Journal of Visual Communication and Image Representation **13**, 195–216 (2002)
doi:10.1006/jvci.2001.0500, available online at http://www.idealibrary.com on IDEAL®

A Fully Global Approach to Image Segmentation via Coupled Curve Evolution Equations[1]

Anthony Yezzi, Jr.

School of ECE, Georgia Institute of Technology, Atlanta, Georgia 30332-0250

and

Andy Tsai and Alan Willsky

Department of EECS, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received January 31, 2000; accepted June 2, 2000

In this paper, we develop a novel region-based approach to snakes designed to optimally separate the values of certain image statistics over a known number of region types. Multiple sets of contours deform according to a coupled set of curve evolution equations derived from a single global cost functional. The resulting active contour model, in contrast to many other edge and region based models, is fully global in that the evolution of each curve depends at all times upon every pixel in the image and is directly coupled to the evolution of every other curve regardless of their mutual proximity. As such evolving contours enjoy a very wide "field of view," endowing the algorithm with a robustness to initial contour placement above and beyond the significant improvement exhibited by other region based snakes over earlier edge based snakes. © 2002 Elsevier Science (USA)

Key Words: active contours; curve evolution; snakes; segmentation; gradient flows.

1. INTRODUCTION

Curve evolution methods have been applied to a variety of problems in image processing and computer vision including image smoothing, morphology, shape modeling, edge-detection, and segmentation. Since the introduction of the snake methodology [10, 27] active contours have become particularly popular for segmentation applications (see [1] and the references therein for more general applications). A number of recent works [5, 6, 21, 29] have sought to combine curve evolution techniques with statistical approaches to segmentation. Much of this work has been motivated by the earlier work of Ronfard [23] as well as Zhu and Yuille [31, 32].

[1] 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.

1047-3203/02 $35.00

In [29, 30], Yezzi *et al*. proposed a new class of active contours, combining curve evolution and statistics in a natural way for images consisting of a known number of region types. In contrast to other snake algorithms, multiple sets of contours, evolving via separate but coupled curve evolution equations, are employed to segment an image into multiple region classes. The coupling between the evolution equations, which does not depend upon the mutual proximities of each set of contours, causes every single pixel in the image to influence the flow of every individual contour. This fully global approach to segmentation is, thus, even more robust to initial contour placement when compared to other region based approaches [6, 23, 31, 32] (in which the evolution of a boundary depends only upon the pixels within regions which share the given boundary) which are already more robust than local edge-based approaches [2, 3, 7, 10, 11, 16, 26–28] (in which the evolution of a curve depends only upon nearby pixels). See [4, 5, 21] for some heterogeneous approaches which blend both region and boundary information. Standard level set implementations (see Osher and Sethian [20, 25] and the references therein) of these coupled curve evolutions involve separate level set functions for each set of curves. This offers a simple mechamism to treat triple points and other multiple junctions. Some other recent level-set, coupled snake algorithms [22, 24], though different, enjoy many of the same advantages as our model.

The key idea behind the coupled curve evolution model presented in [29, 30] was to design curve evolution equations which "pull apart" the values of one or more image statistics (e.g., means, variances, textures) within each region. In this paper, we introduce a new constraint which guarantees that the relevant statistics evolve "away from each other" as opposed to just "apart." This constraint will be treated in Section 4. We first present the original model in Section 2, followed by a comparison in Section 3 with other recent region-based models. In Section 5 we discuss the numerical implementation of both the original and the constrained models using level set techniques [20, 25] followed by simulations in Section 6.

2. COUPLED CURVE EVOLUTION MODELS

In this section, we present the original coupled curve evolution models outlined in [29, 30]. These are based upon gradient flows designed to "pull-apart" the values of two or more image statistics and are therefore useful for segmenting images in which regions are differentiated by the values of a known set of statistics. We start out with the simple case of bimodal imagery in which there are two types of regions, foreground and background. In this case, a single set of curves (and a single statistic) is sufficient to delineate the boundaries between foreground and background. We refer to the resulting gradient flows as *binary flows*. For trimodal imagery, we use two sets of curves to distinguish two different foreground classes from a common background (or vice-versa), explaining our reference to these techniques as *coupled curve evolution models*. We conclude this section by discussing a general approach for more than three region classes (by employing more than two sets of curves).

2.1. Binary Flows

We begin by considering images which consist of just two region types. The most trivial case is that of a binary image $I(x, y)$ consisting of a single 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. Given that $\vec{C}$ 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 separated 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 previous knowledge of I^r or I^c, would be to minimize the following energy:

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

Let $S_u = \int_{R^u} I \, dA$ and $A_u = \int_{R^u} dA$, where R^u denotes the interior of $\vec{C}$. By expressing their gradients (in the space of curves) as $\nabla S_u = I\vec{N}$ and $\nabla A_u = \vec{N}$ (see Appendix 1 in [29] for a derivation), where $\vec{N}$ denotes the outward unit normal of $\vec{C}$, allows us to compute the gradient of $u = S_u/A_u$ as follows:

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

By a similar computation, we may express the gradient of v as

$$\nabla v = -\frac{I - v}{A_v}\vec{N}, \tag{3}$$

where A_v denotes the area of R^v (the exterior of $\vec{C}$). Note the minus sign due to the fact that $\vec{N}$ is the outward normal with respect to R^u and so the outward normal with respect to its complement R^v is $-\vec{N}$. We may use these expressions to compute the gradient flow for E

$$\frac{d\vec{C}}{dt} = -\nabla E = (u - v)\left(\frac{I - u}{A_u} + \frac{I - v}{A_v}\right)\vec{N}, \tag{4}$$

which optimally separates the mean intensities inside and outside the curve $\vec{C}$.

An attractive property of this flow is that it automatically proceeds in the correct direction without using an inflationary term common to earlier snake algorithms. An initial contour encompassing the region R will flow inward toward the boundary; a contour inside R will flow outward; and a contour which overlaps R will flow in both directions toward the boundary. Thus, no prior knowledge of the evolution direction is required, making initial contour placement less restrictive (a property shared by many other region-based snake algorithms [5, 6, 21, 23, 32], constituting a significant improvement over edge-based algorithms).

In a more general bimodal image consisting of light regions and dark regions (i.e., not strictly binary) contaminated by noise (see Fig. 1), a contour may tend to weave around

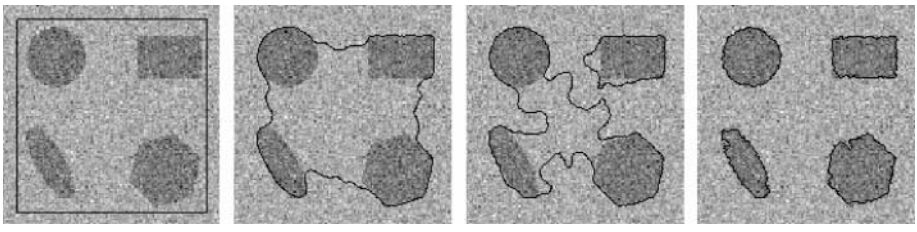

FIG. 1. Flow (6) captures regions with different mean greyscale intensity values.

or encircle extremely small regions due to noise in order to fully separate the interior and exterior means, taking on a fractal appearance in the process. To counter such effects, we follow the lead of Mumford and Shah [17, 18] by adding to our functional (1) a penalty on the arclength of the curve.

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

This geometric penalty regularizes the gradient flow through a curvature-based term

$$\frac{d\vec{C}}{dt} = (u - v)\left(\frac{I - u}{A_u} + \frac{I - v}{A_v}\right)\vec{N} - \alpha\kappa\vec{N} \tag{6}$$

which exerts its strongest influence at contour points with significant curvature. This discourages the contour from wrapping around tiny pieces of noise, with the tradeoff that sharp corners of image objects may be rounded off by the final contour.

We may readily generalize our model to bimodal vector-valued data (e.g., color or multispectral imagery) by employing the following more general energy functional

$$E = -\frac{1}{2}\|u - v\|^2 + \alpha \int_{\vec{C}} ds, \tag{7}$$

where u and v are now vector-valued averages of the data inside and outside the curve respectively. In the case of n-dimensional measurements, $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$, and $I(x, y) = (I_1(x, y), \ldots, I_n(x, y))$, the gradient flow is given by

$$\frac{d\vec{C}}{dt} = \sum_{i=1}^{n}(u_i - v_i)\left(\frac{I_i - u_i}{A_u} + \frac{I_i - v_i}{A_v}\right)\vec{N} - \alpha\kappa\vec{N}. \tag{8}$$

Note that the data components $I_1, \ldots, I_n$ in (8) do not have to represent intensities (as in a color image) but may represent wavelet coefficients, filter responses, or other measurements. The wood grain textures, oriented in two different directions in Fig. 2, for example, were captured via (8) by using two-dimensional unit vectors aligned with the gradient of the greyscale image.

