

# Variational PDE Models in Image Processing

*Tony F. Chan, Jianhong (Jackie) Shen, and Luminita Vese*

**I**mage processing, traditionally an engineering field, has attracted the attention of many mathematicians during the past two decades. From the point of view of vision and cognitive science, image processing is a basic tool used to reconstruct the relative order, geometry, topology, patterns, and dynamics of the three-dimensional (3-D) world from two-dimensional (2-D) images. Therefore, it cannot be merely a historical coincidence that mathematics must meet image processing in this era of digital technology.

The role of mathematics is determined also by the broad range of applications of image processing in contemporary science and technology. These applications include astronomy and aerospace exploration, medical imaging, molecular imaging, computer graphics, human and machine vision, telecommunication, autopiloting, surveillance video, and biometric security identification (such

---

*This paper is based on a plenary presentation given by Tony F. Chan at the 2002 Joint Mathematics Meetings in San Diego. The work was supported in part by NSF grants DMS-9973341 (Chan), DMS-0202565 (Shen), and ITR-0113439 (Vese); by ONR grant N00014-02-1-0015 (Chan); and by NIH grant NIH-P20MH65166 (Chan and Vese).*

*Tony F. Chan is professor of mathematics at the University of California, Los Angeles. His email address is [chan@math.ucla.edu](mailto:chan@math.ucla.edu).*

*Jianhong (Jackie) Shen is assistant professor at the School of Mathematics, University of Minnesota. His email address is [jhshen@math.umn.edu](mailto:jhshen@math.umn.edu).*

*Luminita Vese is assistant professor of mathematics at the University of California, Los Angeles. Her email address is [lvese@math.ucla.edu](mailto:lvese@math.ucla.edu).*

as fingerprints and face identification). All these highly diversified disciplines have made it necessary to develop common mathematical foundations and frameworks for image analysis and processing. Mathematics at all levels must be introduced to address the crucial criteria demanded by this new era—genericity, well-posedness, accuracy, and computational efficiency, just to name a few. In return, image processing has created tremendous opportunities for mathematical modeling, analysis, and computation.

This article gives a broad picture of mathematical image processing through one of the most recent and very successful approaches—the variational PDE (partial differential equation) method. We first discuss two crucial ingredients for image processing: image modeling or representation, and processor modeling. We then focus on the variational PDE method. The backbone of the article consists of two major problems in image processing that we personally have worked on: inpainting and segmentation. By no means, however, do we intend to give a comprehensive review of the entire field of image processing. Many of the authors' articles and preprints related to the subject of this paper can be found online at our group homepage [11], where an extended bibliography is also available.

## **Image Processing as an Input-Output System**

Directly connected to image processing are two dual fields in contemporary computer science: computer vision and computer graphics. Vision (whether machine or human) tries to reconstruct the 3-D world from observed 2-D images, while

graphics pursues the opposite direction by designing suitable 2-D scene images to simulate our 3-D world. Image processing is the crucial middle way connecting the two.

Abstractly, image processing can be considered as an input-output system

$$Q_0 \rightarrow \boxed{\text{Image Processor } \mathcal{T}} \rightarrow Q$$

Here  $\mathcal{T}$  denotes a typical image processor: for example, denoising, deblurring, segmentation, compression, or inpainting. The input data  $Q_0$  can represent an observed or measured single image or image sequence, and the output  $Q = (q_1, q_2, \dots)$  contains all the targeted image *features*.

For example, the human visual system can be considered as a highly involved multilevel image processor  $\mathcal{T}$ . The input  $Q_0$  represents the image sequence that is constantly projected onto the retina. The output vector  $Q$  contains all the major features that are important to our daily life, from the low-level ones such as relative orders, shapes, and grouping rules to high-level feature parameters that help classify or identify various patterns and objects.

Table 1 lists some typical image processing problems.

The two main ingredients of image processing are the input  $Q_0$  and the processor  $\mathcal{T}$ . As a result, the two key issues that have been driving mainstream mathematical research on image processing are (a) the modeling and representation of the input visual data  $Q_0$ , and (b) the modeling of the processing operators  $\mathcal{T}$ . Although the two are independent, they are closely connected to each other by the universal rule in mathematics: the structure and performance of an operator  $\mathcal{T}$  is greatly influenced by how the input functions are modeled or represented.

### Image Modeling and Representation

To efficiently handle and process images, we need first to understand what images really are mathematically and how to represent them. For example, is it adequate to treat them as general  $L^2$  functions or as a subset of  $L^2$  with suitable regularity constraints? Here we briefly outline three major classes of image modeling and representation.

*Random fields modeling.* An observed image  $u_0$  is modeled as the sampling of a random field. For example, the Ising spin model in statistical mechanics can be used to model binary images. More generally, images are modeled by some Gibbs/Markovian random fields [10]. The statistical properties of fields are often established through a filtering technique and learning theory. Random field modeling is the ideal approach for describing natural images with rich texture patterns such as trees and mountains.

$\mathcal{T}$	$Q_0$	$Q$
denoising+ deblurring	$u_0 = Ku + n$	clean & sharp $u$
inpainting	$u_0 _{\Omega \setminus D}$	entire image $u _{\Omega}$
segmentation	$u_0$	"objects" $[u_k, \Omega_k], k = 1, 2, \dots$
scale-space	$u_0$	multiscale images $(u_{\lambda_1}, u_{\lambda_2}, \dots)$
motion estimation	$(u_0^{(1)}, u_0^{(2)}, \dots)$	optical flows $(\vec{v}^{(1)}, \vec{v}^{(2)}, \dots)$

**Table 1. Typical image processors and their inputs and outputs. The symbols represent (1)  $K$ : a blurring kernel, and  $n$ : an additive noise, both assumed in this paper to be linear for simplicity; (2)  $u_0$ : the given noisy or blurred image; (3)  $\Omega$ : the entire image domain, and  $D$ : a subset where image information is missing or inaccessible; (4)  $[u_k, \Omega_k]$ :  $\Omega_k$ 's are the segmented individual "objects", while  $u_k$ 's are their intensity values; (5)  $\lambda_k$ 's are different scales, and  $u_{\lambda}$  can be roughly understood as the projection of the input image at scale  $\lambda$ ; (6)  $u_0^{(n)}$ 's denote the discrete sampling of a continuous "movie"  $u_0(x, t)$  (with some small time step  $h$ ) and  $\vec{v}^{(n)}$ 's are the estimated optical flows (i.e., velocity fields) at each moment.**

*Wavelet Representation.* An image is often acquired from the responses of a collection of microsensors (or photo receptors), either digital or biological. During the past two decades, it has been gradually realized (and experimentally supported) that such local responses can be well approximated by wavelets. This new representation tool has revolutionized our notion of images and their multiscale structures [12]. The new JPEG2000 protocol for image coding and the successful compression of the FBI database of fingerprints are its two most influential applications. The theory is still being actively pushed forward by a new generation of geometric wavelets such as curvelets (Candés and Donoho) and beamlets (Pennec and Mallat).

*Regularity Spaces.* In the linear filtering theory of conventional digital image processing, an image  $u$  is considered to be in the Sobolev space  $H^1(\Omega)$ . The Sobolev model works well for homogeneous regions, but it is insufficient as a global image model, since it "smears" the most important visual cue, namely, edges. Two well-known models have been introduced to recognize the existence of edges. One is the "object-edge" model of Mumford and Shah [13], and the other is the BV image model of Rudin, Osher, and Fatemi [15]. The object-edge model assumes that an ideal image  $u$  consists of disjoint homogeneous object patches  $[u_k, \Omega_k]$  with  $u_k \in H^1(\Omega_k)$  and regular boundaries  $\partial\Omega_k$  (characterized by one-dimensional Hausdorff measure). The BV image model assumes that an ideal image has bounded total variation  $\int_{\Omega} |Du|$ . Regularity-based image models

are generally applicable to images with low texture patterns and without rapidly oscillatory components.

### Modeling of Image Processors

How images are modeled and represented very much determines the way we model image processors. We shall illustrate this viewpoint through the example of denoising  $u = \mathcal{T}u_0$ :  $u_0 = u + n$ , assuming for simplicity that the white noise  $n$  is additive and homogeneous, and there is no blurring involved.

