Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing 26 (2008) 141–163
www.elsevier.com/locate/imavis

Review

A comparative study of deformable contour methods on medical image segmentation

Lei He [a,*], Zhigang Peng [b], Bryan Everding [b], Xun Wang [b], Chia Y. Han [b], Kenneth L. Weiss [c], William G. Wee [b]

[a] *Information Technology Department, Armstrong Atlantic State University, 11935 Abercorn Street, Savannah, GA 31419, USA*
[b] *Electrical & Computer Engineering and Computer Science Department, University of Cincinnati, Cincinnati, OH 45221-0030, USA*
[c] *Department of Psychiatry, University of Cincinnati, Cincinnati, OH 45267-0559, USA*

Received 20 March 2004; received in revised form 15 July 2007; accepted 26 July 2007

Abstract

A comparative study to review eight different deformable contour methods (DCMs) of snakes and level set methods applied to the medical image segmentation is presented. These DCMs are now applied extensively in industrial and medical image applications. The segmentation task that is required for biomedical applications is usually not simple. Critical issues for any practical application of DCMs include complex procedures, multiple parameter selection, and sensitive initial contour location. Guidance on the usage of these methods will be helpful for users, especially those unfamiliar with DCMs, to select suitable approaches in different conditions. This study is to provide such guidance by addressing the critical considerations on a common image test set. The test set of selected images offers different and typical difficult problems encountered in biomedical image segmentation. The studied DCMs are compared using both qualitative and quantitative measures and the comparative results highlight both the strengths and limitations of these methods. The lessons learned from this medical segmentation comparison can also be translated to other image segmentation domains.

Keywords: Medical image segmentation; Deformable contour method; Snake; Level set; Comparative study

Contents

* Corresponding author. Tel.: +1 912 921 7360; fax: +1 912 921 7362.
E-mail address: helei@mail.armstrong.edu (L. He).

0262-8856/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2007.07.010

1. Introduction

Image segmentation is one of the first stages in many image analysis applications. In the domain of biomedical image processing, correct image segmentation would aid physicians greatly in providing visual means for inspection of anatomic structures, identification of disease and tracking of its progress, and even for surgical planning and simulation. The bulk of the earlier segmentation methods are roughly categorized, based on the image features used for segmentation, into two basic approaches, the edge-based and region-based methods [18]. The features used in one approach are usually complementary to the other in segmentation. Therefore, more recent methods, such as the deformable contour methods (DCMs), include the key concepts of both the edge-based and region-based approaches. The DCMs can be sub-categorized according to the various mechanisms used for carrying out the contour deformation process: snakes and level set methods. Each of the mechanisms used in these methods are meant to incorporate a means to handle specific problems and provide the desired features to effectively segment target contours from the image data.

To incorporate powerful, and useful, concepts of energy, force, velocity, and curve constraints, DCM-based solutions are highly analytical and involve extensive numerical computations. All these factors make the solutions less intuitive for the practitioners and hard to compare the methods in terms of their applicability and computational requirements. This paper reviews the strengths and limitations of eight major DCMs and reports on how these methods perform in terms of dealing with some of the toughest image processing challenges in biomedical images. Even though the focus of the comparison is on biomedical segmentation, the lessons learned from the comparison can be translated to other image segmentation domains. In what follows, we shall first provide an overview of the DCM methods before we get into the details of comparative study.

1.1. Snakes

The original snake was introduced by Kass et al. [1], in which the contour deforms to minimize the contour energy that includes the internal energy from the contour and the external energy from the image. Using a variational method, the internal and external energy are converted to internal and external forces to deform the contour. During the contour deformation process, the curvature-based internal force maintains the contour smoothness, while the gradient-based external force attracts the contour to the desired boundaries in the image. The deformation finally stops when the contour reaches an energy minimum (force balance). Note, the original snake was proposed as an interactive method, which requires expert guidance on the snake initialization and the selection of correct deformation parameters. It is important to understand several underlying concepts that identify limitations of the original snake method. First, the magnitude of the external force dies out rapidly when moving away from the image edges or boundaries. This implies that the capture range of the original snake is small. Secondly, image noise can cause the contour to be easily attracted to a local energy minimum, which does not correspond to the ground truth. Therefore, to reach the desired boundary, the initial contour should lie close to the desired boundary to avoid these hazards. Furthermore, the original snake method is a parametric

method and the contour cannot change topology during its deformation process without an additional mechanism. With these limitations, a number of deformable contour methods have been proposed to improve the original snake, such as snake variations [2–4,10,20,22,25,26,37,39,45–47].

In [2–4], different edge-based external forces are proposed to overcome the sensitivity of the initial condition in [1] by enhancing the effect of image edges. The methods in [10] and [37] provide different mechanisms to enable the contour topology to change during the deformation process. Besides the topological constraint, the author of [37] proposes algorithms to apply new physical constraints on the snake in order to control the contour geometry and deformation. The methods of [25,26] deform the contour with the constraint from a priori knowledge of the object shape, which helps the deformable contour avoid being trapped by spurious edges. In [20,22], the authors try to utilize region-based image features or combine them with the edge-based features as the external forces in order to overcome the image noise. However, [20] still requires the initial contour to be close to the desired boundary and it cannot handle the contour topological changes. Starting from multiple seeds, [22] performs image segmentation on the whole image by doing boundary deformation and region merging iteratively; however, it cannot handle contour splitting topology change. In order to relieve the sensitivity to initialization and accurately locate the global minimum, dynamic programming approach was applied in [45–47] to replace the variational method to minimize the contour energy. These approaches also have the advantage of avoiding the estimation of higher order derivatives and improve the numerical stability.

1.2. Level set methods

Different from the snake energy-based framework, the original level set method [5] was first proposed as a numerical technique that tracks an evolving contour. The evolving contour deforms with a speed F that is based on the contour curvature and image features like gradient. The curvature component in the speed keeps the contour propagating smoothly, which performs like the internal energy in snakes. Additionally, an artificial speed term, obtained from the image feature, is synthesized to stop the front, i.e., the contour, at the desired boundary. After the original level set method [5] was proposed, which had no energy minimization, some researchers applied the level set formulation with a contour energy minimization in order to obtain a good convergence. This modification produced a relaxed initial condition requirement and the ability to handle the contour topology changes naturally. The modification resulted in various geodesic deformable models [8,9,11,19,23,24,27–29,32,33].

Starting with [8,27], the authors prove that the minimization of the contour energy is equivalent to the minimization of the contour length weighted by an edge detection function in the Riemannian space. In [28], a review and comparison on the level set methods in [5,8,27] is given and a new geodesic active contour method for multiple objects segmentation is introduced. However, the multiple objects can only be segmented if the initial level sets are close to the objects. The method in [29] tries to locate the global energy minimum of a contour between two end points that are pre-selected on the true boundary. The methods in [32,33,19] integrate the a priori knowledge of the object shape into the level set formulation in order to constrain the contour deformation within an admissible range. In [11,23,24,9], region-based image features are used alone or united with the edge-based features to construct the energy to minimize. Similar to the edge-based methods, the energy minimization is implemented in a level set framework. The method in [23] handles image segmentation using only region-based image features with the assumption that only object and background exist in the image. In [24], the authors focus on supervised texture segmentation and require a priori knowledge of the object texture pattern. In [9], the contour energy minimization is formulated as a constrained optimization problem with a constraint indicating the degree of contour interior homogeneity.

The selected approaches for the study are chosen from the two subsections of deformable contour methods, snakes (balloon snake [2], topology snake [10], distance snake [3], and gradient vector flow snake [4]) and level set methods (original level set [5], geodesic active contour [8], area and length active contour [11], and a constrained optimization [9]). We did not consider other methods incorporating with the a priori knowledge of object shape [15–17,21], random process [12–14], or human interaction [48,49] in the contour deformation processes because they are generally application specific and cannot be compared fairly. Although several surveys on deformable contour methods have been published [18,28,30,31,36,38], to our knowledge, none of them focus on comparing them experimentally, specifically in the area of medical image segmentation. In contrast with other surveys that only describe and compare different approaches qualitatively, this study provides both qualitative and quantitative comparisons. In addition, as a starting point, the lessons learned from this medical segmentation comparison can also be generalized to other image segmentation domains. Our goal is to provide a common platform to compare the deformable contour methods that can be applied automatically in a set of medical images, with the motivation coming from the practical needs of medical users. The remainder of this paper is structured as follows: Section 2 briefly remarks on the variations and nuances of the selected methods involved in the experimental comparisons. Section 3 is an overview of the experimental plan and procedure while Section 4 provides the experimental results and evaluations. Finally, the conclusion of the experiments is drawn in Section 5.

