

# Automatic segmentation of histological structures in mammary gland tissue sections

## R. Fernandez-Gonzalez

Lawrence Berkeley National Laboratory  
Life Sciences Division  
1 Cyclotron Road, Building 84, MS 84-171  
Berkeley, California 94720  
and  
University of California,  
Berkeley & University of California,  
San Francisco Joint  
Graduate Group in Bioengineering  
467 Evans Hall #1762  
Berkeley, California 94720-1762  
E-mail: RFGonzalez@lbl.gov

## T. Deschamps

University of California, Berkeley  
Mathematics Department  
970 Evans Hall #3840  
Berkeley, California 94720-3840  
and  
Lawrence Berkeley National Laboratory  
Mathematics Department  
Berkeley, California 94720

## A. Idica

Lawrence Berkeley National Laboratory  
Life Sciences Division  
1 Cyclotron Road, Building 84, MS 84-171  
Berkeley, California 94720

## R. Malladi

Lawrence Berkeley National Laboratory  
Mathematics Department  
Berkeley, California 94720

## C. Ortiz de Solorzano

Lawrence Berkeley National Laboratory  
Life Sciences Division  
1 Cyclotron Road, Building 84, MS 84-171  
Berkeley, California 94720

## 1 Introduction

The mature mammary gland is a treelike organ made up of three different levels of branching ducts converging at the nipple.<sup>1</sup> The ducts are lined by epithelial cells and end in secretory lobuloalveolar structures, which are the sites of milk production during lactation. In cancer, this hierarchical organization is disrupted by uncontrolled growth of the epithelium, and sometimes by the subsequent invasion of the surrounding tissue.<sup>2</sup> These morphological or structural changes are accompanied by other genetic and epigenetic changes at the cellular level (see Fig. 1). An example of this is seen in ductal carcinoma *in situ* (DCIS), which is a preinvasive form of cancer. In DCIS, the molecular changes common in breast

**Abstract.** Real-time three-dimensional (3-D) reconstruction of epithelial structures in human mammary gland tissue blocks mapped with selected markers would be an extremely helpful tool for diagnosing breast cancer and planning treatment. Besides its clear clinical application, this tool could also shed a great deal of light on the molecular basis of the initiation and progression of breast cancer. We present a framework for real-time segmentation of epithelial structures in two-dimensional (2-D) images of sections of normal and neoplastic mammary gland tissue blocks. Complete 3-D rendering of the tissue can then be done by surface rendering of the structures detected in consecutive sections of the blocks. Paraffin-embedded or frozen tissue blocks are first sliced and sections are stained with hematoxylin and eosin. The sections are then imaged using conventional bright-field microscopy and their background corrected using a phantom image. We then use the fast-marching algorithm to roughly extract the contours of the different morphological structures in the images. The result is then refined with the level-set method, which converges to an accurate (subpixel) solution for the segmentation problem. Finally, our system stacks together the 2-D results obtained in order to reconstruct a 3-D representation of the entire tissue block under study. Our method is illustrated with results from the segmentation of human and mouse mammary gland tissue samples. © 2004 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1699011]

**Keywords:** breast cancer, molecular analysis, automatic segmentation, 3-D reconstruction, fast-marching, level set.

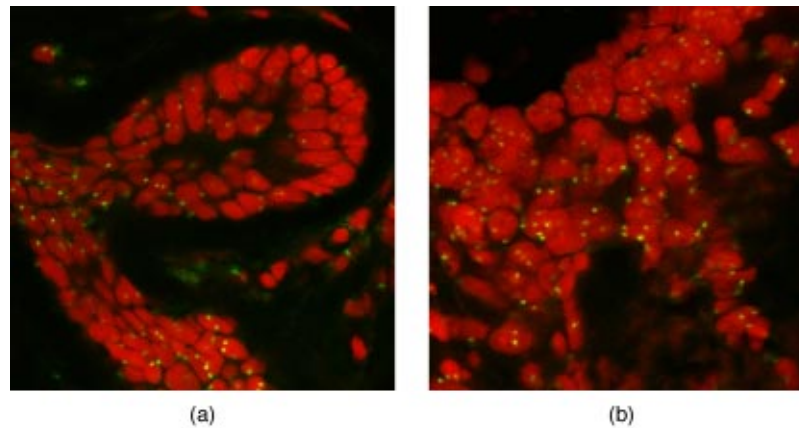
Paper 44006 received Jul. 1, 2003; revised manuscript received Sep. 8, 2003; accepted for publication Sep. 30, 2003. This paper is a revision of a paper presented at the SPIE conference on Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing X, January 2003, San Jose, Calif. The paper presented there appears (unrefereed) in SPIE Proceedings Vol. 4964.

cancer can be observed together with well-defined morphological patterns, such as comedo or cribiform ones, with or without central necrosis.<sup>3,4</sup>

An interesting problem is the quantification of these genetic or molecular changes in the context of the tissue environment where they occur. In this way, by looking at cancer as an organ that is inherently heterogeneous and dynamic, we believe that we will obtain a better understanding of the events that drive the progression of the disease. Therefore, since the normal mammary gland and its neoplastic variants are neither flat nor homogeneous, the approach to this problem should be three-dimensional as well, and take into account the heterogeneity of the gland. However, most classic methods in biology focus on only a single type of abnormality (molecular or morphological) and/or neglect three-dimensionality and heterogeneity.

Although imaging of both tissue structure and function *in vivo* would be extremely desirable, the existing *in vivo* imag-

Address all correspondence to R. Fernandez-Gonzalez, Lawrence Berkeley National Laboratory, Life Sciences Division, 1 Cyclotron Road, Building 84, MS 84-171, Berkeley, California 94720. Tel: 510-486-5359; Email: RFGonzalez@lbl.gov



**Fig. 1** Morphological and genetic alterations in breast cancer. The images show two optical sections of human mammary gland tissue acquired using a confocal laser scanning microscope; nuclei are displayed in red; a probe for a certain DNA sequence in chromosome 17 is shown in green. (a) Normal tissue; as expected, each nucleus contains up to two green signals (up to two copies of chromosome 17 per nucleus in a single optical section). (b) Neoplastic lesion; not only do some nuclei contain more than two copies of the probe, but they also have distinct morphological changes that can be observed in this section.