When images are represented by wavelets, the denoising processor  $\mathcal{T}$  is in some sense “diagonalized” and is equivalent to a simple engineering on the individual wavelet components. This is a celebrated result of Donoho and Johnstone on threshold-based denoising schemes.

Under the statistical/random field modeling of images, the denoising processor  $\mathcal{T}$  becomes MAP (*Maximum A Posteriori*) estimation. By Bayes’s formula, the posterior probability given an observation  $u_0$  is

$$p(u|u_0) = p(u_0|u)p(u)/p(u_0).$$

The denoising processor  $\mathcal{T}$  is achieved by solving the MAP problem  $\max_u p(u|u_0)$ . Therefore, it is important to know not only the random field image model  $p(u)$  but also the mechanism by which  $u_0$  is generated from the ideal image  $u$  (the so-called *generative* data model). The two are crucial for successfully carrying out Bayesian denoising.

Finally, if the ideal image  $u$  is modeled as an element in a regular function space such as  $H^1(\Omega)$  or  $BV(\Omega)$ , then the denoising processor  $\mathcal{T}$  can be realized by a variational optimization. For instance, in the BV image model,  $\mathcal{T}$  is achieved by

$$\min_u \int_{\Omega} |Du| \text{ subject to } \frac{1}{|\Omega|} \int_{\Omega} (u - u_0)^2 dx \leq \sigma^2,$$

where the white noise is assumed to be well approximated by the standard Gaussian  $N(0, \sigma^2)$ . This well-known denoising model, first proposed by Rudin, Osher, and Fatemi, belongs to the more general class of *regularized* data-fitting models.

Just as different coordinate systems that describe a single physical object are related, different formulations of the same image processor are closely interconnected. Again, take denoising for example. It has been shown that the wavelet technique is equivalent to an approximate optimal regularization in certain Besov spaces (Cohen, Dahmen, Daubechies, and DeVore). On the other hand, Bayesian processing and the regularity-based variational approach can also be connected (at least formally) by Gibbs’s formula in statistical mechanics (see (3) in the next section).

### Variational PDE Method

Having briefly introduced the general picture of mathematical image processing, we now focus on the variational PDE method through two processors: inpainting and segmentation.

For the history and a detailed description of current developments of the variational and PDE method in image and vision analysis, see two special issues in *IEEE Trans. Image Processing* [7 (3), 1998] and *J. Visual Comm. Image Rep.* [13 (1/2), 2002] and also two recent monographs [1], [18].

In the variational or “energy”-based models, nonlinear PDEs emerge as one derives formal Euler-Lagrange equations or tries to locate local or global minima by the gradient descent method. Some PDEs can be studied by the viscosity solution approach [8], while many others still remain open to further theoretical investigation.

Compared with other approaches, the variational PDE method has remarkable advantages in both theory and computation. First, it allows one to directly handle and process visually important geometric features such as gradients, tangents, curvatures, and level sets. It can also effectively simulate several visually meaningful dynamic processes, such as linear and nonlinear diffusions and the information transport mechanism. Second, in terms of computation, it can profoundly benefit from the existing wealth of literature on numerical analysis and computational PDEs. For example, various well-designed shock-capturing schemes in Computational Fluid Dynamics (CFD) can be conveniently adapted to edge computation in images.

### Variational Image Inpainting and Interpolation

The word *inpainting* is an artistic synonym for image interpolation; initially it circulated among museum restoration artists who manually restore cracked ancient paintings. The concept of digital inpainting was recently introduced into digital image processing in a paper by Bertalmio, Sapiro, Caselles, and Ballester. Currently, digital inpainting techniques are finding broad applications in image processing, vision analysis, and digital technologies such as image restoration, disocclusion, perceptual image coding, zooming and image super-resolution, error concealment in wireless image transmission, and so on [2], [4], [9]. Figure 1 shows an example of error concealment.

We now discuss the mathematical ideas and methodologies behind variational inpainting techniques. Throughout this section,  $u$  denotes the original complete image on a 2-D domain  $\Omega$ , and  $u_0$  denotes the observed or measured portion of  $u$ , which can be either noisy or blurry, on a subdomain or general subset  $D$ . The goal of inpainting is to recover  $u$  on the entire image domain  $\Omega$

as faithfully as possible from the available data  $u_0$  on  $D$ .

### From Shannon's Theorem to Variational Inpainting

Interpolation is a classical topic in approximation theory, numerical analysis, and signal and image processing. Successful interpolants include polynomials, harmonic waves, radially symmetric functions, finite elements, splines, and wavelets. Despite the diversity of the literature, there exists one most widely recognized result due to Shannon, known as *Shannon's Sampling Theorem*.

**Theorem** (*Shannon's Theorem*). *If a signal  $u(t)$  is band-limited within  $(-\omega, \omega)$ , then*

$$u(t) = \sum_{n=-\infty}^{\infty} u\left(n\frac{\pi}{\omega}\right) \operatorname{sinc}\left(\frac{\omega}{\pi}t - n\right).$$

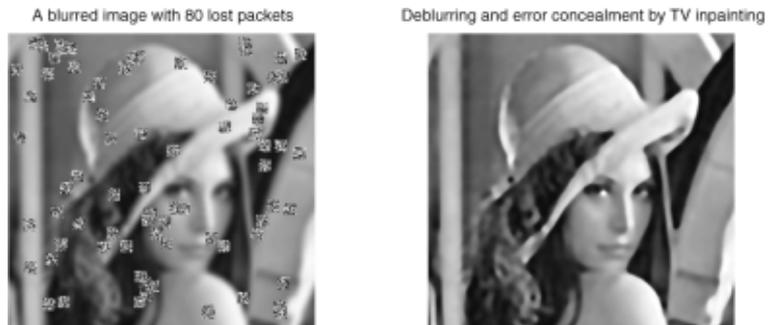
That is, if an analog signal  $u(t)$  (with finite energy or, equivalently, in  $L^2(\mathbb{R})$ ) does not contain any high frequencies, then it can be perfectly interpolated from its properly sampled discrete sequence  $u_0[n] = u(n\pi/\omega)$  (where  $\omega/\pi$  is known as the Nyquist frequency).

All interpolation problems share this "if-then" structure. "If" specifies the space where the target signal  $u$  is sought, while "then" gives the reconstruction or interpolation procedure based on the discrete samples (or, more generally, any partial information about the signal).

Unfortunately, for most real applications in signal and image processing, one cannot expect a closed-form formula as clean as Shannon's. This is due to at least two factors. First, in vision analysis and communication, signals like images are intrinsically not band-limited because of the presence of *edges* (or Heaviside-type singularities). Second, for most real applications, the given incomplete data are often noisy and become blurred during the imaging and transmission processes. Therefore, in the situation of Shannon's Theorem, we are dealing with a class of "bad" signals  $u$  with "unreliable" samples  $u_0$ .

Naturally, for image inpainting, both the "if" and "then" statements in Shannon's Theorem need to be modeled carefully. It turns out that there are two powerful and interdependent frameworks that can carry out this task: one is the variational method, and the other is the Bayesian framework [10].

In the Bayesian approach the "if" statement specifies both the so-called *prior* model and the *data* model. The prior model specifies how images are distributed a priori or, equivalently, which images occur more frequently than others. Probabilistically, it specifies the prior probability  $p(u)$ . Let  $u_0$  denote the incomplete data that are observed, measured, or sampled. Then the second part of "if" is to model how  $u_0$  is generated from  $u$  or



**Figure 1.** TV inpainting for the error concealment of a blurry image.

to specify the conditional probability  $p(u_0|u)$ . Finally, in the Bayesian framework, Shannon's "then" statement is replaced, as indicated earlier, by the *Maximum A Posteriori* (MAP) optimization given by Bayes's formula:

$$(1) \quad \max_u p(u|u_0) = p(u_0|u)p(u)/p(u_0).$$

(It is equivalent to maximizing the product of the prior model and the data model, since the denominator is a fixed normalization constant once  $u_0$  is given.) To summarize, Bayesian inpainting means finding the most probable image given its incomplete and possibly distorted observation.

