Curious Snakes: A Minimum Latency Solution
to the Cluttered Background Problem in Active Contours

Ganesh Sundaramoorthi
University of California, Los Angeles
ganeshs@ucla.edu

Stefano Soatto
University of California, Los Angeles
soatto@ucla.edu

Anthony J. Yezzi
Georgia Institute of Technology
ayezzi@ece.gatech.edu

Abstract

We present a region-based active contour detection algorithm for objects that exhibit relatively homogeneous photometric characteristics (e.g. smooth color or gray levels), embedded in complex background clutter. Current methods either frame this problem in Bayesian classification terms, where precious modeling resources are expended representing the complex background away from decision boundaries, or use heuristics to limit the search to local regions around the object of interest. We propose an adaptive lookout region, whose size depends on the statistics of the data, that are estimated along with the boundary during the detection process. The result is a "curious snake" that explores the outside of the decision boundary only locally to the extent necessary to achieve a good tradeoff between missed detections and narrowest "lookout" region, drawing inspiration from the literature of minimum-latency set-point change detection and robust statistics. This development makes fully automatic detection in complex backgrounds a realistic possibility for active contours, allowing us to exploit their powerful geometric modeling capabilities compared with other approaches used for segmentation of cluttered scenes. To this end, we introduce an automatic initialization method tailored to our model that overcomes one of the primary obstacles in using active contours for fully automatic object detection.

1. Introduction

In [16], Mumford and Shah proposed a model of images with piecewise smooth statistics. While simplistic in terms of its generative power, this model has proven useful for discriminative purposes, and has served as a basis for image segmentation algorithms implemented using Level Set

Methods [17]. Indeed, [6] consider an even simpler model, involving piecewise constant image statistics, that is used in fields ranging from medical image analysis to forensics and entertainment. The power of this method lies in the explicit representation of the null and alternate hypotheses (foreground/background), that compete for the decision without the need for ad-hoc thresholding, a mechanism exploited in *region competition* [25]. This model has also been extended to color [2], texture [22, 19, 20], and motion [9].

The strengths of this model, however, turn into limitations when the underlying assumptions are stretched. Fig. 1 illustrates this phenomenon. The object of interest (called "foreground") often has homogeneous (constant or smooth) statistics, *e.g.* the heart chambers in Fig. 1. However, the remainder of the image (the "background") is certainly not well approximated by a constant gray value. As a result, detecting the boundaries of the chambers, a seemingly easy task, is hampered by the structure of the background, which ends up influencing the boundary more than the characteristics of the object of interest. Of course, any reasonable image (for instance, one that is $\mathbb{L}^2$-integrable) can be approximated arbitrarily well with a collection of constant functions, so the model [6] is still valid, but not with *two* (foreground/background), but rather with a larger number of regions, most of which unrelated to the object of interest (Fig. 1 bottom-right). While techniques to extend [6] to multiple "phases" have been proposed, they are cumbersome, reflecting fact that, for anything more than two regions, the optimization problem associated with [6] is no longer convex [5].[1] The alternative to a fine partition of the image domain linked to simple functions is a coarser partition (*e.g.* foreground/background) with respect to more

---

[1]This reflects the fact that, with only one phase transition, the classification problem is masked as a regression problem: One is not looking for whether the object of interest is in the image. One is *told* so, and the question becomes that of localizing the object, or determining its boundary.

complex functions, say smooth functions, leading us back to the Mumford-Shah functional. In either case, much of the modeling power is devoted to the background (Fig. 1). These difficulties have prompted practitioners to devise heuristics, for instance restricting the domain of the image to a box around the curve, or to a dilation of the curve by an amount proportional to the area within [2]. The shortcomings of these approaches are obvious, as there is no reason why the structure of the boundary of a region should depend on how large the region is. For the purpose of detecting a dis-
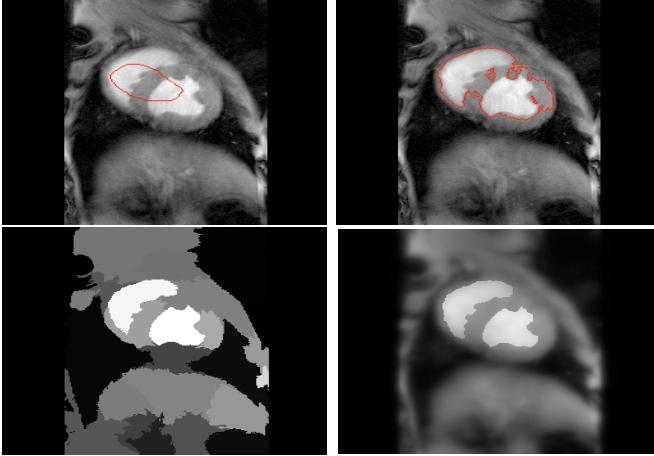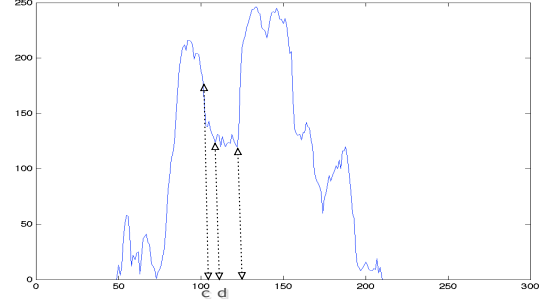

