

# Curious Active Contours

G. Sundaramoorthi      S. Soatto      A. J. Yezzi

March 10, 2009

## Abstract

We present a region-based boundary detection algorithm for objects that exhibit relatively simple 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 the decision boundary, 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.

## 1 Introduction

In [15], Mumford and Shah proposed a model of the image as piecewise smooth statistics. While simplistic in terms of its generative power, this model has proven very useful for discriminative purposes, and has served as a basis for efficient computational algorithms for image segmentation developed in the context of Level Set Methods [16]. Indeed, [6] consider an even simpler model, involving piecewise constant image statistics, that has gained widespread acceptability 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* [26]. This model has also been extended to color [2], and texture [21, 18, 19].

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* regions (foreground/background), but 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].<sup>1</sup> The alternative to a fine partition of the image domain into simple functions is a coarser partition (e.g. foreground/background) with respect to more complex functions, say smooth functions, leading us back to the Mumford-Shah functional. In either case, much of the attention in Fig. 1 is devoted to the background. 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. The shortcomings of these approaches are obvious, as there is no reason why the structure of the boundary of a region (say the edge of a brick wall) should depend on how large the region is. For the purpose of detecting a discontinuity, 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, as argued eloquently by [24]. In our case, the distributions of intensity values specifying the null and alternate hypothesis (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 [10], 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 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, as a way of example, a one-dimensional image, a 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.

---

<sup>1</sup>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 just becomes that of localizing the object, or determining its boundary.

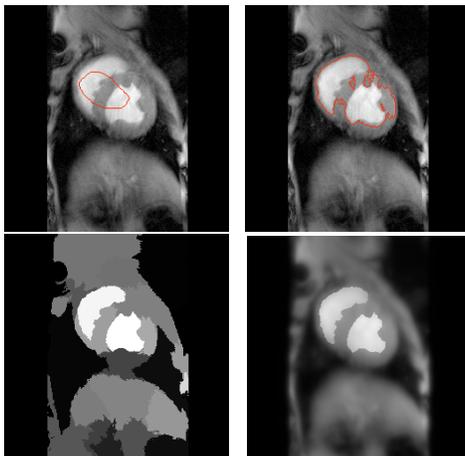


Figure 1: Segmentation of a heart chamber using [6] (top-right, red curve), starting from the initial condition (top-left), is impeded by the fact that the background does not fit the constant model. Extension to multi-phase segmentation (bottom-left, each region is color-coded, and the object of interest corresponds to the white region) is complex and highly non-convex. Extension to more complex models, such as [15] (bottom-right, the foreground region has smooth intensity, displayed in white, and the background region is a smooth function approximating the background – note it is highly blurred) 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.

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  $\sigma$  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  $\sigma \rightarrow 0$ , 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*.<sup>2</sup> We call the ensuing model, which we 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

<sup>2</sup>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 depending only on the standard deviation.

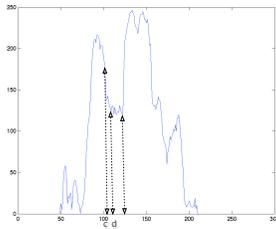


Figure 2: *One scanline from Fig. 1: The detection of the boundary  $c$  should be performed as soon as possible,  $d$ , so as not to have irrelevant background impinge on the decision (past the right-most dashed line).*

as to run into trouble due to the statistics of the background away from the transition of interest.

This seemingly simple modification of the Chan and Vese (C-V) model makes a significant difference in applications, such 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 covering of the image, easily allows detecting multiple transitions without the need for complex interactions of multiple level set functions.

This work naturally relates to a vast body of work in active contours, and in particular in variational image segmentation methods, as we discuss in Section 5.

## 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  $I : D \subset \mathbb{R}^2 \rightarrow [0, 255] \subset \mathbb{R}_+$ ;  $x \mapsto I(x)$ , 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)  $\theta$  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 of error,<sup>3</sup> for both missed detections and false alarms, and then finding a threshold  $\hat{\tau}$  such that

$$\hat{\tau} = \arg \min_{\tau} \underbrace{\int_0^{\tau} \log p(I|\theta) dP(I)}_{\text{missed detections}} + \underbrace{\int_{\tau}^{255} \log p(I|\bar{\theta}) dP(I)}_{\text{false alarms}} \quad (1)$$

<sup>3</sup>An equivalent maximum-likelihood formulation can be derived using likelihood ratios, but that involves continuous products that require some notational effort.