2. Deformable contour methods for comparison

This section briefly introduces the selected set of deformable contour methods, which includes balloon snake, topology snake, distance snake, gradient vector flow snake, original level set, geodesic active contour, area and length active contour, and a constrained optimization method.

2.1. Balloon snake

The original snake [1] is a deformable contour based on minimizing the energy functional

$$E(v) = \int w_1|v'(s)|^2 + w_2|v''(s)|^2 + Q(v(s))\,\mathrm{d}s. \tag{1}$$

In Eq. (1), $v(s) = (x(s), y(s))$ represents the contour in an image $I(x, y)$ with x and y being coordinates, and $s \in [0, 1]$ is a generic parametric coordinate. The component $\int w_1|v'(s)|^2 + w_2|v''(s)|^2\,\mathrm{d}s$ is the internal energy of the contour with w_1 and w_2 being the contour elasticity and rigidity weights. The component $\int Q(v(s))\,\mathrm{d}s$ is the external energy of the contour with $Q(v(s)) = -|\nabla P(v)|^2$, and $P = G_{\sigma_0} * I$ being the Gaussian (variance σ_0) filtered image of the input image I. Using the variational method, it can be proven that the contour satisfies the following Euler–Lagrange equation if E in Eq. (1) reaches a local minimum.

$$-(w_1v')' + (w_2v'')'' = F(v), \quad v(0), v'(0), v(1), v'(1) \text{ being given.} \tag{2}$$

In Eq. (2), $-(w_1v')' + (w_2v'')''$ is the internal force and $F(v)$ is the external force with $F = -\nabla Q(v)$. Since the energy functional is not convex, suppose the initial contour is close to the desired boundary (i.e., a local minimum of energy), the contour evolution process can then be described as

$$\begin{cases} \frac{\partial v}{\partial t} = (w_1v')' - (w_2v'')'' + F(v) \\ v(0,s) = v_0(s), \quad v(t,0) = v_0(0), \quad v(t,1) = v_0(1), \quad v'(t,0) = v_0'(0), \quad v'(t,1) = v_0'(1) \end{cases}, \tag{3}$$

where the contour $v(s) = (x(s), y(s))$ is considered to be a function of time written as $v(s, t) = (x(s, t), y(s, t))$. When $\frac{\partial v}{\partial t} = 0$, contour reaches the desired boundary and stops deformation. Balloon snake [2] enables the initial contour to be located far from the desired boundary by adding a constant force, in the external force, to inflate the contour's growth. The balloon snake uses the modified force function of

$$F(v) = k_1\vec{N} - k\frac{\nabla Q}{|\nabla Q|}. \tag{4}$$

In Eq. (4), $\vec{N}$ is the normal unit vector with magnitude k_1 (inflation force), and k is the external force weight.

2.2. Topology snake

To overcome the topology problem of snakes, McInerney and Terzopoulos [10] designed a set of topology changing rules to be used during the balloon snake's deformation. Using the simplicial decomposition, the image space is partitioned into a set of triangular cells. The snake is then approximated as a polygon with the vertices being the snake contour points on the edges of the triangles. The triangles can be classified as inside, outside, or intersecting with the contour according to their positions relative to the contour. The triangles intersecting with the contour can be used to trace the contour and check the contour topology conveniently during contour deformations. Once contour crossings exist, either the contour can be split or contour segments can be merged together by reconnecting and disconnecting the contour points on the triangles.

2.3. Distance snake

Cohen and Cohen [3] used a finite element method to implement a deformation strategy called the distance snake. Compared with the original snake, the external force field on the image is constructed also as the negative of the external energy gradient, which is the distance from each point to its closest edge points in the image. The new external energy enables a large magnitude for the external force everywhere in the image. Thus, the distance snake has a large capture range, i.e., the initial contour can be located far away from the desired boundary if there are no spurious edges along the way. By using a finite element method, the deformable contour is represented as a continuous curve in the form of weighted sum of local polynomial basis functions. The result has good stability and convergence in the energy minimization process. The distance snake uses the external force function of

$$F = -\nabla Q(v). \tag{5}$$

In Eq. (5), $Q(v) = d(v)$ and $d(v)$ is the smallest normalized Euclidean distance from v to an edge point with the edge point identified by a thresholded gradient.

2.4. Gradient vector flow snake

In [4], Xu and Prince discussed the shortcomings of the original snake and distance snake from the external force field construction. The external force field, for the original and distance snake, is irrotational and based on the contour points and the closest edge points in the contour points' normal direction. This limits the deformation into boundary concavities because there is no external force pointing into the concavities inside. GVF snake constructs a new external force field which is not entirely irrotational, which means the new external force points inside in concavities. Additionally, the magnitudes of the external force are the same over the whole image (field), which means a large capture range for the gradient vector flow (GVF) snake.

The GVF snake has a new external force: $F(v) = (\alpha(x, y), \beta(x, y))$ and $F(v)$ can be obtained by minimizing the energy functional

$$E = \int\int \mu(\alpha_x^2 + \alpha_y^2 + \beta_x^2 + \beta_y^2) + |\nabla f|^2 |F - \nabla f|^2 \, dx \, dy, \quad (6)$$

where f is an edge map of the input image I, and μ is a regularization parameter. Using the variational method, F can be found by solving the following Euler equations

$$\begin{cases} \mu\nabla^2\alpha - (\alpha - f_x)(f_x^2 + f_y^2) = 0 \\ \mu\nabla^2\beta - (\beta - f_y)(f_x^2 + f_y^2) = 0 \end{cases}. \quad (7)$$

In Eq. (7), ∇^2 is the Laplacian operator, $f_x = \frac{\partial f}{\partial x}, f_y = \frac{\partial f}{\partial y}$.

2.5. Original level set

Level set is another important category of deformable contour methods and Malladi and Sethian [5] were the first to propose this method for image segmentation. It originates from the theory of propagating solid/liquid interface (front) with curvature dependent speed. The central idea is to represent a curve as the zero level set of a higher dimensional function (3D surface) with the motion of the curve embedded in the motion of the higher dimensional surface. The motion of the front is matched with the zero level set of a signed distance function, and the resulting partial differential equation for the evolution of the level set function resembles a Hamilton–Jacobi equation. This equation is solved using entropy-satisfying schemes borrowed from the numerical solution of hyperbolic conservation laws, which enable the topological changes, corner and cusp development to be naturally obtained during the front marching process. In the level set scheme, the contour deforms by the function

$$\frac{\partial\psi}{\partial t} = \Gamma|\nabla\psi|, \quad \psi(x, y, t = 0) = \pm d. \quad (8)$$

In Eq. (8), Γ is the contour marching velocity, $\Gamma = g(|G_{\sigma_0} * I|)(c + \kappa)$, where $g(|G_{\sigma_0} * I|) = \frac{1}{1+|\nabla G_{\sigma_0} * I|^2}$, $c > 0$ is a constant, κ is the contour curvature, d is the distance from (x, y) to the initial contour, and ψ is positive or negative when (x, y) is either outside or inside the initial contour.

To speed up the front marching, computationally efficient schemes like the narrow-band method and the fast marching method are proposed in [34,6,7]. The narrow-band method modifies the level set method so that it affects points close to the front, i.e., the points within the narrow-band, instead of the all the points in the image. The fast marching method solves the general static Hamilton–Jacobi equation, which applies in the case of a convex, non-negative function. Starting with an initial position for the front, the method systematically marches only one point of the front outwards per time step instead of the whole front as in the original level set method and the narrow-band method. The contour deforms as

$$\frac{\partial v}{\partial t} = \Gamma\vec{N} \quad (9)$$

with $\vec{N}$ being the contour normal unit vector and Γ being the non-negative velocity function. The fast marching method [7] uses a velocity function as a decreasing function of image gradient (ex. $\Gamma = e^{-\alpha|\nabla G_{\sigma_0} * I|}$, $\alpha > 0$) in order to reach an approximate boundary result.