Figure 2. *One scanline from Fig.* 1*: The detection of the boundary* c *should be performed as soon as possible,* d*, so as not to have irrelevant background impinge on the decision (past the right-most dashed line).*

Figure 1. *Segmentation of a heart chamber using [6] (top-right, red curve), starting from the initial condition (top-left), is impeded by the fact that the background does not fit the constant model. Extension to multi-phase segmentation (bottom-left, each region is color-coded, and the object of interest corresponds to the white region) is complex and highly non-convex. Extension to more complex models, such as [16] (bottom-right) is also laborious (Fig. 4). In both cases, precious modeling and computational resources are expended to capture the structure of the background away from the object of interest.*

continuity, that is a classification task, a generative model is useful only insofar as it determines the classification boundary. Modeling energy expended to represent the distribution away from the decision boundary is all but wasted [24]. In our case, the distributions of intensity values specifying the null and alternate hypotheses (foreground/background) are defined in the co-domain of the image (intensity, color, texture descriptors etc.), but the decision boundary is defined on its domain, $\Omega \subset D$. This is the situation considered in the problem of set-point change detection, which, for one-dimensional causal signals, is treated as an optimal stopping time [11], using the mathematics of filtrations and Martingales.

Robust Statistical methods are designed precisely for this situation, when the null hypothesis (the object of interest)

fits a rather simple model, and one is not interested in characterizing the alternate hypothesis, *i.e.* the statistics of the background, other than the fact that they do *not* fit the foreground model. Thus one could approach the problem in Fig. 1 by seeking some kind of robust detector, which would inevitably involve sensitive thresholds that may not survive the large variability of data customary, for instance, in medical image analysis.

Consider, for example, a one-dimensional scan-line from Fig. 1, where the boundary of interest is c. Clearly, the statistics of the signal far away from c are of little interest in determining the location c (Fig. 2). Therefore, a hypothetical traveler traversing the signal would seek to localize the transition c with the smallest possible latency, $d - c$, for a given level of false alarm rate. For cases when the statistics of the signal are known, the solution to this problem is provided in [23]. Unfortunately, we do not know the statistics. If, on the other hand, d was given to us, but not the statistics, we could use [6]. Unfortunately, we do not know d.

To develop some intuition, let us consider an even simpler version of the one-dimensional example, whereby the statistics before and after the transition are Gaussian, just with different means $\mu, \bar{\mu}$ and standard deviations $\sigma, \bar{\sigma}$. It is immediately clear that if σ is very small and the "jump" $\mu - \bar{\mu}$ is large relative to the standard deviations $\sigma_1 + \sigma_2$, it can be safely detected with a small latency $d - c$. In the limit where $\sigma \to 0$, when the function is continuous, it can be detected instantaneously. If, on the other hand, the deviation of the signal from its mean before the jump is high, it will take longer to integrate the statistics and realize that there has indeed been a transition. This suggests making the area of the "lookout region", which is the region immediately *outside* the object of interest, dependent on the statistics of the image *inside*.[3] We call the ensuing model, which we

---

[2]For example, in the work of [14] the size of the dilated region is a parameter of the algorithm.

[3]In principle, it should depend on the ratio between the gap between the means, and the sum of the standard deviations, but this would lead to a model that is too complex to optimize, so we restrict to the model

describe in the next section, the *"Curious Snake."* It is curious, but cautious, for it peaks over the edge far enough to be sure, but not too far as to run into trouble due to the statistics of the background away from the transition of interest.

1.1. Non-interacting multiple-phase variational region segmentation

This deceivingly simple adaptation makes a significant difference in applications, as exemplified in Fig. 1, where one is interested in objects with relatively simple photometry embedded in complex clutter, without having to spend resources modeling the clutter. We use the plural "objects," because this technique, with multiple regions initialized on a regular tiling of the image, easily allows detecting multiple transitions without the need for complex interactions of multiple level set functions. In fact, our experiments thus far suggest that this is perhaps the most significant and most practical benefit of this approach as it opens up a realistic possibility for *fully automatic segmentation of objects in images with cluttered backgrounds.*

Most of the existing schemes to extend region-based segmentation to multiple phases, including logic combinations of level set functions and recursive partitioning, suffer from interaction among regions, that generates a highly complex residual landscape that renders the optimization highly prone to local minima. As such, these methods are highly sensitive to initial contour placement in the presence of clutter. Even when trying to devise automated methods that seed the entire image domain with small, alternating, regularly spaced initial contours, the sensitivity of multi-phase methods to the choice of the seed size and spacing can exhibit itself in the simplest of images (see Fig. 9).