We may further generalize the bimodal model by letting u and v denote statistics other than means. The image in Fig. 3, for example, contains two regions with identical means but different variances. In this case, the following energy functional

$$E = -\frac{1}{2}\left(\sigma_u^2 - \sigma_v^2\right)^2 + \alpha \int_{\vec{C}} ds \tag{9}$$

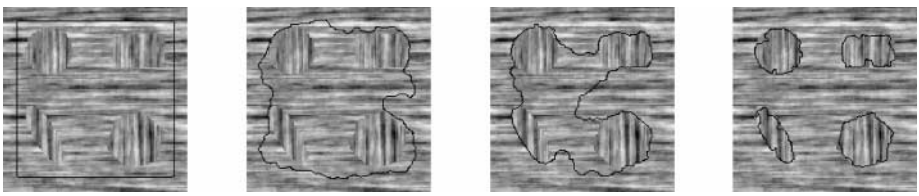

FIG. 2. Flow (8) captures two different wood grain textures with opposite orientations.

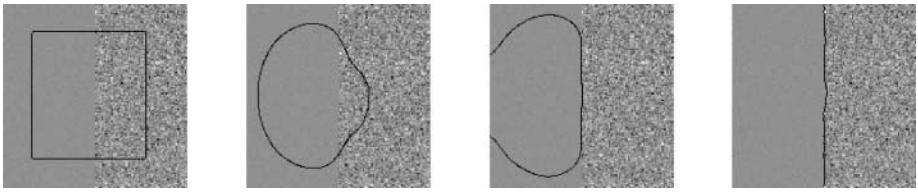

FIG. 3. Flow (10) captures two regions with identical means but different variances.

(where $\sigma_u^2 = \frac{1}{A_u}\int_{R^u}(I-u)^2\,dA$ and $\sigma_v^2 = \frac{1}{A_v}\int_{R^v}(I-v)^2\,dA$ denote the variances inside and outside $\vec{C}$) was used to capture the boundary via the gradient flow

$$\begin{aligned}\frac{d\vec{C}}{dt} &= \left(\sigma_u^2 - \sigma_v^2\right)\left(\nabla\sigma_u^2 - \nabla\sigma_v^2\right) - \alpha\kappa\vec{N} \\ &= \left\{\left(\sigma_u^2 - \sigma_v^2\right)\left(\frac{(I-u)^2 - \sigma_u^2}{A_u} + \frac{(I-v)^2 - \sigma_v^2}{A_v}\right) - \alpha\kappa\right\}\vec{N}. \qquad (10)\end{aligned}$$

2.2. Ternary Flows

We now extend our methodology to the problem of segmenting images with three region classes by utilizing two different sets of curves with separate evolution equations. The flows, however, will be derived from a common energy functional, giving rise to a global coupling regardless of the positions of the curves.

We begin by considering an image $I(x, y)$ whose domain 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 $0 \le \alpha, \beta, \gamma \le 1$ and $\alpha + \beta + \gamma = 1$). Unfortunately, if $I \in \mathbf{R}$, this combination is not unique. Uniqueness requires geometric independence (noncollinearity in this context) which is not possible for three values in $\mathbf{R}$, we must move to $\mathbf{R}^n$ where $n \ge 2$. Let us therefore assume that I is a vector-valued image with at least two components (such as the image in Fig. 4 containing red and green channels) and that $I^a = (I_1^a, \ldots, I_n^a)$, $I^b = (I_1^b, \ldots, I_n^b)$, and $I^c = (I_1^c, \ldots, I_n^c)$ are geometrically independent. We may now represent $u = (u_1, \ldots, u_n)$ as a unique convex combination of these three values. The same situation applies to the average vector intensity v within the interior of a second curve $\vec{C}_v$ and to the average vector 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 other region-based approaches). 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

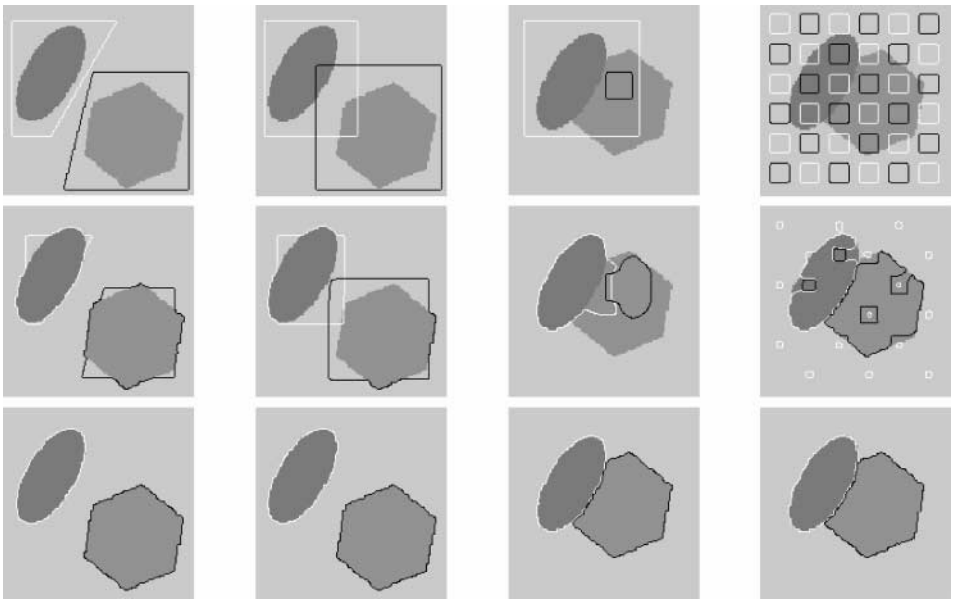

FIG. 4. The coupled ternary flows (13)–(14) drive two different contours toward separate regions, even if the initial contours overlap (b,c). Separate level set functions enable simple capturing of triple points (c,d) and confine topological changes to contours in the same set (d).

without any prior knowledge of I^a, I^b, or I^c by trying to maximize the area of T_{uvw} (given by $\frac{1}{2}\|u-w\|\|v-w\|\sin\theta$, where θ denotes the angle between $u-w$ and $v-w$). To simplify matters, we seek to minimize the following energy functional, utilizing the squared area (which is unsigned and does not explicitly involve θ) and incorporating arclength penalties as we did in the binary case.

$$\begin{aligned} E &= -2\,\text{area}^2(T_{uvw}) + \alpha\left(\int_{\vec{C}_u} ds + \int_{\vec{C}_v} ds\right) \\ 4\,\text{area}^2(T_{uvw}) &= \|u-w\|^2\|v-w\|^2 - ((u-w)\cdot(v-w))^2 \end{aligned} \tag{11}$$

Coupled evolution equations for $\vec{C}_u$ and $\vec{C}_v$ may now be obtained by computing the gradient flow for each curve. Since these computations are rather involved, we will only present some of the key formulas used to obtain the final expressions.

The gradients of the components of u and v may be derived along the lines presented in the beginning of Section 2.1

$$\begin{aligned} \nabla_{\vec{C}_u} u_i &= \frac{I_i - u_i}{A_u}\vec{N}_u, \quad \nabla_{\vec{C}_v} u_i = 0, \\ \nabla_{\vec{C}_v} v_i &= \frac{I_i - v_i}{A_v}\vec{N}_v, \quad \nabla_{\vec{C}_u} v_i = 0, \end{aligned}$$

where $\vec{N}_u$ and $\vec{N}_v$ denote the outward unit normals of $\vec{C}_u$ and $\vec{C}_v$, respectively, while A_u and A_v denote their areas. The gradients of the components of w are more tricky since the interiors of $\vec{C}_u$ and $\vec{C}_v$ may overlap. In this case, the gradient directions are no longer given by smooth, or even continuous, variations of $\vec{C}_u$ and $\vec{C}_v$. The following expressions are

derived in Appendix 2 of [29]

$$\nabla_{\vec{C}_u} w_i = -\frac{I_i - w_i}{A_w}(1 - \chi_v)\vec{N}_u, \quad \nabla_{\vec{C}_v} w_i = -\frac{I_i - w_i}{A_w}(1 - \chi_u)\vec{N}_v,$$

where 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). If we define ∇u, ∇v, and ∇w as vectors of gradient flows for the components of u, v, and w respectively (where ∇ denotes either $\nabla_{\vec{C}_u}$ or $\nabla_{\vec{C}_v}$) we may express the gradient of the area (squared) of T_{uvw} as

$$\nabla(2 \text{ area}^2(T_{uvw})) = \bar{w} \cdot \nabla u + \bar{u} \cdot \nabla v + \bar{v} \cdot \nabla w, \tag{12}$$

where $\bar{u}$, $\bar{v}$, and $\bar{w}$ are functions of u, v, and w defined as follows:

$$\begin{array}{lll} \bar{u} = \tilde{u} - \tilde{v} & \tilde{u} = \hat{u}(\hat{v} \cdot \hat{w}) & \hat{u} = u - v \\ \bar{v} = \tilde{v} - \tilde{w} & \tilde{v} = \hat{v}(\hat{w} \cdot \hat{u}) & \hat{v} = v - w \\ \bar{w} = \tilde{w} - \tilde{u} & \tilde{w} = \hat{w}(\hat{u} \cdot \hat{v}) & \hat{w} = w - u. \end{array}$$

Using expression (12) (and noting that $\nabla_{\vec{C}_u} v = \nabla_{\vec{C}_v} u = 0$) we obtain the following coupled gradient flows (demonstrated in Fig. 4 on a two-channel red/green image)

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

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

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

2.3. More General Coupled Flows

In general, one may develop fully global, coupled flows to segment an image into m different types of regions, where m is arbitrarily large. Generalizing the intuition behind the binary and ternary models, one associates the preferred segmentation with a maximal separation of the values of some statistic over each region. To obtain geometric independence (which allows the desired global coupling), we require a vector-valued statistic U with at least $m - 1$ components and assume that the image to be segmented contains m region classes characterized by distinct, geometrically independent values $U^1, \ldots, U^m$ of this statistic. Given $m - 1$ initial curves (or sets of curves) we may compute m values of this statistic $u^1, \ldots, u^m$ (one value within each curve and the background), which are unique convex combinations of $U^1, \ldots, U^m$. We then evolve the curves in order to maximize the to the volume of the $m - 1$ dimensional simplex whose vertices are given by $u^1, \ldots, u^m$, which will be less than or equal to the volume of the $m - 1$ dimensional simplex whose vertices are given by $U^1, \ldots, U^m$ (with equality holding if and only if the curves have captured m regions in the image). The corresponding gradient flow equations will yield $m - 1$ fully global, coupled curve evolutions.