2.6. Geodesic active contour

Based on the above level set methods, a new level set method called geodesic active contour method, was proposed in [8] which unified the curve evolution approaches with the classical energy minimization methods (snakes). It was proven that the minimization of a simplified contour energy with no second order term in Eq. (1) is equivalent to the minimization of the contour length weighted by an edge detection function in Riemannian space. The flow of the contour can be described as

$$\frac{\partial v}{\partial t} = g(|G_{\sigma_0} * I|)(c + \kappa)\vec{N} - (\nabla g \cdot \vec{N})\vec{N} \quad (10)$$

with $g(|G_{\sigma_0} * I|) = \frac{1}{1+|\nabla G_{\sigma_0} * I|^2}$, c is a constant ($c > 0$) for fast convergence, and κ is the curvature. The embedding surface deformation process using level set implementation is

$$\frac{\partial\psi}{\partial t} = g(|G_\sigma * I|)(c + \kappa)|\nabla\psi| + \nabla g \cdot \nabla\psi \quad (11)$$

with $g(\cdot)$, κ, and c being defined as above.

2.7. Area and length active contour

Siddiqi et al. [11] added a new area function weighted by the edge detection function to the length minimization framework of the geodesic active contour method. This addition strengthened the force attracting the contour to the feature of interest. With this new weighted contour length and contour interior area minimization framework, the deformable contour converges to the desired boundary with a faster speed than the geodesic active contour method. Actually the weighted length minimization flow can be replaced by the weighted area minimization in many cases, which offers significant computational savings. The contour deformation in [11] can be described as

$$\frac{\partial v}{\partial t} = \left\{ g(|G_{\sigma_0} * I|)(c + \kappa) - (\nabla g \cdot \vec{N}) + (g(|G_{\sigma_0} * I|) + \frac{1}{2}\langle v, \nabla g\rangle) \right\}\vec{N}. \quad (12)$$

In the level set marching scheme, the embedding surface deforms as

$$\frac{\partial\psi}{\partial t} = \alpha\{g(|G_{\sigma_0} * I|)|\nabla\psi|(c + \kappa) + \nabla g \cdot \nabla\psi\} + \left\{\frac{1}{2}\mathrm{div}\left[\begin{pmatrix} x \\ y \end{pmatrix} g\right]|\nabla\psi|\right\} \quad (13)$$

with $\mathrm{div}\left[\begin{pmatrix} x \\ y \end{pmatrix} g\right] = 2g + (xg_x + yg_y)$, and α is a constant ($\alpha > 0$).

2.8. Constrained optimization

The constrained optimization method [9] uses a velocity function to deform the contour and an energy function to constrain the deformation. The velocity function is defined as

$$\frac{\partial v}{\partial t} = \left[\frac{1}{1+|\nabla G_{\sigma_0} * I|^2} e^{-\frac{|M(x,y)-I_0|}{\sigma}}(1-\tau N(0,1)) - T_V\right]\vec{N}. \tag{14}$$

Here $\tau = Z_1 e^{-Z_1}$, $0 \leqslant \tau \leqslant 1$, $Z_1 = \frac{|I(x,y)-M(x,y)|}{\sigma_1}$, where $\sigma_1 = 2\sigma$, and σ is a constant ($\sigma > 0$). T_V is a preset threshold for similarity measure, $M(x,y)$ is the average of a 3×3 neighborhood around the contour point (x,y), $N(0,1)$ is a Gaussian perturbation with variance 1 and I_0 is the average contour interior brightness. The contour energy is calculated by $E(v) = \overline{\frac{1}{1+|\nabla G_{\sigma_0} * I|^2}}$, where $\overline{(\cdot)}$ is the average of $(\cdot)$. The procedure uses a modified fast marching method with the interior area being updated and a contour being selected from the lower E from several nearby randomly generated contours at each iteration. The process stops when the input number of iterations is reached and the output is the contour with the smallest energy during the whole deformation process.

3. Experimental setup

The objective of the experiments is to provide insight about both the strength and weakness of the various DCMs when used for medical image segmentation applications. Section 3.1 presents the test set and the typical challenges posed by these images. Section 3.2 presents the evaluation scheme and Section 3.3 covers the implementation details of the parameter settings and tuning, and two other initialization experiments.

3.1. Image test set

A small set of seven 2D images is selected to serve as the representative test set, as shown in Fig. 1. Each test image was one of several similar images or image slices in our database and the experiments were comprehensive within the database. The selection includes representation of MRI, ultrasound, CT, and optical images. The biomedical content includes cell, brain, knee, abdomen and heart at the different standard cross-sectional views. The test images are at different sizes and resolutions. The gray levels are the same for all the test images, i.e., 0–255. Included are also images with added noises (Gaussian, salt and pepper) corrupting the overall quality of the images. Although these types of noise are normally modality dependent or not present due to the high quality of today's imaging devices, it is still interesting to show the performance of the methods in the presence of different types of noise.

The typical difficulties of image processing presented in the test set include: blur or weak edge, strong edge near the missing edge, profile contour in overlapping objects, complex contour shape with accentuated protrusions and concavities, inhomogeneous interior intensity distribution. These are some of the typical challenges that would fail any simple segmentation schemes.

The test set is shown in Fig. 1a–g and the image descriptions are presented in Table 1. Each image contains some bright or dark marks indicating the locations of the reference point for the initial contours used in the experiments. The regions of interest (ROI) in the images are designated by arrows. The "gold standard" segmentations shown beside the test images were selected pixel-by-pixel by a board-certified radiologist (the 6th author, K.L. Weiss), which were used as the expert contours in the comparison with the DCM results. Without the information of the corresponding segmentation methods, the same radiologist subjectively ranked the computer generated results to provide a qualitative "goodness" assessment.

3.2. Evaluation measures

The evaluation of the DCMs is based on two quantitative error measures, ε_1 and ε_2 [35], and one qualitative "goodness" measure (the subjective ranking by the radiologist). The quantitative measure ε_1 is defined as $\varepsilon_1 = 1 - \frac{|TP \cap EP|}{|TP \cup EP|}$ and the ε_2 is defined as $\varepsilon_2 = \max_{X \in TC} \min_{Y \in EC} \mathrm{dist}(X,Y)$. TC denotes the set of pixels belonging to the true boundary and the set of pixels belonging to the extracted boundary is denoted as EC. TP denotes the set of pixels inside the TC and likewise, the set of pixels inside the EC is denoted by EP. The dist(X,Y) function in ε_2 represents the Euclidean distance between the integer coordinates of the pixels X and Y. ε_1 quantifies the overlap of the EC and the TC and it equals to zero if the two contours are identical (complete overlap). ε_1 provides an indicator of the overall goodness of the result, thus a global evaluation measure. ε_2 measures the maximum Euclidean distance between the EC and the TC. It is a local measure, which is useful in determining if the high curvature portions of the contour are extracted.

3.3. Implementation and experimental settings

The implementation of the various methods starts with the identification of all the adjustable parameters for each method. A computer program, written in either C or Matlab, is developed for each method.[1] To speed up the experiments, the original (ex. 256×256) images were cropped to be smaller images (ex. 71×71), with the object of interest in the middle of the trimmed images. The selected methods have all the following common simple characteristics: A

[1] Snake methods are based on the Matlab code from http://iacl.ece.jhu.edu/projects/gvf/ and level set methods are developed in-house with the support from the authors [8,11].

Gaussian (7×7, variance 1) blurring filter is the only preprocessing performed on the image and no post-processing is used; the gradient magnitude images or their thresholded results of the Gaussian blurred images are used as the edge maps; the initial contour is formed as a circle or square (e.g., 3×3 contour) centered around the initial location, a selected image point defined by the user; no a priori information of the object shape or texture pattern is available. The parameters, within each group of deformable contour methods, are very similar and are described in the next subsection. The following subsection discusses the parameter setting and tuning process for each method.