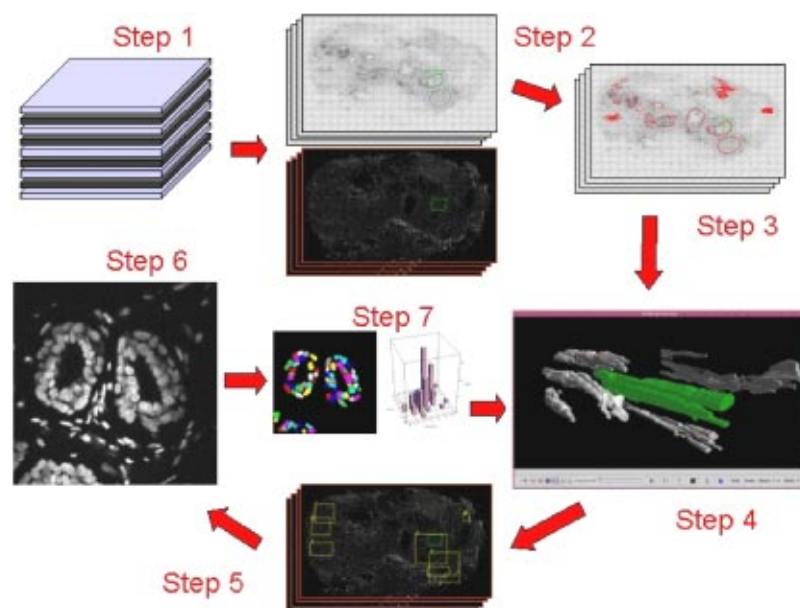
ing methods (x-ray, magnetic resonance imaging, optical tomography, etc.) do not provide the necessary resolution for cell-level molecular analysis. In addition, these methods provide morphological information that can only be indirectly related to the function of the tissue. Consequently, *ex vivo* microscopic analysis of the tissue from flat fixed sections is the routine method in histopathology. However, the limited ability of the human eye to extrapolate and visualize 3-D structures from sequences of 2-D scenes renders this method unsuitable for quantitative 3-D tissue characterization. To overcome these issues, we have developed a system for simultaneous morphological and molecular analysis of thick tissue samples.<sup>5</sup> The system consists of a computer-assisted microscope and a JAVA-based image display, analysis, and visualization program (R3D2). R3D2 allows semiautomatic acquisition, annotation, storage, 3-D reconstruction, and analysis of histological structures (intraductal tumors, normal ducts, blood vessels, etc.) from thick tissue specimens. For this purpose the tissue needs to be embedded in a permanent or semirigid medium after collection. The tissue is then fully sectioned, and the resulting sections are stained in a way that visually highlights the desired structures. In histopathology, hematoxylin and eosin (H&E) is the most common combination of dyes used to observe the morphology of the tissue. In order to image not only structure, but also molecular events, we alternate H&E staining with fluorescent staining (immunofluorescence, fluorescence *in situ* hybridization) of proteins and selected genes in consecutive sections.

This paper focuses on the annotation of the structures of interest on the H&E-stained sections. Manual segmentation has been used before to delineate histological structures.<sup>6–8</sup> In order to build the 3-D reconstruction of the block, we initially annotated each of the interesting features of the tissue on each section by using manually drawn contours. This step constituted a bottleneck in the study of samples. Semiautomatic approaches to segmentation of features of interest in histological sections have also been used,<sup>9</sup> but they still involve too much user interaction to be useful for reconstructing large tissue samples. Automatic segmentation of the structures of

interest is the answer to this problem. We propose an automatic method followed by interactive correction that greatly reduces the amount of interaction required, thus allowing us to use our system for imaging and reconstruction of large, complex tissue structures.

Automatic extraction of contours in 2-D images is usually done with active contour models, originally presented by Kass et al.<sup>10</sup> These methods are based on deforming an initial contour (polygon) toward the boundary of the desired object to extract in an image. The deformation is achieved by minimizing a certain energy function, which is computed by integrating along the contour, terms related to its continuity, and terms related to the pixel values of the area of the image where the contour is defined. That energy function approaches a minimum near the object's boundary, and thus the minimization process drives the curve toward the desired shape.

As an alternative, implicit surface evolution models have been introduced by Malladi et al.<sup>11,12</sup> and Caselles et al.<sup>13</sup> In these models, the curve and surface models evolve under an implicit speed law containing two terms, one that attracts it to the object's boundary and another that is closely related to the regularity of the shape. Specifically, the proposal is to use the level-set approach of Osher and Sethian.<sup>14</sup> This is an interface propagation technique used for a variety of applications, including segmentation. The initial curve is represented here as the zero level set of a higher dimensional function, and the motion of the curve is embedded within the motion of that higher dimensional function. An energy formulation similar to the active contours leads to a minimization process with several advantages. First, the zero level set of the higher dimensional function is allowed to change topology and form sharp corners. Second, geometric quantities such as normal and curvature are easy to extract from the hypersurface. Finally, the method expands straightforwardly to 3-D,<sup>15</sup> but adding a dimension to the problem increases the computational cost associated with the method. The narrow-band approach of Adalsteinsson and Sethian<sup>16</sup> accelerates the level-set flow by considering for computations only a narrow band of pixels around the zero level set. However, in our experience, the



**Fig. 2** Protocol followed on tissue blocks. The different steps (sectioning, annotation, reconstruction, high-magnification acquisition, and molecular analysis) are illustrated. Samples are fully sectioned at  $5\ \mu\text{m}$  (step 1). The odd sections are stained with H&E, the even ones with some kind of fluorescence technique (application dependent). Images are acquired of all the sections (step 2), and the structures of interest are delineated in the H&E-stained ones (step 3). A 3-D reconstruction of the specimen is created from these markings (step 4). From the 3-D reconstruction of the tissue, different areas can be selected for molecular analysis (step 5). The system will take high-magnification images of those areas on the corresponding fluorescent sections (step 6). Image analysis tools can then be used to quantify the presence and distribution of molecular markers in the high-magnification images (step 7).

narrow-band technique does not reduce the computational cost to a reasonable limit, owing to the large size of the images of the sections.

Thus we propose to use the fast-marching method.<sup>17</sup> This method considers monotonically advancing fronts (speed always positive or negative), providing a result very quickly, albeit one that is not as accurate as the one obtained by using the level-set algorithm. Malladi and Sethian<sup>18</sup> showed that the fast-marching method can be used as the initial condition for the slower but more accurate level-set segmentation, obtaining real-time delineation of the structures of interest. A combination of all these tools is the framework we use to reconstruct normal and cancerous ducts in mammary gland tissue sections of DCIS samples.