In this paper we tackle the important problem of fully automatic initialization of multi-phase segmentation schemes head-on. Our goal is to devise a generic initial condition that will, with high probability, yield convergence to the global minimum, consisting of multiple phases – whether the statistics capture intensity, color, texture, or motion.

2. Formalization

In this section we derive our model in steps. First we consider the simple problem of segmenting an image at the pixel level, based on binary classification of its gray levels. We call the image $I : D \subset \mathbb{R}^2 \to [0,\ 255] \subset \mathbb{R}_+;\ x \mapsto I(x)$, and indicate with $\theta \in \mathbb{R}^k$ the parameters of the model, for instance $\theta = \{\mu, \sigma\}$ for the case of a Gaussian model. We indicate the likelihood of the model parameters (foreground model) θ as $p(I|\theta)$. We indicate the alternate hypothesis (background model) with $p(I|\bar{\theta})$. A simple binary classifier can be arrived at by summing the log-probability

_____

depending only on the standard deviation.

of error, for both missed detections and false alarms, and then finding a threshold $\hat{\tau}$ such that

$$\hat{\tau} = \arg\min_{\tau} \int_0^\tau \log p(I|\theta)dP(I) + \int_\tau^{255} \log p(I|\bar{\theta})dP(I)$$
(1)

where $P(I)$ is a measure on the intensity values, which could be uniform in $[0,\ 255]$ if no prior knowledge is available. This basic thresholding model is not very useful for object detection, as it does not enforce spatial continuity, that we know to be relevant from empirical studies on the statistics of natural images [15]. The Mumford-Shah (M-S) model seeks to overcome this limitation, by minimizing the same cost functional above, but placing the decision boundary (threshold) *not* on the gray values, but on the location x instead. For the case of a single scan-line, assumed to start at the center of the object of interest, we have

$$\hat{\tau} = \arg\min_{\tau} \int_0^c \log p(I(x)|\bar{\theta})d\nu(x) + \int_c^d \log p(I(x)|\theta)d\bar{\nu}(x)$$
(2)

where $\nu(x)$ is a measure on the domain D, for instance the uniform measure $d\nu(x) = dx$. If d is fixed, for instance at the boundary of the image domain $d \in \partial D$, then [6] can be used to localize the boundary c as well as to estimate the statistics $\theta, \bar{\theta}$ that are most discriminative. If, on the other hand, the statistics are known, then [23] can be used to find d that yields the smallest latency $d - c$ for a given level of false alarms. For the case of a (two-dimensional) image, the threshold c (decision boundary) is represented by a curve $\partial\Omega$ bounding a region $\Omega \subset D$; the lookout d is represented by a region $\mathcal{D} \supset \Omega$, so the function being minimized, given by the integrals above, reads

$$E \doteq \int_\Omega \log p(I(x)|\bar{\theta})d\nu(x) + \int_{\mathcal{D}\setminus\Omega} \log p(I(x)|\theta)d\bar{\nu}(x) + \Gamma(\Omega)$$
(3)

where Γ denotes a regularizer, for instance the length of $\partial\Omega$. We now focus on the log-probability of error, *i.e.* the integrands above. The probability of missed detection (first integrand) depends on how well the data I at position x fits the background model $\bar{\theta}$. There is a penalty when data in the foreground fits the background model well, and this penalty is integrated on the foreground hypothesis Ω, which is by definition a compact region. False alarms similarly depend on how well the data I at a position x *outside* the foreground region nevertheless fit the foreground model θ, regardless of where x is relative to the foreground. If we keep everything else constant and double the region $\mathcal{D}$, the model above counts twice as many false alarms, contrary to our intuition that a false alarm becomes less likely as we move away from the decision boundary Ω. In other words, instead of characterizing the probability of error (both missed detection and false alarms) based only on photometric prop-

erties I of the data, we wish to account for geometric properties of the data as well, namely proximity to the decision boundary $\partial\Omega$. This can be accounted for in the measure $\nu(x)$.

As suggested in Sect. 1 for the case of a Gaussian model, the probability $p(I|\theta) \propto \exp(-d_\sigma^2(I,\mu))$ is a function of the distance $d_\sigma(I,\mu) = \frac{|I-\mu|}{\sigma}$ from the intensity value to the mean intensity μ; we want to extend it to also be a function of the distance of the point x from the boundary of Ω. In particular, we want the distance from the boundary in one region (foreground/background) to be a function of the statistics of the data in the other, for instance on the background we have

$$d\bar\nu(x) = p(x|\Omega,\theta)dx \qquad (4)$$

and similarly $d\nu(x) = p(x|\Omega,\bar\theta)dx$ where, for the case of a Gaussian model, we have

$$p(x|\Omega,\theta) \propto \exp(-d_\sigma^2(x,\Omega)) \qquad (5)$$

and

$$d_\sigma(x,\Omega) \doteq \min_{y\in\Omega} \frac{\|x-y\|}{\sigma}. \qquad (6)$$

