

A Statistical Approach to Snakes for Bimodal and Trimodal Imagery *

Anthony Yezzi, Jr. Andy Tsai Alan Willsky
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract

In this paper, we describe a new region-based approach to active contours for segmenting images composed of two or three types of regions characterizable by a given statistic. The essential idea is to derive curve evolutions which separate two or more values of a pre-determined set of statistics computed over geometrically determined subsets of the image. Both global and local image information is used to evolve the active contour. Image derivatives, however, are avoided, thereby giving rise to a further degree of noise robustness compared to most edge-based snake algorithms.

1 Introduction

A number of region-based approaches to snakes have been proposed in recent years [1, 2, 8, 9, 12, 13] with the tremendous appeal over edge-based approaches of avoiding computations to explicitly detect edges. Such computations typically require derivative information which is extremely sensitive to noise in the image. These approaches also tend to use both local and global image information whereas most edge-based approaches rely primarily on local information around the active contours.

This paper presents a new class of region-based active contour models that assume an image consists of a finite number of regions, characterizable by a pre-determined set of features (e.g. means, variances, textures) which may be inferred or estimated from the image data. Curve evolution equations are derived by computing the gradient directions of energy functionals which favor a maximal separation of these features. Introducing a penalty on the length of the active contours gives rise to a class of geometrically constrained clustering algorithms in which data elements are grouped both by value and by mutual proximity. We should point out that the approach presented here

shares common aspects with the region competition approach of Zhu-Yuille [12, 13]. The relationship between the two approaches is discussed in detail in [11].

In contrast to [12, 13] and most other region based approaches to segmentation, we operate on a set of independent curves and use these curves to define a set of regions as opposed to operating on a set of regions and using their boundaries to define a set of curves. The difference may sound subtle at first, but the key benefits of working directly with curves come at the implementation level. In particular, we are able to utilize level set techniques (see Osher and Sethian [7, 10] and the references therein) to implement the flows presented in this paper.

The remainder of this paper is a shortened version of [11], to which we refer the interested reader for all of the details (including the level set implementations) and a much more complete set of references to other curve evolution techniques for segmentation.

2 Binary Flows

In this section, we present gradient flows designed to segment bimodal images via an evolving curve. The flow presented in the first part of this section for simple binary intensity images is offered for the purpose of illustration (since such images are already segmented) and to develop an intuition for more general flows designed for less trivial forms of bimodal imagery.

2.1 Flows for binary images

We begin with the assumption that the domain of an image $I(x, y)$ consists of a foreground region R of intensity I^r and a complementary background region R^c of intensity $I^c \neq I^r$. We wish to determine an evolution that will continuously attract any initial closed curve \vec{C} toward the boundary ∂R of R .

Since an arbitrary closed curve over the domain of I will enclose some portion of R and some portion of R^c , the average intensities u and v inside and outside the curve respectively are bounded above and below by I^r and I^c . Consequently, using the distance between u and v to measure how well \vec{C} has sepa-

*This work was supported by ONR grant N00014-91-J-1004, by subcontract GC123919NGD from Boston University under the AFOSR Multidisciplinary Research Program on Reduced Signature Target Recognition, and by ARO grant DAAH04-96-1-0494 through Washington University.

rated the foreground from the background will ensure an upper-bound of $|I^r - I^c|$ that is uniquely attained when $\vec{C} = \partial R$. A related strategy, which also assumes no previous knowledge of I^r or I^c , would be to descend along the following quadratic energy functional:

$$E = -\frac{1}{2}(u - v)^2. \quad (1)$$

Letting $S_u = \int_{R^u} I dA$ and $A_u = \int_{R^u} dA$, where R^u denotes the interior of \vec{C} , and expressing their first variations as $\nabla S_u = I \vec{N}$ and $\nabla A_u = \vec{N}$ (see [11] for details), where \vec{N} denotes the outward unit normal of \vec{C} , allows us to compute the first variation of $u = S_u/A_u$ as follows:

$$\nabla u = \frac{A_u \nabla S_u - S_u \nabla A_u}{A_u^2} = \frac{A_u I - S_u}{A_u^2} \vec{N} = \frac{I - u}{A_u} \vec{N}.$$

We may use this expression (and a similar expression for ∇v) to compute the following gradient descent curve evolution

$$\frac{d\vec{C}}{dt} = -\nabla E = (u - v) \left(\frac{I - u}{A_u} + \frac{I - v}{A_v} \right) \vec{N}, \quad (2)$$

yielding a flow that pulls apart the mean intensities inside and outside the curve as fast as possible.