This paper is organized as follows. Section 2 describes the general tissue handling and image acquisition protocols that we use, as well as the theoretical basis of the segmentation scheme; Sec. 3 shows the results of applying our method to histological tissue sections; and Sec. 4 discusses the results and suggests some future developments and improvements to our approach.

## 2 Methodology

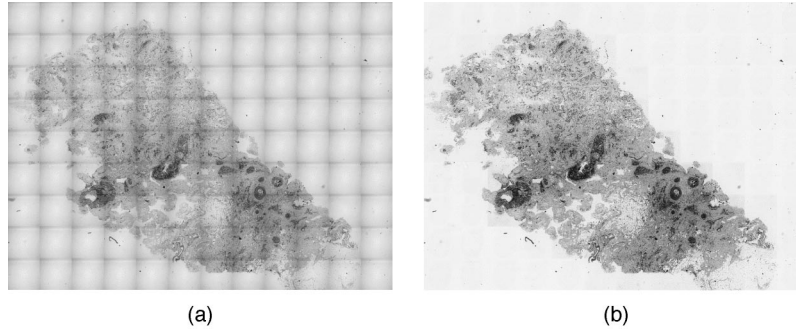
### 2.1 Tissue Processing and Imaging

The tissue processing and staining protocol used is illustrated in Fig. 2. Tissue blocks of 4- to 5-mm thickness were sliced into  $5\ \mu\text{m}$  (thin) sections (step 1). The odd sections were stained with H&E to obtain morphological information at both the cytological (single cell) and architectural (organ) levels. The even sections were stained using some fluorescence technique (e.g., immunocytochemistry, fluorescence *in situ*

hybridization), depending on the molecular phenomena that we wanted to study. Describing the acquisition and analysis of the fluorescent images is out of the scope of this paper, which focuses on the segmentation of epithelial structures in the H&E-stained sections. Therefore the rest of this section describes only the protocol used for the H&E-stained sections. Low magnification ( $2.5\times$ ), panoramic images of all the sections were automatically acquired using a motorized Zeiss Axioplan I microscope coupled with a monochrome XilliX Microimager CCD camera (step 2). To create these large panoramic images, the system scanned the entire section, taking single-field-of-view snapshots and tiling them together into single, whole-view images of the sections. The required sequence of microscope movements and camera operations is produced by an application running on the Sun Ultra 10 workstation that controls the camera and all moving parts of the microscope.

Next we annotated the structures of interest (ducts, lymph nodes, tumors) in the images of the H&E-stained sections (step 3). These annotated structures were used to produce a three-dimensional model of each tissue block (step 4).

These four initial steps are the focus of this paper. However, to better understand the rationale for the process, we briefly describe the final three steps, which are out of the intended scope for this paper. The user can choose to revisit areas in the three-dimensional rendition of the organ, based on their morphology. This can be done on the H&E sections or on the intermediate fluorescent sections. To do so, the system requires the user to acquire high-magnification ( $40$  to  $100\times$ ) images of the corresponding section(s) (steps 5 and 6). If the high-magnification images are taken from the intermediate



**Fig. 3** Background correction. (a) Image of a section belonging to a human case. The background pattern created by the acquisition method is readily noticeable. (b) Same image after background correction.

immunofluorescent sections, quantification routines can be run on these new images (step 7).

Manually annotating all the relevant morphological structures in all the sections of fully sectioned tissue blocks is feasible but, for all purposes, impractical because of the tremendous human effort needed. Although it might be the most accurate and reliable approach, manual annotation is not possible except when reconstructing small, very simple tissue volumes. As a result, we have developed automatic methods that eliminate or greatly reduce human interaction, thus making the reconstruction of complex systems possible. The following discussion describes our approach.

## 2.2 Segmentation

### 2.2.1 Background removal

The algorithm that we use to acquire an image of an entire section creates a mosaic from a set of snapshots (one per field of view). This approach gives rise to a background pattern across the image [Fig. 3(a)] involving relatively large gradients in between elements of the mosaic. The objects of interest often span several fields of view, and since our segmentation approach depends largely on the gradients of the image, we need to eliminate the background pattern in order to obtain good segmentation results. This can be done by performing a set of arithmetic operations on the “mosaic” image, known as background compensation.

First we need a phantom, that is, an image of an empty field of view taken under the same illumination conditions and microscope configuration that we used to acquire the initial image. Since most of the images that we acquire have an empty frame in the upper left corner, it is simple to choose that frame as our phantom for the corresponding section. After normalizing the pixel values in the phantom, and for each frame in the entire image, we divide the value at each pixel by the value at the corresponding pixel in the phantom frame. The resulting image is background-corrected as shown in Fig. 3(b), and it is a better input for our segmentation algorithms.

### 2.2.2 Preliminary segmentation: fast-marching method

Consider a monotonically advancing 2-D front  $C$  with a speed  $F$  that is always positive in the normal direction, starting from an initial point  $p_0$ ,

$$\frac{\partial C}{\partial t} = F \mathbf{n}. \quad (1)$$

This equation drives the evolution of a front starting from an infinitesimal circular shape around  $p_0$  until each point  $p$  inside the image domain is visited and assigned a crossing time  $U(p)$ , which is the time  $t$  at which the front reaches point  $p$ .

The gradient of the arrival time is inversely proportional to the speed function, and thus we have a form of the eikonal equation

$$|\nabla U|F = 1 \quad \text{and} \quad U(p_0) = 0. \quad (2)$$

One way to solve Eq. (2) is to use upwind finite-difference schemes and iterate the solution in time.<sup>15</sup> In other words, the scheme relies on one-sided differences that look in the upwind direction of the moving front, thereby choosing the correct viscosity solution, namely

$$[\max(u - U_{i-1,j}, u - U_{i+1,j}, 0)]^2 + [\max(u - U_{i,j-1}, u - U_{i,j+1}, 0)]^2 = \frac{1}{F_{i,j}^2}. \quad (3)$$

The key to solving this equation rapidly is to observe that the information propagates from smaller values of  $U$  to larger values in the upwind difference structure in Eq. (3). The idea is to construct the time surface, one piece at a time, by only considering the “frontier” points; we detail the fast-marching method in Table 1.