For the case of missed detection, since the region Ω is by assumption compact and the dispersion $\bar\sigma$ of the background is large, in general $d\nu(x) = p(x|\mathcal{D}\setminus\Omega,\bar\theta)dx$ will be essentially constant, and therefore we simply take $d\nu(x) = dx$. Thus, the first term of the error functional E above remains

$$E_{\text{inside}} = \int_\Omega \log p(I(x)|\bar\theta)dx. \qquad (7)$$

For the background, however, we have

$$E_{\text{outside}} = \int_{\mathcal{D}\setminus\Omega} \log p(I(x)|\theta)p(x|\Omega,\theta)dx \qquad (8)$$

plus the usual regularizer $E_{\text{reg}} = \Gamma(\Omega)$, leading to

$$\hat\Omega = \arg\min_{\Omega,\lambda_1,\lambda_2} E = E_{\text{inside}} + \lambda_1 E_{\text{outside}} + \lambda_2 E_{\text{reg}}. \qquad (9)$$

A simpler version of this functional can be arrived at following the rationale of [6]; assuming Gaussian densities, the energies above become $E_{\text{inside}} = \int_\Omega -\left(\frac{I-\bar\mu}{\bar\sigma}\right)^2 dx$, and $E_{\text{outside}} = \int_{\mathcal{D}\setminus\Omega} -\left(\frac{I-\mu}{\sigma}\right)^2 dx$, minimizing which is equivalent to minimizing $E_{\text{inside}} = \int_\Omega \left(\frac{I-\mu}{\sigma}\right)^2 dx$, and $E_{\text{outside}} = \int_{\mathcal{D}\setminus\Omega} \left(\frac{I-\bar\mu}{\bar\sigma}\right)^2 dx$. Note that we have switched the sign, and as a consequence the roles of the inside and outside statistics have swapped. This is equivalent to assuming that the probability $p(x|\Omega,\theta)$ is uniform in $\mathcal{D}\setminus\Omega$:

$$p(x|\Omega,\theta) \propto \chi_{\mathcal{D}\setminus\Omega}(x) \qquad (10)$$

If we assume $\sigma = \bar\sigma = 1$, we obtain the model of [6]. In our case, rather than fixing $\mathcal{D}$, we allow it to change as a dilation (lookout) of Ω proportional to the statistics of the image inside:

$$\mathcal{D} = \Delta_{\sigma(I|\Omega)}\Omega \qquad (11)$$

where Δ_σ denotes a dilation by σ and, for the case of a Gaussian model,

$$\sigma^2(I|\Omega) = \frac{\int_\Omega |I-\mu|^2 dx}{\int_\Omega dx}. \qquad (12)$$

Therefore, while the "inside" term of [6] remains the same, the "outside" term is now controlled by the statistics of the image inside:

$$E(\Omega,\mu,\bar\mu,\lambda_1,\lambda_2) \doteq \int_\Omega |I(x)-\mu|^2 dx +$$
$$+ \lambda_1 \int_{\Delta_{\sigma(I|\Omega)}\Omega\setminus\Omega} |I(x)-\bar\mu|^2 dx + \lambda_2\Gamma(\Omega). \qquad (13)$$

The minimization of this functional involves computing its variation with respect to the unknowns $\Omega,\mu,\bar\mu$, which we do in Sect. 3.

In the next section we discuss the details of the minimization of the model (13). The optimization of a more general model that separates local histograms that we use for more complex data types is more involved but conceptually poses no challenges.

3. Implementation

To solve the optimization problem (13), we implement an (infinite-dimensional) gradient flow, corresponding to a partial differential equation (PDE), that evolves an initial contour towards a fixed point, corresponding to a (local) minimum of (13), as customary in the active contour framework. The PDE is implemented on a discrete grid using Level Set methods [17]. To this end, we need to compute the first variation of the functional (13) with respect to changes of the boundary of the region Ω, and perform an incremental update in the (opposite) direction of the gradient. To this end, we call $C \doteq \partial\Omega$ the contour, and use t to indicate the iteration. Therefore, $C = C(x,t)$ evolves over "time," with only changes along the (outward) normal direction $N \in \mathbb{S}^1$ affecting the deformation of the contour, in a way that is proportional to $-\nabla E$, so that at the fixed point $\nabla E = 0$, i.e. the first-order optimality conditions are satisfied. We call s the arc-length parameterization of the contour C.