The variational approach resembles the Bayesian methodology, but now everything is expressed deterministically. The Bayesian prior model  $p(u)$  becomes the specification of the regularity of an image  $u$ , while the data model  $p(u_0|u)$  now measures how well the observation  $u_0$  fits if the original image is indeed  $u$ . Regularity is enforced through "energy" functionals: for example, the Sobolev norm  $E[u] = \int_{\Omega} |\nabla u|^2 dx$ , the total variation (TV) model  $E[u] = \int_{\Omega} |Du|$  of Rudin, Osher, and Fatemi, and the Mumford-Shah free-boundary model  $E[u, \Gamma] = \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \beta H^1(\Gamma)$ , where  $H^1$  denotes the one-dimensional Hausdorff measure. The quality of data fitting  $u \rightarrow u_0$  is often judged by an error measure  $E[u_0|u]$ . For instance, the *least square* measure prevails in the literature due to the genericity of Gaussian-type noise and the *Central Limit Theorem*:  $E[u_0|u] = \frac{1}{|D|} \int_D (Tu - u_0)^2 dx$ , where  $D$  is the domain on which  $u_0$  has been sampled or measured,  $|D|$  is its area (or cardinality for the discrete case), and  $T$  denotes any linear or nonlinear image processor (such as blurring and diffusion). In this variational setting, Shannon's "then" statement becomes a constrained optimization problem:

$$\min E[u] \quad \text{over all } u \text{ such that } E[u_0|u] \leq \sigma^2.$$

Here  $\sigma^2$  denotes the variance of the white noise, which is assumed to be known by proper statistical



Figure 2. Mumford-Shah inpainting for text removal.

estimators. Equivalently, the model solves the following unconstrained problem using Lagrange multipliers (e.g., Chambolle and Lions):

$$(2) \quad \min_u E[u] + \lambda E[u_0|u].$$

Generally,  $\lambda$  expresses the balance between regularity and fitting. In summary, variational inpainting searches for the most “regular” image that best fits the given observation.

The Bayesian approach is more universal in the sense of allowing general statistical prior and data models, and it is powerful for restoring both artificial images and natural images (or *textures*). But to learn the prior model and the data model is usually quite expensive. The variational approach is ideal for dealing with regularity and geometry and tends to work best for man-made indoor and outdoor scenes and images with low textures. The two approaches (1) and (2) can be at least formally unified under Gibbs’s formula in statistical mechanics:

$$(3) \quad E[\cdot] \propto -\beta \log p(\cdot), \quad \text{or} \quad p(\cdot) \propto e^{-E[\cdot]/\beta},$$

where  $\beta = kT$  is the product of the Boltzmann constant and temperature, and  $\propto$  means equality up to a multiplicative or additive constant. (However, the definability of a rigorous probability measure over “all” images is highly nontrivial because of the multiscale nature of images. Recent efforts can be found in the work of Mumford and Gidas.)

### Variational Inpainting Based on Geometric Image Models

In a typical image inpainting problem,  $u_0$  denotes the observed or measured incomplete portion of a clean “good” image  $u$  on the entire image domain  $\Omega$ . A simplified but already very powerful data model in various digital applications is blurring followed by noise degradation and spatial restriction:

$$u_0|_D = (Ku + n)_D,$$

where  $K$  is a continuous blurring kernel, often assumed to be linear or even shift-invariant, and  $n$  is an additive white noise field assumed to be close to Gaussian for simplicity. The information  $u_0|_{\Omega \setminus D}$  is missing or inaccessible. The goal of inpainting is to reconstruct  $u$  as faithfully as possible from  $u_0|_D$ .

The data model is explicitly given by

$$(4) \quad E[u_0|u, D] = \frac{1}{|D|} \int_D (Ku - u_0)^2 dx.$$

Therefore, from the variational point of view, the quality of an inpainting model crucially depends on the prior model or the regularity energy  $E[u]$ .

The TV prior model  $E[u] = \int_\Omega |Du|$  was first introduced into image processing by Rudin, Osher, Fatemi in [15]. Unlike the Sobolev image model  $E_2[u] = \int_\Omega |\nabla u|^2 dx$ , the TV model recognizes one of the most important vision features, the “edges”. For example, for a cartoon image  $u$  showing the night sky ( $u = 0$ ) with a full bright moon ( $u = 1$ ), the Sobolev energy blows up, while the TV energy  $\int_\Omega |Du|$  equals the perimeter of the moon, which is finite. Therefore, in combination with the data model (4), the variational TV inpainting model minimizes

$$(5) \quad E_{IV}[u|u_0, D] = \alpha \int_\Omega |Du| + \lambda \int_D (Ku - u_0)^2 dx.$$

The admissible space is  $BV(\Omega)$ , the Banach space of all functions with bounded variation. It is very similar to the celebrated TV restoration model of Rudin, Osher, and Fatemi [15]. In fact, the beauty and power of the model exactly lie in the provision of a unified framework for denoising, deblurring, and image reconstruction from incomplete data. Figure 1 displays the computational output of the model applied to a blurry image with simulated random packet loss due to the transmission failure of a network.

The second well-known prior model is the object-edge model of Mumford and Shah [13]. The edge set  $\Gamma$  is now explicitly singled out, unlike in the TV model, and an image  $u$  is understood as a combination of both the geometric feature  $\Gamma$  and the piecewise smooth “objects”  $u_i$  on all the connected components  $\Omega_i$  of  $\Omega \setminus \Gamma$ . Thus in both the Bayesian and the variational languages, the prior model consists of two parts (applying (3) for the transition between probability and “energy”):

$$p(u, \Gamma) = p(u|\Gamma)p(\Gamma) \quad \text{and} \\ E[u, \Gamma] = E[u|\Gamma] + E[\Gamma].$$

In the Mumford-Shah model the edge regularity is specified by  $E[\Gamma] = H^1(\Gamma)$ , the one-dimensional Hausdorff measure, or in most computational applications,  $E[\Gamma] = \text{length}(\Gamma)$ , assuming that  $\Gamma$  is Lipschitz. The smoothness of the “objects” is naturally characterized by the ordinary Sobolev norm:  $E[u|\Gamma] = \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx$ . Therefore, in combination with the data model (4), the variational inpainting model based on the Mumford-Shah prior is given by

$$(6) \quad \inf_{u, \Gamma} E_{\text{ms}}[u, \Gamma | u_0, D] = \alpha \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \beta H^1(\Gamma) + \lambda \int_D (Ku - u_0)^2 dx.$$

Figure 2 shows one application of this model for text removal. Notice that edges are preserved and smooth regions remain smooth.

Numerous applications have demonstrated that, for classical applications in denoising, deblurring, and segmentation, both the TV and the Mumford-Shah models perform sufficiently well even by the high standard of human vision. But inpainting does have special characteristics. We have demonstrated in [2], [4], [9] that for large-scale inpainting problems, high-order image models which incorporate the curvature information become necessary for more faithful visual effects.

The key to high-order geometric image models is Euler’s elastica curve model:

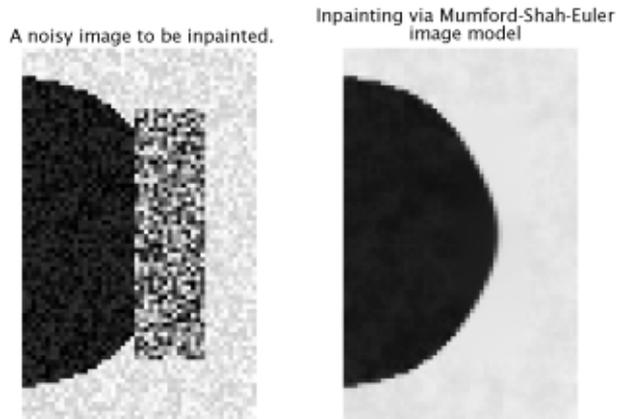
$$e[y] = \int_y (a + b\kappa^2) ds, \quad a, b > 0,$$

where  $\kappa$  denotes the scalar curvature. Birkhoff and de Boor called it the “nonlinear spline” model in approximation theory. It was first introduced into computer vision by Mumford. Unlike straight lines (for which  $b = 0$ ), the elastica model allows smooth curves because of the curvature term, which is important for computer vision and computer graphics.