3.3.1. Parameters

The snakes have five parameters in common: elasticity (E), rigidity (R), viscosity (V), external force (X), and deformation step (DS). The E and R parameters are associated with the internal force variables w_1 and w_2 in the original snake model Eq. (1). The V and DS parameters are used in iteratively updating contour location (i.e., deformation), thus not explicitly included in the deformation equations. The combination of the E and R parameters allows the contour to maintain smoothness during the deformation process. Decreasing E or R will result in corners and self-intersections in the deforming contour, while increasing them too much will shrink the contour to a line or point. The V parameter is a weight parameter to adjust the "viscosity" used in the updating of the contour position. Increasing V will slow down the contour deformation process (i.e., a large number of DS is then required to finish the process) and make the deformation more stable. In the experiments, the V was limited to be smaller than 20 to avoid a very large DS number. The X parameter is a weighting variable applied to the external force and determines the strength of the effect

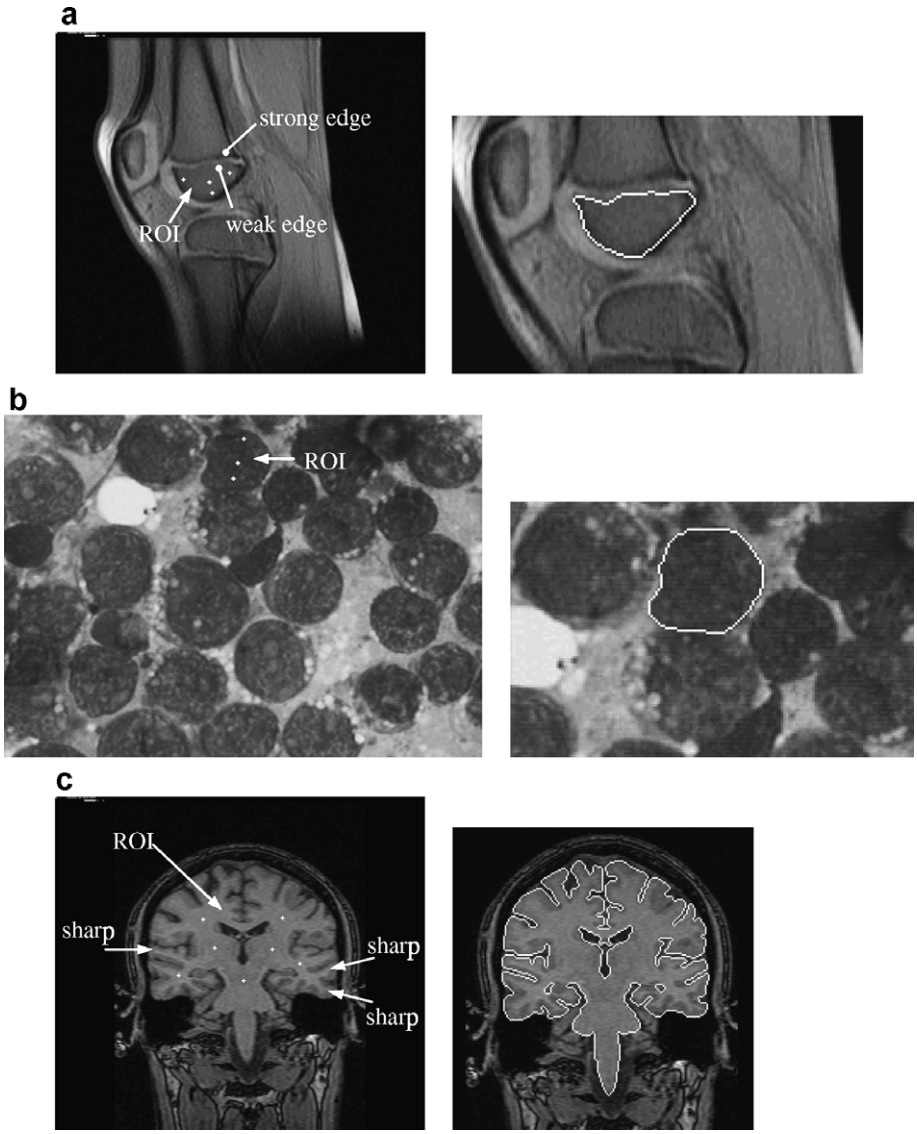

Fig. 1. (a) MRI knee image and the expert contour. (b) Light microscopy blood cell image and the expert contour. (c) MRI brain image to extract the brain–CSF interface and the expert contour. (d) MRI brain image to extract the corpus callosum and the expert contour. (e) Ultrasound pig heart image to extract left ventricular inner contour (blood volume) and the expert contour. (f) CT kidney image with salt and pepper noise (10%). (g) CT kidney image with Gaussian noise (variance 3000) and the kidney expert contour.

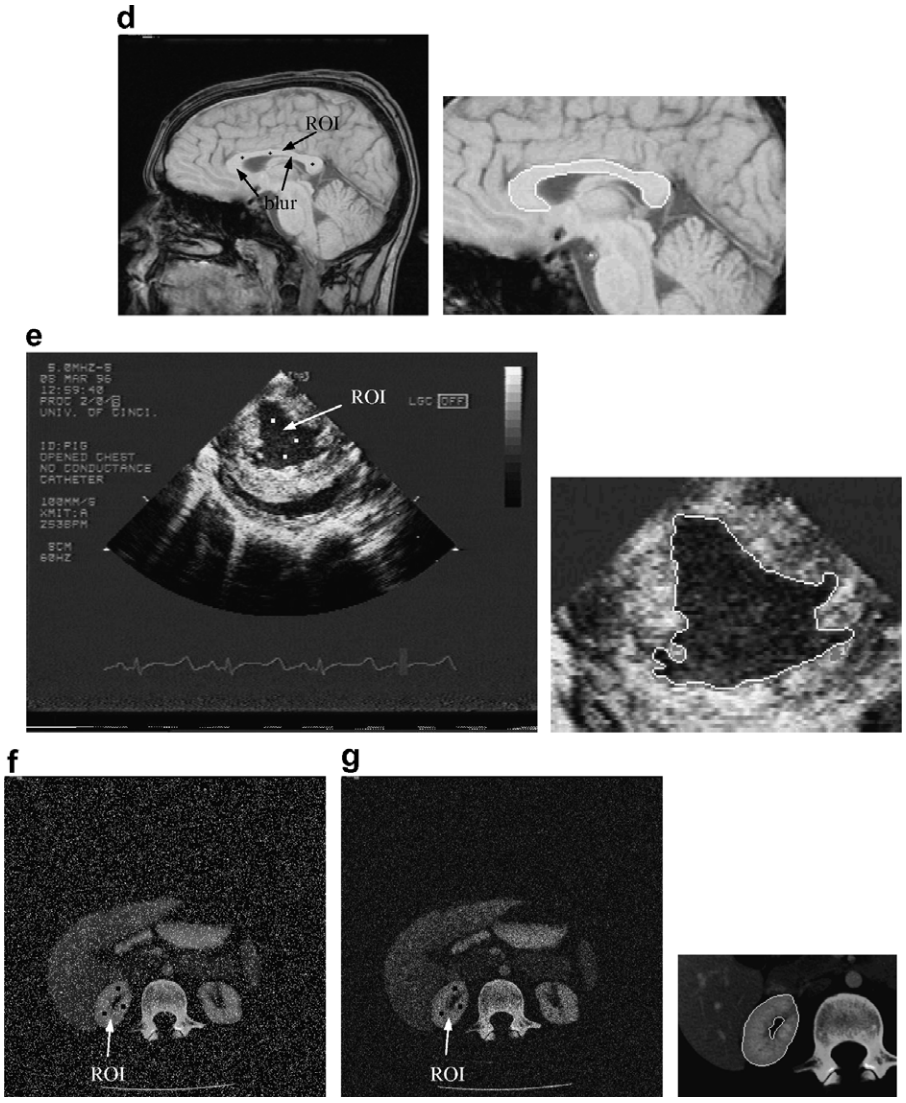

Fig. 1 (*continued*)

of the image features that make up the external force. For example, k in the balloon snake Eq. (4) is a X parameter. The DS parameter is the maximum number of contour deformations that is permitted.

The balloon and topology snake have an added parameter, K_1 inflation parameter, which is an additive constant to the external force. The K_1 parameter, k_1 in Eq. (4), allows these two snakes to march over spurious noise points and its value is usually larger than the external force weight. For the distance snake, the GT (gradient threshold) parameter is a threshold that is applied to the gradient magnitude of the input image to generate binary edge map. Increasing GT value will keep only strong edge points on the edge map. The binary image is used in calculating the image features for the distance snake external force. The GVF snake has two additional parameters, GVF N (iteration) and GVF U (regularization). The GVF N is an internal loop counter used to calculate the external force field (α, β) in the image, Eq. (7). Generally it is proportional to the image size, i.e., a large image size requires a large GVF N. The GVF regularization parameter, μ in Eq. (7), has a correlation with the noise level of the image. The higher the image noise the larger μ should be.