Note that in solving Eq. (3), only alive points are considered. This means that for each point, the calculation is made using the current values of  $U$  at the neighbors, and not estimates at other trial points. Considering the neighbors of the grid point  $(i,j)$  in 4-connectedness, we designate  $\{A_1, A_2\}$  and  $\{B_1, B_2\}$  as the two couples of opposite neighbors so that we get the ordering  $U(A_1) \leq U(A_2)$ ,  $U(B_1) \leq U(B_2)$ , and  $U(A_1) \leq U(B_1)$ . Since we have  $u \geq U(B_1) \geq U(A_1)$ , we can derive

$$[u - U(A_1)]^2 + [u - U(B_1)]^2 = \frac{1}{F_{i,j}^2}. \quad (4)$$

Computing the discriminant ( $\Delta$ ) of Eq. (4), we complete the steps described in Table 2. Thus the algorithm needs only one



**Table 1** Fast-marching algorithm.

Algorithm for 2-D Fast Marching
● Definitions:
Alive set: all grid points where the action value $U$ has been reached and will not be changed;
Trial set: next grid points (4-connectedness neighbors) to be examined. An estimate $U$ of $U$ has been computed using Eq. (3) from Alive points only (i.e., from $U$ );
Far set: all other grid points, where there is no estimate for $U$ yet;
● Initialization:
Alive set: reduced to the starting point $p_0$ , with $U(p_0) = U(p_0) = 0$ ;
Trial set: reduced to the four neighbors $p$ of $p_0$ with initial value $U(p) = 1/F(p)$ ( $U(p) = \infty$ );
Far set: all other grid points, with $U = \infty$ ;
● Loop:
Let $p = (i_{\min}, j_{\min})$ be the trial point with the smallest action $U$ ;
Move it from the trial to the alive set (i.e., $U(p) = U_{i_{\min}, j_{\min}}$ is frozen);
For each neighbor $(i, j)$ (4-connectedness in 2-D) of $(i_{\min}, j_{\min})$ :
If $(i, j)$ is far, add it to the trial set and compute $U_{i, j}$ using Eq. (3);
If $(i, j)$ is trial, update the action $U_{i, j}$ using Eq. (3).

pass over the image to find a solution. To execute all the operations that we just described in the minimum amount of time, the trial points are stored in a min-heap data structure.<sup>17</sup> Since the complexity of changing the value of one element of

**Table 2** Solving the upwind scheme locally.

1. ● If $\Delta \geq 0$ , $u$ should be the largest solution of Eq. (4);
If the hypothesis $u > U(B_1)$ is wrong, go to 2;
If this value is larger than $U(B_1)$ , this is the solution;
● If $\Delta < 0$ , $B_1$ has an action too large to influence the solution. It means that $u > U(B_1)$ is false. Go to 2;
Simple calculus can replace case 1 by the test:
If $1/F_{i, j} > U(B_1) - U(A_1)$ , $u = U(B_1) + U(A_1) + \{21/F_{i, j}^2 - [U(B_1) - U(A_1)]^2\}^{1/2}/2$ is the largest solution of Eq. (4) else go to 2;
2. Considering that we have $u < U(B_1)$ and $u \geq U(A_1)$ , we finally have $u = U(A_1) + 1/F_{i, j}$ .

the heap is bounded by a worst-case processing time of  $[O(\log_2 N)]$ , the total algorithm has a complexity of  $O(N \log_2 N)$  on a grid with  $N$  nodes.

Finally, we define the speed of propagation in the normal direction as a decreasing function of the gradient  $|\nabla I(x)|$ , that is, a function that is very small near large image gradients (i.e., possible edges) and large when the brightness level is constant:

$$F(x) = \exp - \alpha |\nabla I(x)|, \quad \alpha > 0, \quad (5)$$

where  $\alpha$  is the edge strength, or the weight that we give to the presence of a gradient in order to slow down the front. Depending on this value, the speed function falls to zero more or less rapidly, and thus it could stop a few grid points away from the real edge. Also, variations in the gradient along the boundary can cause inaccurate results. False gradients caused by noise can be avoided using an edge-preserving smoothing scheme on the image as a preprocessing step; see Ref. 19.

The user can run the fast-marching method from a given set of initial points (mouse clicks on the background of the images). Alternatively, the user can decide to segment only the structures within a manually defined rectangular region of interest. If the region is too big, subsampling can be used so that the segmentation process is not too slow. However, this option must be used carefully, since subsampling smooths the boundaries of the objects present in the image and can completely obliterate smaller structures. The resulting contour (or contours if we are segmenting several objects at the same time) will provide an excellent initial condition for the level-set method.

Putting all of these elements together, we are able to obtain a good approximation of the shape of the object that we are trying to segment (Fig. 5). In order to improve the final result, we propose to run a few iterations of the level-set method using the result of the fast-marching method as the initial condition.

### 2.2.3 Final segmentation: level-set method

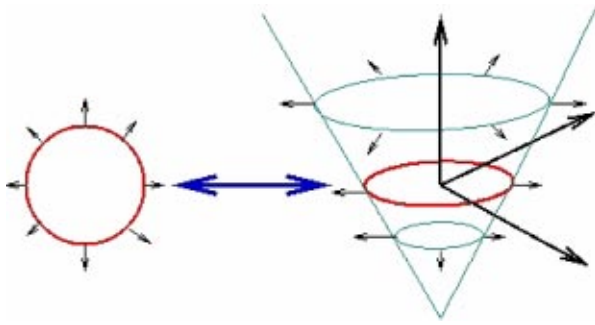
Once we have obtained a good approximation of the shape of the object using the fast-marching algorithm, we can afford to use the more computationally expensive level-set method to improve the result of the segmentation. The essential idea here is to embed our marching front as the zero level set of a higher dimensional function. In our case, we take that function to be  $\phi(x) = \pm d$ , where  $d$  is the signed distance from  $x$  to the front (see Fig. 4), assigning negative distances for pixels inside the evolving curve, and positive distances for pixels outside it.

Following the arguments in Ref. 14, we obtain the following evolution equation:

$$\phi_t + F(x, y) |\nabla \phi| = 0, \quad (6)$$

where  $F$  is again the speed in the normal direction. This is an initial-value partial differential equation, since it describes the evolution of the solution on the basis of an initial condition defined as  $\phi(x, t=0) = \phi_0$ . As pointed out earlier, the level-set approach offers several advantages:

- The zero level set of the function can change topology and form sharp corners.



**Fig. 4** Basic concept behind the level-set method. The marching front is embedded as the zero level set of a higher dimensional function.