By imposing the elastica energy on each individual level line of  $u$  (at least symbolically or by assuming that  $u$  is regular enough), we obtain the so-called elastica image model:

$$(7) \quad \begin{aligned} E_{\text{el}}[u] &= \int_{-\infty}^{\infty} e[u \equiv \lambda] d\lambda \\ &= \int_{-\infty}^{\infty} \int_{u=\lambda} (a + b\kappa^2) ds d\lambda \\ &= \int_{\Omega} (a + b\kappa^2) |\nabla u| dx. \end{aligned}$$

In the last integrand the curvature is given by  $\kappa = \nabla \cdot [\nabla u / |\nabla u|]$ . (Notice that in the absence of the curvature term, the above formula is exactly the *co-area* formula for smooth functions (e.g., Giusti). This elastica prior model was first studied for inpainting by Masnou and Morel, and by Chan, Kang,



**Figure 3. Smooth inpainting by the Mumford-Shah-Euler model.**

and Shen [2], and as expected it improves the TV inpainting model.

Similarly, the Mumford-Shah image model  $E_{\text{ms}}$  can be improved by replacing the length energy by Euler’s elastica energy:

$$E_{\text{mse}}[u, \Gamma] = \alpha \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + e[\Gamma].$$

This was first applied to image inpainting by Esedoglu and Shen [9]. Figure 3 shows one example of applying this image prior model to the inpainting of an occluded disk. Both the TV and Mumford-Shah inpainting models would complete the interpolation with a straight-line edge and introduce visible corners as a result. The elastica model restores the smooth boundary.

The improved performance of curvature-based models comes at a price in terms of both theory and computation. The existence and uniqueness of the TV and Mumford-Shah inpainting models can be studied in a fashion similar to the classical restoration and segmentation problems. But theoretical study on high-order models is only beginning. The difficulty lies in the involvement of the second-order geometric feature of curvature and in the identification of a proper function space to study the models. Secondly, in terms of computation, the calculus of variation on the curvature term leads to fourth-order highly nonlinear PDEs, whose fast and efficient numerical solution imposes a tremendous challenge.

We conclude this section with a brief discussion of computation, especially for the TV and Mumford-Shah inpaintings.

For the TV inpainting model  $E_{\text{tv}}$ , the Euler-Lagrange equation is formally (or assuming that  $u$  is in the Sobolev space  $W^{1,1}$ ) given by

$$(8) \quad -\nabla \cdot \left[ \frac{\nabla u}{|\nabla u|} \right] + \mu K^* \chi_D (Ku - u_0) = 0.$$

Here  $K^*$  denotes the adjoint of the linear blurring kernel  $K$ , the multiplier  $\chi_D(x)$  is the indicator of  $D$ , and  $\mu = 2\lambda/\alpha$ . The boundary condition along  $\partial\Omega$

is Neumann adiabatic to eliminate any boundary contribution during the integration-by-parts process. This nonlinear PDE can be solved iteratively by the freezing technique: if  $u^{(n)}$  denotes the current inpainting at step  $n$ , then the updated inpainting  $u^{(n+1)}$  solves the linearized PDE

$$-\nabla \cdot \left[ \frac{\nabla u^{(n+1)}}{|\nabla u^{(n)}|} \right] + \mu K^* \chi_D(Ku^{(n+1)} - u_0) = 0.$$

In practice the intermediate diffusivity coefficient  $1/|\nabla u^{(n)}|$  is often modified to  $1/\sqrt{|\nabla u^{(n)}|^2 + \epsilon^2}$  for some small conditioning parameter  $\epsilon$  or by the mandatory ceiling and flooring between  $\epsilon$  and  $1/\epsilon$ . The convergence of such algorithms has been well studied in the literature (e.g., Chambolle and Lions, and Dobson and Vogel). There are also many other possible techniques in the literature for solving (8) (e.g., Vogel and Oman, and Chan, Mulet, and Golub). We need only to relate (8) to the conventional TV restoration case.

The computation of the Mumford-Shah inpainting model is also very interesting. For inpainting, unlike segmentation, one's direct interest is only in  $u$ , not in  $\Gamma$ . Such understanding makes the  $\Gamma$ -convergence approximation theory perfect for inpainting. According to Ambrosio and Tortorelli, by introducing an edge signature function  $z(x) \in [0, 1]$ ,  $x \in \Omega$ , and having  $E[u|\Gamma] = \alpha \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx$  replaced by  $E[u|z] = \alpha \int_{\Omega} z^2 |\nabla u|^2 dx$ , one can approximate the length energy in the Mumford-Shah model by a quadratic integral in  $z$  (up to a constant multiplier):

$$E_\epsilon[z] = \beta \int_{\Omega} \left( \frac{\epsilon |\nabla z|^2}{2} + \frac{(z-1)^2}{2\epsilon} \right) dx, \quad \epsilon \ll 1.$$

Thus the Mumford-Shah inpainting model is approximated by

$$E_\epsilon[u, z|u_0, D] = E[u|z] + E_\epsilon[z] + \lambda E[u_0|u, D],$$

which is a quadratic integral in both  $u$  and  $z$ ! It leads to a coupled system of linear elliptic-type PDEs in both  $u$  and the edge signature  $z$ , which can be solved efficiently using any numerical elliptic solver. The example in Figure 2 was computed by this scheme.

Finally, we mention some of the major applications of the inpainting and geometric image interpolation models developed above. These include digital zooming, primal-sketch-based perceptual image coding, error concealment for wireless image transmission, and progressive disocclusion in computer vision [2], [4], [9]. Extensions to color or more general hyperspectral images and nonflat image features (i.e., ones that live on Riemannian manifolds) are also currently being studied in the literature. Other approaches to the inpainting problem can be found in the papers by Bertalmio, Sapiro,

Caselles, and Ballester, and by Bertalmio, Bertozzi, and Sapiro. In particular, it has been interestingly found in the latter paper that the earlier PDE model by Ballester, Bertalmio, Caselles, Sapiro, and Verdera is closely related to the stream function-vorticity equation in fluid dynamics.

## Variational Level Set Image Segmentation

Images are the proper 2-D projections of the 3-D world containing various objects. To successfully reconstruct the 3-D world, at least approximately, the first crucial step is to identify the regions in images that correspond to individual objects. This is the well-known problem of image segmentation. It has broad applications in a variety of important fields such as computer vision and medical image processing.

Denote by  $u_0$  an observed image on a 2-D Lipschitz open and bounded domain  $\Omega$ . Segmentation means finding a visually meaningful edge set  $\Gamma$  that leads to a complete partition of  $\Omega$ . Each connected component  $\Omega_i$  of  $\Omega \setminus \Gamma$  should correspond to at most one real physical object or pattern in our 3-D world, for example, the white matter in brain images or the abnormal tissues in organs. In some applications, one is interested also in the clean image patches  $u_i$  on each  $\Omega_i$  of the segmentation, since  $u_0$  is often noisy.

Therefore, there are two crucial ingredients in the mathematical modeling and computation of the segmentation problem. The first is how to formulate a model that appropriately combines the effects of both the edge set  $\Gamma$  and its segmented regions  $\{\Omega_i, i = 1, 2, \dots\}$ . The other is to find the most efficient way to represent the geometry of both the edge set and the regions and to represent the segmentation model as a result. This of course reflects the general philosophy in the introduction.

In the variational PDE approach, these two issues have found good answers in the literature: for the first, the celebrated segmentation model of Mumford and Shah [13] and for the second, the level-set representation technology of Osher and Sethian [14]. In what follows we detail our recent efforts in advancing the application of the level-set technology to various Mumford-Shah-related image segmentation models. Much of the work can be found in our papers (e.g., [3], [5], [17], [19] and many more on our group homepage [11]) and also in related works by Yezzi, Tsai, and Willsky; Paragios and Deriche; Zhu and Yuille; and Cohen, Bardinet, and Ayache [6], [7].

We start with a novel active-contour model whose formulation is independent of intensity edges defined by the gradients, in contrast to most conventional ones in the literature. We then explain how this model can be efficiently computed based on the multiphase level-set method. In the second part we extend these results to the level-set formulation and

computation of the general Mumford-Shah segmentation model for piecewise-smooth images. In the last part we present our recent work on extending the previous models to logical operations on multichannel image objects.

### Active Contours without Edges and Multiphase Level Sets

The active contour is a powerful tool in image and vision analysis for boundary detection and object segmentation. The key idea is to evolve a curve so that it eventually stops along the object edges of the given image  $u_0$ . The curve evolution is controlled by two sorts of energies: the internal energy defining the regularity of the curve and the external energy determined by the given image  $u_0$ . The latter is often called the feature-driven energy.