3. COMPARISON WITH OTHER REGION BASED METHODS

In this section we compare and contrast our coupled curve evolution models with other region based approaches to segmentation. A number of approaches that combine both region and edge-based information have been considered recently [4, 5, 21, 22]. Since our models are strictly region based we will restrict our comparisons to other approaches that are strictly region based as well.

Some early region-based strategies for active contours were put forth by Ronfard in [23]. Shortly thereafter Zhu and Yuille, building on the work of Leclerc [13], proposed an MDL-based (minimum description length) system consisting of a smoothing term based on Euclidean curve shortening and a data driven term to affect the motion of a point on the common boundary of two regions via a likelihood force. Using known parametric forms for the probability density functions within each region, their algorithm consists of two alternating steps: first, fix the evolving regions and choose maximum likelihood estimates for the density parameters; next, us these estimates to evolve the region boundaries via the resulting likelihood forces. For the special case of two regions and Gaussian densities with equal variances, their algorithm leads to the following curve evolution for the shared region boundary $\vec{C}$

$$\frac{d\vec{C}}{dt} = (u - v)(I - u + I - v)\vec{N} - \alpha\kappa N. \tag{15}$$

Flow (15) also arises as the gradient flow for a special case of the Mumford–Shah functional in which piecewise constant (as opposed to piecewise smooth) functions are chosen to model the image data. This is known as the minimum partition problem and is the motivation behind the recent work of Chan and Vese in [6] who consider variants of the following energy functional,

$$E_2 = \int_{inside(\vec{C})} |I - c_1|^2 \, dA + \int_{outside(\vec{C})} |I - c_2|^2 \, dA + \textit{arclength-term}, \tag{16}$$

where c_1 and c_2 are constants which are chosen to best approximate the image data inside and outside a partitioning curve $\vec{C}$.

Note that flow (15), although different, has a similar structure to the binary flow (6) developed in Section 2.1. Initially this may seem a bit strange since the MDL based models of Zhu and Yuille region competition and the Mumford–Shah based models of Chan and Vese are quite different from our models designed to deterministically "pull apart" vector-valued statistics. However, the similarity of the resulting flows only arises in the case of exactly two regions. When more regions are involved, we will see that the formats of these flows diverge considerably.

Note that the boundaries between three regions can no longer be described by a single curve. For this reason, the Zhu–Yuille region competition algorithm (for more than just two regions) is not amenable to standard level set implementations. This is probably the reason why Chan and Vese, who do in fact use level set techniques, restrict their discussion in [6] to two region types (foreground and background) even though their Mumford–Shah based energy functional can easily be extended to more than two regions. The three-region version of (16) would be given by

$$E_3 = \int_{R_1} |I - c_1|^2 \, dA + \int_{R_2} |I - c_2|^2 \, dA + \int_{R_3} |I - c_3|^2 \, dA + \textit{arclength-term}. \tag{17}$$

Our coupled curve evolution models, in contrast to region competition, treat curves, not regions, as the objects to be deformed. Region statistics are then computed over the interiors of each curve (and their mutual exterior). This distinction leads to some interesting consequences. First, regions are allowed to overlap during the evolution process as demonstrated in Fig. 4. Second, by embedding each set of curves into separate level set functions we are able to handle triple points and other multiple junctions with no additional effort. The key is that more than one curve runs through every junction, each with its own tangent vector and curvature. Thus, the curve evolution equations remain well-defined at these junctions, allowing us to employ standard level set techniques instead of specialized multiphase variants.

When three or more regions are involved, our coupled curve evolution models, are also more global than other region based models (including the hybrid models [4, 5, 21] which incorporate edge information). In most other models the evolution of the boundary between two regions depends only upon the image data inside the two given regions. In coupled curve evolution models, the evolution of any such boundary depends upon all of the image data, not just the data within neighboring regions. In the coupled ternary flows (13)–(14) for example, we see that the evolution of curve $\vec{C}_u$ depends upon the statistics u and w within its interior and within the background region, but also upon the statistics v inside the second curve $\vec{C}_v$ regardless of its proximity to $\vec{C}_u$, and vice-versa. These coupled curve evolution equations therefore comprise a truly global model for segmentation.

Consider the three-region version of the region competition flow (15) corresponding to the energy functional (17) above. The evolution equations for the region boundaries are as follows (if we ignore the curve shortening term),

$$\begin{aligned}
\vec{C}_t &= (u-w)\cdot[(I-u)+(I-w)]\vec{N}_u, \quad \text{between } R_u \text{ and } R_w \\
\vec{C}_t &= (v-w)\cdot[(I-v)+(I-w)]\vec{N}_v, \quad \text{between } R_v \text{ and } R_w \\
\vec{C}_t &= (u-v)\cdot[(I-u)+(I-v)]\vec{N}_u, \quad \text{between } R_u \text{ and } R_v,
\end{aligned} \tag{18}$$

where $\vec{N}_u$ and $\vec{N}_v$ denote unit normals pointing away from regions R_u and R_v and where u, v, and w denote the means inside R_u, R_v, and R_w, respectively. If we use this model to segment an image with three different regions we must be careful not to place both initial contours near the same region. If we do, both contours may be attracted to that region (over-segmenting one region) and will completely miss one of the other regions (under-segmenting the background) due to the lack of coupling between the evolving curves. This is demonstrated in Fig. 5. A similar phenomenon would occur with other region based methods, including [4, 5, 21, 23].

Even if we modify the implementation of the original region competition algorithm, employing the same energy functional (17) but utilizing separate contours with separate

FIG. 5. The sensitivity of the region competition flow (18) to initial contour placement is shown by placing two contours near the same object. Both contours flow into the green object.

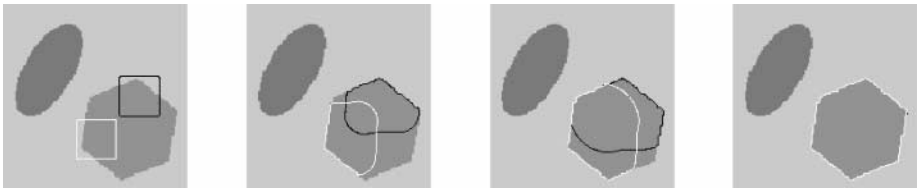

FIG. 6. A multiple level set implementation of the region competition flow via (19)–(20) still does not prevent both contours from flowing into the same object.

evolution equations as we have done in our coupled curve evolution model, we must still initialize carefully. In this case the evolution equations become

$$\frac{\partial \vec{C}_u}{\partial t} = \nabla_{\vec{C}_u} E = \{u \cdot (2I - u) - w \cdot (2I - w)(1 - \chi_v)\}\vec{N}_u \tag{19}$$

$$\frac{\partial \vec{C}_v}{\partial t} = \nabla_{\vec{C}_v} E = \{v \cdot (2I - v) - w \cdot (2I - w)(1 - \chi_u)\}\vec{N}_v, \tag{20}$$

where $\vec{C}_u$ and $\vec{C}_v$ denote two distinct curves. We may now employ level set techniques to this "modified" region competition algorithm (which also corresponds to the three region generalization of the Chan–Vese algorithm) However, both contours still flow toward the same object as illustrated in Fig. 6, they just overlap each other instead of splitting the object in two. The only coupling between flows (19)–(20) is through the background statistic w. The coupling in the ternary flows (13)–(14) from Section 2.2, however, involves all three statistics u, v, and w. This fully global coupling prevents the same two initial contours from flowing into the same region as seen in Fig. 7. The blue contour is forced out of the green object as the red contour flows into it, even though the contours do not initially touch each other. We therefore see how geometrically independent statistics are exploited our algorithm for this added robustness to initial contour placement.