- A discrete grid can be used together with finite differences to approximate the solution.
- Intrinsic geometric quantities such as normal and curvature can be easily extracted from the higher dimensional function.
- Everything extends directly to 3-D.

To mold the initial condition (in our case the result of the fast-marching method) into the desired shape, we use two force terms. By substituting these two terms in the motion equation<sup>18,20</sup> we get:

$$\phi_t - g(1 - \epsilon\kappa)|\nabla\phi| - \beta\nabla g \times \nabla\phi = 0, \quad \epsilon > 0, \quad \beta > 0, \quad (7)$$

where  $g$  is an edge indicator function defined by the speed of the front [Eq. (5)] in the eikonal Eq. (2);  $\kappa$  is the curvature of the expanding front, and  $\phi_t$  is the unknown that we are trying to compute. As before, the image  $I(x)$  can be preprocessed using an edge-preserving smoothing scheme.

The second term of the equation has two components. The first one slows the surface in the neighborhood of high gradients (edges), while the second one (motion by curvature) provides regularity to the curve. The parameter  $\epsilon$  is the weight of the motion by curvature term, and determines the strength of its regulatory effect: the bigger we make  $\epsilon$ , the more we limit

the possibility of obtaining sharp corners and irregular contours. In practice, an intermediate value of  $\epsilon$  provides a good trade off between contour smoothing and accuracy.

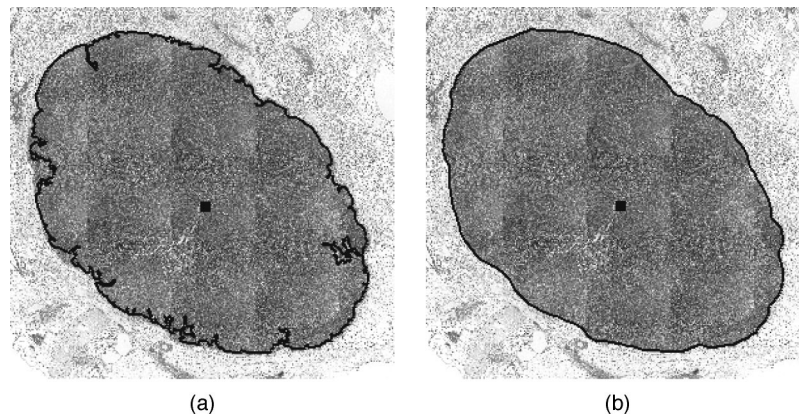
Finally, the last term of the equation attracts the front to the object's boundaries. It aligns all the level sets with the ideal gradient, which would be a step function centered at the point of maximum gradient in the original image.  $\beta$  is the weight of the advection of the front by the edge vector field  $\nabla g$ . It determines the strength of the attraction of the front to the edges.

At times, for very small objects, it is possible to use the level-set method from the initial point. However, for any type of morphological structure, we can use the result of the eikonal Eq. (2),  $U(x,y)$ , as the initial condition for the level-set algorithm:  $\phi(x,y;t=0) = U(x,y)$ . Then by solving Eq. (7) for a few time steps using the narrow-band approach, we obtain an accurate, real-time segmentation of the desired object (see Fig. 5).

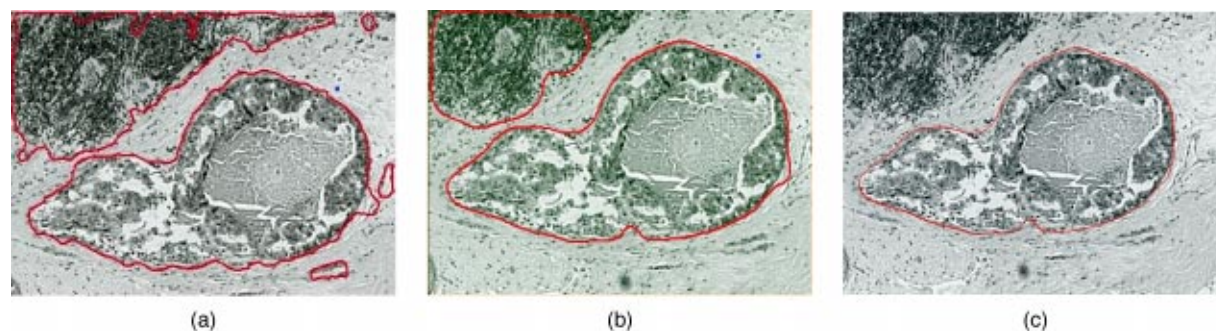
### 3 Results

In this section we consider the problem of reconstructing DCIS areas in a tissue biopsy of a cancerous human breast, as well as a group of normal ducts through an entire mouse mammary gland. The tissue samples were sliced for a total of 55 sections in the human tissue and 206 sections in the mouse one. We used all of the sections for the reconstruction of the human case and 40 sections for reconstruction of the mouse case. Manually delineating all the structures in every section is an extremely time-consuming process. To automatically segment those structures using the framework described in Sec. 2, we begin by defining a region-of-interest (ROI) where we will run the segmentation algorithm. This ROI can be extended to cover the entire section, but considering the size of the images, it is wise to use subsampling in order to run the algorithm in real time. The level of subsampling can be determined by the user; greater subsampling can be used on large ROIs without compromising the resolution and accuracy of the final segmentation.

After defining the ROI, we have to tune the different parameters of the segmentation process, particularly  $\alpha$  [Eq. (5)]



**Fig. 5** The segmentation of a lymph node in a mouse mammary gland section is shown in black. (a) The result of the fast-marching method; it provides a good approximation to the boundaries of the lymph node; the blue point in the middle of the lymph node is the initial contour from which the algorithm was run. (b) The result of using the level-set method after the fast-marching method; the final contour is more accurate and smoother. In both cases the images were subsampled in the x and y directions to be able to run the segmentation in real time.



**Fig. 6** Segmentation of a DCIS tumor in human mammary gland tissue. (a) Segmentation using just the level set method. The blue dot represents the initial contour. (b) Results of the full segmentation with blue lines delineating tumor masses. (c) Results after editing using the interactive tools provided by the system.

and  $\epsilon$  [Eq. (7)]. This is done by the user based on the default values provided by the algorithm and the type of object that he or she is trying to segment. However, we have observed that a particular set of parameters is frequently good enough to segment similar structures (i.e., all the tumors, all the ducts, all the lymph nodes) throughout all the sections of a particular tissue block. Thus the user only needs to modify the parameters the first time that he or she tries to segment a new type of morphological element in the tissue.