In almost all classical active-contour models, the feature-driven energies rely heavily on the gradient feature  $|\nabla u_0|$  or on its smoothed version  $|\nabla G_\sigma * u_0|$ , where  $G_\sigma$  denotes a Gaussian kernel with a small variance  $\sigma$ . They work well for detecting gradient-defined edges but fail for more general classes of edges such as the boundary of a nebula in some astronomical images or the top image in Figure 4.

Our new model, *active contours without edges*, first introduced in [5], is independent of the gradient information and therefore can handle more general types of edges. The model is to minimize the energy

$$(9) \quad E_2[c_1, c_2, \Gamma | u_0] = \int_{\text{int}(\Gamma)} |u_0(x) - c_1|^2 dx + \int_{\text{ext}(\Gamma)} |u_0(x) - c_2|^2 dx + \nu |\Gamma|,$$

where  $\nu$  denotes a given positive weight, the  $c$ 's are unknown constants,  $\text{int}(\Gamma)$  and  $\text{ext}(\Gamma)$  denote the interior and exterior of  $\Gamma$ , and  $|\Gamma|$  is its length. The subscript 2 in  $E_2$  indicates that it deals with two-phase images, i.e., ones whose "objects" can be completely indexed by the interior and exterior of  $\Gamma$ .

In the level-set formulation of Osher and Sethian [14],  $\Gamma$  is embedded as the zero-level set  $\{\phi = 0\}$  of a Lipschitz continuous function  $\phi : \Omega \rightarrow \mathbb{R}$ . Consequently,  $\{\phi > 0\}$  and  $\{\phi < 0\}$  define the interior  $\Omega^+$  and exterior  $\Omega^-$  of the curve. (The level-set approach is computationally superior to other curve representations, because it lets one directly work on a fixed rectangular grid and it allows automatic topological changes such as merging and breaking.) Denote by  $H$  the 1-dimensional Heaviside function:  $H(z) = 1$  if  $z \geq 0$  and 0 if  $z < 0$ . Then the energy in our model becomes

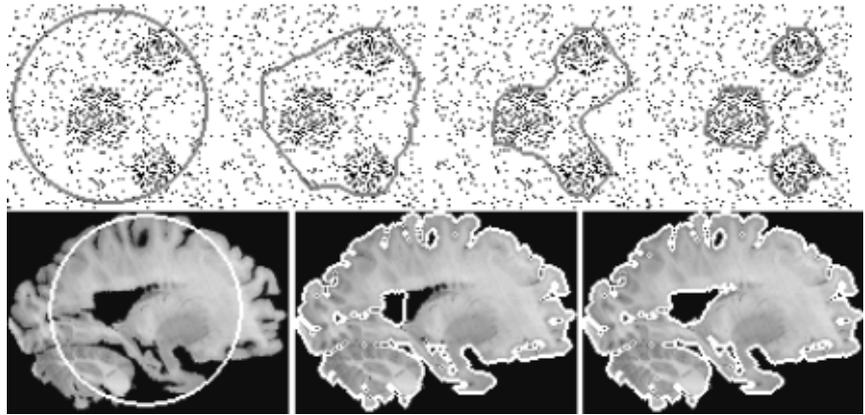


Figure 4. Top: Detection of a simulated minefield by our new active-contour model. Bottom: Segmentation of an MRI brain image. Notice that the interior boundaries are automatically detected.

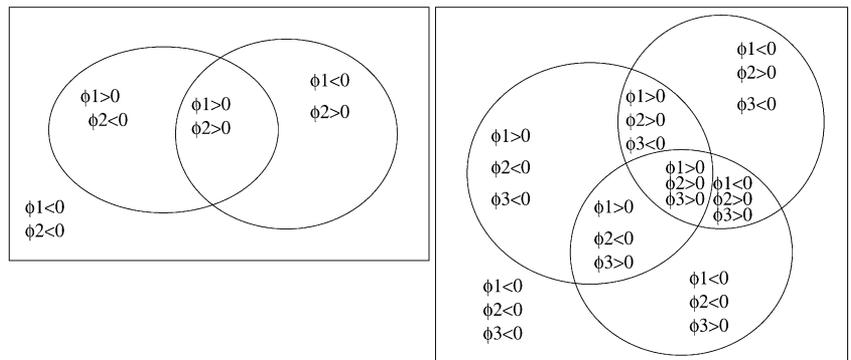


Figure 5. Left: Two curves given by  $\phi_1 = 0$  and  $\phi_2 = 0$  partition the domain into four regions based on indicator vector  $(\text{sign}(\phi_1), \text{sign}(\phi_2))$ . Right: Three curves given by  $\phi_1 = 0$ ,  $\phi_2 = 0$ , and  $\phi_3 = 0$  partition the domain into eight regions based on the triple  $(\text{sign}(\phi_1), \text{sign}(\phi_2), \text{sign}(\phi_3))$ .

$$E_2[c_1, c_2, \phi | u_0] = \int_{\Omega} |u_0(x) - c_1|^2 H(\phi) dx + \int_{\Omega} |u_0(x) - c_2|^2 (1 - H(\phi)) dx + \nu \int_{\Omega} |\nabla H(\phi)| dx.$$

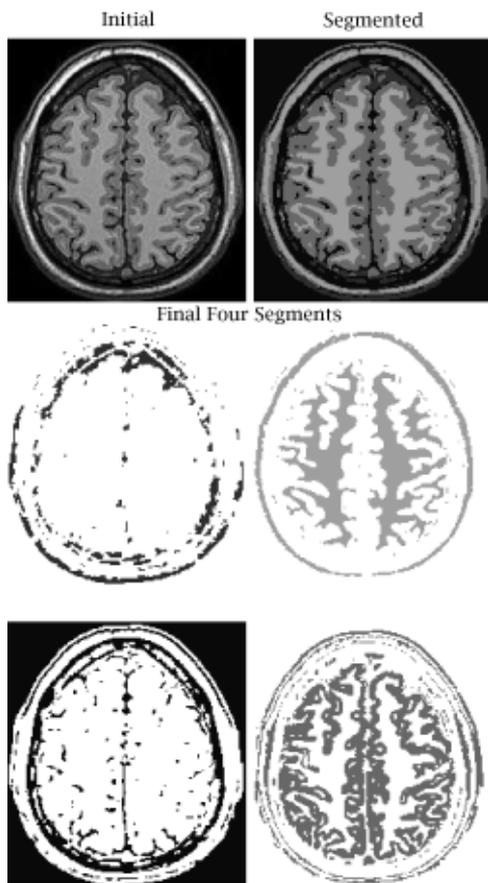
Minimizing  $E_2[c_1, c_2, \phi | u_0]$  with respect to  $c_1$ ,  $c_2$ , and  $\phi$  leads to the Euler-Lagrange equation:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[ \nu \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - |u_0 - c_1|^2 + |u_0 - c_2|^2 \right],$$

$$c_1(t) = \frac{\int_{\Omega} u_0(x) H(\phi(x)) dx}{\int_{\Omega} H(\phi(x)) dx},$$

$$c_2(t) = \frac{\int_{\Omega} u_0(x) (1 - H(\phi(x))) dx}{\int_{\Omega} (1 - H(\phi(x))) dx},$$

with a suitable initial guess  $\phi(0, x) = \phi_0(x)$ . In numerical implementations the Heaviside function  $H(z)$  is often regularized by some  $H_\varepsilon(z)$  in  $C^1(\mathbb{R})$  that converges as  $\varepsilon \rightarrow 0$  to  $H(z)$  in some suitable sense. As a result, the Dirac function  $\delta(z)$  in the last



**Figure 6. The original and segmented images (top row), and the final four segments (the rest).**

equation is regularized to  $\delta_\varepsilon(z) = H'_\varepsilon(z)$ . We have discovered in [5] that a carefully designed approximation scheme can even allow interior contours to emerge, a challenging task for most conventional algorithms. Also notice that the length term in the energy has led to the mean-curvature motion.

The model performs as an active contour in the class of piecewise-constant images taking only two values; it looks for a two-phase segmentation of a given image. The internal energy is defined by the length, while the external energy is independent of the gradient  $|\nabla u_0|$ . Defining the segmented image by  $u(x) = c_1 H(\phi(x)) + c_2(1 - H(\phi(x)))$ , we realize that the energy model is exactly the Mumford-Shah segmentation model [13] restricted to the class of piecewise-constant images. However, our model was initially developed from the active-contour point of view.