The derivation of the first variation of E is standard, for instance [6], except for the derivative of E_{outside}, and specifically for its dependency of the domain Ω, now represented by the boundary C. So, we focus our attention on

that term, which for convenience we write as

$$\int f(x)\mathbb{H}(\underbrace{\sigma(I|C) - d(x,C)}_{\doteq F(C,x)})dx \qquad (14)$$

where $f(x) \doteq (I(x) - \bar{\mu})^2$, $\mathbb{H}$ is the Heaviside function, $d(x,C)$ is the distance function from the point $x \in \mathbb{R}^2$ to the contour C, and $\sigma(I|C)$ is defined in (12), except that the dependency on the region Ω is now written, with a slight abuse of notation, in terms of its boundary $C = \partial\Omega$. The first variation is computed by taking the total derivative of E with respect to time, which in turn depends on the partial derivative of C with respect to time $C_t \doteq \frac{\partial C}{\partial t}$, and consists of two terms. The first is standard:

$$\int_C \langle C_t, -f(C(s))\mathbb{H}(F(C,C(s)))N\rangle ds \qquad (15)$$

the second is

$$\int_{\mathcal{D}\backslash\Omega} f(x)\mathbb{H}'(F(C,x)) \quad \underbrace{\frac{d}{dt}F(C,x)}_{\int_C \langle C_t, \nabla_C F(C,x)(s)\rangle ds} \quad dx =$$

$$= \int_C \langle C_t, \int_{\mathcal{D}\backslash\Omega} f(x)\mathbb{H}'(F(C,x))\nabla_C F(C,x)(s)dx\rangle ds. \quad (16)$$

Consequently, the gradient flow is given by

$$C_t = f(C)N + \left(\int_{\mathcal{D}\backslash\Omega} f(x)\mathbb{H}'(F(C,x))\nabla_C F(C,x)(s)dx\right). \qquad (17)$$

A straightforward but suboptimal implementation can be arrived at by discounting the second term, and only focusing on the first.

Note that the first term, $f(C)$ can be written in terms of an optimal constant value μ^*, which can be found by taking the derivative of $f(x)\mathcal{H}(d(x,C) - \sigma(I|C))$, yielding $(I - \mu^*)\mathbb{H}(d(x,C) - \sigma(I|C)) = 0$, that brings the dependency of the statistics of the image *inside* the region Ω into the first variation of the component of the cost functional *outside* of it. This is the model we test empirically in the experimental section that follows.

3.1. Automatic initialization procedure

One of the most important advantages of this new model is its ability to robustly segment multiple distinct objects within a cluttered image by exploiting the fact that separate contours evolving within an image domain utilize independent "look ahead" regions (even if these overlap) and therefore do not have to be artificially coupled in their evolutions as do active contours implemented via multi-phase methods. This greatly reduces their dependence on the initialization and allows us to pair an automatic initialization procedure with our model. It is precisely this highly practical benefit that allows us to take seriously the idea of utilizing this class of active contour for fully automatic detection of one or more objects amidst complex background clutter, a possibility that is not realistic for most other active contour models.

In a two-phase variational segmentation scheme, where strong convexity results are available, one can tile the image with a regular sampling of contours, and verify empirically that a two-phase scheme converges to the "correct" solution with high frequency. Our goal is to devise a similar generic initialization scheme for multi-phase segmentation, by providing seed initializations consisting of an *initial position* and *initial velocity*.

To this end, we tile the image with "seeds", consisting of a set of small regularly-spaced circular contours. This is the "initial position", whereby each seed is a point in the residual surface we are trying to minimize. We expect that some of these seeds will fall all within target regions, and some will intersect region boundaries. However, we assume that for each target region there is at least one seed that is fully contained within (this implicitly imposes a lower bound on the size of the regions we are willing to consider).

The seeds that fall all within a target region will be, by definition, on a flat portion of the residual surface. So, together with an initial position, we must endow seeds with an "initial velocity". This is a direction in the space of curves, that corresponds to a variation of the initial seed. The most obvious direction is the one corresponding to an expansion of the seed which, in the absence of damping, corresponds to a constant inflationary force. While this is the simplest choice, more elaborate initial velocities can be implemented.

Each seed is thus evolved independently, so there is no interaction among phases, and the algorithm can be run in parallel for each initialization. Multiple seeds will converge to the same salient target regions (all seeds that are contained in a target region should converge to the same solution) while isolated seeds may converge to spurious local minima. All final regions are put through a voting scheme to select those with consistent convergence, which yields the final multi-phase segmentation.

4. Experiments

In the first synthetic example we illustrate the behavior of our approach depending on the statistics of the data. The common heuristic of tying the lookout region to the size of the foreground fails since the probability of detection of discontinuities between two regions depends on the statistics of the images on both sides of the decision boundary, not on how large these regions are. Tying the lookout region on each side to the statistics on the other side, on the other hand, scales nicely with the levels of noise as well as with the size of the regions (Fig. 3). In the second experiment we test our generalization of the Chan and Vese model (13) on the same data of Fig. 1. In Fig. 4, the contour adapts nicely
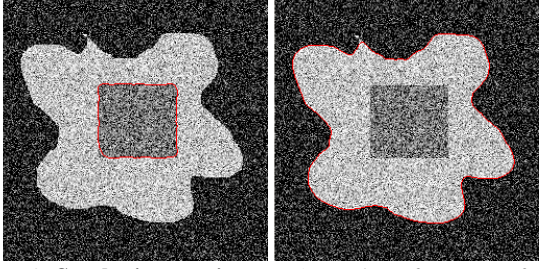
Figure 3. **Synthetic experiment:** *A number of squares of different sizes, with varying intensity gaps from the immediate background (light) and a randomized far background (dark), with additive noise of varying standard deviation. Initialization is randomized within three classes: The entire initial contour is contained within the target square, contains it, or intersects its boundary. Convergence is considered successful when the contour captures the boundaries of the square, with accuracy measured by the set-symmetric difference between the "true" and the estimated regions. Failure is declared when the contour evolves to capture the structure of the background (right), and the percentage of occurrences measures the robustness of the algorithm to initialization. Quantitative results are reported in Figs. 6–8.*

