PARTIAL DIFFERENTIAL EQUATIONS-BASED SEGMENTATION FOR RADIOTHERAPY TREATMENT PLANNING

Frédéric Gibou

Department of Computer Science and Department of Mechanical Engineering
University of California at Santa Barbara, CA 93106-5070

Doron Levy

Department of Mathematics, Stanford University
Stanford, CA 94305-2125

Carlos Cárdenas

Siemens Medical Solutions
Med SW West, 755 College Road East, Princeton, NJ 08540

Pingyu Liu

Department of Radiation Oncology, Stanford University
Stanford, CA 94305

Arthur Boyer

Department of Radiation Oncology, Stanford University
Stanford, CA 94305

(Communicated by Yang Kuang)

ABSTRACT. The purpose of this study is to develop automatic algorithms for the segmentation phase of radiotherapy treatment planning. We develop new image processing techniques that are based on solving a partial differential equation for the evolution of the curve that identifies the segmented organ. The velocity function is based on the piecewise Mumford-Shah functional. Our method incorporates information about the target organ into classical segmentation algorithms. This information, which is given in terms of a three-dimensional wireframe representation of the organ, serves as an initial guess for the segmentation algorithm. We check the performance of the new algorithm on eight data sets of three different organs: rectum, bladder, and kidney. The results of the automatic segmentation were compared with a manual segmentation of each data set by radiation oncology faculty and residents. The quality of the automatic segmentation was measured with the "κ-statistics", and with a count of over- and undersegmented frames, and was shown in most cases to be very close to the manual segmentation of the same data. A typical segmentation of an organ with sixty slices takes less than ten seconds on a Pentium IV laptop.

1. **Introduction.** Three-dimensional conformal radiotherapy (3DCRT) and intensity-modulated radiation therapy (IMRT) are being widely developed and implemented for clinical applications. These procedures depend on intense use of patient imaging. The availability of spiral computerized tomography (CT) x-ray scanners has made practical the acquisition of large patient image sets consisting of around one hundred reconstructed planes. Commercially available computer systems provide sophisticated treatment-planning capabilities, including non-coplanar 3DCRT and IMRT. Other imaging techniques, such as contrast-enhanced magnetic resonance imaging (MRI), magnetic resonance spectroscopy (MRS), and positron emission tomography (PET), are being investigated to provide more specific information regarding the biological activity of the target volume. Most frequently these three-dimensional studies are fused with a treatment-planning CT to transfer this target volume data onto the treatment-planning CT. Using these treatment-planning systems, the radiation oncologist can prescribe dose distributions that conform closely to tumor target volumes. With computerized treatment planning, it is a standard practice to reduce the dose that neighboring normal anatomical structures receive during the course of the radiotherapy procedure. However, implementing this technology represents a time-consuming part of the process that most centers have adopted. Typically, these structures are segmented on workstations by drawing closed contours around the cross sections of the anatomy as perceived by the operator in axial CT reconstructions. Software tools supporting this procedure are provided in most commercial treatment-planning systems. These tools use the current state-of-the-art image display and graphic interaction techniques. Nevertheless, the segmentation process is still subjective and time consuming.

Investigators in the field of computer vision have carried out intensive studies of image segmentation techniques, developing a wide variety of segmentation methods intended for a broad range of imaging applications. John Canny [1] introduced the first-edge detection method, which was based on local gradients. Techniques such as seeded region growing [2] or split-and-merge [3], using Markov random-field modeling, were the first region-based methods. A good review on region-growing algorithms can be found in [4].

Deformable models, which are based on functional minimization, were introduced to the computer-vision [5] and computer-graphics [6, 7] communities in the late 1980s by Demetri Terzopoulos et al.. An example that has gained popularity is known as "snakes" [8]. Researchers have improved on the original "snakes" to control the difficulties associated with their sensitivity to initial conditions. For example L. Cohen [9] proposed a balloon force model that uses an internal inflation term. Vicent Caselles et al. [10] proposed the first level-set formulation in a non variational setting (see also [11]). The ability of the level-set method [12] to handle complex topological changes automatically alleviates the difficulties associated with topological transformations. Caselles et al. [13] and Satyanad Kichenassamy et al. [14] introduced geodesic active contours. This geometric and variational model in level-set form shows the connection to the original "snakes." Geodesic active contours have been extended to incorporate shape information and probabilistic estimates (e.g., see [15], [16], and the references therein). The interested reader is referred to the book by Guillermo Sapiro [17] and the survey by McInerney and Terzopoulos [18].

One drawback of gradient-based methods is that the user must correctly guess a threshold for the gradient. Moreover, since the boundaries of an object are

not necessarily defined by the same gradient jump, such methods often lead to "leakage" through the boundary. That is, the curve often does not stop at the correct location of the boundary of an organ. Based on the Mumford-Shah minimal partition functional [19], Tony Chan and Luminita Vese [20] (see also [21]) proposed a new level-set model for active contours to detect objects whose boundaries are not necessarily defined by a gradient. This algorithm does not suffer from the limitation associated with traditional gradient-based methods but is limited by the computational expense inherent to solving the proposed nonlinear parabolic partial differential equation(PDE). Recently, Frederic Gibou and Ron Fedkiw [22] proposed a new hybrid algorithm that benefits from the simplicity and efficiency of k-Means [23] while preserving the robustness of level-set algorithms, and they showed that three-dimensional objects can be segmented in real time. We would like to emphasize that standard level-set algorithms, which do not utilize additional information on the target organ, will typically fail to properly segment such images because they cannot stop moving the level set when the boundaries are ill defined.

The purpose of this work is to develop new image-processing techniques that are based on modifying existing curve evolution algorithms by taking into account prior knowledge on the target organ. In these studies, the target of the segmentation procedure was a normal tissue structure. We have excluded from this investigation the more difficult problem of automatically segmenting clinical target volumes (CTVs). However, these techniques may have potential for CTV segmentation in MRI and PET studies. The ultimate goal is to automate the segmentation phase in radiotherapy treatment planning to the point that segmentation is not an excessively time-consuming part of the process.

Our approach has two stages. First, a three-dimensional wireframe representation of a model organ is manually placed on top of the CT data set. This stage is then followed by a fully automatic segmentation step in which we implement the piecewise Mumford-Shah energy [19] with the fast algorithm of Gibou and Fedkiw [22]. Our study focuses on CT data sets, but the method can be equally applied to different data set types (e.g MRI, PET, etc.)

This paper is organized as follows: The segmentation algorithm is described in section 2. We start in section 2.1 by discussing the manual segmentation procedure. Our segmentation strategy is presented in section 2.2. We describe both phases of the segmentation process: the manipulation of a model organ and the automatic segmentation algorithm. Statistical tools for comparing the automated segmentation algorithm with the manual segmentation algorithm are briefly reviewed in section 3. The results of the implementation of our ideas on eight different data sets of three organs are presented in section 4. Some concluding remarks are provided in section 5.