2.2 Binary images with additive noise

Our previous model may be easily modified to handle grayscale images which are well-approximated by binary images plus additive white noise so long as the contaminating noise is zero-mean. Clearly, since our cost functional is based on *average* intensity values inside and outside the evolving contour, zero-mean noise away from the contour will not have a significant effect on its evolution. This is not the case for noise in the vicinity of the contour. The contour may end up weaving around or encircling extremely small regions due to noise in order to gain tiny decreases in the cost functional, causing the contour to appear fractal. To counter these effects, we follow the philosophy of Mumford and Shah [5, 6] by incorporating a geometric constraint on the evolving contour via an additional term in the energy functional (1) which penalizes its arclength (analogous to the internal energy term of the original snake formulation in [3]). Doing so yields the following new energy

$$E = -\frac{1}{2}(u - v)^2 + \alpha \int_{\vec{C}} ds, \quad (3)$$

where $\alpha \geq 0$ and s represents the arclength parameter of \vec{C} . Since the gradient direction for length is given

by $\kappa \vec{N}$, where κ denotes the signed curvature of \vec{C} , the corresponding gradient descent on E is given by

$$\frac{d\vec{C}}{dt} = (u - v) \left(\frac{I - u}{A_u} + \frac{I - v}{A_v} \right) \vec{N} - \alpha \kappa \vec{N}. \quad (4)$$

The influence of the second term in this flow is most strongly felt where the magnitude of the curvature is very large. This helps prevent the contour from wrapping around tiny pieces of noise, with the tradeoff that sharp corners in the underlying binary image may be rounded off by the final contour.

2.3 More general binary flows

Until now we have used the term binary to suggest two separate scalar intensities (greylevels). We may readily generalize our results to the vector-valued case by employing the following more general energy functional:

$$E = -\frac{1}{2} \|u - v\|^2 + \alpha \int_{\vec{C}} ds, \quad (5)$$

where $u = (u_1, \dots, u_n)$ and $v = (v_1, \dots, v_n)$ now denote average values of some vector-valued measurement $I(x, y) = (I_1(x, y), \dots, I_n(x, y))$ inside and outside the curve respectively. The corresponding gradient flow is given by

$$\frac{d\vec{C}}{dt} = \sum_{i=1}^n (u_i - v_i) \left(\frac{I_i - u_i}{A_u} + \frac{I_i - v_i}{A_v} \right) \vec{N} - \alpha \kappa \vec{N}. \quad (6)$$

The measurements I_1, \dots, I_n do not necessarily have to represent image intensities (as in a color image) but may represent wavelet coefficients from a grayscale image or other forms of multi-spectral measurements. With this observation, one could segment an image consisting of two different textures using (6) so long as a distinguishing “texture vector” can be derived.

Finally, the binary approach may be further generalized by considering statistics other than means. The basic idea is to formulate a set of statistics which distinguish the foreground and background regions from each other and then derive curve evolutions to “pull them apart”.

Suppose, for example, that an image consists of two regions with identical means but different variances. In this case, one could descend along the following energy functional

$$E = -\frac{1}{2} (\sigma_u^2 - \sigma_v^2)^2 + \alpha \int_{\vec{C}} ds, \quad (7)$$

where σ_u^2 and σ_v^2 denote the sample variances inside and outside \vec{C} , using the corresponding gradient flow:
$$\frac{d\vec{C}}{dt} = \left\{ (\sigma_u^2 - \sigma_v^2) \left(\frac{(I - u)^2 - \sigma_u^2}{A_u} + \frac{(I - v)^2 - \sigma_v^2}{A_v} \right) - \alpha \kappa \right\} \vec{N}. \quad (8)$$

3 Generalization

In this section, we generalize the methodology of Section 2 to develop flows for segmenting trimodal or more general forms of multimodal imagery. The *binary flows* (2), (4), (6), and (8) of Section 2 partition an image domain into exactly two regions. These regions may be multiply connected, consisting of many individual *subregions*; however, the evolving contour distinguishes just two *classes* of regions at any given time.