4. GLOBAL CONSTRAINTS

We begin this section by observing the performance of our models on three images. The two rows in Fig. 8 illustrate successful uses of the binary flow (6) and the coupled ternary flows (13)–(14) to segment a noisy bimodal image of an Apache helicopter and a trimodal color image of a coal train. The last frames depict the final segmentations and final mean intensities. In Fig. 9 we see the famous 1950 Oregon U.F.O. image, which is clearly

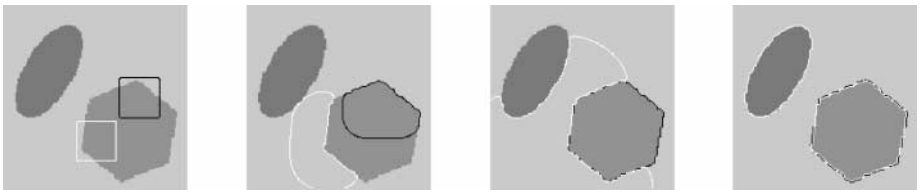

FIG. 7. The coupled ternary flows (13)–(14) do not allow both contours to flow into the green object. As one curve moves in, the other is forced out, even before they touch each other.

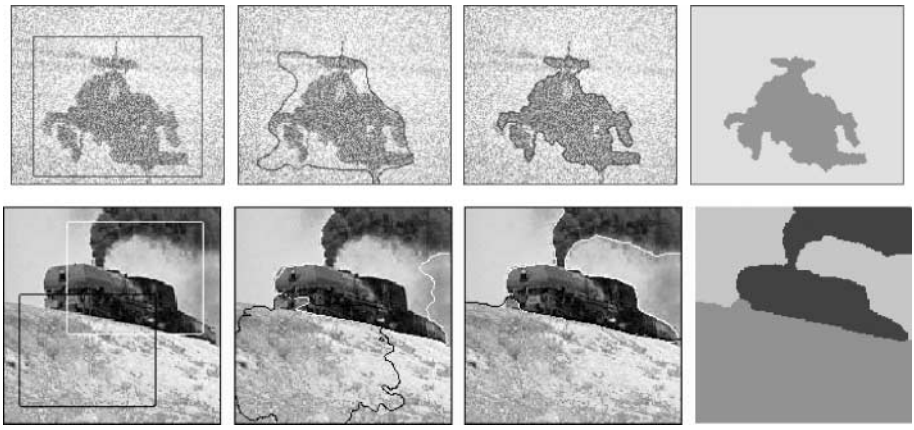

FIG. 8. (top) Original binary and (bottom) ternary flows successfully segmenting a (top) bimodal and (bottom) trimodal image.

bimodal and much less noisy than the helicopter image in Fig. 8. However, the binary flow (6), illustrated in the top row, does not capture the boundary of the object, but only a few of the dark pixels inside.

The unexpected behavior in the first row of Fig. 9 may be explained by noting that means both inside and outside the curve decrease as it moves past the boundary and into the interior of the object. However, the more rapid decrease of the interior mean (as the curve converges onto the darkest pixels inside the object) outweighs the slower decrease in the exterior mean so that overall, the two means move further apart, just as the binary model dictates. If we constrain the two means to move in opposite directions, we avoid this problem as shown in the bottom row of Fig. 9. In the following section, we discuss the implementation of this constraint.

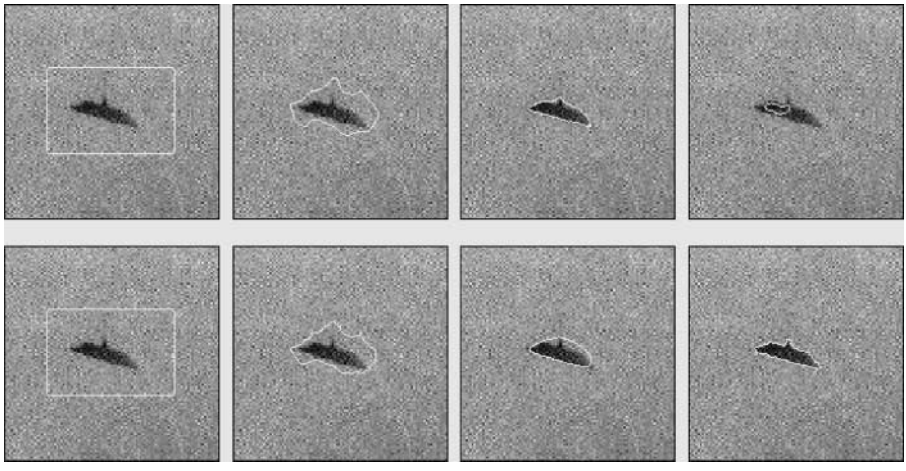

FIG. 9. (bottom) A constrained binary flow outperforming the (top) original binary flow.

4.1. Constrained Binary Flows

Let $\frac{d\vec{C}}{dt}$ and $\frac{d\vec{C}'}{dt}$ denote two admissible variations for a curve $\vec{C}$ taken from the Hilbert space of all smooth curve evolutions with the following inner product:

$$\left\langle \frac{d\vec{C}}{dt}, \frac{d\vec{C}'}{dt} \right\rangle = \oint_{\vec{C}} \left(\frac{d\vec{C}}{dt} \cdot \frac{d\vec{C}'}{dt} \right) ds. \tag{21}$$

We may express time derivatives of the statistics u and v under a binary flow as

$$u' = \langle \nabla u, -\nabla E \rangle = (u - v)(\Gamma(u, u) - \Gamma(u, v)) \tag{22}$$

$$v' = \langle \nabla v, -\nabla E \rangle = (u - v)(\Gamma(v, u) - \Gamma(v, v)), \tag{23}$$

where $-\nabla E$ denotes the gradient (descent) flow and where $\Gamma(u, v)$ denotes the inner product $\langle \nabla u, \nabla v \rangle$. If u and v represent mean intensities, ∇u and ∇v are given by (2) and (3), respectively, allowing us to express the inner products as

$$\begin{aligned} \Gamma(u, u) &= \frac{1}{A_u A_u} \oint_{\vec{C}} (I - u)(I - u)\, ds \\ \Gamma(u, v) &= \frac{-1}{A_u A_v} \oint_{\vec{C}} (I - u)(I - v)\, ds \\ \Gamma(v, v) &= \frac{1}{A_v A_v} \oint_{\vec{C}} (I - v)(I - v)\, ds. \end{aligned} \tag{24}$$

The statistics move in opposite directions when $u'v' < 0$ or, equivalently, when

$$\Gamma(u, v)(\Gamma(u, u) + \Gamma(v, v)) < \Gamma(u, u)\Gamma(v, v) + \Gamma^2(u, v). \tag{25}$$

We use the unconstrained binary flow $-\nabla E$ so long as (25) is satisfied. Otherwise we keep the statistic u or v fixed (depending upon which one is moving the wrong way) by subtracting from the evolution the component along the gradient direction for that statistic. This leads to the following constrained binary flows,

$$\frac{d\vec{C}}{dt} = -\nabla E + \frac{\langle \nabla E, \nabla u \rangle}{\langle \nabla u, \nabla u \rangle} \nabla u = \frac{u - v}{A_v}((I - u) + \gamma_u (I - v))\vec{N} \tag{26}$$

$$\frac{d\vec{C}}{dt} = -\nabla E + \frac{\langle \nabla E, \nabla v \rangle}{\langle \nabla v, \nabla v \rangle} \nabla v = \frac{u - v}{A_u}((I - v) + \gamma_v (I - u))\vec{N} \tag{27}$$

which preserve the value of u (26) or v (27). The constants γ_u and γ_v are given by

$$\gamma_u = \frac{\oint_{\vec{C}} (I - u)(I - v)\, ds}{\oint_{\vec{C}} (I - u)(I - u)\, ds} \quad \text{and} \quad \gamma_v = \frac{\oint_{\vec{C}} (I - u)(I - v)\, ds}{\oint_{\vec{C}} (I - v)(I - v)\, ds}. \tag{28}$$

4.2. A More General Constraint

For the binary flows, we want the statistics u and v to move in opposite directions. To generalize this constraint to models, such as the ternary model, which involve more than two statistics we must not think in terms of "oppoiste directions" but as follows. In the binary case, we want each statistic to evolve in a direction that would increase the distance between

the two statistics if the other statistic remained fixed. In the general case, when n statistics are involved, we want each statistic to evolve in a direction that would increase the volume of the $n-1$ dimensional simplex (whose vertices are given by the values of the n statistics, see Section 2.3), if the other $n-1$ statistics remained fixed. An equivalent geometric statement is that each vertex must move away from the opposite face of the simplex.

4.3. Constrained Ternary Flows

Here we apply the general constraint, outlined above, to our ternary model, which involves three statistics u, v, and w. In this case, the constraint is that u' (the time derivative of u) must be directed away from the line connecting v and w (to increase the area of the triangle T_{uvw} if v and w were held fixed). We employ similar constraints for v' and w'. First, we derive expressions for u', v', and w'.