based on its local context, regardless of the complexity of the background far away from it. Note, however, that in this example many variants of the simple model of [6], for instance the full M-S model (Fig. 1 bottom-right) gives satisfactory results, albeit at an increased computational cost. In the third experiment, in Fig. 5, we try a challenging ex-
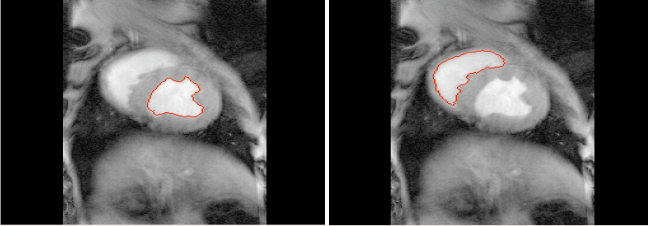

Figure 4. **Heart chambers:** *Although the C-V model fails (Fig. 1), the full M-S model can successfully detect the boundaries of the heart chambers, at the cost of expending most of the modeling efforts on the background (Fig. 1 bottom). Our generalization of the C-V model (13), on the other hand, only focuses on an outer neighborhood of the boundary, controlled by the statistics of the object of interest.*

ample where even the general M-S model, as well as the standard C-V model, fails to detect the boundaries of the flatworm (left). Our model (13), on the other hand, successfully detects it despite the complex background and significant variation in the intensity gap along the boundary. Fig. 4 also highlights the fact that our model allows dealing with multiple regions in a straightforward way that does not involve logic combinations of level set functions. A systematic covering of the image with multiple initializations yields multiple convergent runs to each region of interest,
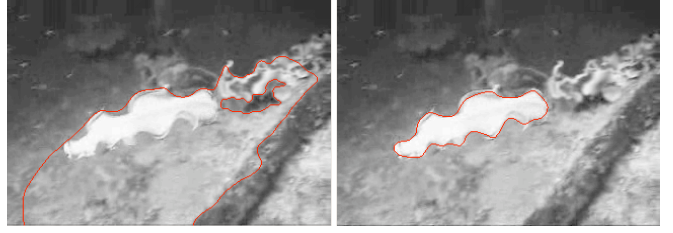

Figure 5. **Flatworm:** *The C-V model, as well as the full M-S model, fail to detect the boundaries of the flatworm. Our model (13), however, successfully detects it despite the complex background (right).*

		Inside Init.	Outside Init.	In/Out Init.
	C-S	1130 (21)	830 (16)	882 (17)
$\sigma = 0$	C-V	3440 (53)	2820 (44)	3976 (62)
	M-S	1580 (1216)	1140 (877)	1208 (930)
	C-S	808 (15)	304 (6)	1062 (20)
$\sigma = 0.01$	C-V	1610 (25)	1020 (16)	2900 (45)
	M-S	906 (697)	610 (469)	1260 (970)
	C-S	888 (17)	551 (10)	1202 (23)
$\sigma = 0.05$	C-V	710 (11)	1722 (27)	3480 (54)
	M-S	1180 (908)	915 (704)	1670 (1285)
	C-S	970 (18)	522 (10)	1470 (28)
$\sigma = 0.1$	C-V	730 (11)	1877 (29)	3570 (55)
	M-S	1181 (909)	730 (562)	1800 (1386)

Figure 6. **Computational Cost**: Number of iterations and processor time per iterations in milliseconds in parenthesis. M-S refers to the Mumford-Shah model, C-V to the Chan-Vese model, and C-S to the "Curious Snake" model presented here. The table shows that our approach is competitive both in terms of number of iterations to reach a specified accuracy, as well as in the computational cost per iteration. The figures are averaged over 10 trials per each configuration, with initial condition starting all inside, all outside, or partially overlapping the target region.

so the two chambers are detected individually.

In order to arrive at a quantitative comparison between the M-S model, the C-V model, and the model (13), we consider the experiment in Fig. 3, and vary the size of the target, the noise level, the gap between the means, and the initialization, which is randomized within three classes: All inside the target, all outside, and partially overlapping the boundary. We measure accuracy by the set-symmetric difference between the true region Ω and the estimated one $\hat{\Omega}$, normalized by the area of Ω. We measure robustness by the percentage of runs that converge to within a 10% accuracy, starting from each of the three classes of randomized initial conditions. Finally, we evaluate computational complexity by measuring the number of iterations necessary to reach a 1% accuracy, conditioned on convergence, as well as the computational complexity of each iteration. The results are summarized in the tables in Figs. 6–8.

