, Shen J (2005) Image processing and analysis. Variational, PDE, wavelet, and stochastic methods. SIAM, Philadelphia

33. Chan TF, Wong CK (1998) Total variation blind deconvolution. IEEE Trans Image Process 7: 370–375

34. Chung G, Vese LA (2005) Energy minimization based segmentation and denoising using a multi-layer level set approach. Lecture Notes in Comput Sci 3757:439–455

35. Chung G, Vese LA (2009) Image segmentation using a multilayer level-set approach. Computing Visual Sci 12(6):267–285

36. Cohen L, Bardinet E, Ayache N (1993) Surface reconstruction using active contour models. In SPIE '93 conference on geometric methods in computer vision, San Diego, July 1993

37. Cohen LD (1997) Avoiding local minima for deformable curves in image analysis. In: Le Méhauté A, Rabut C, Schumaker LL (eds) Curves and Surfaces with applications in CAGD, pp 77–84

38. Dal Maso G (1993) An introduction to Γ-convergence. Progress in nonlinear differential equations and their applications. Birkhäuser, Boston

39. David G (2005) Singular sets of minimizers for the Mumford-Shah functional. Birkhäuser Verlag, Basel

40. Evans LC (1998) Partial differential equations. American Mathematical Society, Providence, Rhode Island

41. Evans LC, Gariepy RF (1992) Measure theory and fine properties of functions. CRC Press

42. Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesean restoration of images. IEEE TPAMI 6:721–741

43. Gilboa G, Osher S (2007) Nonlocal linear image regularization and supervised segmentation. SIAM MMS 6(2):595–630

44. Gilboa G, Osher S (2008) Nonlocal operators with applications to image processing. Multiscale Model Simulat 7(3):1005–1028

45. Huber PJ (1981) Robust statistics. Wiley, New York

46. Jung M, Chung G, Sundaramoorthi G, Vese LA, Yuille AL (2009) Sobolev gradients and joint variational image segmentation, denoising and deblurring. In: IS&T/SPIE on electronic imaging, vol 7246 of Computational imaging VII, pp 72460I–1–72460I–13

47. Jung M, Vese LA (2009) Nonlocal variational image deblurring models in the presence of gaussian or impulse noise. In: International conference on Scale Space and Variational Methods in Computer Vision (SSVM' 09), vol 5567 of LNCS, pp 402–413

48. Kim J, Tsai A, Cetin M, Willsky AS (2002) A curve evolution-based variational approach to simultaneous image restoration and segmentation. In: Proceedings of IEEE international conference on image processing, vol 1, pp 109–112

49. Koepfler G, Lopez C, Morel JM (1994) A multiscale algorithm for image segmentation by variational methods. SIAM J Numer Anal 31(1):282–299

50. Kundur D, Hatzinakos D (1996) Blind image deconvolution. Signal Process Mag 13:43–64

51. Kundur D, Hatzinakos D (1996) Blind image deconvolution revisited. Signal Process Mag 13:61–63

52. Larsen CJ (1998) A new proof of regularity for two-shaded image segmentations. Manuscripta Math 96:247–262

53. Leonardi GP, Tamanini I (1998) On minimizing partitions with infinitely many components. Ann Univ Ferrara - Sez. VII - Sc. Mat XLIV:41–57

54. Li C, Kao C-Y, Gore JC, Ding Z (2007) Implicit active contours driven by local binary fitting energy. In IEEE conference on computer vision and pattern recognition (CVPR), CVPR'07

55. Dal Maso G, Morel JM, Solimini S (1989) Variational approach in image processing – existence and approximation properties. C R Acad Sci Paris Sér. I Math 308(19):549–554

56. Dal Maso G, Morel JM, Solimini S (1992) A variational method in image segmentation - existence and approximation properties. Acta Mathem 168(1–2):89–151

57. Massari U, Tamanini I (1993) On the finiteness of optimal partitions. Ann Univ Ferrara – Sez VII - Sc Mat XXXIX:167–185

58. Modica L (1987) The gradient theory of phase transitions and the minimal interface criterion. Arch Rational Mech Anal 98:123–142