2. The segmentation algorithm.

2.1. **Manual segmentation.** A rich sample of CT data sets was accessible to the investigators. To build a model repository of segmented organs for each structure, specific organs were manually segmented. Special care was taken to follow the surfaces smoothly across reconstruction planes and at the superior and inferior extremes of structures. By viewing the structures after segmentation as a three-dimensional structure from several angles, the investigators could verify the quality of the segmentations. For each CT data set, several manual segmentations of the same organ were acquired. The manual segmentations of each organ were compared

using the κ-statistic [24] as a measure of the variance of the manual segmentations associated with that organ and a count of over- and undersegmentation. The hand-drawn polygons that bound the structure surface were saved with each CT data set as well.

2.2. **A segmentation strategy.** We propose a new level-set segmentation algo-rithm that takes into account prior knowledge of the organ to be segmented. The key idea is to incorporate the structure of the target organ into the segmentation, while processing three-dimensional data. The segmentation is done in two steps. First, a radiation oncologist or a member of the experienced planning staff manually places a three-dimensional wireframe that closely represents the organ of interest inside a multiplanar CT reconstruction. This procedure is supported by a library of standard organs that is part of our segmentation and visualization framework called VolVisT. An example of multiplanar CT reconstruction is given in Figure 1. In the second step, the actual segmentation is performed. This part uses geometric PDE-based image segmentation techniques and is fully automatic. The benefit of this approach is that the human time required during the manual segmentation process can be reduced drastically while still retaining the desired accuracy.

The goal is to determine a closed contour that surrounds and defines a given anatomical structure. Such closed contours are routinely produced by a manual selection of the vertices overlaid on a sequence of axial CT reconstruction planes by an expert operator. The set of multiple closed contours defined in contiguous axial planes constitutes the segmentation of the structure. The automated strategy described here seeks to find such contours by means of a computer algorithm that encloses the same anatomical features that would be selected by the majority of human observers using manual methods. We assume that the organ we seek to segment has an associated model in the model data set. The model data set consists of a CT study and a set of closed contours that encapsulate the organ in the study. The target data set is the CT study containing the organ to be segmented.
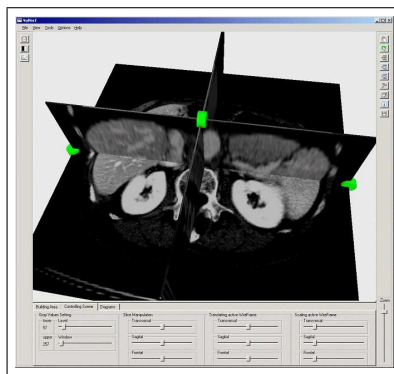

FIGURE 1. Multiplanar reconstruction of the CT data inside the VolVisT framework.

2.3. **Step I: Manipulating a model organ.** The first step in the procedure is for the human operator to use manual affine transformations (stretching, translation, rotation, and scaling) to place a generic wireframe representation of the organ around the target organ. Generally, the shape of the target organ will be different
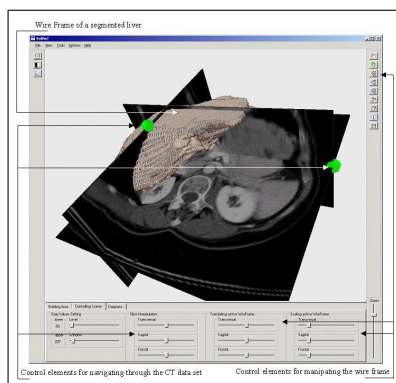
FIGURE 2. Example of the wireframe hull of a segmented liver in VolVisT.

from the shape of the model wireframe. To conform the wireframe to the target organ, one must use some type of deformable warping to match the target organ shape. The strategy proposed here is designed to achieve such deformable warping.

The program framework developed for this step, VolVisT, allows the operator to view axial, saggital, and transverse planes simultaneously. The viewing planes can be moved across the study and the entire display can be translated, rotated and scaled to provide the operator with a view of any set of orthogonal planes from any desired angle (see Fig. 1). The wireframe can be saved in association with the original study so as to serve as a model or as a reference for the structure. The wireframe can be recalled by the operator and displayed within a different study. Provisions are made to move and scale the model wireframe by means of affine transformation parameters and to view it overlaid in the cut planes in three spatial dimensions. An example is given in Figure 2.

At the completion of the manual affine transformation, the vertices of the wireframe have been transformed into the coordinate system of the target study and lie in the vicinity of the target organ. The contours making up the wireframe for the structure are represented by a stack of hexagons. The wireframe is manipulated based on the correspondence to a feature (in the gray-scale distribution) identified by the operator when the model contour was drawn. This feature was identified using the operator's experience with anatomy and CT images. This step is critically dependent on the quality of the model library as well as the complexity of the organ.

We stress that, even though it is not important to place the wireframe exactly on the organ, the following two guidelines should be followed in order for the automatic segmentation to be successful: First, the wireframe should envelope the target organ; that is, the segmented contour must not enclose an area that is outside the wireframe. Second, special attention should be given only to regions where a single image does not provide with a clean separation between organs (see, e.g., the bladder/prostate example in Fig. 3).

2.4. **Step II: An automatic segmentation algorithm.** The second step of the segmentation process is to transform the topology given by the wireframe in such a way as to conform to the target organ. Therefore, the process is reduced to evolving a curve (or surface) in two (or three) spatial dimensions from a given topology to
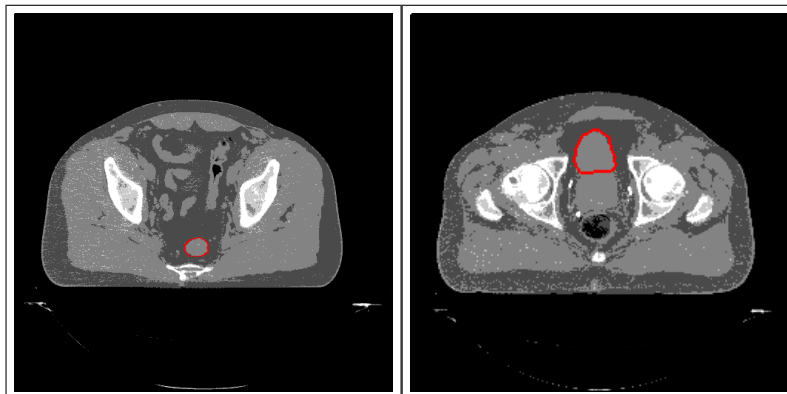
FIGURE 3. Segmentation of the rectum (left) and the bladder (right).

the desired one. Methods that represent curves or surfaces and that carry out their evolutions under a prescribed velocity field are numerous. The representation we choose in this work is the level-set method of Osher and Sethian [12].