The level set methods have three common parameters and they are: iteration step (IS), inflation factor (IF), and deformation step (DS). The IS is related to the discrete implementation of the level set contour deformation process and has a value less than one. Decreasing IS will result in a slow deformation process but a more stable deformation. In the experiments, the IS was fixed to be 0.01 for all methods. Thus it was not considered as a parameter in the comparison. The IF is an additive variable, c in Eqs. (8), (10)–(13), to the image feature used in calculating the velocity, which plays as the ''balloon force'' in the balloon snake. Increasing IF can speed the deformation process and march the contour over the weak edges. The DS is the number of iterations or deformation steps permitted. The level set methods are not sensitive to the parameters IF and DS, i.e., they have very large ranges (e.g., 100–10,000, see Table 3) for the parameters to obtain acceptable results. If the IF increases

Table 1
Test image descriptions

Image and description	Size	Segmentation goal	Challenges
Fig. 1a: A midline sagittal MRI knee image	256×256	The femoral condyle (top portion of the knee), i.e., to delineate the top segment of the contour that separates the semicircular portion of the femur from the stem	There is a weak edge segment (indicated by the solid circle on a line labeled with "weak edge") along the middle top boundary and next to a strong spurious edge (indicated by the solid circle on a line labeled with "strong edge"), while the left and right portions of the femoral condyle are rather darker than the middle region. This prevents the deformable contour to reach the real boundary on the two sides before it flows out from the top
Fig. 1b: A light microscopy blood cell image	320×240	The cell(s) boundary	Large range of contrast in the interior regions of multiple overlapping target objects
Fig. 1c: A coronal MRI brain image	256×256	The sulci contour, i.e., the interface between brain substances (composed of gray and white matters) and CSF	Complex contour shape with deep concavities and sharp protrusions (three examples are indicated by arrows labeled with "sharp"), and inhomogeneous interior
Fig. 1d: A midline sagittal MRI brain image	256×256	The corpus callosum	The very blur contour segments at the lower left corner and middle bottom region, as indicated by the arrows labeled with "blur"
Fig. 1e: An echocardiographic image of a pig heart	640×480	The pig heart	Complex shape, gaps (missing edges along the contour due to signal drop out), and non-homogeneous interior, which are typical for cross-sectional left ventricular (LV) echo images
Fig. 1f: An axial CT image of the abdomen including cross-sections of stomach, kidney, and vertebra	512×512	The kidney	Salt and pepper noise (10%)
Fig. 1g: Same as Fig. 1f.	512×512	The kidney	Gaussian noise of variance 3000

too much, the deformable contour may not stable. Therefore, in the experiments, the IF was selected to avoid a very fast or a very slow deformation process.

The AF (area–length factor), used in the area and length active contour, is a variable used to weight the area and length functionals, α in Eq. (13). In practice, Dr. Siddiqi recommended that the area component should be scaled comparably to the length component and therefore a constant scaling factor was incorporated into Eq. (13) before the area component. Additionally, an inflation factor (IF) was added to increase the deformation speed. The constrained optimization method uses a SM (interior smoothness) parameter, σ in Eq. (14), with a value that is related to the smoothness of the interior of the desired contour boundary. A larger variance of the desired boundary interior requires a larger SM parameter value.

Tables 2 and 3 outline the effects of the parameters for the snake and level set methods. The tables also list the value ranges of the parameters used to process the test set images. The process to determine the parameter value ranges is described in Section 3.3.2, Step 2. For each parameter, its value range has some variations for different images. Therefore, for each parameter, the union of the ranges for different testing images is used as an approximation of the value range in the tables. As indicated in Section 3.1, each test image serves as a representative sample for a small group of similar images in our database. After repeating the parameter tuning process to other images in each group, comparable results are obtained, indicating the robustness of our parameter tuning scheme.

3.3.2. Parameter setting and tuning

Once the parameters were identified and encoded, the next question is how to determine and adjust the parameter values to obtain good results with small errors compared with the expert contours. In practice, for a new set of images with similar characteristics, some preliminary training and testing are normally required. In the experiments, we used a coarse to fine scheme on the image test set to achieve the "best"[2] parameter set based on both the qualitative and quantitative error measures. The following steps outline the process used to determine parameter settings and tuning of the deformable contour methods under consideration. Even though the deformation step (DS) is considered a parameter, it could not be varied independently and it was adjusted accordingly with the variations of the

[2] In practice, it is generally a range of values, not a single value, for each parameter that derive the results close to the expert contour, thus any parameter values from the range can be considered the "best", or near best. The middle value in the range was chosen in our experiments for consistency in determining the range for each parameter.

Table 2
Snake method parameter effects and value ranges

Parameters	↓	↑	Balloon/topology	Distance	GVF
E, Elasticity	Contour develops corners	Contour shrinks to a line or point	0–0.5	0–2.77	0–11.59
R, Rigidity	Develop corners	Contour shrinks to a line or point	0–1.88	0–13.93	0–24
V, Viscosity	Faster deformation, less smooth contour	Slower deformation, smoother contour	0–20.0	0.12–10.0	0.3–15.0
X, Image Feature Strength	Reduce image edge effects, contour cross over weak edges	Increase image edge effects, stops contour at edges	1.2–25.0	0.1–8.56	1.02–16
K_1, Inflation Force	Slow deformation	Speed deformation	0.08–0.45	N/A	N/A
DS, Deformation Steps	Shorter deformation	Longer deformation	50–12,000	30–900	10–400
GVF N, Iteration Number	For low image noise	For large image noise	N/A	N/A	80–200
GVF U, Regularization Factor	For low image noise	For large image noise	N/A	N/A	0.02–0.21
GT, Image Gradient Threshold	Keep weak edges	Keep strong edges	N/A	0.02–0.18	N/A

Table 3
Level set method parameter effects and value ranges

Parameters	↓	↑	Level set	Geodesic	Area/length	Constrained opt.
IF, Inflation Factor	Slower deformation	Faster deformation cross weak edges	2–20	4–20	4–20	N/A
DS, Deformation Steps	Shorter deformation	Longer deformation	100–15,000	500–20,000	600–20,000	600–7000
AF, Area Length Factor	Smaller length factor	Larger length factor	N/A	N/A	1	N/A
SM, Interior Smoothness	For smaller interior variance	For larger interior variance	N/A	N/A	N/A	4–20

other parameters during the parameter setting and tuning stages.

(1) Initial parameter setting: The initial parameter settings were obtained by manually selecting the parameters, based on experience, and iterating the process until a visually acceptable result was obtained.
(2) Parameter range determination: Once the initial values of the parameters were identified, the following procedures were used to determine their variation ranges.
   (I) Except for the selected adjustable parameter and the deformation step parameter, fix all other parameters to their initial values or the middle value of their acceptable ranges;
   (II) Adjust the selected parameter value and rerun the program. For example, the inflation force parameters of the level set methods were adjusted to be 0.01, 0.1, 0.5, 1, 2, 5, 10, 100, 200, etc.[3]
   (III) Evaluate the resultant contour by visually comparing with the expert contour and rejecting the adjusted value when a large variation[4] occurs between the resultant contour and the expert contour.
   (IV) Determine the suitable range for the parameter in which the resultant contour is visually acceptable.
   (V) Repeat steps (I)–(IV) for the other parameters with the assumption that the parameters are independent, thus the order in choosing parameters does not matter. The resultant ranges of the parameters are obtained by doubling the acceptable range to increase the search range for the optimal parameter. Intuitively, if the parameters were dependent on each other then the resultant contours would not converge or provide accurate results. However, in practice this is not the case and the treatment of the parameters as independent provides accurate results.
(3) Parameter fine tuning: The parameters are fine-tuned based on the acceptable ranges determined above.
   (I) Assign all parameters to the middle value of their testing range.
   (II) Select one parameter to be free and choose uniformly incremented values[5] in its range to adjust the free parameter.
   (III) Rerun the programs with adjusted parameter.
   (IV) Evaluate resultant contour by computing the quantitative error measures described in Section 3.2.

[3] The steps to determine the value range of a parameter were used in a multi-scale fashion, i.e., start with a coarse resolution consisting of large "jump" between step values to determine the approximate range, then a higher resolution with smaller step values was used to locate the range limits. In fact, the process to determine the parameter range is not sensitive to the step values.