If θ_i denotes the ith component of one of the n-dimensional statistics u, v, or w, its time derivative θ_i' under the coupled flows $-\nabla_{\vec{C}_u} E$ and $-\nabla_{\vec{C}_v} E$ given by (13) and (14) may be computed as

$$\theta_i' = \theta_i'^u + \theta_i'^v = -\left\langle \nabla_{\vec{C}_u}\theta_i, \nabla_{\vec{C}_u} E\right\rangle - \left\langle \nabla_{\vec{C}_v}\theta_i, \nabla_{\vec{C}_v} E\right\rangle,$$

where $\theta_i'^u$ and $\theta_i'^v$ denote the derivatives of θ_i strictly due to the evolutions of $\vec{C}_u$ and $\vec{C}_v$ respectively. If we ignore the curvature terms in (13–14), we obtain

$$\theta_i'^u = \sum_{j=1}^{n} \left\langle \nabla_{\vec{C}_u}\theta_i, \bar{w}_j \nabla_{\vec{C}_u} u_j + \bar{v}_j \nabla_{\vec{C}_u} w_j \right\rangle$$

$$\theta_i'^v = \sum_{j=1}^{n} \left\langle \nabla_{\vec{C}_v}\theta_i, \bar{u}_j \nabla_{\vec{C}_v} v_j + \bar{v}_j \nabla_{\vec{C}_v} w_j \right\rangle$$

$$u'^u = \Gamma_{\vec{C}_u}(u,u)\,\bar{w} + \Gamma_{\vec{C}_u}(u,w)\,\bar{v}, \quad v'^v = \Gamma_{\vec{C}_v}(v,v)\,\bar{u} + \Gamma_{\vec{C}_v}(v,w)\,\bar{v}$$

$$w'^u = \Gamma_{\vec{C}_u}(w,u)\,\bar{w} + \Gamma_{\vec{C}_u}(w,w)\,\bar{v}, \quad w'^v = \Gamma_{\vec{C}_v}(w,v)\,\bar{u} + \Gamma_{\vec{C}_v}(w,w)\,\bar{v},$$

where $\Gamma_{\vec{C}_u}$ and $\Gamma_{\vec{C}_v}$ denote matrices whose entries are given by the inner product of the gradients of the respective components of each statistic. For example, the entry $\Gamma^{ij}_{\vec{C}_u}(u,w)$ in the ith row and jth column of $\Gamma_{\vec{C}_u}(u,w)$ would be given by $\langle \nabla_{\vec{C}_u} u_i, \nabla_{\vec{C}_u} w_j\rangle$. In the case that u, v, and w represent means of a vector-valued image $I = (I_1, \ldots, I_n)$, the Γ matrices are given by

$$\Gamma_{\vec{C}_u}(u,u) = \frac{1}{A_u A_u} \oint_{\vec{C}_u} (I-u)(I-u)^T\, ds,$$

$$\Gamma_{\vec{C}_v}(v,v) = \frac{1}{A_v A_v} \oint_{\vec{C}_v} (I-v)(I-v)^T\, ds,$$

$$\Gamma_{\vec{C}_u}(u,w) = \frac{-1}{A_u A_w} \oint_{\vec{C}_u} (I-u)(I-w)^T(1-\chi_v)\, ds,$$

$$\Gamma_{\vec{C}_v}(v,w) = \frac{-1}{A_v A_w} \oint_{\vec{C}_v} (I-v)(I-w)^T(1-\chi_u)\, ds,$$

$$\Gamma_{\vec{C}_u}(w,w) = \frac{1}{A_w A_w} \oint_{\vec{C}_u} (I-w)(I-w)^T(1-\chi_v)\, ds,$$

$$\Gamma_{\vec{C}_v}(w,w) = \frac{1}{A_w A_w} \oint_{\vec{C}_v} (I-w)(I-w)^T(1-\chi_u)\, ds.$$

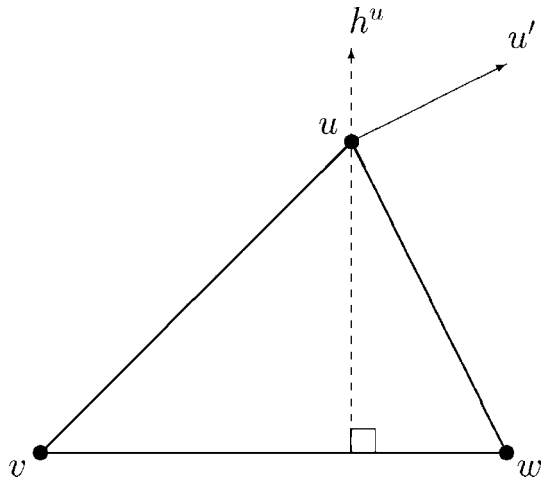

FIG. 10. Visualization of ternary constraint.

We now state our constraints more precisely as (refer to Fig. 10)

$$u' \cdot h^u > 0, \quad v' \cdot h^v > 0, \quad w' \cdot h^w > 0 \quad (u' = u'^u, v' = v'^v, w' = w'^u + w'^v) \tag{29}$$

where u', v', and w' denote the derivatives of u, v, and w, and where h^u, h^v, and h^w denote vectors perpendicular to the edges of the triangle T_{uvw} (represented by $\hat{v} = v - w$, $\hat{w} = w - u$, and $\hat{u} = u - v$) and directed toward the vertices u, v, and w when starting from the opposite edges. The following vectors constitute such a set.

$$\begin{aligned} h^u &= u(\hat{v} \cdot \hat{v}) + v(\hat{v} \cdot \hat{w}) + w(\hat{v} \cdot \hat{u}) \\ h^v &= v(\hat{w} \cdot \hat{w}) + w(\hat{w} \cdot \hat{u}) + u(\hat{w} \cdot \hat{v}) \\ h^w &= w(\hat{u} \cdot \hat{u}) = u(\hat{u} \cdot \hat{v}) + v(\hat{u} \cdot \hat{w}). \end{aligned} \tag{30}$$

As in the binary case, we start out using the unconstrained flows $-\nabla_{\vec{C}_u} E$ and $-\nabla_{\vec{C}_u} E$ given by (13) and (14) so long as conditions (29) are satisfied. Otherwise we use one or both of the following constrained ternary flows which restrict the motion of the problematic statistic so that its distance from the opposite edge of the triangle T_{uvw} remains fixed (and therefore does not alter the area of T_{uvw}).

$$\frac{d\vec{C}_u}{dt} = -\nabla_{\vec{C}_u} E + \begin{cases} \gamma_u^u \nabla_{\vec{C}_u}(u \cdot h^u), & u'^u \cdot h^u < 0 \\ \gamma_w^u \nabla_{\vec{C}_u}(w \cdot h^w), & w'^u \cdot h^w < 0 \end{cases} \tag{31}$$

$$\frac{d\vec{C}_v}{dt} = -\nabla_{\vec{C}_v} E + \begin{cases} \gamma_v^v \nabla_{\vec{C}_v}(v \cdot h^v), & v'^v \cdot h^v < 0 \\ \gamma_w^v \nabla_{\vec{C}_v}(w \cdot h^w), & w'^v \cdot h^w < 0. \end{cases} \tag{32}$$

The constants and the gradient terms in (31) and (32) are defined as follows:

$$\gamma_u^u = \frac{\left\langle \nabla_{\vec{C}_u}(u \cdot h^u), \nabla_{\vec{C}_u} E \right\rangle}{\left\langle \nabla_{\vec{C}_u}(u \cdot h^u), \nabla_{\vec{C}_u}(u \cdot h^u) \right\rangle} = \frac{h^u \cdot u'^u}{h^u \cdot \Gamma_{\vec{C}_u}(u, u) h^u}$$

$$\gamma_w^u = \frac{\left\langle \nabla_{\vec{C}_u}(w \cdot h^w), \nabla_{\vec{C}_u} E \right\rangle}{\left\langle \nabla_{\vec{C}_u}(w \cdot h^w), \nabla_{\vec{C}_u}(w \cdot h^w) \right\rangle} = \frac{h^w \cdot w'^u}{h^w \cdot \Gamma_{\vec{C}_u}(w, w) h^w}$$

$$\gamma_v^v = \frac{\left\langle \nabla_{\vec{C}_v}(v \cdot h^v), \nabla_{\vec{C}_v} E \right\rangle}{\left\langle \nabla_{\vec{C}_v}(v \cdot h^v), \nabla_{\vec{C}_v}(v \cdot h^v) \right\rangle} = \frac{h^v \cdot v'^v}{h^v \cdot \Gamma_{\vec{C}_v}(v, v) h^v}$$

$$\gamma_w^v = \frac{\left\langle \nabla_{\vec{C}_v}(w \cdot h^w), \nabla_{\vec{C}_v} E \right\rangle}{\left\langle \nabla_{\vec{C}_v}(w \cdot h^w), \nabla_{\vec{C}_v}(w \cdot h^w) \right\rangle} = \frac{h^w \cdot w'^v}{h^w \cdot \Gamma_{\vec{C}_v}(w, w) h^w}$$

$$\nabla_{\vec{C}_u}(u \cdot h^u) = \sum_{i=1}^{n} h_i^u \nabla_{\vec{C}_u} u_i = \sum_{i=1}^{n} h_i^u \frac{I_i - u_i}{A_u} \vec{N}_u$$

$$\nabla_{\vec{C}_u}(w \cdot h^w) = \sum_{i=1}^{n} h_i^w \nabla_{\vec{C}_u} w_i = -(1 - \chi_v) \sum_{i=1}^{n} h_i^w \frac{I_i - w_i}{A_w} \vec{N}_u$$

$$\nabla_{\vec{C}_v}(v \cdot h^v) = \sum_{i=1}^{n} h_i^v \nabla_{\vec{C}_v} v_i = \sum_{i=1}^{n} h_i^v \frac{I_i - v_i}{A_v} \vec{N}_v$$

$$\nabla_{\vec{C}_v}(w \cdot h^w) = \sum_{i=1}^{n} h_i^w \nabla_{\vec{C}_v} w_i = -(1 - \chi_u) \sum_{i=1}^{n} h_i^w \frac{I_i - w_i}{A_w} \vec{N}_v.$$