Level-set methods provide a natural mathematical representation for dealing with complicated structures and offer great advantages when the curve or surface undergoes complex topological changes such as pinching and merging. An added benefit is that geometrical features such as the normal direction, the curvature, the area, and the volume of an object can be easily computed from the information that already exists in the level-set representation.

With level-set methods, the boundary of an organ, which is a curve in two dimensions and a surface in three dimensions, is chosen to be implicitly represented as the zero level set of a smooth function ϕ. Instead of deforming the contour directly, we evolve the function ϕ with respect to a fictitious time t. At any time instance, we can read back the contour (which represents the boundary of the organ) by extracting the zero level set of ϕ. The evolution of the level set under the velocity field $\vec{S} = S_n \cdot \vec{n} + S_\tau \cdot \vec{\tau}$, where S_n and S_τ are respectively the normal and tangential component of the velocity field, satisfies a nonlinear Hamilton-Jacobi PDE of the form

$$\frac{\partial \phi}{\partial t} + S_n |\nabla \phi| = 0. \tag{1}$$

Note that the speed function S_n can depend on ϕ and its derivatives, and therefore, it can depend, for example, on the mean curvature or on the normal to the contour. Often, it is of practical interest to define ϕ as the signed distance function to the object in consideration, since it yields robust numerical results. To keep the values of ϕ close to those of a signed distance function (i.e., $|\nabla \phi| = 1$), the level set is reinitialized using the fast marching method [25, 26]. The normal ($\vec{n} = \nabla \phi / |\nabla \phi|$) and the mean curvature ($\nabla \cdot \vec{n}$) to the level set are computed with standard central differencing. Solutions of the level-set advection equation (1) are approximated in this work using HJ-WENO upwind schemes [27]. These schemes combine a monotone numerical flux (which we choose to be the local Lax-Friedrichs flux) with a fifth-order weighted essentially non oscillatory (WENO) reconstruction. For more details on the level-set method and its numerical implementation, see [28, 29].

2.4.1. *An intensity-based speed function.* The fundamental and most intricate ingredient in the construction of a successful level-set segmentation algorithm is the definition of the speed function S_n of equation (1). The goal is to define a speed function such that the initial contour given by the wireframe wraps the contour of the target organ in the steady-state limit, for which $S_n \to 0$. In principle, one can define sophisticated velocity fields. For example, it is possible to construct a velocity field that is based on boundary detection (BD), mutual information (MI), geometric properties (G) (such as the curvature or the normal map), probabilistic models (P), texture (T), and so forth. Hence, in general we can assume

$$S_n = f(S_{BD}, S_{MI}, S_G, S_P, S_T, ...). \tag{2}$$

Also, the velocity sources can be combined in different ways. For example, defining $S_n = S_{BD} + S_{MI}$ requires both criteria to be satisfied, while ($S_n = S_{BD}$ or S_{MI}) requires that only one criterion be satisfied. In this work, we use a velocity that is defined from the piecewise-constant Mumford-Shah functional [19].

A typical CT scan obtained in radiotherapy treatment planning is illustrated in Figure 3. Organs can be either distinctly separated from each other (see Fig. 3, left-hand image) or can bound other organs (see Fig. 3, right-hand image). For the sake of clarity, we will first describe our level-set approach in the ideal case where the target organ is clearly separated from other organs, and then comment on how to adjust the algorithm to the second setup. First, we define a window around the target organ. Focusing on the restriction of the CT scan to that window, it becomes apparent to the human eye where the organ lies; that is, one can make a clear distinction between the organ and its background. A more quantitative description of this process is to say that one seeks to separate the image into regions with respect to their respective average intensity value. The piecewise Mumford-Shah functional offers a framework to do precisely this. In this case, the velocity field is defined as

$$S_n(\vec{x}, t) = (I(\vec{x}) - c_{in}(t))^2 - (I(\vec{x}) - c_{out}(t))^2 - \mu \nabla \cdot \left(\frac{\nabla \phi(\vec{x}, t)}{|\nabla \phi(\vec{x}, t)|} \right), \tag{3}$$

where $I(\vec{x})$ defines the image intensity map at the voxel location $\vec{x}$. The image mean intensity value outside ϕ, c_{out}, and inside ϕ, c_{in}, are computed using

$$c_{out}(t) = \frac{\int_\Omega I(\vec{x}) H(\phi(\vec{x}, t)) d\Omega}{\int_\Omega H(\phi(\vec{x}, t)) d\Omega}, \quad c_{in}(t) = \frac{\int_\Omega I(\vec{x})(1 - H(\phi(\vec{x}, t))) d\Omega}{\int_\Omega (1 - H(\phi(\vec{x}, t))) d\Omega}. \tag{4}$$

Here, $H(\phi)$ is the Heaviside function; that is, $H(\phi) = 1$ if $\phi > 0$, and $H(\phi) = 0$ otherwise. The last term in equation (3) is the curvature and is a regularizing term that helps processing noisy CT scans. For more details, see [20]. The advantage in this formulation is that we avoid using gradient-based information, which eliminate the need to tune threshold parameters.

In [22], Gibou and Fedkiw proposed a fast hybrid algorithm that draws on the efficiency of standard clustering algorithms and the robustness of level-set methods. We have used this method and obtained similar results, as in our previous work, with the obvious gain of efficiency. Typically, it takes few seconds to segment a complete volume, which is orders of magnitude faster than with the standard implementation of [20]. Within the algorithm of [22], the notion of curvature can be implemented in the same fashion as is traditionally done with level-set methods. However, in cases such as the segmentation of CT scans, a preprocessing step of a nonlinear diffusion algorithm [30] is enough to treat noisy images, and

consequently, the motion by mean curvature term can be ignored all together. To optimize the computational efficiency of this preprocessing step, the authors used an implicit backward Euler time-integration scheme on a linearized approximation of the nonlinear diffusion equation at each time step.