Two typical numerical outputs of the model are displayed in Figure 4. The top row shows that our model can segment and detect objects without clear gradient edges. The bottom one shows that it can also capture complicated boundaries and interior contours.

For more complicated situations where multiple objects occlude each other and multiphase edges such as T-junctions emerge, the above two-phase

active-contour model is insufficient, and we need to introduce multiple level-set functions. Therefore, we have generalized the above framework to multiphase active contours or, equivalently, the piecewise-constant Mumford-Shah segmentation with multiphase regions:

$$(10) \quad \inf_{u, \Gamma} E_{\text{ms}}[u, \Gamma | u_0] = \sum_i \int_{\Omega_i} |u_0 - c_i|^2 dx + \nu |\Gamma|.$$

Here the  $\Omega_i$ 's denote the connected components of  $\Omega \setminus \Gamma$ , and  $u = c_i$  on  $\Omega_i$ . Notice that  $\Gamma$  can now be a general set of edge curves, including for example the T-junction class.

Generally, consider  $m$  level-set functions  $\phi_i : \Omega \rightarrow \mathbb{R}$ . The union of the zero-level sets of the  $\phi_i$  represents the edges in the segmented image. Using these  $m$  level-set functions, one can define up to  $n = 2^m$  phases, which form a disjoint and complete partitioning of  $\Omega$ . Therefore, each point  $x \in \Omega$  belongs to one and only one phase. In particular, there is no vacuum or overlap among the phases. This is an important advantage compared with the classical multiphase representation, where a level-set function is associated to each phase and more level-set functions are needed as a result. Figure 5 shows two typical examples of multiphase partitioning corresponding to  $m = 2$  and  $m = 3$ .

We now illustrate the multiphase level-set approach through the example of  $n = 4$  and  $m = 2$ . Let  $c = (c_{11}, c_{10}, c_{01}, c_{00})$  denote a constant vector and  $\Phi = (\phi_1, \phi_2)$  the two-phase level-set vector. Then we are looking for an ideal image  $u$  in the form of

$$u = c_{11}H(\phi_1)H(\phi_2) + c_{10}H(\phi_1)(1-H(\phi_2)) + c_{01}(1-H(\phi_1))H(\phi_2) + c_{00}(1-H(\phi_1))(1-H(\phi_2)).$$

The Mumford-Shah segmentation energy becomes

$$(11) \quad E_4[c, \Phi | u_0] = \int_{\Omega} |u_0(x) - c_{11}|^2 H(\phi_1)H(\phi_2) dx + \int_{\Omega} |u_0(x) - c_{10}|^2 H(\phi_1)(1 - H(\phi_2)) dx + \int_{\Omega} |u_0(x) - c_{01}|^2 (1 - H(\phi_1))H(\phi_2) dx + \int_{\Omega} |u_0(x) - c_{00}|^2 (1 - H(\phi_1))(1 - H(\phi_2)) dx + \nu \int_{\Omega} |\nabla H(\phi_1)| dx + \nu \int_{\Omega} |\nabla H(\phi_2)| dx.$$

Its minimization leads to the Euler-Lagrange equations. First, with  $\Phi$  fixed, the  $c$  minimizer can be explicitly worked out as before:

$$c_{ij}(t) = \text{average of } u_0 \text{ on} \\ \{(2i - 1)\phi_1 > 0, (2j - 1)\phi_2 > 0\}, \\ i, j = 0, 1.$$

In turn, this new  $c$  information leads to the Euler-Lagrange equations for  $\Phi$ :

$$\frac{\partial \phi_1}{\partial t} = \delta(\phi_1) \left[ \nu \operatorname{div} \left( \frac{\nabla \phi_1}{|\nabla \phi_1|} \right) \right. \\ \left. - \left( (u_0 - c_{11})^2 - (u_0 - c_{01})^2 \right) H(\phi_2) \right. \\ \left. - \left( (u_0 - c_{10})^2 - (u_0 - c_{00})^2 \right) (1 - H(\phi_2)) \right],$$

$$\frac{\partial \phi_2}{\partial t} = \delta(\phi_2) \left[ \nu \operatorname{div} \left( \frac{\nabla \phi_2}{|\nabla \phi_2|} \right) \right. \\ \left. - \left( (u_0 - c_{11})^2 - (u_0 - c_{01})^2 \right) H(\phi_1) \right. \\ \left. - \left( (u_0 - c_{10})^2 - (u_0 - c_{00})^2 \right) (1 - H(\phi_1)) \right].$$

Notice that the equations are governed both by the mean curvatures and by jumps of the data-energy terms across the boundary.

Figure 6 shows an application of the model to the medical analysis of a brain image. Displayed are the final segmented image and its associated four phases. Our model successfully identifies and segments the white and the gray matters.

Recently the above models and algorithms have been extended to multichannel, volumetric, and texture images (e.g., Chan, Sandberg, and Vese [3]). Let us give a little more detail about texture segmentation from our work. Texture images are general images of natural scenes, such as grasslands, beaches, rocks, mountains, and human body tissues. They typically carry certain coherent structures in scales, orientations, and local frequencies. To segment texture images using the above models, we first apply Gabor's filters to extract these coherent structures. The filter responses create a new vectorial (or multichannel) feature image in the form of  $U(x) = (u_\alpha(x), u_\beta(x), \dots, u_\gamma(x))$ , where the Greek letters stand for the filter signatures, and typically each takes a value of (scale, orientation, local frequency). We then apply the vectorial active-contour-without-edges model to the segmentation of  $U$ . Figure 7 shows one typical example.

#### Piecewise-Smooth Mumford-Shah Segmentation

The most general Mumford-Shah piecewise-smooth segmentation [13] is defined by

$$(12) \quad \inf_{u, \Gamma} E_{\text{ms}}[u, \Gamma | u_0] = \int_{\Omega} |u - u_0|^2 dx \\ + \mu \int_{\Omega \setminus \Gamma} |\nabla u|^2 dx + \nu |\Gamma|,$$

where  $\mu$  and  $\nu$  are positive parameters. It allows the segmented "objects" to have smoothly varying intensities instead of being strictly constant. We

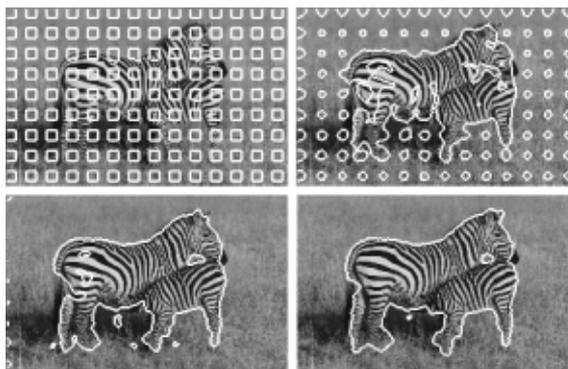


Figure 7. An example of texture segmentation (at increasing times).

now show how to carry out the model based on the multiphase level-set approach [5]. As before, we start with the two-phase situation where a single level-set function  $\phi$  is sufficient, followed by the more general multiphase case.

In the two-phase situation, the ideal image  $u$  is segmented to  $u^\pm$  by the level-set function  $\phi$ :

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

We assume that both  $u^+$  and  $u^-$  are  $C^1$  functions up to the boundary  $\{\phi = 0\}$ . Substituting this expression into (12), we obtain

$$(13) \quad E[u^+, u^-, \phi | u_0] = \int_{\Omega} |u^+ - u_0|^2 H(\phi) dx \\ + \int_{\Omega} |u^- - u_0|^2 (1 - H(\phi)) dx \\ + \mu \int_{\Omega} |\nabla u^+|^2 H(\phi) dx \\ + \mu \int_{\Omega} |\nabla u^-|^2 (1 - H(\phi)) dx + \nu \int_{\Omega} |\nabla H(\phi)|.$$

First, with  $\phi$  fixed, the variation on  $E[u^+, u^-, \phi | u_0]$  leads to the two Euler-Lagrange equations for  $u^\pm$  separately:

$$(14) \quad u^\pm - u_0 = \mu \Delta u^\pm \quad \text{on } \pm \phi > 0, \\ \frac{\partial u^\pm}{\partial \vec{n}} = 0 \quad \text{on } \{\phi = 0\}.$$

(Here  $\pm$  takes either of the values  $+$  and  $-$ , but uniformly across the formula.) They act as denoising operators on the homogeneous regions only. Notice that no smoothing is done *across* the boundary  $\{\phi = 0\}$ , which is very important in image analysis.