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 [14]. 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

$$\hat{\tau} = \arg \min_{\tau} \underbrace{\int_0^c \log p(I(x)|\bar{\theta}) d\nu(x)}_{\text{missed detections}} + \underbrace{\int_c^d \log p(I(x)|\theta) d\bar{\nu}(x)}_{\text{false alarms}} \quad (2)$$

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

$$E \doteq \underbrace{\int_{\Omega} \log p(I(x)|\bar{\theta}) d\nu(x)}_{\text{missed detections}} + \underbrace{\int_{\mathcal{D} \setminus \Omega} \log p(I(x)|\theta) d\bar{\nu}(x)}_{\text{false alarms}} + \Gamma(\Omega) \quad (3)$$

where  $\Gamma$  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  $\Omega$ , which is by definition a compact region. The false alarms similarly depend on how well the data  $I$  at a position  $x$  *outside* the foreground region nevertheless fit the foreground model  $\theta$ , 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  $\Omega$ . In other words, instead of characterizing the probability of error (both missed detection and false alarms) based only on photometric properties  $I$  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  $\mu$ ; we want to extend it to also be a function of the distance of the point  $x$  from the boundary of  $\Omega$ . 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

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

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

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

and

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

For the case of missed detection, since the region  $\Omega$  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

$$E_{\text{inside}} = \int_{\Omega} \underbrace{\log p(I(x)|\bar{\theta})}_{\text{pr. miss. det.}} dx. \quad (7)$$

For the background, however, we have

$$E_{\text{outside}} = \int_{\mathcal{D}\setminus\Omega} \underbrace{\log p(I(x)|\theta)p(x|\Omega, \theta)}_{\text{pr. false alarm}} dx \quad (8)$$

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

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

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  $\mathcal{D}\setminus\Omega$ :

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

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  $\Omega$  proportional to the statistics of the image inside:

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

where  $\Delta_\sigma$  denotes a dilation by  $\sigma$  and, for the case of a Gaussian model,

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

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

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

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

An alternate model that is related to [6] consists of simply pulling the means of the image inside and outside  $\Omega$  as far apart as possible. This corresponds to

$$\max |\mu - \bar{\mu}| \quad (14)$$

where  $\mu = \int_{\Omega} I(x) dx / |\Omega|$  and  $\bar{\mu} = \int_{\mathcal{D} \setminus \Omega} I(x) dx / |\mathcal{D} \setminus \Omega|$ , and yields a minimizing flow that is similar to that of [6], except for a normalization factor: In the first case one has  $\frac{d}{dt} \partial \Omega = (\mu - \bar{\mu})[(I - \mu) + (I - \bar{\mu})]$ , whereas in the second case  $\frac{d}{dt} \partial \Omega = (\mu - \bar{\mu}) \left[ \frac{I - \mu}{|\Omega|} + \frac{I - \bar{\mu}}{|\mathcal{D} \setminus \Omega|} \right]$ . An additional modification of this model can be obtained by forcing the means not only to be well separated, but also to move in opposite directions during the evolution; [25] elucidates the differences between these models, that in the case of grayscale (non binary) images can be substantial.

The advantage of (14) is that it lends itself to easy generalizations, obtained by separating, rather than the means, other statistics. For instance, one can separate the entire distributions, by maximizing the  $\chi^2$ , Bhattacharyya or any other distance, or Kullback-Liebler’s divergence, and controlling the dilation, rather than by the standard deviation, by the entropy of the inside region. For the case of Kullback-Liebler, one has

$$\max_{\Omega} \text{KL} \left( p(I|_{\Omega}); p(I|_{\mathcal{D} \setminus \Omega}) \right) \quad | \quad \Gamma(\Omega) \leq \epsilon \quad (15)$$

where  $\mathcal{D}$  is defined as in (11), except that now

$$\sigma^2(I|\Omega) \doteq H(p(I|_{\Omega})) \quad (16)$$

is the entropy the histogram of the image restricted to  $\Omega$ . In the next section we discuss the details of the minimization of the model (13), that is the simplest rendition of the general model (15) for the case of Gaussian statistics. The optimization of (15) is significantly more involved and outside the scope of this manuscript. It is introduced primarily to illustrate the fact that the model (13) is not restricted to Gaussian statistics, but instead admits rather wide generalizations.