Perona and Malik [30] introduced the use of nonlinear diffusion for denoising images while keeping the image edges intact. This denoising is achieved by solving the nonlinear PDE

$$\frac{\partial I(\vec{x}, t)}{\partial t} = \nabla \cdot (g(|\nabla I|)\nabla I), \tag{5}$$

where $I(\vec{x}, t)$ defines the image intensity map at the voxel location $\vec{x}$ and at the fictitious time t, and g is an edge-stopping function chosen such that $\lim_{s \to \infty} g(s) = 0$ so that diffusion stops at the location of large gradients. Based on the original function proposed in [30], we choose

$$g(s) = \frac{\nu}{1 + \frac{s^2}{K^2}},$$

where the threshold parameter K tunes the edge-stopping sensitivity on the image gradient, and ν controls the length scale. In our work, ν replaces the curvature coefficient μ of equation (3) and plays the same role; that is, large values of ν translate into more denoising. The amount of nonlinear diffusion can be tuned with a single parameter (ν) in real time, allowing the user to decide on the appropriate value. Since the amount of noise in typical CT scans is small, it is easy to calibrate the preprocessing step once and for all, for example, by taking a large enough coefficient. Moreover, since the nonlinear diffusion algorithm preserves the edges, one need not tune the coefficient to produce the minimum amount of denoising necessary. Instead, one has the leeway of choosing a quite large coefficient, insuring the preprocessing will work in *all* CT scans encountered in practice. Therefore, we have used the same value of the smoothing coefficients in *all* of our test cases ($K = 7$ and $nu = 1$), leading to a *parameter-free* algorithm.

Consider a mesh superimposed on the image in such a way that each pixel corresponds to a grid node (x_i, y_j). Assume that the distance between two pixels in the x-direction is $\triangle x$ ($\triangle y$ in the y-direction) and assume a fixed time step $\triangle t$. The image is then processed by solving equation (5), with the following fully conservative semi-implicit numerical discretization:

$$I_{i,j}^{n+1} - \triangle t(\nabla \cdot (g^n \nabla I^{n+1}))_{i,j} = I_{i,j}^n,$$

where $I_{i,j}^n = I(x_i, y_j, n\triangle t)$. The equation is linearized by evaluating g at time n; that is, $g^n = g(|\nabla I^n|)$. We use standard central difference approximations for the derivatives; for example,

$$((gI_x)_x)_{i,j} = \frac{g_{i+1/2,j}(I_x)_{i+1/2,j} - g_{i-1/2,j}(I_x)_{i-1/2,j}}{\triangle x},$$

where

$$(I_x)_{i+1/2,j} = \frac{I_{i+1,j} - I_{i,j}}{\triangle x}, \quad (I_y)_{i+1/2,j} = \frac{I_{i+1,j+1} - I_{i+1,j-1} + I_{i,j+1} - I_{i,j-1}}{4\triangle y},$$

$$g_{i+1/2,j} = g\left(|\nabla I|_{i+1/2,j}\right), \quad |\nabla I| = \sqrt{I_x^2 + I_y^2},$$

and so forth. The resulting system of equations for the unknowns $I_{i,j}^{n+1}$ is both symmetric and linear and thus can be solved using a robust and fast iterative solver.

We use a preconditioned conjugate gradient method with a modified Cholesky preconditioner; see, for example, [31]. Moreover, this linear system need not be solved exactly but for only a few iterations until the residual is reasonably small (in our investigation we observed that only one iteration is needed). Indeed, it is well known that the conjugate gradient algorithm for solving a linear system can be related to a steepest descent minimization algorithm for optimization indicating that one can obtain useful results without iterating to convergence. We emphasize that only one iteration is sufficient as a preprocessing step to obtain the desired denoising effect.

After this preprocessing step, we want to solve equation (1) with the velocity S_n given by (3). As observed in [22], the preprocessing step allows one to ignore the third term in the right-hand side of equation (3). To evolve ϕ to the desired result one should replace the PDE (1) with the following ordinary differential equation (ODE):

$$\frac{d\phi}{dt} = (I - c_{out})^2 - (I - c_{in})^2, \tag{6}$$

where c_{in} and c_{out} are defined in equation (4) and change as ϕ evolves in time. Finally, since only the steady-state result is of interest, one can solve (6) with arbitrarily large time steps, ending with an algorithm that resembles k-Means [23].

The process described above can lead to the correct segmentation of the organ, except in regions where it is adjacent to another organ with similar average intensity values. For example, see the bladder in Figure 3 (right-hand image). Characteristic to this organ is its well-defined contour except in the region where it bounds the prostate. The strategy we adopt is to determine the boundary by carefully placing the wireframe in the region where the boundary of the bladder is not clearly characterized by jumps in the image intensity. Since this region occupies a small percentage of the organ and since its shape is not complex, it is easy and fast to place such a wireframe accurately in three spatial dimensions, as described above. We note that one needs only to focus on this region and that the rest of the wireframe is arbitrary (typically one will not change the shape of the original library organ). The presence of the user-placed wireframe defines a window that takes into account the fact that the organ to be segmented can bound others. In such a way, the segmentation will also be accurate in regions where the edge is not clearly defined. Standard level-set algorithms that do not use additional information on the target organ will typically fail to properly segment such images, because they cannot stop moving the level set when the boundaries are ill defined.

We note that the main advantage of seeking to separate the image with respect to its mean intensity value is that this process is parameter free after selection of the amount of smoothing necessary to deal with the presence of noise in the CT scans. Hence, it is unnecessary for the user to correctly guess parameters like those used in threshold algorithms, for instance, offering a *truly automatic* approach. Also, this algorithm is fast enough to allow the user to adjust the segmentation results if necessary, possibly assisted by postprocessing tools.

3. **Statistical evaluation of the results.**

3.1. κ-**statistics.** To evaluate our results, we use the κ-statistical test described in [24]. With this measure we can compare segmentations made in different ways while quantifying the similarity between different methods.

Assuming that we are interested in comparing an automatic segmentation and a manual segmentation, the κ-value is defined as

$$\kappa = \frac{\sum_{i=1}^{2} P_{i,i} - \sum_{i=1}^{2} P_{i,T} P_{T,i}}{1 - \sum_{i=1}^{2} P_{i,T} P_{T,i}}. \tag{7}$$

Here, $P_{1,1}$ represents the normalized number of voxels that are common between both segmentation methods in the volume of interest, and $P_{2,2}$ represents the number of voxels that are common between the areas that were not segmented by either methods. The values $P_{1,1}$ and $P_{2,2}$ represent the most frequently used index of agreement and are known as the "overall proportion of agreement." From these two values, we subtract the proportion of voxels that are included in the target volume by one segmentation method but are missing from the other ($P_{1,T}$ and $P_{T,1}$), as well as the proportion of voxels outside of the volume that appear in one segmentation method and are missing from the other ($P_{2,T}$ and $P_{T,2}$).

The resulting expression is normalized such that values of κ between 0 and 1 represent a measure of similarity between the volumes. A value of 1 indicates that the volumes are identical. Our scale has been used in the literature [24] and assumes an excellent correlation between the volumes when the κ-value lies between 0.80 and 1. A κ-value between 0.4 and 0.80 shows a fair to good correlation and a κ-value smaller than 0.4 demonstrates a poor correlation between the volumes. The κ-statistical test was integrated into the VolVisT framework.