In the first part of this section, we present a framework for handling *ternary flows*, which partition an image domain into three different region classes. Later, the approach is generalized for an arbitrary number of classes.

3.1 Ternary flows

We begin our discussion of ternary flows by assuming (for now) that the domain of an image $I(x, y)$ consists of two disjoint, simply connected, foreground regions R^a and R^b and a background region R^c (the complement of $R^a \cup R^b$) with mutually distinct intensities I^a , I^b , and I^c , respectively. A closed curve \vec{C}_u in the domain of I will generally enclose some portion of each region; thus, the average intensity u inside \vec{C}_u can be written as a convex combination of I^a , I^b , I^c (i.e. $u = \alpha I^a + \beta I^b + \gamma I^c$ where $0 \leq \alpha, \beta, \gamma \leq 1$ and $\alpha + \beta + \gamma = 1$). Unfortunately, if I takes its values in \mathbf{R} , there is no unique convex combination since any three points in \mathbf{R} are obviously collinear. This poses a problem since the algorithm we are about to present relies upon geometrically independent¹ statistics to distinguish the regions R^a , R^b , and R^c .

To be geometrically independent I^a , I^b , and I^c must belong to \mathbf{R}^2 or a higher dimensional space. Accordingly, assume that I is a vector-valued image with vectors in \mathbf{R}^2 and that $I^a = (I_1^a, I_2^a)$, $I^b = (I_1^b, I_2^b)$, and $I^c = (I_1^c, I_2^c)$ are geometrically independent. We may now represent $u = (u_1, u_2)$ as a unique convex combination of these three values. The same situation applies to the average intensity v within the interior of a second curve \vec{C}_v and to the average intensity w within the mutual exterior of \vec{C}_u and \vec{C}_v . Our segmentation goal is to construct coupled flows that will continuously attract \vec{C}_u toward one of the boundaries ∂R^a or ∂R^b (of R^a and R^b respectively) while simultaneously attracting \vec{C}_v toward the other.

By virtue of their geometric independence, I^a , I^b , and I^c form the vertices of a triangle T_{abc} . As convex combinations of these three values, u , v , and w lie within this triangle, forming another triangle T_{uvw} completely contained in T_{abc} . (This is true even if the

interiors of \vec{C}_u and \vec{C}_v overlap, providing a flexibility to our approach that is not provided by region competition in which evolving regions must be disjoint.) As such, the area of the triangle T_{uvw} will always be less than or equal to the area of the triangle T_{abc} , with equality holding if and only if $\vec{C}_u = \partial R^a$ and $\vec{C}_v = \partial R^b$ or vice-versa. We may therefore attract \vec{C}_u and \vec{C}_v toward the desired boundaries without any prior knowledge of I^a , I^b , or I^c by trying to maximize the area of T_{uvw} using the following tri-quadratic energy functional:

$$E = -\frac{1}{2} \det^2(u - w, v - w) = -2 \text{ area}^2(T_{uvw}). \quad (9)$$

If u , v , and w are geometrically independent, then $u - w$ and $v - w$ are linearly independent and therefore yield a nonzero determinant.

By computing the partial variations $\nabla_{\vec{C}_u} E$ and $\nabla_{\vec{C}_v} E$ of this energy functional with respect to \vec{C}_u and \vec{C}_v , we may derive the following pair of coupled gradient descent equations (a detailed derivation appears in [11]).