After selecting the parameters of the flow, initial points are defined inside (to find the internal contour) or outside (to find the external contour) the structures of interest. In most cases one computer mouse click is enough, although large images may require several evenly distributed clicks. The value of  $U(x)$  at these points is set to zero as in Eq. (2), and the fast-marching algorithm of Table 1 is executed. When this method finishes, the final  $U(x)$  function is passed as the initial condition to the level-set motion Eq. (7). We then iterate this equation for a few steps. This segmentation scheme provides a result in less than 1 s for images whose size (after subsampling, if any) is around 2 kbytes (e.g.,  $512 \times 512$  pixels) running on a Sun Ultra 10 workstation with 1 Gbytes of RAM.

### 3.1 Human Case Segmentation

Figure 6 displays the result obtained with the combination of the fast-marching and the level-set methods in a tissue biopsy from a patient with an intraductal carcinoma. Figure 6(b)

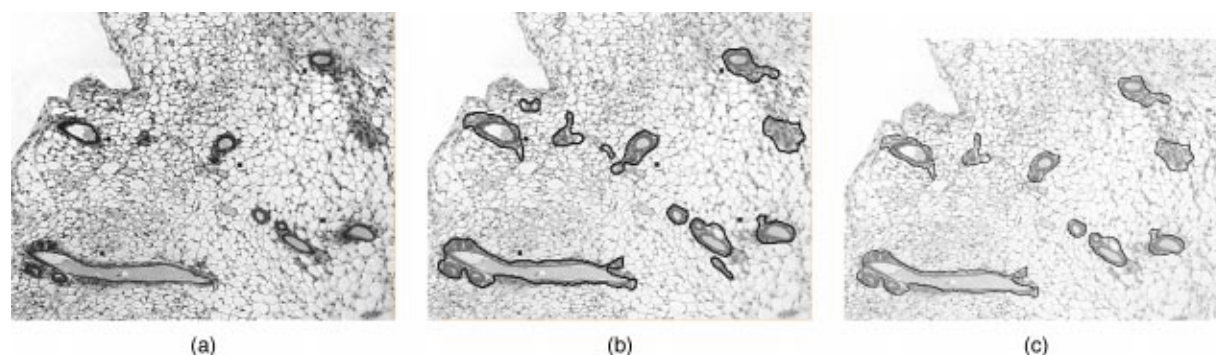
shows the segmentation of a tumor mass in a human tissue block. The results of the segmentation can be edited and removed with the interactive tools provided by our system [see Fig. 6(c)]. For this segmentation, an area was selected around the structures of interest and no subsampling was used. The initial contours are represented by blue points in Fig. 6(a).

### 3.2 Mouse Case Segmentation

In Fig. 7 we can see the segmentation of the external contours of several ducts in a particular area of a mouse mammary gland. In this case we also run the algorithm on an ROI on one of the sections with no subsampling factor. Figure 7(a) displays the points where we initialized the contours. Figures 7(b) and 7(c) show the results of the segmentation before and after interactive correction of the results, respectively.

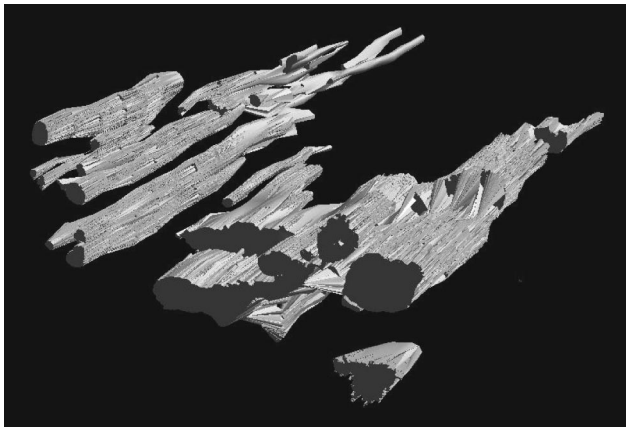
### 3.3 3-D Reconstruction

Finally the segmented shapes are connected (manually) between sections, and 3-D reconstructions of the samples are built. Figure 8 shows a reconstruction of the tumors contained in the human tissue block, including the tumor shown in Fig. 6. Increasing the “motion by curvature” term, as described in Sec. 2, can reduce surface noise. In Fig. 9 we can see the reconstruction of the normal ducts segmented in the images of the mouse mammary gland (see one of the corresponding 2-D segmentations in Fig. 7).

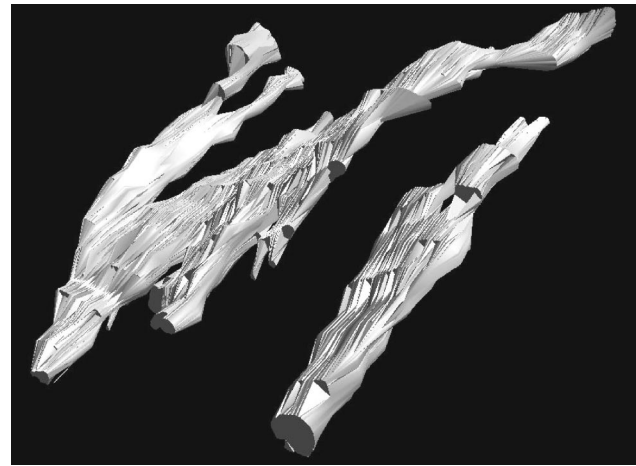


**Fig. 7** Segmentation of normal ducts in a mouse mammary gland tissue sample. (a) Initial data. (b) Results of the segmentation (red lines delineate normal ducts). (c) Results after editing using the interactive tools provided by the system.





**Fig. 8** 3-D reconstruction of tumors in a human mammary gland tissue block. Tumor masses are rendered as gray volumes. The scene was stretched ten times in the z direction to obtain a better view.



**Fig. 9** 3-D reconstruction of normal ducts in a mouse mammary gland. The ducts are rendered as gray volumes. A single duct and its branches can be traced throughout the gland. The z direction was not stretched in this case.

### 3.4 Further Examples

Figures 10 and 11 show two more examples of the results that can be obtained with our segmentation approach. In both cases the initial ROI was subsampled by a factor of 2 in both the x and y directions. The same segmentation parameters ( $\alpha$  and  $\epsilon$ ) were used for both examples.