3.2. **Over- and undersegmentation.** The κ-statistics described above encapsulates in a single parameter a measure of the segmentation success. To separate the slices that are within an acceptable range from those that will need post processing, we seek for each volume the number of slices that are over- or undersegmented. More precisely, consider a particular slice of a volume with its corresponding manual segmentation and the segmentation obtained with our approach. We define as *false positive* (FP) the number of pixels that are considered inside the organ by our approach but outside the organ by the manual segmentation. Likewise, we define as *false negative* (FN) the number of pixels that are considered outside the organ by our approach but inside the organ by the manual segmentation. Finally, we define as *true positive* (TP) the number of pixels that are considered inside the organ by both segmentations. The oversegmentation (S^+) and undersegmentation (S^-) are defined as follows:

$$S^+ = \frac{FP}{TP + FN}, \quad S^- = \frac{FN}{TP + FN}. \tag{8}$$

Then, for each volume, we introduce a threshold parameter T and identified how many slices are oversegmented (i.e., $S^+ \leq T$) and how many slices are undersegmented (i.e., $S^+ \leq T$) with respect to T.

4. **Results.** Examples of the results obtained with the segmentation of the three different organs are shown in Figures 4 through 6. In each of these figures, the outline contour is the result of our automatic segmentation.

We test our segmentation algorithm on 8 randomly chosen data sets: three kidneys, two rectums, and three bladders. Each data set contains thirty to eighty frames of CT scans of a given patient. For each data set, we compare the results of our algorithm to a manual segmentation by experienced radiation oncologists, which is taken as our gold standard. Using the two different measures described

above to compare the segmentations, we succeeded in determining the similarities and differences between both approaches.

The first comparison was made using the κ-statistics; the results are shown in Table 1. The κ-values of the first three data sets, corresponding to three different kidneys, are larger than .93, reflecting a very high similarity between the manual and automatic segmentations. For the rectum, in data sets 4 and 5, the κ-values are respectively .86 and .93, which again illustrates a high similarity between the different segmentation methods. The last three data sets in the table show the κ-value for three different bladders.

Although the κ-values in sets 6 and 8 are lower than others, they are still above .75 and thus are very close to the highest level of our scale.

TABLE 1. Results for the κ-statistics comparing our approach to manual segmentation

Set	Organ	κ-value
1	Kidney	0.94
2	Kidney	0.93
3	Kidney	0.94
4	Rectum	0.86
5	Rectum	0.93
6	Bladder	0.79
7	Bladder	0.90
8	Bladder	0.77

The κ-statistics are a global measure of similarity between different segmentations. However, they do not create a complete description of image comparison. For example, they provide no estimate of how many frames are oversegmented or undersegmented, which is considered more important in practice. In this case, since the manual segmentation provides a gold standard, we can obtain more detailed information about specific frames. Specifically, we measured the number of pixels that were identified as part of the segmented organ by our automatic approach and those that were not, and we compared them to the gold standard. This gives a handle on undersegmentation and oversegmentation and allows us to determine how many slices in one data set are acceptable and how many will need to be postprocessed.

The result of these measurements is illustrated in Figure 7. Each graph, which is associated with a specific data set, gives a percentage of slices that were over- and undersegmented for a given threshold (T) according to formulas (8). Graphs 1 through 3 represent the results obtained for the three kidneys. Less than 15% of the slices are oversegmented by more than 10%. For kidney 1, 77% of the slices are not undersegmented by more than 40%; for kidney 2, 84% of the slices are not undersegmented by more than 20%; for the kidney 3, 70% of the slices are not undersegmented by more than 30%. Graphs 4 through 5 illustrate the results for the two rectums. For rectum 4, 14% of the slices are oversegmented by more than 20%, but most of the slices are undersegmented. For rectum 5, 60% of the slices are not oversegmented by more than 40%, and 75% of the slices are not undersegmented by more than 40%. The last three graphs depict the results for the bladder. For bladder 6, 74% of the slices are not oversegmented by more than

20%, and 85% of the slices are not undersegmented by more than 50%; for bladder 7, none of the slices are oversegmented by more than 10% but most of the slices are undersegmented. Finally, for bladder 8, most of the slices are oversegmented, but 69% of the slices are undersegmented by no more than 30%.
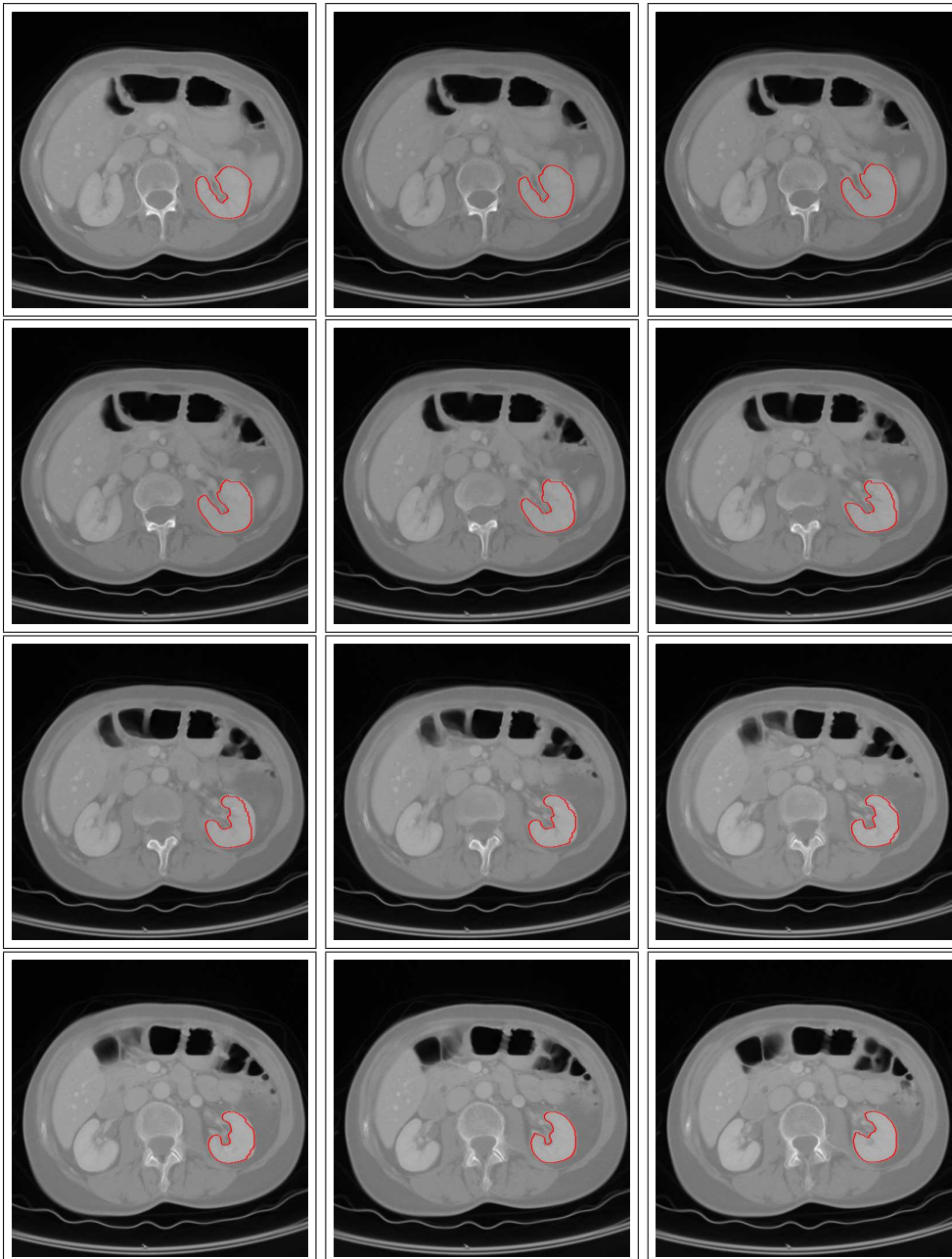