### 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 [16]. To this end, we need to compute the first variation of the functional (13) with respect to changes of the boundary of the region  $\Omega$ , 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  $E_{\text{outside}}$ , and specifically for its dependency of the domain  $\Omega$ , now represented by the boundary  $C$ . So, we focus our attention on that term, which for convenience we write as

$$\int f(x) \mathbb{H}(\underbrace{\sigma(I|C) - d(x, C)}_{\doteq F(C, x)}) dx \quad (17)$$

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  $\Omega$  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:

$$\int_C \langle C_t, -f(C(s)) \mathbb{H}(F(C, C(s))) N \rangle ds \quad (18)$$

the second is

$$\begin{aligned} \int_{\mathcal{D} \setminus \Omega} f(x) \mathbb{H}'(F(C, x)) \underbrace{\frac{d}{dt} F(C, x)}_{\int_C \langle C_t, \nabla_C F(C, x)(s) \rangle ds} dx = \\ = \int_C \langle C_t, \int_{\mathcal{D} \setminus \Omega} f(x) \mathbb{H}'(F(C, x)) \nabla_C F(C, x)(s) dx \rangle ds. \end{aligned} \quad (19)$$

Consequently, the gradient flow is given by

$$C_t = f(C)N + \left( \int_{\mathcal{D} \setminus \Omega} f(x) \mathbb{H}'(F(C, x)) \nabla_C F(C, x)(s) dx \right). \quad (20)$$

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  $\mu^*$ , 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  $\Omega$  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.

## 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

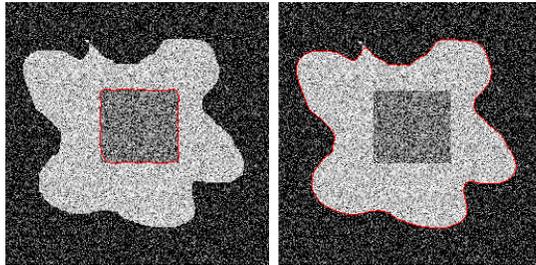


Figure 3: **Synthetic experiment:** A number of squares of different sizes, with varying intensity gaps from the immediate background (light) and a randomized far background (dark), with additive noise of varying standard deviation. Initialization is randomized within three classes: The entire initial 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 escapes 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.

Vese model (13) on the same data of Fig. 1. As one can see in Fig. 4, the contour adapts nicely 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 example where even the general M-S model, as well as the standard C-V model with all the latest bells and whistles [22], fails to detect the boundaries of the flatworm (left).

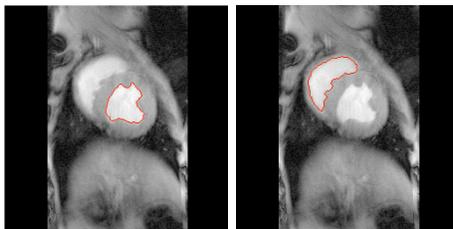


Figure 4: **Heart chambers:** Although the C-V model fails (Fig. 1), the full M-S model can successfully detect the boundaries of the heart chambers, at the cost of expending most of the modeling efforts on the background (Fig. 1 bottom). Our generalization of the C-V model (13), on the other hand, only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

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 re-

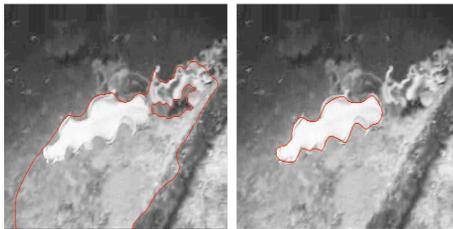


Figure 5: **Flatworm:** Even the C-V model with the Sobolev metric, as well as the full M-S model, fail to detect the boundaries of the flatworm. Our model (15), however, successfully detects it despite the complex background (right).

gions 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, so the two chambers are detected individually.

In order to arrive at a quantitative comparison between the M-S model, the C-V model with the Sobolev metric, 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  $\Omega$  and the estimated one  $\hat{\Omega}$ , normalized by the area of  $\Omega$ . 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

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

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.

## 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 that of tying 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, as well as common practice in minimum-latency optimal stopping time detection. We have presented both a simple version of the model (13), generalizing [6] under the assumption of Gaussian statistics, and a more sophisticated one (15) based on general distributions and on an information-theoretic measure of uncertainty.

The models we have presented, like any other model, are neither *right* nor *wrong*, to paraphrase the statistician Box. They just reflect different modeling assumptions, and can therefore be *useful* to the extent in which real data and application problems reflect these assumptions. This can only be validated empirically, which we do, by showing that on common detection/localization/segmentation problems arising in medical imaging and tracking, our model can successfully discount background statistics away from the decision boundary. The C-V