$$\begin{aligned} \frac{d\vec{C}_u}{dt} = & (u_1 v_2 - u_1 w_2 + v_1 w_2 - v_1 u_2 + w_1 u_2 - w_1 v_2) \times \\ & \left\{ (v_2 - w_2) \frac{I_1 - u_1}{A_u} - (v_1 - w_1) \frac{I_2 - u_2}{A_u} - \right. \\ & \left. (u_2 - v_2) \frac{I_1 - w_1}{A_w} (1 - \chi_v) + (u_1 - v_1) \frac{I_2 - w_2}{A_w} (1 - \chi_v) \right\} \vec{N}_u \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{d\vec{C}_v}{dt} = & (u_1 v_2 - u_1 w_2 + v_1 w_2 - v_1 u_2 + w_1 u_2 - w_1 v_2) \times \\ & \left\{ (w_2 - u_2) \frac{I_1 - v_1}{A_v} - (w_1 - u_1) \frac{I_2 - v_2}{A_v} - \right. \\ & \left. (u_2 - v_2) \frac{I_1 - w_1}{A_w} (1 - \chi_u) + (u_1 - v_1) \frac{I_2 - w_2}{A_w} (1 - \chi_u) \right\} \vec{N}_v \end{aligned} \quad (11)$$

where \vec{N}_u and \vec{N}_v denote the outward unit normals of \vec{C}_u and \vec{C}_v and χ_u and χ_v denote the characteristic functions over R^u and R^v (the interiors of \vec{C}_u and \vec{C}_v respectively)

When R^u and R^v are disjoint, the evolution of each curve is not directly tied to the other curve. The coupling, arising from the common set of parameters u , v , and w , is indirect. The characteristic functions χ_u and χ_v yield a more direct coupling when the curves overlap. Nevertheless, in both cases, each curve evolves as a separate entity, enabling the use of curve evolution rather than region-based methods. Level set implementations in particular allow automatic merging and splitting of initial contours (see [11] for a discussion of these implementations and their advantages).

Note that the evolution of each curve depends upon statistics computed over every region in the image. In

¹Noncollinear in this context.

this sense, (10) and (11) comprise a truly global model for segmentation. On the other hand, the need for vector-valued statistics imposes a restriction on the types of data acceptable to our algorithm. The need for a vector-valued statistic, however, does not necessarily require vector-valued data. Ternary flows may be applied to greyscale images, for example, by considering both means and variances.

3.2 More general ternary flows

We now modify the flows (10) and (11) and their associated energy functional (9) to handle more general forms of trimodal imagery.

First, we allow the vector-valued data I to take its values in \mathbf{R}^n where $n \geq 2$ as opposed to just \mathbf{R}^2 . Unfortunately, the determinant in (9) no longer makes sense when $n > 2$. However, three noncollinear points, $I^a, I^b, I^c \in \mathbf{R}^n$ still comprise a triangle in \mathbf{R}^n , and $u, v, w \in \mathbf{R}^n$, as convex combinations of these values, will always lie inside this triangle (within the context of its two-dimensional plane). We may therefore generalize the ternary energy functional with the same goal of maximizing the area of the triangle T_{uvw} :

$$|\text{area}(T_{uvw})| = \frac{1}{2} \|u - w\| \|v - w\| \sin \theta$$

$$4 \text{ area}^2(T_{uvw}) = \|u - w\|^2 \|v - w\|^2 - ((u - w) \cdot (v - w))^2$$

where θ denotes the angle between $u - w$ and $v - w$.

Next, we attach a geometric penalty on the lengths of \vec{C}_u and \vec{C}_v (as in Section 2.2) to handle the presence of zero-mean noise in the image. In general one may penalize the two lengths differently; here we consider an equal penalty and rewrite (9) more generally as

$$E = -2 \text{ area}^2(T_{uvw}) + \alpha \left(\int_{\vec{C}_u} ds + \int_{\vec{C}_v} ds \right) \quad (12)$$

where $\alpha \geq 0$. We now use the previous expression to compute the variation of the first term

$$\nabla(2 \text{ area}^2(T_{uvw})) = \{ \bar{w} \cdot \nabla u + \bar{u} \cdot \nabla v + \bar{v} \cdot \nabla w \} \vec{N}$$

with the following definitions:

$$\nabla u = (\nabla u_1 \cdot \vec{N}, \dots, \nabla u_n \cdot \vec{N})$$

(likewise for ∇v and ∇w)

$$\begin{aligned} \bar{u} &= \tilde{u} - \tilde{v}, & \tilde{u} &= \hat{u}(\hat{v} \cdot \hat{w}), & \hat{u} &= u - v \\ \bar{v} &= \tilde{v} - \tilde{w}, & \tilde{v} &= \hat{v}(\hat{w} \cdot \hat{u}), & \hat{v} &= v - w \\ \bar{w} &= \tilde{w} - \tilde{u}, & \tilde{w} &= \hat{w}(\hat{u} \cdot \hat{v}), & \hat{w} &= w - u \end{aligned}$$