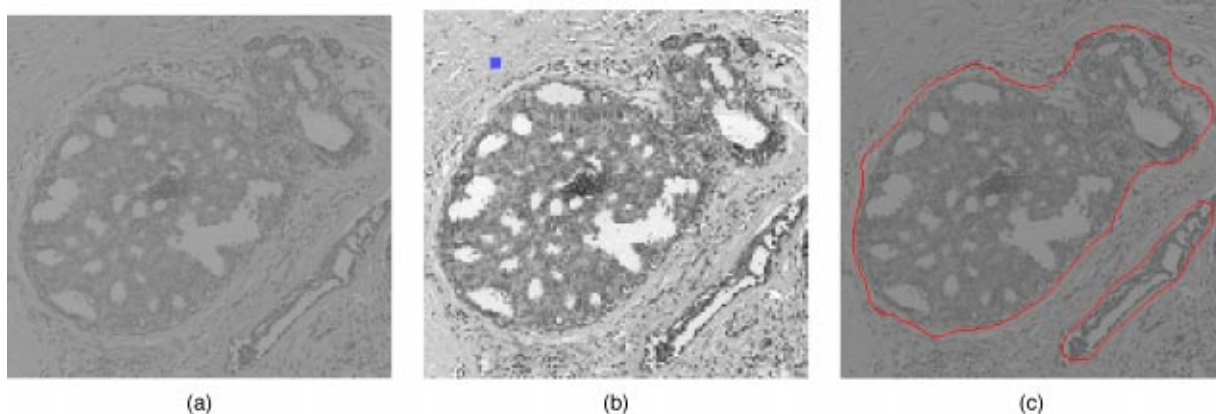
Figure 10(a) shows a DCIS lesion together with a normal duct in a human tissue section. Both structures were segmented at the same time using a single initial point, as can be seen in Fig. 10(b). Figure 10(c) shows the results incorporated to the full resolution image.

In Fig. 11(a), a terminal ductal lobular unit (TDLU) can be observed. These are lobuloalveolar structures where milk is produced during lactation in the human breast. The multiple alveoli that form the TDLU, together with the presence of a ductal part, make automatic segmentation of this type of structure a difficult task. However, after subsampling the image, our algorithm is able to find a contour that surrounds the entire structure [Figs. 11(b) and 11(c)], thus allowing its 3-D reconstruction.

## 4 Discussion

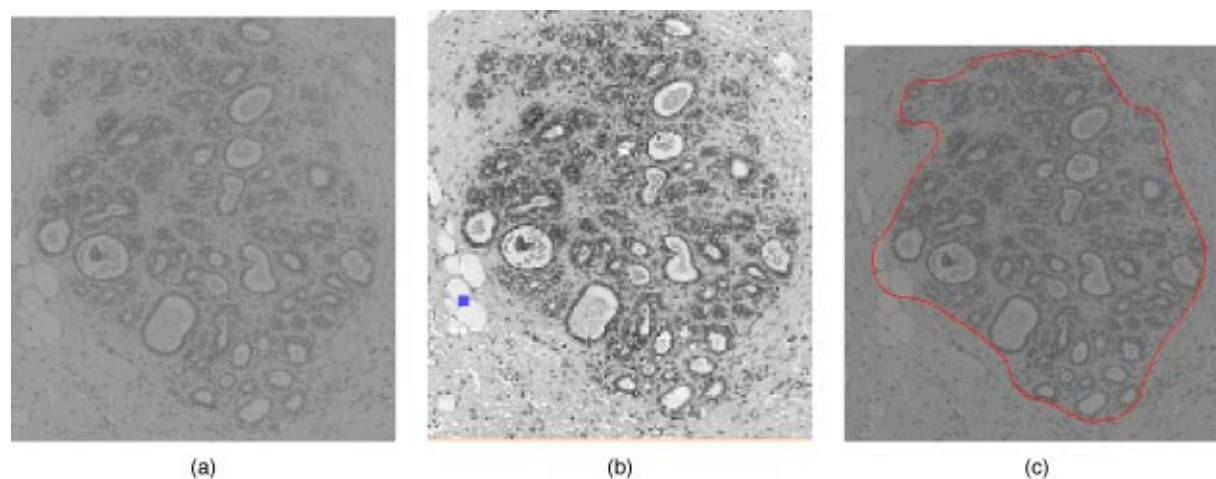
We have developed a microscopy system that semiautomatically reconstructs histological structures from fully sectioned tissue samples. However, the interaction required to operate the system is quite intensive, limiting the scope of its application to small tissue volumes or to studies not requiring high throughput analysis of the samples. In this paper we have presented a method that reduces the time and interaction needed to build the 3-D reconstruction of a tissue block, thus increasing the potential throughput of the system and therefore allowing us to use this approach for the reconstruction of large, complex specimens. To achieve this goal we have combined image processing techniques and two well-established schemes for interface propagation: the fast-marching method and the level-set method.

Our approach starts by correcting the background of the images. This is an important step, since the background pattern generated during image acquisition modifies the gradient



**Fig. 10** Segmentation of a different DCIS tumor in human mammary gland tissue. (a) The duct on the-left contains a DCIS lesion with a necrotic center; the one on the right is normal. (b) The ROI was subsampled by a factor of 2 in both the x and y directions; the blue dot represents the initial seed. (c) Results of the segmentation on the full-resolution image.





**Fig. 11** Segmentation of a terminal ductal lobular unit in a human mammary gland. (a) The duct on the left contains a section of a TDLU, one of the sites of milk production in the human breast. (b) The ROI was subsampled by a factor of 2 in both the x and y directions; the blue dot represents the initial seed. (c) Results of the segmentation on the full-resolution image.

of the image, and the speed function that we use for interface propagation depends on that gradient. Once the background has been corrected, we run the fast-marching method. This technique provides a good approximation of the boundaries of the objects that we are trying to segment in a very short time, since it assumes monotonic speed functions (always positive or always negative). We then use the approximation provided by the fast-marching method as the initial condition for the level-set method. This more computationally expensive algorithm is run for just a few steps, enough to fit the front to the contours of the structures of interest, but not as many as to make the segmentation too time-consuming.