[4] In practice, manual visual comparison is good enough to detect noticeable differences between the extracted contours with the expert contour, i.e., the visual difference can be easily identified by a people is considered as a "large variation". In these cases, the manual comparison results are in agreement with the two quantitative error measurements.

[5] As mentioned above, the parameter range was determined using a multi-scale approach. To fine-tune the results, a uniform distributed set of 10 points is placed across the resultant parameter range. The point with the best result is chosen as the optimal value. In practice, it was found that one pass of fine-tuning for each parameter was sufficient.

(V) Derive an error curve or chart from the quantitative measures (e.g., Figs. 11 and 12 in Section 4.9).
(VI) The value for the best resultant contour is determined as the optimal value for this parameter.
(VII) Repeat the process for all other parameters.

In practice, the above fine-tuning process can be repeated for finer resolution if needed. For issues of location sensitivity, the above process was repeated at several locations (as shown in Fig. 1) inside the object of interest. The initial contour locations were generally selected to be uniformly distributed within the object, and to cover different difficulties. The parameter ranges for different initial locations were based on the coarse first initial location parameter results. From the coarse parameter ranges, the parameter values, for each different initial starting point, were fine-tuned as described above. For example, for the knee with a relatively simple shape, there were four initial locations (Fig. 1a), which are in the smooth region (middle location), rather dark regions (left and right locations) and blurry region (top location). For the brain with a quite complex shape, we applied seven initial locations uniformly distributed in the shape (Fig. 1c). The parameter tuning is based on the test images in Fig. 1, which were selected from several similar image slices in our database. The derived parameters in Table 2 and 3 also provide good results on other similar database images.

3.3.3. Different initializations

To broaden the scope of the DCM test, several other experiments were conducted. These excursions from the base experiment, the initial contour being formed as a circle around the user selected location, included two variations of initial contour locations and sizes, as illustrated in Section 4.8. In the experiments, the initial contours were approximated as polygons crossing the desired object boundary and surrounding the object, respectively. The contours either grew or shrank to extract the object boundary. The parameter setting and tuning process, for these experiments, used the same methods described in Section 3.3.2. The objective of these experiments was to test the sensitivity of the DCMs on different initial conditions and results are briefly stated in Section 4.8.

4. Experimental results

Due to the space limitations, only a partial set of representative images (see Fig. 1) and their results are presented. The parameter ranges described in Tables 2 and 3 are derived from the selected set of image results, which are presented in this section. As indicated in Section 3.3.1, the derived parameters in Tables 2 and 3 also provide comparable results on other database images. The results of each test image are first presented with the corresponding parameters listed in the subsequent tables, then followed by an overall comparative result table, Table 10, using the different evaluation measures. For a clear visual comparison of these results to the expert manual segmentations, the expert contours in Fig. 1 are displayed again in the first image of each figure (Figs. 2–8a). A discussion summarizing the findings is given at the end of this section.

4.1. MRI knee

The goal of the MRI knee experiment was to determine which method could segment an object that has diverse contrast in the region within the target boundary and a weak edge. In this test image, near both the left and right top corners of the region are darker in intensity and weak edge is located on the top center of the target contour. The results are displayed in Fig. 2b–h and the parameters to derive these results are listed in Table 4. The balloon snake has the same result as the topology snake, shown in Fig. 2b, because there is no contour topology change in the deformation process. All the results, except the balloon snake, show a crossing over of the blur segment in the upper middle region of the femoral condyle and most of the contours could not reach into the upper left and right corner. Four initial points were used for the test, as shown in Fig. 1a. The balloon/topology snake, original level set, geodesic active contour, area and length active contour, and constrained optimization method are generally not sensitive to the initial starting locations if deforming in a homogeneous region with a clear boundary. However, as describe in Section 3.1, the left and right portions of the femoral condyle are rather darker than the middle region, thus these methods had best results with the two initial locations both at the middle section. The best results for the distance snake and the GVF snake occurred with the upper middle initial location and are shown in Fig. 2c and d, respectively. Both of them required a bigger initial contour (e.g., double and triple the initial contour radius) than others in order to catch the attraction forces from the edge points in all directions. Otherwise, the whole contour will be pulled towards only a part of the boundary. For the two non-central initial locations, the distance snake and GVF snake had very poor results even when using a larger initial contour size (e.g., double, triple, four and five times of the initial contour radius) due to the above reason. From the visual observation of the radiologist, the balloon snake, distance snake and GVF snake methods were determined to give the best qualitative results, i.e., most similar to the expert in Fig. 2a.

4.2. Blood cells

The blood cell segmentation exemplifies the case of overlapping segments and non-homogenous regions. Three initial contour starting locations, upper, middle, and lower, see Fig. 1b, were tested. In Fig. 3b, the balloon/topology

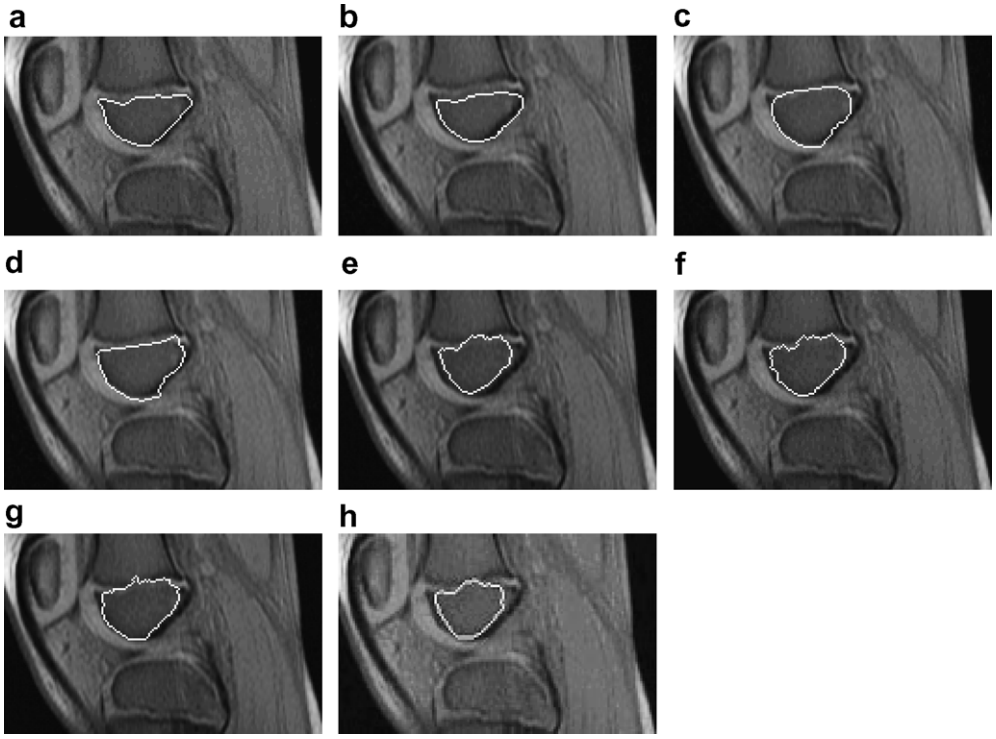

Fig. 2. (a) Knee expert contour. (b) Balloon/topology snake. (c) Distance snake. (d) GVF snake. (e) Original level set. (f) Geodesic active contour. (g) Area and length active contour. (h) Constrained optimization.