FIGURE 4. Segmentation results for the left kidney (data set 3). The initial wireframe is a cylinder englobing the organ.
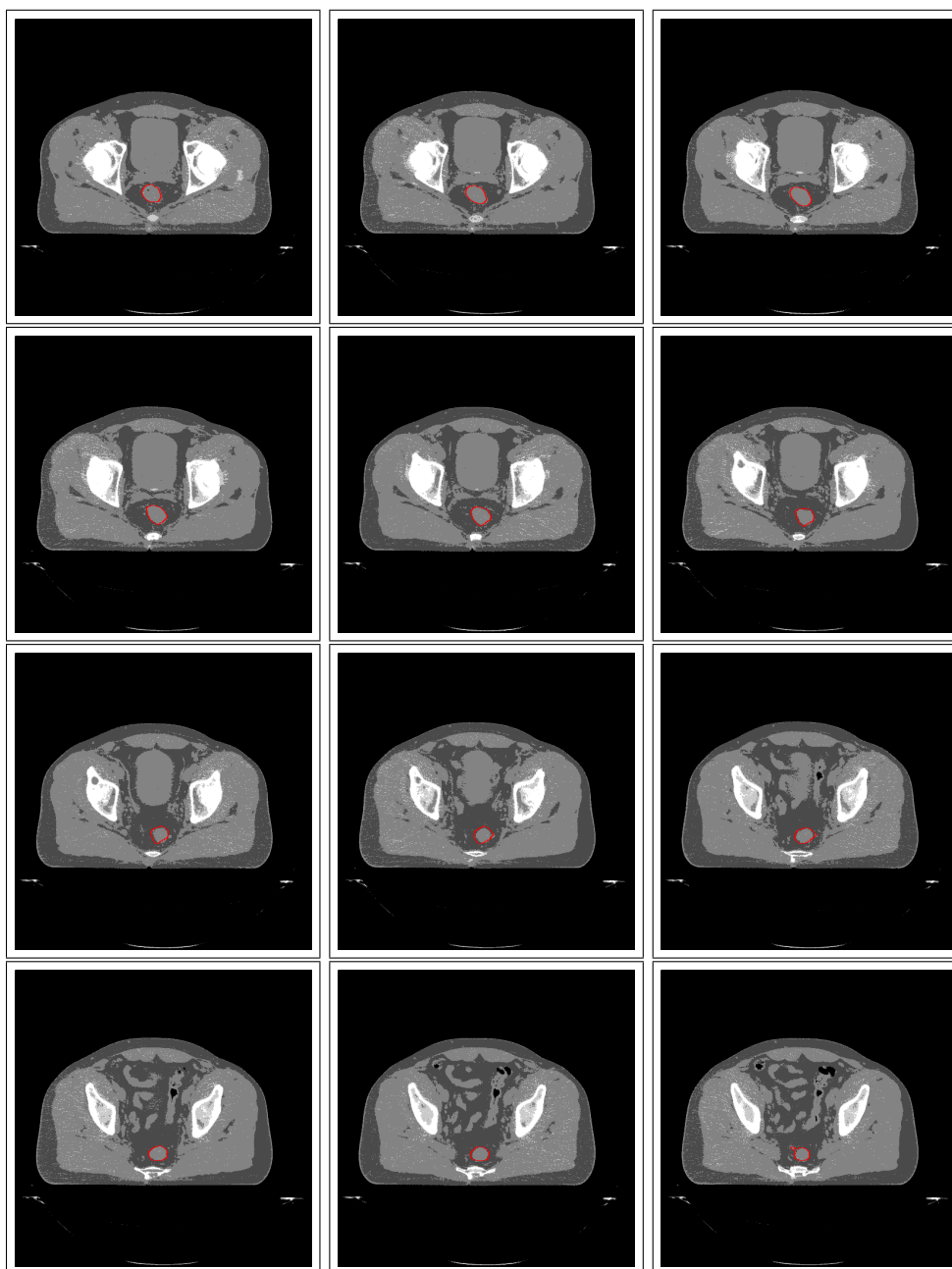
FIGURE 5. Segmentation results for the rectum (data set 4). The
initial wireframe is a cylinder englobing the organ.