5. NUMERICAL IMPLEMENTATIONS

In this section we address the numerical issues involved in implementing the coupled curve evolutions presented in this paper. There are basically two problems that must be addressed. First, contours may change their topology as they evolve. Second, the data-driven terms in these flows do not generally admit smooth solutions for all time. If we omit the regularizing, curvature-driven term, singularities such as shocks may occur. At such points, the evolving curve is no longer differentiable, and our flow no longer admits a classical solution. If we reintroduce the curvature-driven term immediately before such singularities occur and determine the limit of the resulting smooth solutions as we allow the weight on the curvature term to vanish, we obtain a unique, physically meaningful weak solution. This is known as the *entropy solution* in the theory of hyperbolic conservation laws [14, 19] or the *viscosity solution* in the more general theory of Hamilton–Jacobi equations [8, 9, 15]. The level set methods of Osher and Sethian [20, 25], used with upwind, monotone differencing schemes, [16, 19, 25] offer a natural implementation of these solutions while automatically handling topological changes.

The basic idea of the level set approach is to embed the contour as the level set of a graph $\Psi : \mathbf{R}^2 \to \mathbf{R}$ and then evolve Ψ so that the level curve moves according to the prescribed flow. The curve may develop singularities and change topology while Ψ itself remains regular. The correct evolution for Ψ is found by solving

$$\frac{d\Psi}{dt} = \Psi_t + \nabla\Psi \cdot \frac{d\vec{C}}{dt} = 0$$

so that the interface (the level set of interest) maintains a constant value as Ψ evolves. The level set implementation of (6), for example, would look like

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

If we define $\vec{C}$ as the set where $\Psi = 0$ and let $\Psi < 0$ inside $\vec{C}$ and $\Psi > 0$ outside $\vec{C}$, then

$\vec{N} = \nabla\Psi/\|\nabla\Psi\|$ and $\kappa = \nabla \cdot (\nabla\Psi/\|\nabla\Psi\|)$, allowing us to write

$$\Psi_t = \left\{(v-u)\left(\frac{I-u}{A_u} + \frac{I-v}{A_v}\right) + \alpha\nabla\cdot\left(\frac{\nabla\Psi}{\|\nabla\Psi\|}\right)\right\}\|\nabla\Psi\|. \tag{33}$$

The level surface evolution shown above must be interpreted carefully since the data driven term should depend only upon the values of I along the curve itself (the zero level set). Thus, for other level sets, we must utilize the values of I from a nearby points on the the zero level set. Various techniques for this *extension* process are described in [25]. To keep the computational cost to a minimum, we employ localized level set techniques known as narrowband methods [25] which only require computations (and extensions) at the grid points near the evolving contour (allowing all of our simulations to run in less than ten seconds on a Sparc-10).

When more than one set of curves is involved (such as in the ternary model) we generate distinct level set functions $\Psi^1, \Psi^2, \ldots$ for each set of curves, allowing curves embedded in the same level set function to merge and split, but not allowing such topological transitions to occur between curves embedded in different level set functions (this feature is illustrated in Fig. 4 and Fig. 13). The coupled ternary flows (13)–(14), for example, are implemented through the following level set equations,

$$\frac{d\Psi^u}{dt} = \left\{\sum_{i=1}^{n}\left(-\bar{w}_i\frac{I_i-u_i}{A_u} + \bar{v}_i(1-\chi_v)\frac{I_i-w_i}{A_w}\right) + \alpha\nabla\cdot\frac{\nabla\Psi^u}{\|\nabla\Psi^u\|}\right\}\|\nabla\Psi^u\| \tag{34}$$

$$\frac{d\Psi^v}{dt} = \left\{\sum_{i=1}^{n}\left(-\bar{u}_i\frac{I_i-v_i}{A_v} + \bar{v}_i(1-\chi_u)\frac{I_i-w_i}{A_w}\right) + \alpha\nabla\cdot\frac{\nabla\Psi^v}{\|\nabla\Psi^v\|}\right\}\|\nabla\Psi^v\| \tag{35}$$

where $\vec{C}_u$ and $\vec{C}_v$ are represented by the zero level sets of Ψ^u and Ψ^v, respectively.

Finally, we address the computation of the line integrals used in the constrained flows of Section 4 and outline a simple strategy in the level set framework. Consider for example the inner product $\Gamma(u,u)$ in (24) whose expression we repeat below

$$\Gamma(u,u) = \frac{1}{A_u A_u}\oint_{\vec{C}}(I-u)(I-u)\,ds. \tag{36}$$

This contour integral is difficult to evaluate in the level set framework since we do not have a parameterized curve available to "follow." One could attempt to extract an ordered set of contour points, being careful to separate the points into distinct sets for each connected component of the zero level set. This could be rather costly, however, and would become exceedingly difficult during topological transitions. We may side-step almost all of these difficulties by rewriting the integral (36) as

$$\Gamma(u,u) = \frac{L(\vec{C})}{A_u A_u}\overline{\{(I-u)^2\}} = \frac{L(\vec{C})}{A_u A_u}\left[\overline{\{I^2\}} - 2u\overline{\{I\}} + u^2\right], \tag{37}$$

where $L(\vec{C})$ denotes the length of $\vec{C}$ and where $\overline{\{I^2\}}$ and $\overline{\{I\}}$ denote the average values of I^2 and I, respectively, along the curve $\vec{C}$. The latter two quantities may be approximated by computing the average values of I^2 and I over the set of grid points where a sign change in the level set function occurs (i.e., positive points that are neighbored by negative points

and vice-versa). This does not require us to extract ordered contour points nor does it care about topology. The computation of $L(\vec{C})$ is not so easy. Fortunately, however, we do not have to compute it. This is because $L(\vec{C})$ becomes a common factor in $\Gamma(u, u)$, $\Gamma(u, v)$ and $\Gamma(v, v)$ when written in the manner of (37), and may be factored out of the inequality (25) and the ratios used to express the constants γ_u and γ_v in the constrained binary flows (26) and (27). Similar strategies may be employed for the constrained ternary model.

6. SIMULATIONS

In this section we demonstrate the performance of our coupled curve evolution models on a number of bimodal and trimodal images. We include images which are segmented successfully by the original models from Section 2 as well as images for which the original models fail, requiring the constrained models from Section 4.

The original, unconstrained binary flow (6) is used in Fig. 11, to capture a scorpion sitting on a rock and a teapot. Note that the boundary of the rock, which constitutes a very strong edge in the first image, initially attracts part of the contour, but is eventually ignored as the evolution proceeds. The flow also ignores the shadow of the teapot handle (which is darker than the four specular reflections on the teapot surface, posing a problem for simple thresholding). The synthetic aperture radar (SAR) image of a forest's tree line in Fig. 12 constitutes a differnt type of bimodal image. The forest region in the lower left and the grassy region in the upper right have approximately the same grayscale mean but different variances. Flow (10), therefore, is able to segment the image quite successfully by separating variances rather than means.

The original, unconstrained ternary flows (13) and (14) are demonstrated in Fig. 13 on three different images. In the first two images, we begin with two overlapping curves which eventually move apart to capture different regions in each image. Note in particular that the final segmentation of the first image contains many triple points where the letters of "Hollywood" (captured by the green curve) come in contact with both the terrain (captured by the red curve) and the sky (the mutual background of both curves). The next two images technically consist of four regions apiece. In both images, however, two regions are characterized by a similar color (the light sand and the cloudy sky in the golf green; and the dark leaves and the shadows in the Windsor Castle), allowing us to use the ternary model (which lumps the two similar regions together). In the third image, we see a compelling demonstrating of the advantages of using two separate level set functions. Here, we initialized with not just two contours, but two sets of contours. The red contours were allowed to merge together to capture the castle and the yellow contours were allowed to merge together to capture the sky. However, such topological transitions were not allowed between a red contour and a yellow contour. As such, the outer edge of the castle is captured twice, both by the red and yellow contours, allowing us to see a quadruple point in the final segmentation where this edge forms a cusp making contact with the dark region on the left, the blue region on the right, and two separate components of the grey region in the middle.