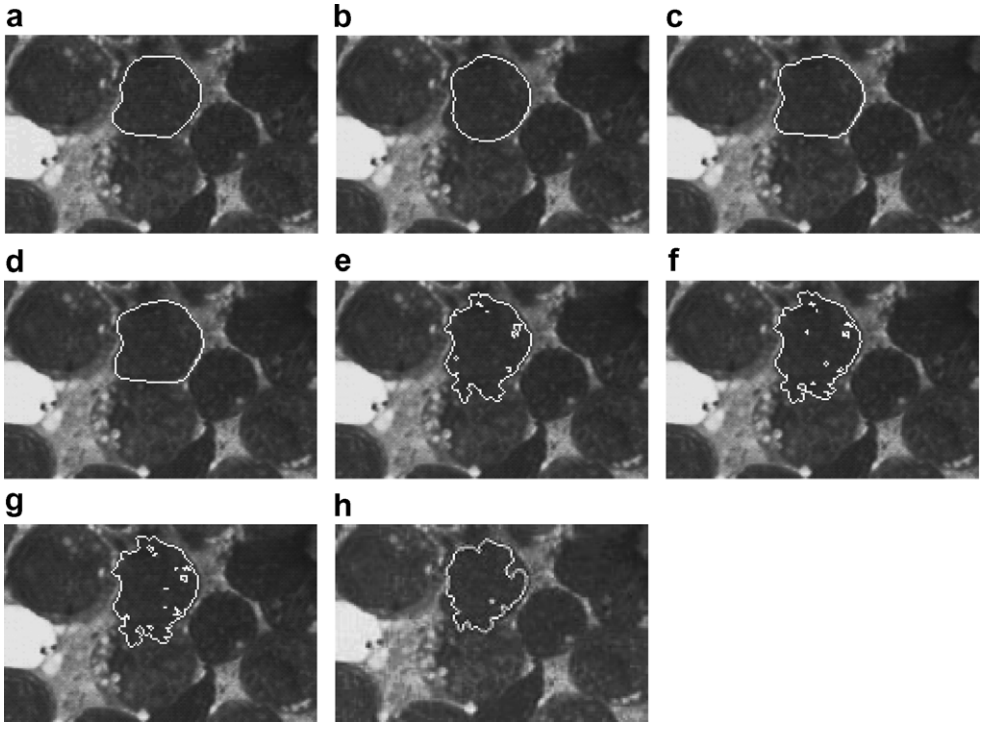

Fig. 3. (a) Cell expert contour. (b) Balloon/topology snake. (c) Distance snake. (d) GVF snake. (e) Original level set. (f) Geodesic active contour. (g) Area and length active contour. (h) Constrained optimization.

snake result is shown using middle initial location. Due to the overlapping on both the top and bottom regions of the cell, these two methods as well as the distance and GVF snakes were very sensitive to the initial contour location and the result of both upper and lower initial locations caused the contour to flow out of the expected boundary or shrink to be a line segment sticking to edges. The distance snake and GVF snake results, when middle location

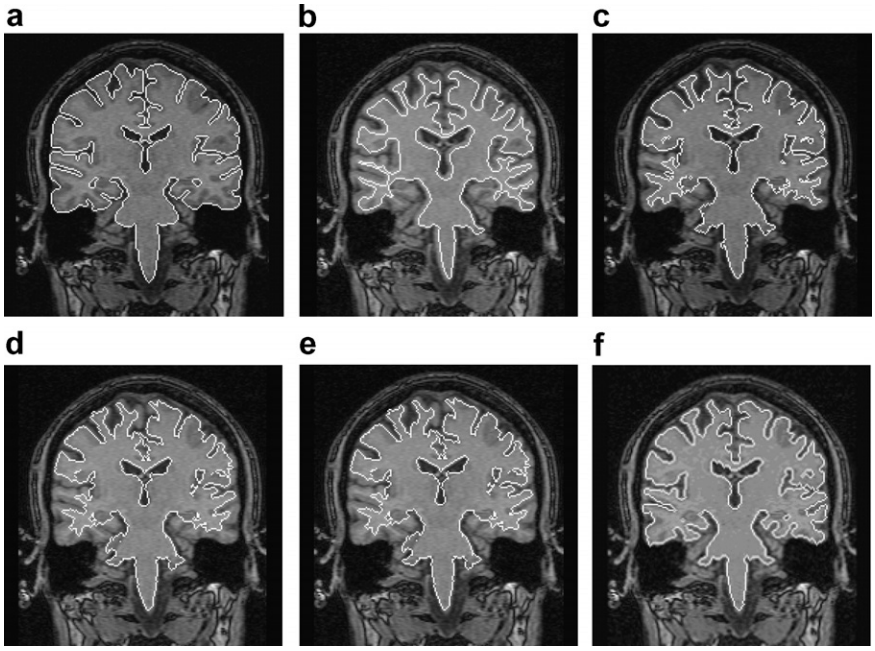

Fig. 4. (a) Sulci expert contour. (b) Topology snake. (c) Original level set. (d) Geodesic active contour. (e) Area and length active contour. (f) Constrained optimization.

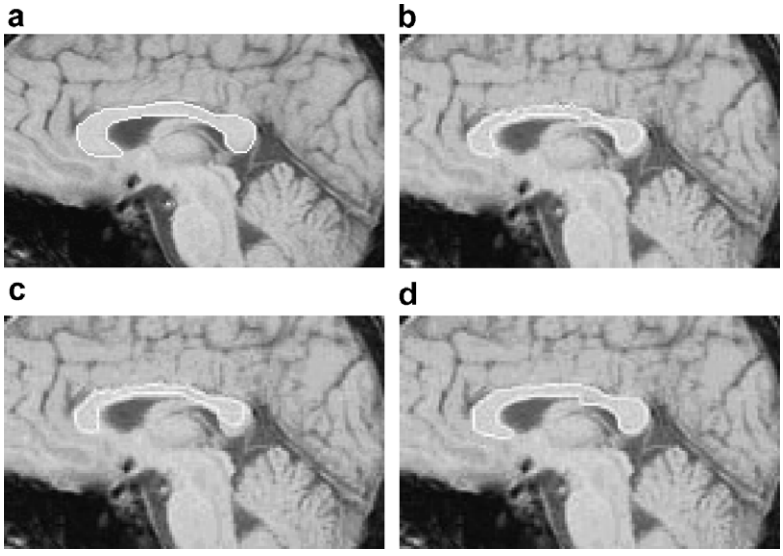

Fig. 5. (a) Corpus callosum expert contour. (b) Geodesic active contour. (c) Area and length active contour. (d) Constrained optimization.

was used, were displayed in Fig. 3c and d. All four of the level set group of methods yielded comparable results as shown in Fig. 3e–h. They are not sensitive to the initial starting location and the other parameters (IF or AF). However, with a fixed IF, AF, or SM value, the contour will flow out from the overlapped regions if we allow a large DS number in the deformation, which is resulted from the multiple local energy minima due to the missing edges in the overlapping regions. The distance snake and GVF snake were determined to obtain the best qualitative result and the parameters of different methods to derive the results in Fig. 3 are listed in Table 5.

4.3. MRI brain (sulci contour)

The coronal MRI brain image presents an example of complex contour shape and tests the deformable contour methods' ability to handle this type of difficulty. The goal of the segmentation was to extract the brain–CSF interface contour (sulci) and seven initial locations were tried, as shown in Fig. 1c. It can also be seen that the sulci contour is not topologically equivalent to the shape of the initial contour. Thus the balloon snake failed to provide any meaningful result due to incapability for topology change. So did the distance and GVF snakes as they produced

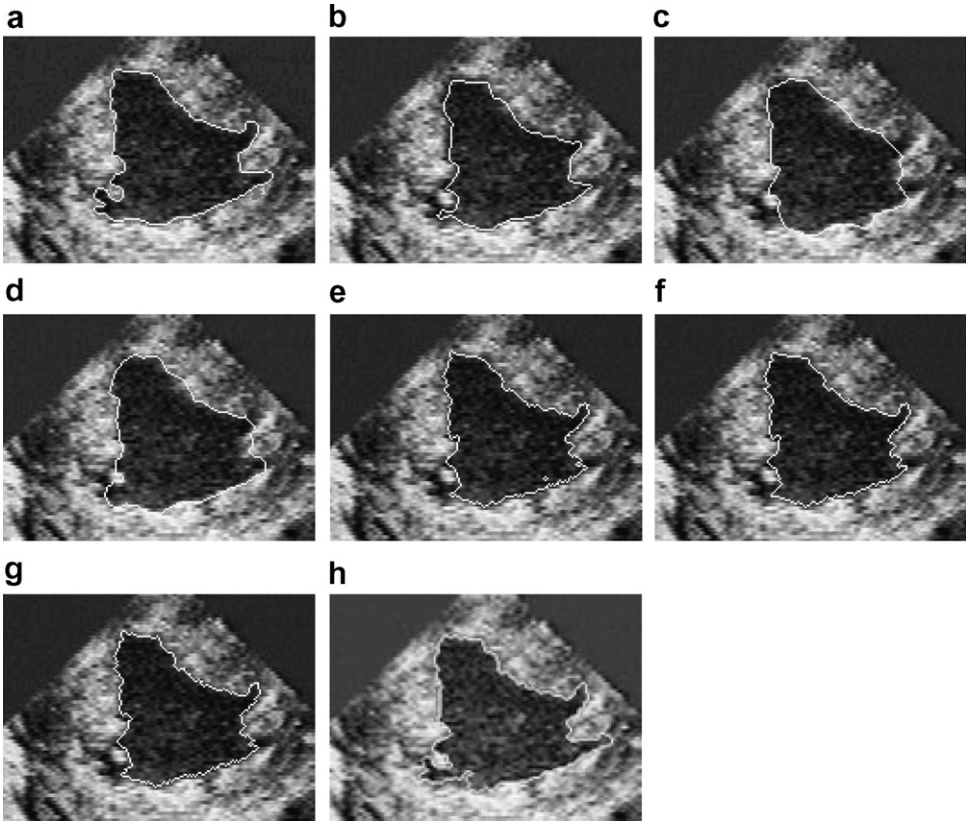

Fig. 6. (a) Pig heart expert contour. (b) Balloon/topology snake. (c) Distance snake. (d) GVF snake. (e) Original level set. (f) Geodesic active contour. (g) Area and length active contour. (h) Constrained optimization.