Though this approach is very useful, it can still be improved. The most accurate segmentations (and reconstructions) are obtained on full-resolution images. However, for some large structures, such as lymph nodes, or when trying to delineate multiple elements at the same time (for example, a group of ducts), segmentation on a full-resolution image is not real time, and can take up to 1 min. Using subsampling takes the segmentation execution time back to real time at the expense of some accuracy loss. Also, in areas with a lot of texture in the tissue around the ducts (stroma), tuning the parameters of the algorithm ( $\alpha$  and  $\epsilon$ ) becomes more difficult. At times this process can take a few trials, since the expanding front tends to get trapped in high gradients that do not correspond to the boundaries of the feature that we are trying to segment, but to the texture of the stroma. Finally, once all the structures of interest have been segmented, the user still needs to manually connect them between sections. This constitutes a new bottleneck in the tissue analysis process.

For these reasons, we are currently working on a time step-independent scheme that is expected to be faster than the current one. To improve the accuracy of the results, we have developed an edge-preserving smoothing algorithm based on the Beltrami flow,<sup>19</sup> which can replace the Gaussian smoothing currently used before executing the segmentation methods. This algorithm eliminates false gradients that are due to noise, while enhancing gradients that are due to the object's boundaries, thus allowing the front to fit the boundaries of the

object more accurately. Finally, the level-set approach is readily extensible to 3-D. The ability to segment 3-D structures of interest versus 2-D ones would save the process of connecting the segmented 2-D contours from section to section, thus improving the analysis time. Also, since geometric properties can be easily extracted from the higher dimensional function used in the level-set algorithm, we could readily obtain some information about the extracted volume from the segmentation algorithm itself. We are also working on a fully interactive segmentation method based on the model of the intelligent scissors,<sup>21</sup> because the selection of the seed points for segmentation is something that cannot be automated, owing to the high variability in the shapes of object's to be segmented.

In conclusion, we have presented a real-time method for automatic segmentation of morphological structures in mammary gland tissue sections. It is precisely the delineation of those structures that required the heaviest user interaction in our sample analysis protocol. Therefore, the automatic approach to segmentation that we describe here represents a first step toward real-time reconstruction and analysis of mammary gland samples. Achieving that goal would allow us to accelerate our studies on the biological basis of human breast cancer. Moreover, obtaining real-time reconstruction and analysis of samples from our system would be useful for pathological diagnosis in a clinical environment; 3-D renderings of all the morphological structures in a mammary gland biopsy could be mapped with the distribution of particular markers of breast cancer within a few hours of extracting the tissue from the patient. From an intrasurgical point of view, the renderings would prove—tissue processing and stain permitting—an important tool in the evaluation of breast tumors and their margins.

### Acknowledgments

The U.S. Army Medical Research Materiel Command under grants DAMD17-00-1-0306 and DAMD17-00-1-0227, the Lawrence Berkeley National Laboratory Directed Research

and Development Program, and the California Breast Cancer Research Program, through project #8WB-0150, supported this work. This work was also supported by the Director, Office of Science, Office of Advanced Scientific Research, Mathematical, Information, and Computational Sciences Division, U.S. Department of Energy under contract no. DE-AC03-76SF00098.

## References

1. L. Hennighausen and G. W. Robinson, "Think globally, act locally: the making of a mouse mammary gland," *Genes Dev.* **12**(4), 449–455 (1998).
2. L. M. Franks and N. M. Teich, *Introduction to the Cellular and Molecular Biology of Cancer*, Oxford University Press, Oxford (1997).
3. T. Tot, L. Tabar, and P. B. Dean, *Practical Breast Pathology*, 1st ed. Falun Central Hospital, Thieme (2002).
4. R. D. Cardiff and S. R. Wellings, "The comparative pathology of human and mouse mammary glands," *J. Mammary Gland Biol. Neoplasia* **4**(1), 105–122 (1999).
5. R. Fernandez-Gonzalez, A. Jones, E. Garcia-Rodriguez, P. Y. Chen, A. Idica, M. H. Barcellos-Hoff, and C. Ortiz de Solorzano, "A system for combined three-dimensional morphological and molecular analysis of thick tissue samples," *Microsc. Res. Tech.* **59**(6), 522–530 (2002).
6. T. Ohtake, R. Abe, I. Kimijima, T. Kukushima, A. Tsuchiya, K. Hoshi, and H. Wakasa, "Computer graphic three-dimensional reconstruction of the mammary duct-lobular systems," *Cancer* **76**(1), 32–45 (1995).
7. D. F. Moffat and J. J. Going, "Three-dimensional anatomy of complete duct systems in human breast: pathological and developmental implications," *J. Clin. Path.* **49**, 48–52 (1996).
8. T. Ohtake, I. Kimijima, T. Fukushima, M. Yasuda, K. Sekikawa, S. Takenoshita, and R. Abe, "Computer-assisted complete three-dimensional reconstruction of the mammary gland ductal/lobular systems," *Cancer* **91**(12), (2001).
9. F. Manconi, R. Markham, G. Cox, E. Kable, and I. S. Fraser, "Computer-generated, three-dimensional reconstruction of histological parallel serial sections displaying microvascular and glandular structures in human endometrium," *Micron* **32**, 449–453 (2001).
10. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comp. Vis.* **1**(4), 321–331 (1988).
11. R. Malladi, J. A. Sethian, and B. C. Vemuri, "A topology-independent shape modeling scheme," *Proc. SPIE* **2031**, 246–258 (1993).
12. R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modelling with front propagation: a level set approach," *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(2), 158–175 (1995).
13. V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours," *Numer. Math.* **66**, 1–31 (1993).
14. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on the Hamilton-Jacobi formulation," *J. Comput. Phys.* **79**, 12–49 (1988).
15. J. A. Sethian, *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision and Materials Sciences*, 2nd ed. Cambridge University Press, University of California, Berkeley (1999).
16. D. Adalsteinsson and J. A. Sethian, "A fast level set method for propagating interfaces," *J. Comput. Phys.* **118**, 269–277 (1995).
17. J. A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proc. Natl. Acad. Sci. U.S.A.* **93**(4), 1591–1595 (1996).
18. R. Malladi and J. A. Sethian, "A real-time algorithm for medical shape recovery," in *Proceedings-of-the IEEE International Conference on Computer Vision (ICCV'98)*, pp. 304–310 (Jan. 1998).
19. N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Trans. Image Process.* **7**(3), 310–318 (1998).
20. C. Ortiz de Solorzano, R. Malladi, S. A. Lelievre, and S. J. Lockett, "Segmentation of nuclei and cells using membrane-related protein markers," *J. Microsc.* **201**(3), 404 (2001).
21. W. Barrett and E. Mortensen, "Interactive live-wire boundary extraction," *Med. Image Anal.* **1**(4), 331–341 (1997).