5. **Conclusion.** In all test cases, the results obtained with this approach are
promising. We see a close similarity between the manual and the automatic seg-
mentations. The quantitative measurements of under- and oversegmentation show
that only in a few cases there is a need to postprocess the results produced by
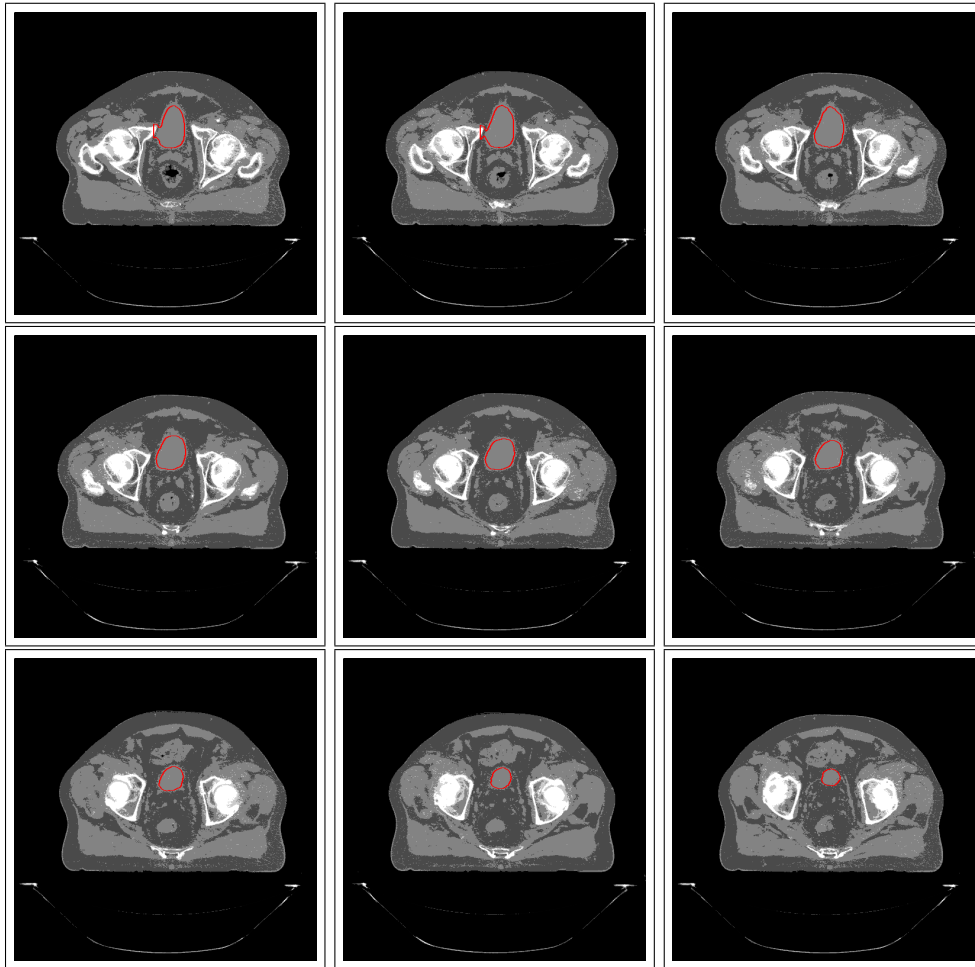
FIGURE 6. Segmentation results for the bladder (data set 7). The initial wireframe is a cylinder englobing the organ.

our algorithm. In Figure 8, we show few examples where our algorithm produces results that need to be modified. In some cases this can be done with standard image-processing tools. In other cases, we hope to extend the algorithm so that the results are improved.

The quality of the segmentation differs from organ to organ. For example, the results are very good for kidneys, whereas the bladders are consistently under-segmented. In our present approach we use the same algorithm (which is parameter free) for all organs. In the future we plan to modify the algorithm to integrate additional information about the organs. Such information could be used in the future to improve our algorithm. In addition, one expects to improve the algorithm by having different wireframe models for each organ (depending on different parameters, such as gender and age). Our current library of wireframe models is small, and we are in the process of building a more general database. With such an extended library, our approach will be more effective.
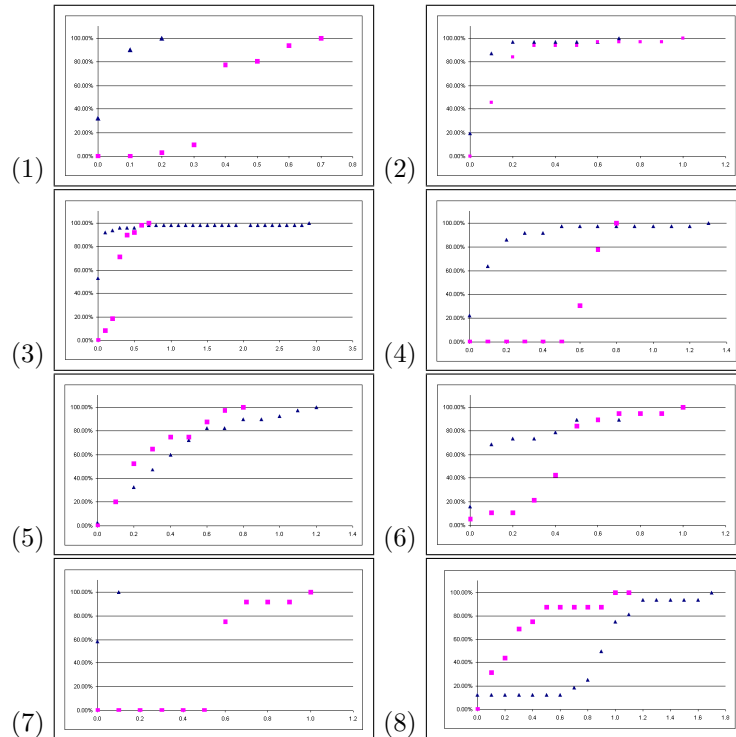
FIGURE 7. Comparison of the semi-automatic and manual segmentations for data sets 1–8. The y-axis represents the percentage of oversegmented (triangles) and undersegmented (squares) reconstruction planes as defined by equation (8). The x-axis represents the threshold parameter T in formulas (8).

In some cases, the results of the automatic segmentation differ from the manual segmentation because of the decision of the radiation oncologist to mark an area larger than the organ itself. At present, our algorithm does not take into account various topological constraints and safety margins that are part of the manual segmentation procedure. For example, in the case of the rectum, our approach tends to segment the internal wall (septum) of the rectum, while a manual segmentation will typically identify the external wall as the boundary of the organ.

Depending on the initial placement of the wireframe and the similarity with the data to be segmented, the placement and adaptation of the wireframe to the data set could become difficult. For example, the image in the first column and first row of Figure 4 is undersegmented due to a misplacement of the wireframe. Placing the wireframe requires expertise and training. Therefore, it would be desirable to have a variety of wireframe models in the repository and methods that will assist the user to place them correctly.