In the top row of Fig. 14 we see an example where the unconstrained binary flow (6) fails. Instead of capturing the outline of the dog, the curve is attracted toward a very bright area of fur on the hind leg. Both means inside and outside the curve increase, while nevertheless moving further apart, as the bright region is captured. In the bottom row we see the constrained binary flow yielding a more desirable segmentation with the same initial contour. Figure 15 illustrates a reverse scenario in which the unconstrained flow performs

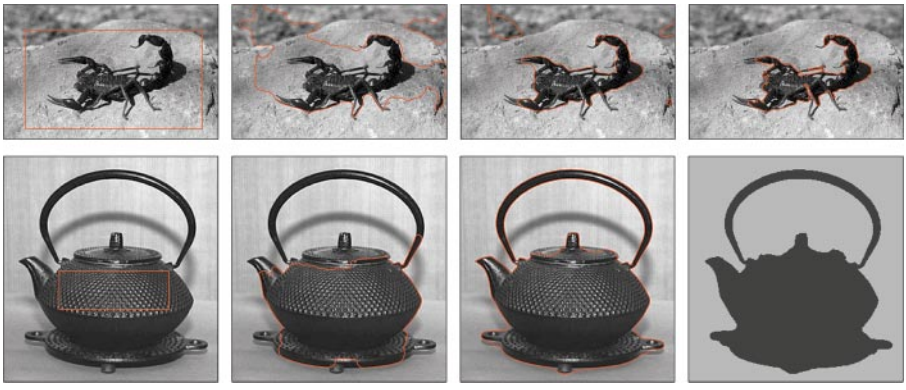

FIG. 11. The unconstrained binary flow (6) based on means applied to 2 bimodal images.

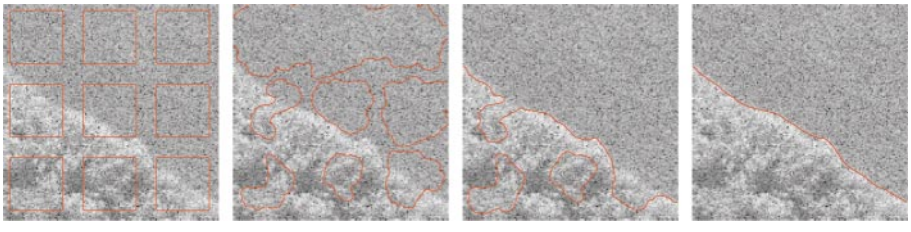

FIG. 12. The unconstrained binary flow (10) based on variances applied to a SAR image.

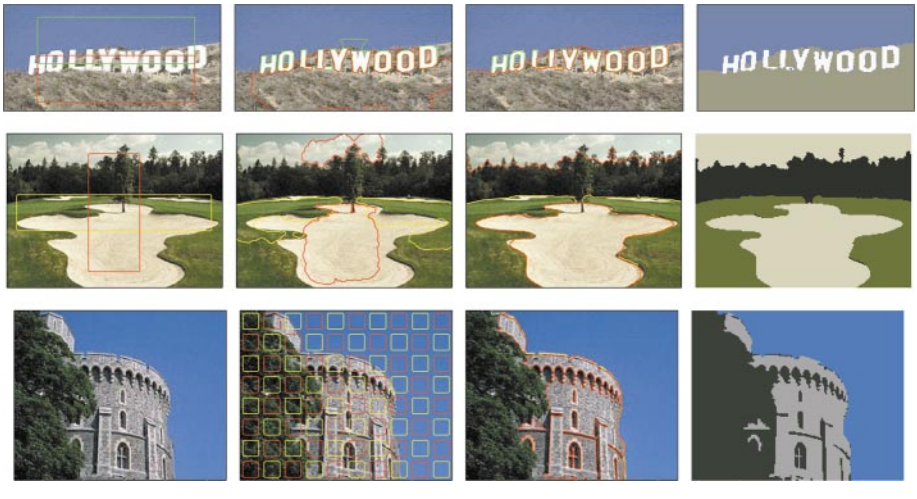

FIG. 13. The original ternary flows (13) and (14) applied to three different color images.

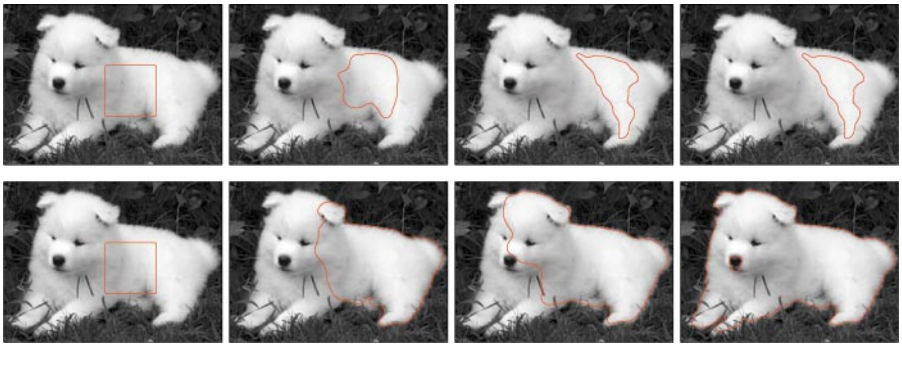

FIG. 14. (bottom) A constrained binary flow outperforming the (top) unconstrained flow.

FIG. 15. (bottom) A constrained binary flow outperforming the (top) unconstrained flow.

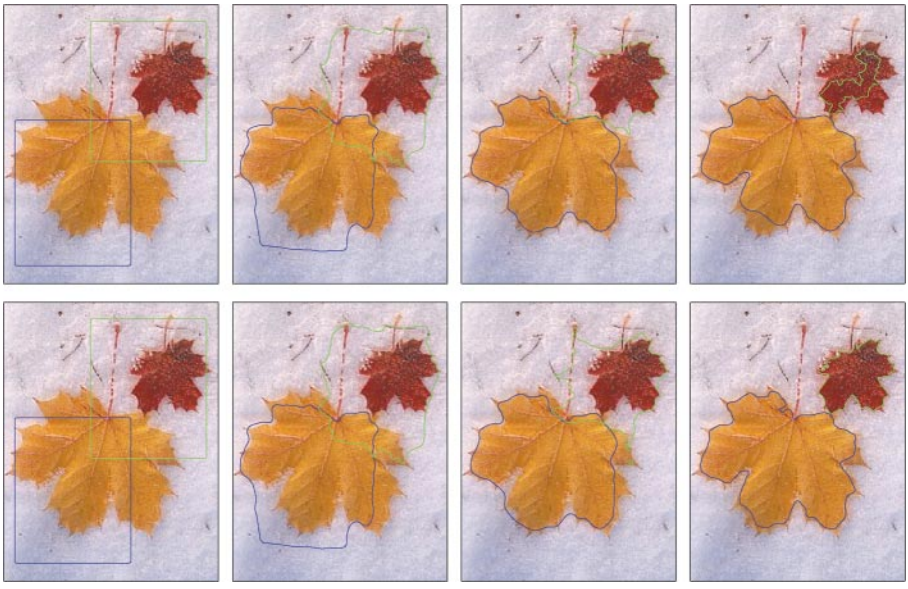

FIG. 16. (bottom) Constrained ternary flows outperforming the (top) unconstrained flows.

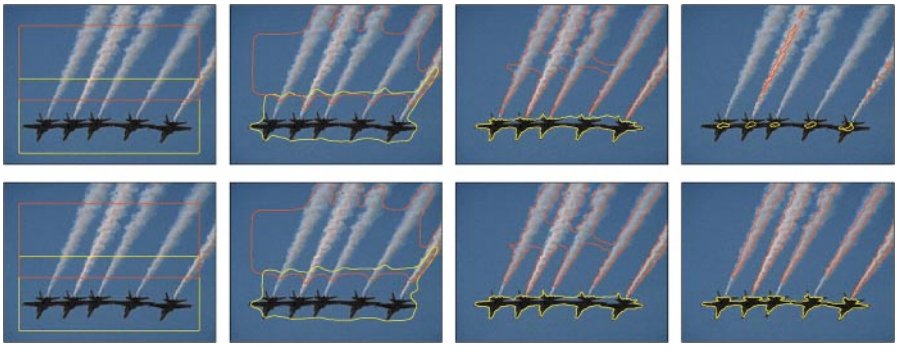

FIG. 17. (bottom) Constrained ternary flows outperforming the (top) unconstrained flows.

poorly as well. In this case the curve is attracted toward a very dark area of fur on the bear (verses the very bright area of fur in Fig. 14) as illustrated in the top row, causing both means inside and outside the curve to decrease as the dark region is captured. In the bottom row we see the constrained binary flow yielding a more desirable result once again.

Finally, we illustrate the use of the constrained ternary flows to segment trimodal images for which the unconstrained ternary flows (13) and (14) fail. In the top row of Fig. 16, we see the unconstrained flows capturing the larger leaf but collapsing into the interior of the smaller leaf. In the bottom row the constrained flows perform much better by capturing the boundary of the smaller leaf (and also do a better job of capturing the boundary of larger leaf). In the top row of Fig. 17, we see the unconstrained ternary flows collapsing into the interiors of the airplanes and the contrails. In the bottom row, we see the constrained flows doing a much better job.