59. Modica L, Mortola S (1977) Un esempio di γ-convergenza. Boll Un Mat Ital B(5)(14):285–299

60. Mohieddine R, Vese LA (2010) Open curve level set formulations for the Mumford and Shah segmentation model. UCLA C.A.M. Report, pp 10–33

61. Morel J-M, Solimini S (1995) Variational methods in image segmentation. Birkhäuser, Boston

62. Mumford D, Shah J (1985) Boundary detection by minimizing functionals. In Proceedings of IEEE conference on computer vision and pattern recognition, pp 22–26

63. Mumford D, Shah J (1989) Boundary detection by minimizing functionals. In: Ullman S, Richards W (eds) Image understanding. Springer, Berlin, pp 19–43

64. Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. Comm Pure Appl Math 42:577–685

65. Neuberger JW (1997) Sobolev gradients and differential equations. Springer lecture notes in mathematics, vol 1670

66. Nikolova M (2002) Minimizers of cost-functions involving nonsmooth data-fidelity terms: application to the processing of outliers. SIAM J Numer Anal 40:965–994

67. Nikolova M (2004) A variational approach to remove outliers and impulse noise. J Math Imaging Vis 20:99–120

68. Osher S, Burger M, Goldfarb D, Xu J, Yin W (2005) An iterative regularization method for total variation based image restoration. SIAM MMS 4:460–489

69. Osher S, Sethian JA (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation. J Comput Phys 79:12–49

70. Osher SJ, Fedkiw RP (2002) Level set methods and dynamic implicit surfaces. Springer, New York

71. Renka RJ (2006) A Simple Explanation of the Sobolev Gradient Method. (online manuscript at http: / /www.cse.unt.edu /~ renka/papers/sobolev.pdf)

72. Richardson WB (2008) Sobolev gradient preconditioning for image processing PDEs. Commun Numer Meth Eng 24:493–504

73. Rudin L, Osher S (1994) Total variation based image restoration with free local constraints. In: Proceedings of IEEE international conference on image processing, vol 1, Austin, pp 31–35

74. Rudin LI, Osher S, Fatemi E (1992) Non linear total variation based noise removal algorithms. Physica D 60:259–268

75. Samson C, Blanc-Féraud L, Aubert G, Zerubia J (1999) Multiphase evolution and variational image classification. Technical Report 3662, INRIA Sophia Antipolis

76. Sethian JA (1996) Level set methods. Evolving interfaces in geometry, fluid mechanics, computer vision, and materials science. Cambridge University Press

77. Sethian JA (1999) Level set methods and fast marching methods. Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science. Cambridge University Press, Cambridge

78. Shah J (1996) A common framework for curve evolution, segmentation and anisotropic diffusion. In: Proceedings of IEEE conference on computer vision and pattern recognition, pp 136–142

79. Smereka P (2000) Spiral crystal growth. Physica D 138:282–301

80. Tamanini I (1996) Optimal approximation by piecewise constant functions. In: Progress in nonlinear differential equations and their applications, vol 25. Birkhäuser Verlag, Basel, pp 73–85

81. Tamanini I, Congedo G (1996) Optimal segmentation of unbounded functions. Rend Sem Mat Univ Padova 95:153–174

82. Tikhonov AN, Arsenin V (1977) Solutions of ill-posed problems. Winston, Washington

83. Tsai A, Yezzi A, Willsky A (2001) Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. IEEE Trans Image Process 10(8):1169–1186

84. Vese LA, Chan TF (2002) A multiphase level set framework for image segmentation using the Mumford and Shah model. Int J Comput Vis 50(3):271–293

85. Vogel CR, Oman ME (1998) Fast, robust total variation-based reconstruction of noisy, blurred images. IEEE Trans Image Process 7:813–824

86. Weisstein EW Minimal residual method. MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/MinimalResidualMethod.html.

87. You Y, Kaveh M (1996) A regularization approach to joint blur identification and image restoration. IEEE Trans Image Process 5:416–428

88. Zhao HK, Chan T, Merriman B, Osher S (1996) A variational level set approach to multiphase motion. J Comput Phys 127:179–195