In fig. 9 we show a case where multiphase level sets with different tiling initializations (different box centers and box lengths) on a simple synthetic image can lead to a correct or

		Inside Init.	Outside Init.	In/Out Init.
	C-S	0.0000	0.0000	0.0000
$\sigma = 0$	C-V	0.3640	0.3634	0.3990
	M-S	0.0000	0.0000	0.0000
	C-S	0.0001	0.0005	0.0006
$\sigma = 0.01$	C-V	0.3705	0.3804	0.3990
	M-S	0.0000	0.0002	0.0008
	C-S	0.0004	0.0002	0.0010
$\sigma = 0.05$	C-V	0.3742	0.3634	0.4210
	M-S	0.0003	0.0002	0.0007
	C-S	0.0010	0.0012	0.0043
$\sigma = 0.1$	C-V	0.3842	0.3910	0.4337
	M-S	0.0015	0.0017	0.0032

Figure 7. **Accuracy:** The normalized set-symmetric difference between the estimated region and the contour. Zero means perfect matching, 1 means that the region is disjoint. The figures are averaged over 10 trials starting from each initial configuration. Our approach performs nearly as well as $M - S$, but at a fraction of the computational cost.

		Inside Init.	Outside Init.	In/Out Init.
	C-S	100	100	100
$\sigma = 0$	C-V	100	100	100
	M-S	100	100	100
	C-S	100	100	100
$\sigma = 0.01$	C-V	100	90	100
	M-S	100	100	100
	C-S	90	100	100
$\sigma = 0.05$	C-V	80	70	80
	M-S	100	100	100
	C-S	90	90	90
$\sigma = 0.1$	C-V	70	60	70
	M-S	90	90	90

Figure 8. **Robustness:** The percentage of trials where the contour converged to within 10% accuracy (Fig. 7). M-S performs best, since it relies on a global model. C-V performs the worst; although, in principle, it relies on an equally universal model (piecewise constant statistics), limiting the model to two phases causes significant convergence problems in the presence of complex background clutter.
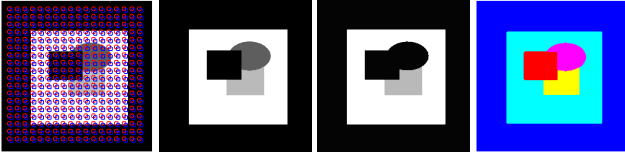

Figure 9. **Non-Interacting Multiphases:** Typical multiphase level set initialization with overlapping zero level sets (left). Depending on the size and spacing of the regions, multiphase levels sets could converge to the correct (2nd image) or incorrect (3rd image) solution even in the case of a noiseless synthetic image. Our algorithm correctly detects all regions (right).

incorrect segmentation, whereas our automatic initialization algorithm consistently converges to the correct solution.

In the last experiment (fig. 10), we show that multi-phase level sets are not robust to generic tiling typically used in the literature, whereas our algorithm consistently converges
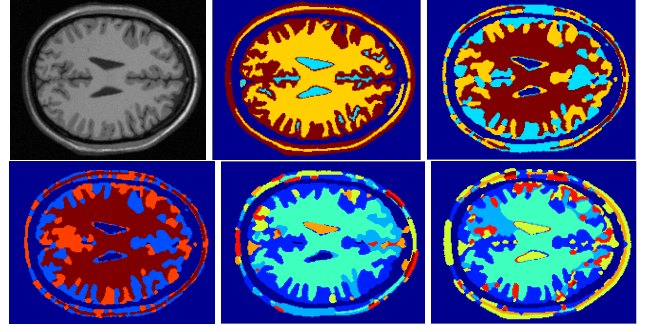

Figure 10. **Non-Interacting Multiphases:** Phantom brain image used in medical imaging (top left). Our automatic algorithm (top-middle) correctly detects gray matter/white matter with below 10% error, while 4 phase multiphase fails (top right and bottom left) 40% of the time (error above 10%) and fails 100% of time (two bottom right) with 8 phases. Initialization for multiphase level sets is done with densely placed overlapping boxes of differing size/spacing.

from generic initialization. We perform the experiment on a phantom brain image typical in medical imaging where the goal is to separate white matter from gray matter. We initialize the multiphase algorithm with a dense tiling of squares. We randomize the initialization by varying the size (3-15 pixels) and spacing (2-5 pixels) of the squares. In 40% of the cases with 4 phases, the multiphase algorithm correctly detects the gray/white matter with error (set-symmetric difference) less than 10%. In the case of 8 phases (3 level sets), multiphase level sets always fails (error greater than 10%). Fig. 10 shows a representative sample of failures. Also, if the number of phases is unknown, one would wish that a larger number of level set functions would converge to a solution where all phases are captured and some overlap. Unfortunately, this does not usually happen. This makes multi-phase algorithms impractical in real scenarios where there is no such thing as a "true number of phases" and one would want a segmentation algorithm to be robust with respect to singular perturbations (changes in the number of phases).