Since $\nabla_{\vec{C}_u} v = \nabla_{\vec{C}_v} u = 0$ the gradient descent equations for E become

$$\frac{d\vec{C}_u}{dt} = \left\{ \sum_{i=1}^n \left(\bar{w}_i \frac{I_i - u_i}{A_u} - \bar{v}_i (1 - \chi_v) \frac{I_i - w_i}{A_w} \right) - \alpha \kappa_u \right\} \vec{N}_u \quad (13)$$

$$\frac{d\vec{C}_v}{dt} = \left\{ \sum_{i=1}^n \left(\bar{u}_i \frac{I_i - v_i}{A_v} - \bar{v}_i (1 - \chi_u) \frac{I_i - w_i}{A_w} \right) - \alpha \kappa_v \right\} \vec{N}_v \quad (14)$$

where κ_u and κ_v denote the signed curvatures of \vec{C}_u and \vec{C}_v respectively.

3.3 Segmenting more than three regions

In general one may wish to partition an image domain into m different types of regions, where m is an arbitrarily large number. By adhering to the same philosophy of associating the preferred segmentation with a maximal separation of some statistic over each region, a vector-valued statistic, U , with at least $m - 1$ components would be required. If the m distinct values, U^1, \dots, U^m , of this statistic constitute a set of geometrically independent points in the preferred segmentation of the image, and if the statistic is chosen such that an arbitrary segmentation yields values u^1, \dots, u^m , which are convex combinations of U^1, \dots, U^m (which is the case if we are considering means of a vector-valued image) then the natural energy functional will relate to the volume of the $m - 1$ dimensional simplex whose vertices are given by u^1, \dots, u^m . The corresponding gradient flow equations will yield a coupled evolution of $m - 1$ curves which tend to maximize the volume of this simplex, with the interiors of each curve representing $m - 1$ regions and their mutual exteriors representing the m 'th region.

4 Simulations

In this section, we demonstrate the performance of binary and ternary flows on real data.

Flow (4) is used in Fig. 1, to segment a microscopic image of red blood cells, providing a compelling demonstration of the topological transitions allowed by its level set implementation (see [7, 10, 4, 11] for details). A single initial contour appears in the first frame; the multiple steady state contours and the resulting segmentation, showing the steady state mean intensity values, appear in the last two frames.

The synthetic aperture radar (SAR) image of a forest's tree line in Fig. 2 constitutes a bimodal image of a rather different nature. Means cannot be used here to distinguish one region from the other. The forest region in the lower left half of the image and the grassy region in the upper right half of the image give rise to two different textures with approximately the same greyscale mean, but with different variances. Flow

(8), therefore, is able to segment the image quite successfully by separating variances rather than means.

Means and variances may also be used together in this methodology as components of a two-dimensional vector which must be chosen to minimize the energy functional (5). However, due to the dissimilarity between these two statistics, their first variations have different forms. Thus the gradient flow equation is not given by (6) but by a hybrid flow using the sum of the image-driven terms of (4) and (8). Such a flow was used to capture the tadpole dermal cells in the optical coherence tomography (OCT) image of Fig. 3.

The coupled ternary flows (13) and (14) were used to segment the clouds, the sky, and the B-2 bomber from the color image shown in Fig. 4. The final segmentation, showing the steady state mean color intensity values, appears in the last frame.

5 Conclusions

We have presented a novel statistical approach to snakes for the segmentation of images which are known *a priori* to consist of a given number of regions distinguishable by a given set of statistics. The resulting gradient flows, derived from *deterministic* considerations, were designed to essentially pull the values of these statistics as far apart as the data in a given image would allow, subject to geometric constraints on the length of the active contour(s).

Two key attractions of the flows in this paper were a natural use of both local and global information in the image and a deliberate avoidance of differential operators for detecting edges. In addition, our adherence to separate (although coupled) curve evolution equations enabled the use of level set techniques in the implementation of our flows.