Next, keeping the functions  $u^+$  and  $u^-$  fixed and minimizing  $E[u^+, u^-, \phi | u_0]$  with respect to  $\phi$ , we obtain the motion of the zero-level set:

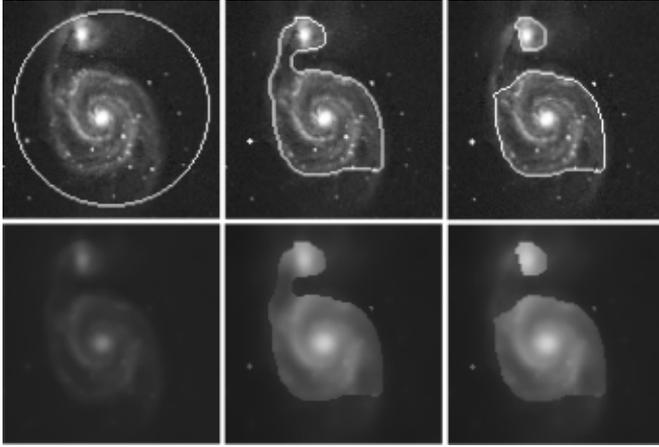


Figure 8. Numerical result from the piecewise-smooth Mumford-Shah level-set algorithm with one level-set function.

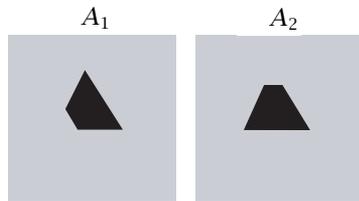


Figure 9. A synthetic example of an object in two different channels. Notice that the lower left corner of  $A_1$  and the upper corner of  $A_2$  are missing.

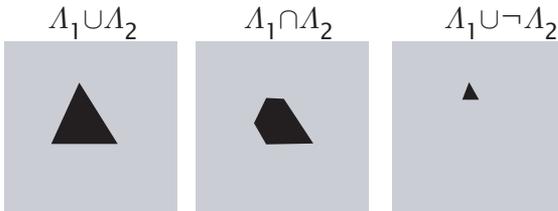


Figure 10. Different logical combinations for the sample image: the union, the intersection, and the differentiation.

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[ \nu \nabla \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - (|u^+ - u_0|^2 + \mu |\nabla u^+|^2 - |u^- u_0|^2 - \mu |\nabla u^-|^2) \right],$$

with some initial guess  $\phi(t = 0, x)$ . The above equation is actually computed at least near a narrow band of the zero-level set. As a result, computationally we have to continuously extend both  $u^+$  and  $u^-$  from their original domains  $\{\pm \phi > 0\}$  to a suitable neighborhood of the zero-level set  $\{\phi = 0\}$ . Figure 8 displays an application of the model in astronomical image analysis. Although the nebula itself does not seem to be a smooth object, the piecewise-smooth model can still correctly capture the main features.

As in the previous section, there are cases where the boundaries forming a complete partition of the image cannot be represented by a single level-set function. Then one has to turn to the multi-phase approach. In our papers, thanks to the planar Four-Color Theorem, we have been able to conclude that *two* level-set functions are sufficient for all multi-phase partition problems.

By the Four-Color Theorem one can color all the regions in a partition using only four colors, so that any two adjacent regions are color distinguishable. Identifying a phase with one color, we see that two level-set functions  $\phi_1$  and  $\phi_2$  are sufficient to produce four “colors”:  $\{\pm \phi_1 > 0, \pm \phi_2 > 0\}$ . Therefore, they can completely segment a general image with a multiphase boundary set  $\Gamma$  given by  $\{\phi_1 = 0\}$  or  $\{\phi_2 = 0\}$ . As before, we do not have the problems of “overlapping” or “vacuum” as in the works by Zhao, Chan, Merriman, and Osher. Note that in this formulation, generally each “color” can still have many isolated components. Therefore, the segmentation is complete only after one applies an extra step of the well-known topological processor for finding the connected components of an open set.

In this four-phase formulation, the ideal image  $u$  is segmented into four disjoint but complete parts  $u^{\pm\pm}$ , each defined by one of the four phases:

$$\{\pm \phi_1 > 0, \pm \phi_2 > 0\}.$$

Overall, by using the Heaviside function, we obtain the following synthesis formula:

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

for all  $x \in \Omega$ . We can express the energy function of  $u$  and  $\Phi = (\phi_1, \phi_2)$  in a similar way and derive the corresponding Euler-Lagrange equations.

Notice the remarkable feature of this single model, which includes both the original energy formulation and the elliptic and evolutionary PDEs: it naturally combines all three image processors—active contour, segmentation, and denoising.

### Logic Operators for Multichannel Image Segmentation

In a multichannel image  $u(x) = (u_1(x), u_2(x), \dots, u_n(x))$ , a single physical object can leave different traces in different channels. For example, Figure 9 shows a two-channel image containing a triangle that is, however, incomplete in each individual channel. For this example, most conventional segmentation models for multichannel images (e.g., Guichard, Sapiro, Zhu and Yuille) would output the complete triangle, i.e., the union of both channels. The union is just one of the several possible logical operations for multichannel images. For example, the intersection and the

differentiation are also very common in applications, as illustrated in Figure 10.

In this section we outline our recent efforts in developing logical segmentation schemes for multi-channel images based on the active-contour-without-edges model [16].

First, we define two logical variables to encode the information inside and outside the contour  $\Gamma$  separately for each channel  $i$ :

$$z_i^{in}(u_0^i, x, \Gamma) = \begin{cases} 1, & \text{if } x \text{ is inside } \Gamma \text{ and not on} \\ & \text{the object,} \\ 0, & \text{otherwise;} \end{cases}$$

$$z_i^{out}(u_0^i, x, \Gamma) = \begin{cases} 1 & \text{if } x \text{ is outside } \Gamma \text{ and on} \\ & \text{the object,} \\ 0 & \text{otherwise.} \end{cases}$$

Such different treatments are motivated by the energy minimization formulation. Intuitively speaking, in order for the active contour  $\Gamma$  to evolve and eventually capture the *exact* boundary of the targeted logical object, the energy should be designed so that both partial capture and overcapture lead to high energies (corresponding to  $z_i^{out} = 1$  and  $z_i^{in} = 1$  separately). Imagine that the target object is tumor tissue: then in terms of decision theory, over and partial captures correspond respectively to *false alarms* and *misses*. Both are to be penalized.

In practice we do not have precise information of “the object” to be segmented. One possible way to approximate  $z_i^{in}$  and  $z_i^{out}$  is based on the interior ( $\Omega^+$ ) and exterior ( $\Omega^-$ ) averages  $c_i^\pm$  in channel  $i$ :

$$z_i^{in}(u_0^i, x, \Gamma) = \frac{|u_0^i(x) - c_i^+|^2}{\max_{y \in \Omega^+} |u_0^i(y) - c_i^+|^2},$$

for  $x \in \Omega^+$ , and

$$z_i^{out}(u_0^i, x, \Gamma) = \frac{|u_0^i(x) - c_i^-|^2}{\max_{y \in \Omega^-} |u_0^i(y) - c_i^-|^2},$$

for  $x \in \Omega^-$ .

The desired truth table can then be described using the  $z_i^{in}$ 's and  $z_i^{out}$ 's. Table 2 shows three examples of logical operations for the two-channel case. Notice that “true” is represented by 0 inside  $\Gamma$ . The method is designed so as to encourage energy minimization when the contour tries to capture the targeted object inside. Also note that the “ $z_i^{in}$ ” terms and the “ $z_i^{out}$ ” terms play asymmetric but complementary roles. For example, the union  $A_1 \cup A_2$  corresponds to the union of the “in” terms and the *intersection* of the “out” terms. Similarly, the intersection  $A_1 \cap A_2$  corresponds to the intersection of the “in” terms and the *union* of the “out” terms.