One of the greatest advantages of our algorithm compared to other existing algorithms is that the automatic part works in real time. A typical segmentation of an organ with sixty slices takes less than ten seconds on a Pentium IV laptop. This efficient implementation will allow us to add different constraints while keeping the process in real time, offering a practical tool for assisting the radiotherapist.
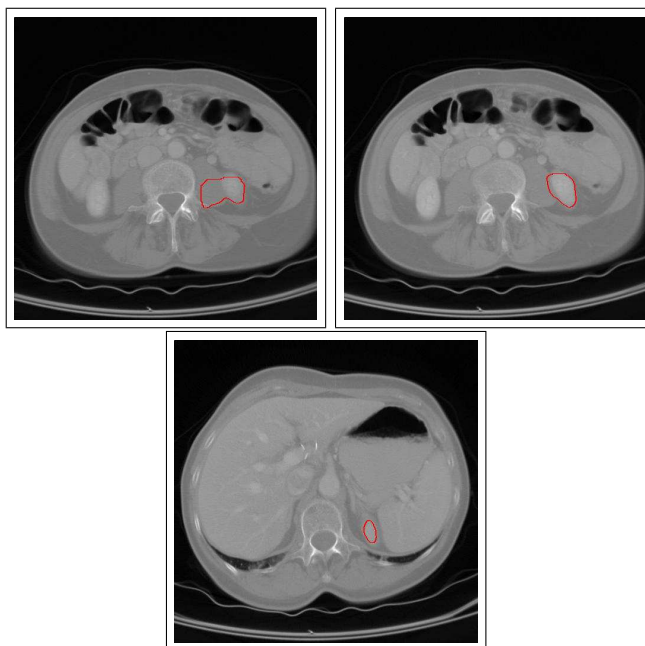
FIGURE 8. Illustration of some "pathological" cases of over- and undersegmentations. Top left: 280% of the organ is oversegmented and 6% is under-segmented. This is a typical example where the segmentation fails. Top right: 27% of the organ is over-segmented and 5% is undersegmented. Bottom: 0% of the organ is overseg- mented and 62% is undersegmented. Note that even in this case our segmentation occupies the core part of the organ and a simple dilation postprocessing would be sufficient to correct the underseg- mentation.

We attribute the pathological results to the convergence to the wrong local mini- mum. However, the algorithm is already quite successful since only a few slices will have to be "retouched" by the physician. Finally, we would like to stress that this is an initial study in which we examined the performance of our new algorithm on a limited number of data sets and without resorting to sophisticated postprocessing. An extensive study is required to evolve this algorithm to a clinical status.

REFERENCES

[1] Canny, J.F. (1986). A computational approach to edge detection. *Trans. Pattern Anal. Ma- chine Intel.* 8(1):679–98.
[2] Adams, R. and L. Bischof. (1984). Seeded region growing. *IEEE Trans. Pattern Anal. Ma- chine Intel.* 16(1):641–46.
[3] Pavlidis, T. (1982). Algorithms for graphics and image processing. Rockvill, MD: Computer Science Press.

[4] Zucker, S.W. (2001). Relaxation labeling: 25 years and still iterating. In: Foundations of Image Understanding, ed. L. S. Davis, 289–322. Boston: Kluwer Academic Pub.

[5] Terzopoulos, D., A. Witkin and M. Kass. (1988). Constraints on deformable models: recovering 3D shape and nonrigid motion. *Artificial Intelligence* 36(1):91–123.

[6] Terzopoulos, D., J. Platt, A. Barr, et al. (1987). Elastically deformable models. *SIGGRAPH*, ACM Press/ACM SIGGRAPH, Comp. Graphics Proc.; 205–14.

[7] Terzopoulos, D. and K. Fleischer. (1988). Deformable models. *The Visual Computer* 4(6):306–31.

[8] Kass, M., A. Witkin and D. Terzopoulos. (1987). Snakes: active contour models. *Int. J. Comp. Vision* 1(4):321–31.

[9] Cohen, L.D. (1991). On active contour models and balloons. *CVGIP: Image Understanding* 53(2):211–18.

[10] Caselles, V., F. Catte, T. Coll, et al. (1993). A geometric model for active contours in image processing. *Numerische Mathematik* 66:1–31.

[11] Malladi, R., J.A. Sethian and B.C. Vemuri. (1995). Shape modeling with front propagation: a level set approach. *IEEE Trans. Pattern Anal. Machine Intel.* 17(2):158–75.

[12] Osher, S. and J.A. Sethian. (1998). Front propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation. *J. Comput. Phys.* 79(1):12–49.

[13] Caselles, V., R. Kimmel and G. Sapiro. (1997). Geodesic active contours. *Int. J. Comp. Vision* 22(1):61–79.

[14] Kichenassamy, S., A. Kumar, P. Olver, et al. (1995). Gradient flows and geometric active contour models. *ICCV*.

[15] Paragios, N. and R. Deriche. (1998). Geodesic active regions for texture segmentation. INRIA Report no. 3440.

[16] Leventon, M., W. Grimson and O. Faugeras. (2000). Statistical shape influence in geodesic active contour. *Comp. Vision Patt. Recon.*.

[17] Sapiro, G. (2001). Geometric partial differential equations and image analysis. New York: Cambridge University Press.

[18] McInerney, T. and D. Terzopoulos. (1996). Deformable models in medical image analysis: a survey. *Med. Image Anal.* 1(2):91–108.

[19] Mumford, D. and J. Shah. (1989). Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. in Pure and Applied Math.* 42(1):577–685.

[20] Chan, T. and L. Vese. (2001). Active contour without edges. *IEEE Trans. Image Proc.* 10(2):266–77.

[21] Yezzi, A., A. Tsai and A. Willsky. (2002). A fully global approach to image segmentation via coupled curve evolution Equations. *J. Visual Comm. Image Rep.* 13:195–216.

[22] Gibou, F. and R. Fedkiw. (2005). A fast hybrid k-Means level set algorithm for segmentation. *Proc. Fourth Annual Hawaii Int. Conf. Stat. Math.* 281–91; and *Stanford Report* Nov. 2002,

[23] MacQueen, J. (1967) Some methods for classification and analysis of multivariate observation. *Proc. Fifth Berkeley Symp. Math. Stat. Prob.* 281–97.

[24] Fleiss, J.L. (1982). Statistical methods for rates and proportions. New York: John Wiley and Sons; 213–36.

[25] Tsitsiklis, J. (1995). Efficient algorithms for globally optimal trajectories. *IEEE Trans. Auto. Control* 40:1528–38.

[26] Sethian, J.A. (1996). A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci.* 93(4):1591–95.

[27] Jiang, G.S. and D. Peng. (2000). Weighted ENO schemes for Hamilton-Jacobi equation. *SIAM J. Sci. Comp.* 21:2126–43.

[28] Osher, S. and R. Fedkiw. (2002). Level set methods and dynamic implicit surfaces. New York: Springer-Verlag.

[29] Sethian, J.A. (1999). Level set methods and fast marching methods. 2nd ed. Cambridge, UK: Cambridge University Press.

[30] Perona, P. and J. Malik. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Patt. Anal. Machine Intel.* 12(1):629–39.

[31] Saad, Y. (1996). Iterative methods for sparse linear systems. Boston: PWS publishing company.

E-mail address: fgibou@engineering.ucsb.edu
E-mail address: dlevy@math.stanford.edu
E-mail address: Carlos.E.Cardenas@siemens.com
E-mail address: pliu@stanford.edu
E-mail address: boyer@reyes.stanford.edu