To summarize, we have outlined a very general curve evolution approach to segmentation that clusters pixels in an image based upon both geometric and statistical considerations. The performance of our algorithm depends upon how well the chosen set of statistics is able to distinguish the various regions within a given image. Specifically, we have demonstrated the use of means, and variances as the discriminating statistics.

References

- [1] A. Chakraborty and J. Duncan, "Game-Theoretic Integration for Image Segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 1, pp. 12–30, Jan. 1999.
- [2] A. Chakraborty, L. Staib, and J. Duncan, "Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information," *IEEE Trans. Medical Imaging*, vol. 15, no. 6, pp. 859–870, Dec. 1996.
- [3] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. Journal of Computer Vision*, vol. 1, pp. 321–331, 1987.
- [4] R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation: a level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, 1995.
- [5] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications in Pure and Applied Mathematics*, vol. 42, no. 4, 1989.
- [6] D. Mumford and J. Shah, "Boundary detection by minimizing functionals," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 1985.
- [7] S. Osher and J. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.
- [8] N. Paragios and R. Deriche, "Geodesic Active Regions for Texture Segmentation," Research Report 3440, INRIA, France, 1998.
- [9] R. Ronfard, "Region-Based Strategies for Active Contour Models," *Int. J. Computer Vision*, vol. 13, no. 2, pp. 229–251, 1994.
- [10] J. Sethian, *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science*, Cambridge University Press, 1996.
- [11] A. Yezzi, A. Tsai, and A. Willsky, "A Statistical Approach to Curve Evolution for Image Segmentation," LIDS Technical Report, January, 1999 (submitted to *IEEE Trans. on Pattern Analysis and Machine Intelligence*).
- [12] S. C. Zhu, T. S. Lee, and A. L. Yuille, "Region Competition: Unifying snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *Proc. of ICCV*, pp. 416–423, 1995.
- [13] S. Zhu and A. Yuille, "Region Competition: Unifying snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996.

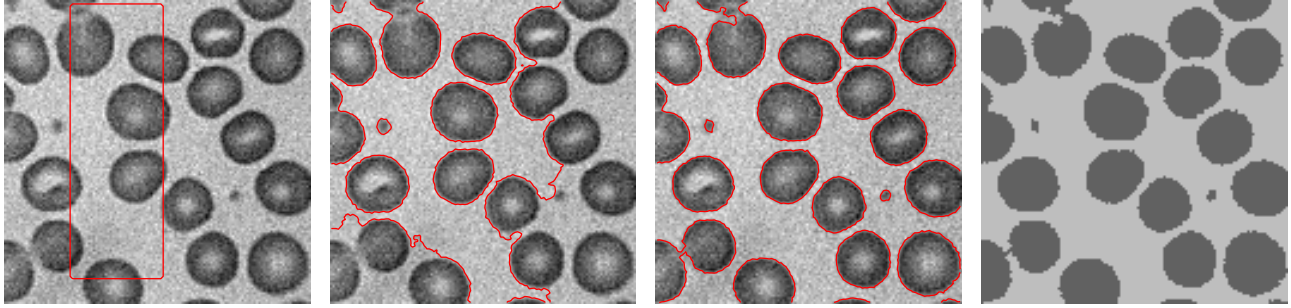


Figure 1: Multiple red blood cells are captured by a single contour using flow (4).