7. CONCLUSIONS

We have presented in this paper a fully global approach to snakes for the segmentation of images which consist of a known number of region types distinguishable by certain statistics. This approach was derived based upon the *deterministic* principle of maximally separating the values of these statistics within a set of curves and the background subject to geometric constraints on the length of each active contour. This led to a coupled curve evolution model in which the evolution of each curve depended very naturally upon every single pixel in the image and was tied to the evolution of every other curve through a common set of statistics. By introducing additional global constraints (Section 4) on the evolution of the statistics, we were able to extend the applicability of these models to a larger class of imagery.

A key attraction of these techniques is the fully global nature of the flows. This endows the algorithm with an extensive capture range, exhibiting more robustness to initial contour placement when compared to other region based snake algorithms. We are also able to capture triple points and other multiple junctions with no addditional effort by employing separate, distinct sets of curves.

To summarize, we have outlined a fully global, coupled 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 and can be greatly improved in a number of cases by using the global constraints presented in Section 4. We point out that these algorithms are more specialized than many other snake techniques which do not require the same amount of prior information. However, our techniques exploit this information, in a way that other methods do not, to yield improved performance and robustness.

REFERENCES

1. A. Blake and A. Yuille, *Active Vision*, MIT Press, Cambridge, MA, 1992.
2. V. Caselles, F. Catte, T. Coll, and F. Dibos, A geometric model for active contours in image processing, *Numer. Math.* **66**, 1993, 1–31.
3. V. Caselles, R. Kimmel, and G. Sapiro, Geodesic snakes, *Int. J. Comput. Vision*, 1998.
4. A. Chakraborty and J. Duncan, Game-theoretic integration for image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.* **21**, 1999, 12–30.
5. A. Chakraborty, L. Staib, and J. Duncan, Deformable boundary finding in medical images by integrating gradient and region information, *IEEE Trans. Medical Imaging* **15**, 1996, 859–870.

6. T. Chan and L. Vese, An active contours model without edges, *Int. Conf. Scale-Space Theories in Computer Vision*, 1999, pp. 141–151.
7. L. Cohen, On active contour models and balloons, *CVGIP: Image Understanding* **53**, 1991, 211–218.
8. M. Crandall, H. Ishii, and P. Lions, Users guide to viscosity solutions of second order partial differential equations, *Bull. Am. Math. Soc.* **27**, 1992, 1–67.
9. W. Fleming and H. Soner, *Controlled Markov Processes and Viscosity Solutions*, Springer-Verlag, New York, 1993.
10. M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active contour models, *Int. J. Comput. Vision* **1**, 1987, 321–331.
11. S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, Conformal curvature flows: From phase transitions to active vision, *Arch. Rational Mech. Anal.* **134**, 1996, 275–301.
12. K. Siddiqi, Y. Lauziere, A. Tannenbaum, and S. Zucker, Area and length minimizing flows for segmentation, *IEEE Trans. Image Process.* **7**, 1998, 433–444.
13. Y. Leclerc, Constructing stable descriptions for image partitioning, *Int. J. Comput. Vision* **3**, 1989, 73–102.
14. R. J. LeVeque, *Numerical Methods for Conservation Laws*, Birkhäuser, Boston, 1992.
15. P. L. Lions, *Generalized Solutions of Hamilton–Jacobi Equations*, Pitman, Boston, 1982.
16. R. Malladi, J. Sethian, and B. Vemuri, Shape modeling with front propagation: A level set approach, *IEEE Trans. Pattern Anal. Machine Intell.* **17**, 1995, 158–175.
17. D. Mumford and J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Comm. Pure Appl. Math.* **42**(4), 1989.
18. D. Mumford and J. Shah, Boundary detection by minimizing functionals, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 1985.
19. S. Osher, Riemann solvers, the entropy condition, and difference approximations, *SIAM J. Numer. Anal.* **21**, 1984, 217–235.
20. S. Osher and J. Sethian, Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations, *J. Comput. Phys.* **79**, 1988, 12–49.
21. N. Paragios and R. Deriche, Geodesic active regions for supervised texture segmentation, *Proceedings of ICCV*, Sept. 1999, Corfu, Greece.
22. N. Paragios and R. Deriche, Coupled geodesic active regions for image segmentation: A level set approach, *Proceedings of ECCV*, June 2000, Dublin, Ireland.
23. R. Ronfard, Region-based strategies for active contour models, *Int. J. Comput. Vision* **13**, 1994, 229–251.
24. C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia, A level set method for image classification, *Int. Conf. Scale-Space Theories in Computer Vision*, 1999, pp. 306–317.
25. J. Sethian, *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science*, Cambridge University Press, Cambridge, MA, 1996.
26. H. Tek and B. Kimia, Image segmentation by reaction diffusion bubbles, *Proc. Int. Conf. Computer Vision*, 1995, pp. 156–162.
27. D. Terzopoulos and A. Witkin, Constraints on deformable models: Recovering shape and non-rigid motion, *Artificial Intell.* **36**, 1988, 91–123.
28. A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, A geometric snake model for segmentation of medical imagery, *IEEE Trans. Medical Imaging* **16**, 1997, 199–209.
29. A. Yezzi, A. Tsai, and A. Willsky, *Fully Global, Coupled Curve Evolution Equations for Image Segmentation*, LIDS Technical Report, January, 1999.
30. A. Yezzi, A. Tsai, and A. Willsky, A statistical approach to image segmentation for bimodal and trimodal imagery, *Proceedings of ICCV*, September, 1999.
31. S. Zhu, T. Lee, and A. Yuille, Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, *Proc. of ICCV*, 1995, pp. 416–423.
32. S. Zhu and A. Yuille, Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.* **18**, 1996, 884–900.

ANTHONY YEZZI was born in Gainsville, Florida in 1972 and obtained his Ph.D. in 1997 through the Department of Electrical Engineering at the University of Minnesota. After completing postdoctoral research in

the Laboratory for Information and Decision Systems (LIDS) at the Massacusetts Institute of Technology, Dr. Yezzi began his current position in 1999 at the Georgia Institute of Technology as an assistant professor of electrical and computer engineering with an adjunct appointment in biomedical engineering. Dr. Yezzi has also consulted for a number of medical imaging companies including GE, Picker, and VTI and has been an IEEE member since 1999. His research lies primarily within the fields of image processing and computer vision. He has worked and continues to work on a variety of problems within these fields including image denoising, edge detection, segmentation and grouping, shape analysis, multiframe stereo reconstruction, tracking, and registration. Much of his work is motivated by and directed toward problems in medical imaging. Two central themes of his research in general are curve/surface evolution theory from differential geometry and partial differential equations.

ANDY TSAI received his B.S. in electrical engineering from the University of California at San Diego in 1993 and his S.M., E.E., and Ph.D. in electrical engineering from the Massachusetts Institute of Technology in 1995, 1999, and 2000, respectively. He then continued his research as a postdoctoral research associate at the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Currently, he is a third-year medical student at Harvard Medical School. His research interests fall within the fields of computer vision and image processing.

ALAN S. WILLSKY received the S.B. and Ph.D. from the Massachusetts Institute of Technology in 1969 and 1973, respectively. He joined the MIT faculty in 1973 and his present position is as the Edwin S. Webster Professor of Electrical Engineering. From 1974 to 1981 he served as Assistant Director of the MIT Laboratory for Information and Decision Systems. He is also a Founder and member of the board of directors of Alphatech, Inc., and is currently a member of the U.S. Air Forc Scientific Advisotry Board. He has held visiting positions at Imperial College, London, l'Universite de Paris-Sud, Paris, and the Institut de Recherche eninformatique et Systèmes Aleatoires, Rennes. He is the author of the research monograph "Digitial Signal Processing and Control and Estimation" and is the co-author of the undergraduate text "Signals and Systems." His research interests are in the development and application of advanced methods of estimation and statistical signal and image processing. Methods he has develoed have been successfully applied in a wide variety of applications including failure detection in high-performance aircraft, advanced surveillance and tracking systems, electrocardiogram analysis, computerized tomography, and remote sensing. Dr. Willsky was awarded the 1979 Alfred Nobel Prize by the ASCE and the 1980 Browder J. Thompson Memorial Prize Award by the IEEE for a paper excerpted from his monograph. In 1975 he received the Donald P. Eckman Award from the American Automatic Control Council. He was program chairman for the 17th IEEE Conference on Decision and Control and has been an associate editor of several journals and a special guest editor for several special issues, and has served as a member of the Board of Governors and as Vice President for Technical Affairs of the IEEE Control Systems Society. He has given plenary and keynote lectures at a number of major scientific meetings, including the 20th IEEE Conference on Decision and Control, the 1991 SIAM Conference on Appplied Linear Algebra, the 1992 Inaugural Workshop for the National Centre for Robust and Adaptive Systems, Canberra, Austrailia, the 1992 INTRIA 25th Anniversary Symposium in Paris, the 1993 IEEE Symposium on Image and Multidimensional Signal Processing in Cannes, and the 1997 Wavelet Applications in Signal and Image Processing Conference.