5. Discussion

We have presented a semi-local region-based segmentation model that generalizes that of [6] to an adaptive lookout region. The basic idea is to tie the size of the region *not* to the size of the object of interest, but to the fitness of its model, following a robust-statistical perspective. Our assumption is that, for each object of interest for which a statistical model can be easily specified, the detection of its boundary depends on a violation of this model, a hypothesis that can be tested locally, with the locality controlled by the deviation of the data from the model. This affords us the added benefit of simultaneously detecting multiple objects

in an image, by initializing several regions on a covering of the image, without having to manage logic combinations of level set functions [21, 8].

Our results naturally relate to the wealth of research on active contour models, pioneered by [12, 1], imported into the framework of geometric variational optimization by [3]. Contributions that are particularly relevant in the context of our paper include [13, 18, 4, 7], as well as various "unilateral" region segmentation approaches based on fast-marching methods [10].

Our implementation neglects some terms of the optimizing flow, that is therefore only a sub-optimal solution of (13). Investigating efficient approaches to minimizing the full energy functional is the subject of future investigation, along with an empirical characterization of the loss from optimality in the simplified algorithm we have reported in Sect. 3.

Our technique is subject to the same limitations of any region-based active contour model: The model is based on the first-order optimality conditions, which are only necessary but not sufficient for a global minimum.

Acknowledgments

References

[1] A. Blake and M. Isard. *Active contours*. Springer, 1998. 8

[2] P. Blomgren and T. Chan. Color TV: total variation methods for restoration of vector-valuedimages. *IEEE Transactions on Image Processing*, 7(3):304–309, 1998. 1

[3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proceedings of the IEEE Int. Conf. on Computer Vision*, pages 694–699, Cambridge, MA, USA, June 1995. 8

[4] A. Chakraborty, L. Staib, and J. Duncan. Deformable boundary finding in medical images by integrating gradient and region information. *IEEE Transactions on Medical Imaging*, 15(6):859–870, 1996. 8

[5] T. Chan and S. Esedoglu. Aspects of Total Variation Regularized L1 Function Approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817, 2005. 1

[6] T. Chan and L. Vese. An active contours model without edges. In *Int. Conf. Scale-Space Theories in Computer Vision*, pages 141–151, 1999. 1, 2, 3, 4, 6, 7

[7] G. Charpiat, R. Keriven, J. Pons, and O. Faugeras. Designing spatially coherent minimizing flows for variational problems based on active contours. In *ICCV*, 2005. 8

[8] D. Cremers, N. Sochen, and C. Schnoerr. Towards recognition-based variational segmentation using shape priors and dynamic labeling. In *Intl. Conf. on Scale-Space Theories in Computer Vision*, pages 388–400, June 2003. 8

[9] D. Cremers, and S. Soatto. Motion competition: a variational approach to piecewise parametric motion segmentation. In *Intl. J. on Comp. Vision*, pages 249–265, May 2005. 1

[10] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast geodesic active contours. *IEEE Transactions on Image Processing*, 10(10):1467–1475, 2001. 8

[11] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1988. 2

[12] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int. J. of Comp. Vis.*, 1(4):321–331, 1987. 8

[13] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Trans. on Patt. Anal. and Mach. Intel.*, (17):158–175, 1995. 8

[14] J. Mille and L. Cohen. A local normal-based region term for active contours. *EMMCVPR*, 2009. 2

[15] D. Mumford and B. Gidas. Stochastic models for generic images. *Quarterly of Appl. Math.*, 54(1):85–111, 2001. 3

[16] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. on Pure and Appl. Math.*, 42:577–685, 1989. 1, 2

[17] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi equations. *J. of Comp. Physics*, 79:12–49, 1988. 1, 4

[18] N. Paragios and R. Deriche. Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision. *International Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics*, 13(2):249–268, June 2002. 8

[19] N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46(3):223, 2002. 1

[20] M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. Tech. rep., INRIA, January 2003. RR n. 4695. 1

[21] B. Sandberg and T. Chan. A logic framework for active contours on multi-channel images. *Journal of Visual Communication and Image Representation*, 16(3):333–358, 2005. 8

[22] B. Sandberg, T. Chan, and L. Vese. A level-set and Gabor-based active contour algorithm for segmenting textured images. *UCLA CAM report*, pages 02–39, 2002. 1

[23] A. Tartakovsky and V. Veeravalli. General asymptotic Bayesian theory of quickest change detection. *Theory of Prob. and Appl.*, 49(3):458–497, 2005. 2, 3

[24] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995. 2

[25] S. Zhu, T. Lee, and A. Yuille. Region competition: Unifying snakes, region growing, energy /bayes/mdl for multi-band image segmentation. In *Int. Conf. on Computer Vision*, pages 416–423, 1995. 1