We then design continuous objective functions to smoothly interpolate the binary truth table. This

Truth table for the two-channel case							
	$z_1^{in}$	$z_2^{in}$	$z_1^{out}$	$z_2^{out}$	$A_1 \cup A_2$	$A_1 \cap A_2$	$A_1 \cap \neg A_2$
x inside $\Gamma$ (or $x \in \Omega^+$ )	1	1	0	0	1	1	1
	1	0	0	0	0	1	1
	0	1	0	0	0	1	0
	0	0	0	0	0	0	1
x outside $\Gamma$ (or $x \in \Omega^-$ )	0	0	1	1	1	1	0
	0	0	1	0	1	0	1
	0	0	0	1	1	0	0
	0	0	0	0	0	0	0

**Table 2. The truth table for two channels. Notice that inside  $\Gamma$  “true” is represented by 0. It is designed so as to encourage the contour to enclose the targeted logical object at a lower energy cost.**

is because in practice, as mentioned above, the  $z$ 's are approximated and take continuous values. For example, possible interpolants for the union and intersection are

$$f_{A_1 \cup A_2}(x) = \sqrt{z_1^{in}(x)z_2^{in}(x)} + \left(1 - \sqrt{(1 - z_1^{out}(x))(1 - z_2^{out}(x))}\right),$$

$$f_{A_1 \cap A_2}(x) = 1 - \sqrt{(1 - z_1^{in}(x))(1 - z_2^{in}(x))} + \sqrt{z_1^{out}(x)z_2^{out}(x)}.$$

The square roots are taken to keep the functions of the same order as the original scalar models. It is straightforward to extend the two-channel case to more general  $n$ -channel ones.

The energy functional  $E$  for the logical objective function  $f$  can be expressed by the level set function  $\phi$ . Generally, as just shown above, the objective function can be separated into two parts,

$$f = f(z_1^{in}, z_1^{out}, \dots, z_n^{in}, z_n^{out}) = f_{in}(z_1^{in}, \dots, z_n^{in}) + f_{out}(z_1^{out}, \dots, z_n^{out}).$$

The energy functional is then defined by

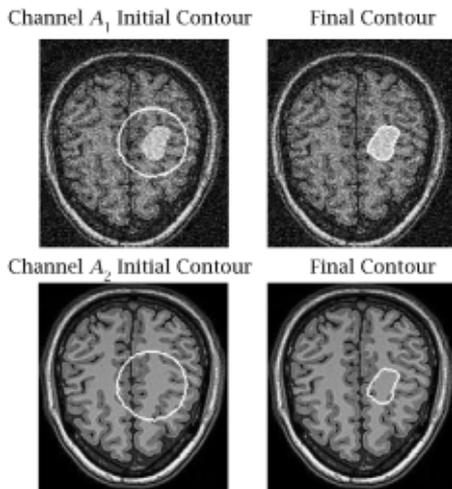
$$E[\phi|c^+, c^-] = \mu \text{length}(\phi = 0) + \lambda \int_{\Omega} [f_{in}(z_1^{in}, \dots, z_n^{in})H(\phi) + f_{out}(z_1^{out}, \dots, z_n^{out})(1 - H(\phi))] dx.$$

Here each  $c^\pm = (c_1^\pm, \dots, c_n^\pm)$  is in fact a multi-channel vector. The associated Euler-Lagrange equation is similar to the scalar model:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[ \mu \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \lambda (f_{in}(z_1^{in}, \dots, z_n^{in}) - f_{out}(z_1^{out}, \dots, z_n^{out})) \right],$$

with suitable boundary conditions as before. Even though the form often looks complicated for a typical application, its implementation is very similar to that of the scalar model.

Numerical results support our above efforts. Figure 9 shows two different occlusions of a triangle.



**Figure 11. Region-based logical model on a medical image. In the first channel,  $A_1$ , the noisy image has a “brain tumor”, while channel  $A_2$  does not. The goal is to spot the tumor that is in channel  $A_1$  but not in  $A_2$ , i.e., the differentiation  $A_1 \cap \neg A_2$ . In the right-hand column we observe that the tumor has been successfully captured.**

## Conclusion

In this article we have discussed some recent developments in one successful approach to mathematical image and vision analysis, the variational PDE method. Besides the inpainting and segmentation problems discussed here, some other problems for which this method is well suited are adaptive image enhancement and scale-space theory, geometric processing of curves and surfaces, optical flows of motion pictures, and dynamic object tracking. Advantages of the method include faithful modeling and processing of vision geometry and its related visual optimization, effective simulation of dynamic visual processes such as selective diffusion and information transport, and close interaction with the rich literature of numerical analysis and computational PDEs. This subject shows that mathematics has a key role to play in addressing real-world problems in science and technology. Some challenges for the future are further theoretical study on the variational and PDE models developed in recent years, more intrinsic integration with stochastic modeling and applied harmonic analysis such as geometric wavelets, and more systematic investigation on the computation and numerical analysis of geometry-based variational optimizations and PDEs.

## Acknowledgments

Jackie Shen thanks his former advisors Gilbert Strang and Stan Osher for their constant support and inspiration.

## References

[1] G. AUBERT and P. KORNPÖBST, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus*

We are able to successfully recover the union, the intersection, and the differentiation of the objects in Figure 10 using our model. In Figure 11 we have a two-channel image of the brain. In one channel we have a “tumor” with some noise, while the other channel is clear. The images are not registered. We want to find  $A_1 \cap \neg A_2$  so that the tumor can be observed. This happens to be a very complicated example, as there are a lot of features and textures. However, the model finds the tumor successfully.

*of Variations*, Appl. Math. Sci., vol. 147, Springer-Verlag, 2001.

[2] T. F. CHAN, S.-H. KANG, and J. SHEN, Euler’s elastica and curvature based inpaintings, *SIAM J. Appl. Math.* (2002), in press.

[3] T. F. CHAN, B. SANDBERG, and L. VESE, Active contours without edges for vector-valued images, *J. Visual Comm. Image Rep.* **11** (1999), 130–41.

[4] T. F. CHAN and J. SHEN, Mathematical models for local nontexture inpaintings, *SIAM J. Appl. Math.* **62** (2001), 1019–43.

[5] T. F. CHAN and L. A. VESE, Active contours without edges, *IEEE Trans. Image Process.* **10** (2001), 266–77.

[6] L. COHEN, Avoiding local minima for deformable curves in image analysis, *Proc. 3rd. Internat. Conf. on Curves and Surfaces (Chamonix-Mont Blanc, 1996)*, Vol. 1, (A. Le Méhauté, C. Rabut, and L. L. Schumaker, eds.), Vanderbilt Univ. Press, Nashville, TN, 1997, pp. 77–84.

[7] L. COHEN, E. BARDINET, and N. AYACHE, Surface reconstruction using active contour models, *Proc. SPIE Conf. on Geometric Methods in Computer Vision*, SPIE, Bellingham, WA, 1993.

[8] M. G. CRANDALL, H. ISHII, and P. L. LIONS, User’s guide to viscosity solutions of second order partial differential equations, *Bull. Amer. Math. Soc. (N.S.)* **27** (1992), 1–67.

[9] S. ESEDOGLU and J. SHEN, Digital inpainting based on the Mumford-Shah-Euler image model, *European J. Appl. Math.* **13** (2002), 353–70.

[10] S. GEMAN and D. GEMAN, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.* **6** (1984), 721–41.

[11] UCLA Imagers homepage, <http://www.math.ucla.edu/~imagers/>.

[12] S. MALLAT, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[13] D. MUMFORD and J. SHAH, Optimal approximations by piecewise smooth functions and associated variational problems, *Comm. Pure Applied. Math.* **42** (1989), 577–685.

[14] S. OSHER and J. SETHIAN, Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulation, *J. Comput. Phys.* **79** (1988), 12–49.

[15] L. RUDIN, S. OSHER, and E. FATEMI, Nonlinear total variation based noise removal algorithms, *Physica D* **60** (1992), 259–68.

[16] B. SANDBERG and T. F. CHAN, Logic operations for active contours on multi-channel images, UCLA Department of Mathematics CAM Report 02-12, 2002.

[17] B. SANDBERG, T. F. CHAN, and L. VESE, A level-set and Gabor-based active contour algorithm for segmenting textured images, UCLA Department of Mathematics CAM report 02-39, 2002.

[18] G. SAPIRO, *Geometric Partial Differential Equations and Image Processing*, Cambridge University Press, 2001.

[19] L. A. VESE and T. F. CHAN, A multiphase level set framework for image segmentation using the Mumford and Shah model, UCLA Department of Mathematics CAM Report 01-25, to appear in *Internat. J. Comp. Vision* (2001).