		Inside Init.	Outside Init.	In/Out Init.
$\sigma = 0$	C-S	0.0000	0.0000	0.0000
	C-V	0.3640	0.3634	0.3990
	M-S	0.0000	0.0000	0.0000
$\sigma = 0.01$	C-S	0.0001	0.0005	0.0006
	C-V	0.3705	0.3804	0.3990
	M-S	0.0000	0.0002	0.0008
$\sigma = 0.05$	C-S	0.0004	0.0002	0.0010
	C-V	0.3742	0.3634	0.4210
	M-S	0.0003	0.0002	0.0007
$\sigma = 0.1$	C-S	0.0010	0.0012	0.0043
	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.
$\sigma = 0$	C-S	100	100	100
	C-V	100	100	100
	M-S	100	100	100
$\sigma = 0.01$	C-S	100	100	100
	C-V	100	90	100
	M-S	100	100	100
$\sigma = 0.05$	C-S	90	100	100
	C-V	80	70	80
	M-S	100	100	100
$\sigma = 0.1$	C-S	90	90	90
	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.

model assumes that the scene exhibits piecewise constant statistics, which can be done without loss of generality if the image is in  $\mathbb{L}^2(D)$  and one is willing to let the number of regions grow large, thus expending precious modeling power to represent the background. The M-S model assumes that the scene exhibits piecewise smooth statistics, which again can be done with no loss of generality if the image is in  $\mathbb{L}^2(D)$ , and trades off the benefit of segmenting the image into fewer larger regions with the cost of a more involved computation, again spending modeling resources on the background.

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 cover of the image, without having to manage logical combinations of level set functions [20, 8], or other global models that, unavoidably, yield non-convex optimization problems [5].

Our results naturally relate to the wealth of research on active contour models, pioneered by [11, 1], imported into the framework of geometric variational optimization by [3, 12]. Contributions that are particularly relevant in the context of our paper include [13, 17, 4, 7], as well as various “unilateral” region segmentation approaches based on fast-marching methods [9].

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, and extensions to more general models such as (15).

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. Furthermore, because of the simplifications of the flow, our scheme is not even guaranteed to converge to a local minimum, although in all the experiments we have performed we have always experienced convergence to a fixed point, and very rarely have noticed convergence to local minima, especially when using regularized metrics such as the Sobolev metric. In particular, improvement from the C-V is significant in all circumstances, both in terms of accuracy, robustness and computational efficiency. It is certainly possible to construct examples that break down the model, so our approach can only be validated in the field. Our initial experiments indicate promise in this direction.

## References

- [1] A. Blake and M. Isard. *Active contours*. Springer Verlag, 1998. 13
- [2] P. Blomgren and T. Chan. Color TV: total variation methods for restoration of vector-valued images. *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. **13**
- [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. **13**
- [5] T. Chan and S. Esedoglu. Aspects of Total Variation Regularized L’Function Approximation. *SIAM Journal on Applied Mathematics*, 65(5):1817, 2005. **2, 13**
- [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, 5, 6, 7, 8, 9, 11**
- [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. **13**
- [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. **13**
- [9] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky. Fast geodesic active contours. *IEEE Transactions on Image Processing*, 10(10):1467–1475, 2001. **13**
- [10] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1988. **2**
- [11] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int. J. of Computer Vision*, 1(4):321–331, 1987. **13**
- [12] S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. Yezzi. Analysis of planar shape influence in geodesic active contours. In *Int. Conf. Comp. Vis.*, pages 810–815, 1995. **13**
- [13] R. Malladi, J. Sethian, and B. Vemuri. Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (17):158–175, 1995. **13**
- [14] D. Mumford and B. Gidas. Stochastic models for generic images. *Quarterly of Applied Mathematics*, 54(1):85–111, 2001. **5**
- [15] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. on Pure and Applied Mathematics*, 42:577–685, 1989. **1, 3**
- [16] 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, 8**
- [17] 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. **13**
- [18] 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**
- [19] M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. Technical report, INRIA, Sophia Antipolis, January 2003. RR n. 4695. **1**

- [20] 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. [13](#)
- [21] B. Sandberg, T. Chan, and L. Vese. A level-set and Gabor-based active contour algorithm for segmenting textured images. *UCLA Department of Mathematics CAM report*, pages 02–39, 2002. [1](#)
- [22] G. Sundaramoorthi, A. Yezzi, and A. Mennucci. Sobolev active contours. In *VLSM*, pages 109–120, 2005. [9](#)
- [23] A. Tartakovsky and V. Veeravalli. General asymptotic Bayesian theory of quickest change detection. *Theory of Probability and its Applications*, 49(3):458–497, 2005. [2](#), [5](#)
- [24] V. N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995. [2](#)
- [25] A. J. Yezzi, A. Tsai, and A. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *Proc. of Intl. Conf. on Computer Vision*, volume 2, pages 898–903, 1999. [7](#)
- [26] 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](#)