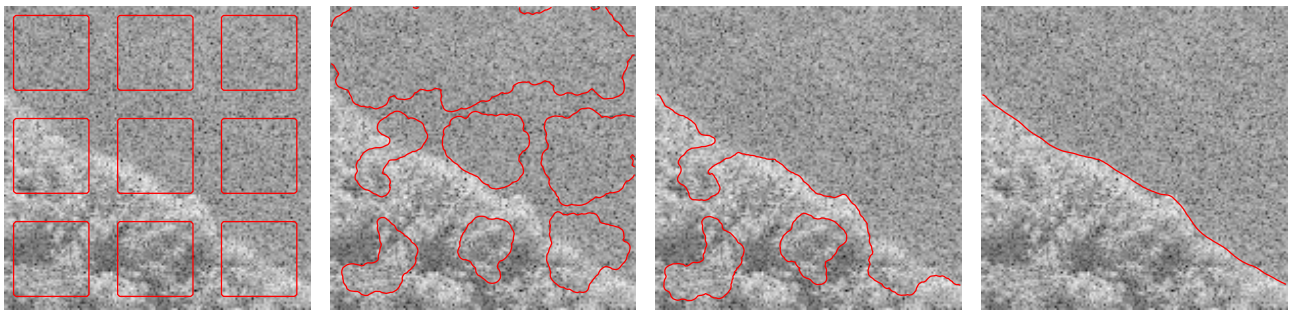


Figure 2: The tree line shown here is captured using flow (8) to separate variances.

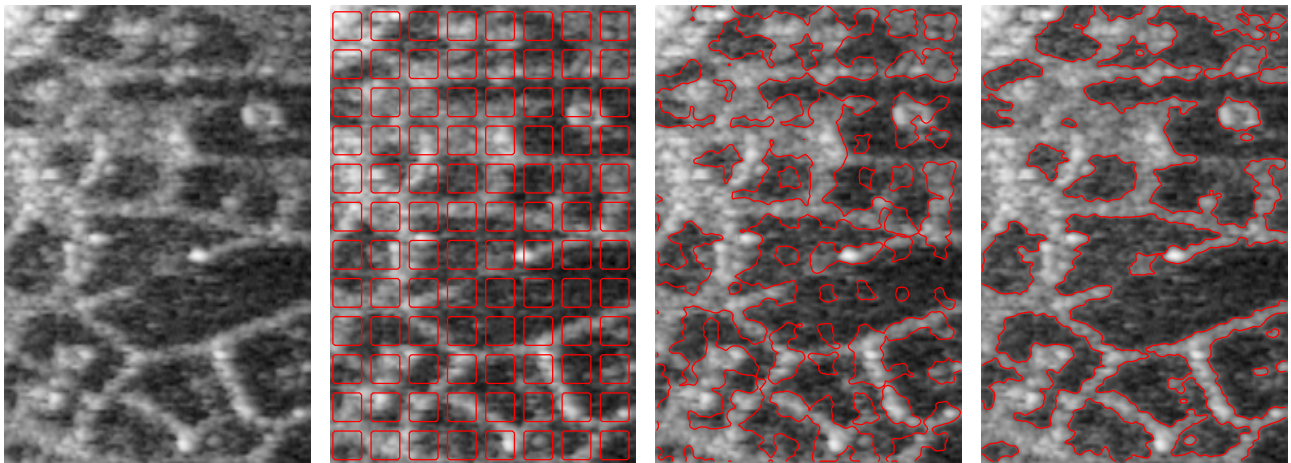


Figure 3: Both means and variances are used to segment this OCT image of tadpole cells (image courtesy of S. Boppart and J. Fujimoto of MIT, and appears in *Nature Medicine*, vol. 4, pp. 861–865, July 1998).

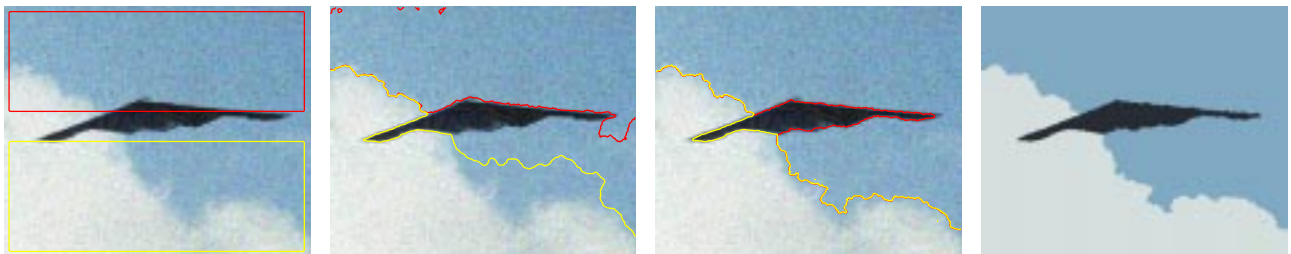


Figure 4: The B-2 bomber, clouds, and sky are captured by coupled flows (13) and (14).