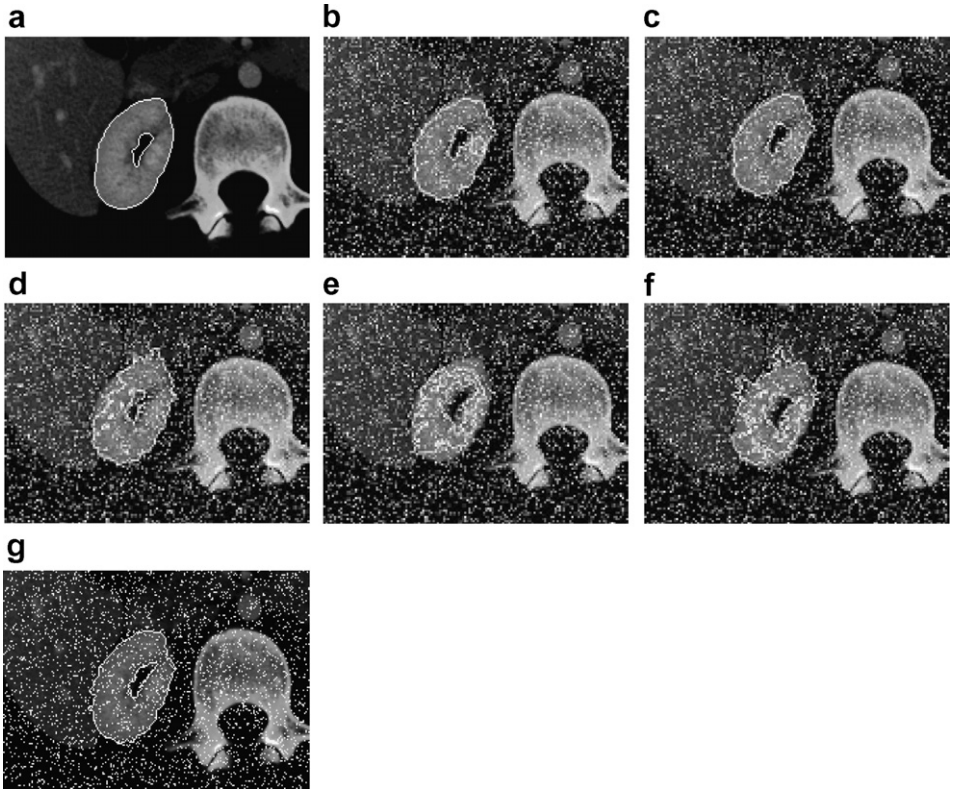

Fig. 7. (a) Kidney expert contour. (b) Balloon snake. (c) Topology snake. (d) Original level set. (e) Geodesic active contour. (f) Area and length active contour. (g) Constrained optimization.

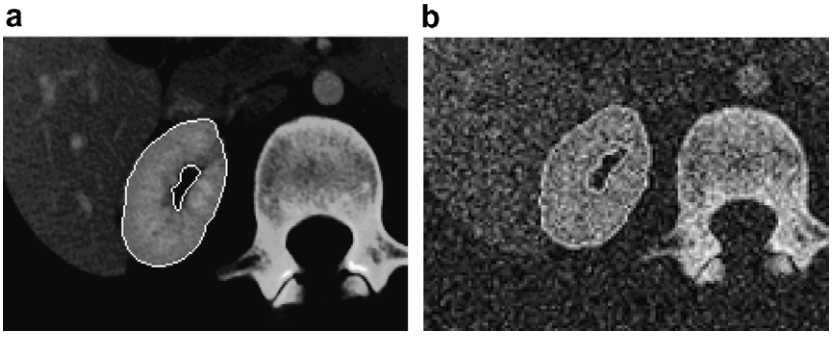

Fig. 8. (a) Kidney expert contour. (b) Constrained optimization.

Table 4
Parameter settings for Fig. 2 results

	Balloon/topology	GVF	Distance	Level set	Geodesic	Area/length	Constrained opt.
E	0.44	5.07	0.378	–	–	–	–
R	1.128	0	2.157	–	–	–	–
V	6.056	14.492	2	–	–	–	–
X	24.049	11.638	0.105	–	–	–	–
K_1	0.28	–	–	–	–	–	–
GVF N	–	0.025	–	–	–	–	–
GVF U	–	120	–	–	–	–	–
GT	–	–	0.025	–	–	–	–
IF	–	–	–	8	4	3	–
AF	–	–	–	–	–	1	–
SM	–	–	–	–	–	–	20
DS	600	30	900	140	390	650	600

Table 5
Parameter settings for Fig. 3 results

	Balloon/topology	GVF	Distance	Level set	Geodesic	Area/length	Constrained opt.
E	0.182	3.2	0.62	–	–	–	–
R	1.05	24	4.435	–	–	–	–
V	3.042	7.568	2.096	–	–	–	–
X	5	6.924	2.834	–	–	–	–
K_1	0.2	–	–	–	–	–	–
GVF N	–	0.057	–	–	–	–	–
GVF U	–	80	–	–	–	–	
GT	–	–	0.049	–	–	–	–
IF	–	–	–	10	10	6	–
AF	–	–	–	–	–	1	–
SM	–	–	–	–	–	–	20
DS	430	60	60	1500	2500	4000	1100

small or incomplete contours. With simple initial contour, both the distance and GVF snakes cannot catch the attraction forces from all the edge points in the external force field. The topology snake, being insensitive to the initial starting location, did a good job of extracting the gray–white interface instead of the sulci contour and the result is displayed in Fig. 4b. If further deformation is allowed, the contour will flow out from some gaps (e.g., left and right bottom regions of the sulci contour) before it reaches the correct sulci contour in other locations. Once again, the group of level set methods had similar results, which can be seen in Fig. 4c–f. With these methods the starting location caused little variance in the final contour. The best result for the original level set, geodesic and area and length active contour methods show that these methods cannot obtain small sharp protrusions (see Fig. 1c) in the contour segments at the lower left and lower right sides of the brain. Experiments also showed that further DS or iterations resulted in the contour crossing over the weak edge segments. The constrained optimization method provided the best qualitative results for this difficult problem. The parameters of different methods to derive the results in Fig. 4 are listed in Table 6.

Table 6
Parameter settings for Fig. 4 results

	Topology	Level set	Geodesic	Area/length	Constrained opt.
E	0.01	–	–	–	–
R	0.05	–	–	–	–
V	0.8	–	–	–	–
X	6	–	–	–	–
K_1	0.15	–	–	–	–
GVF N	–	–	–	–	–
GVF U	–	–	–	–	
GT	–	–	–	–	–
IF	–	10	10	10	–
AF	–	–	–	1	–
SM	–	–	–	–	4
DS	1000	9000	11,800	12,000	7000

4.4. MRI brain (corpus callosum contour)

The segmentation of the corpus callosum in the midline sagittal MRI brain image provides a difficult challenge because it has large area of blur boundary segments, which results in big gaps in the external force field. Three initial starting locations of Fig. 1d were used for test. All four of the snake methods, the balloon/topology, GVF, and distance snake, failed to segment the corpus callosum properly due to the insufficient attraction forces from the gaps in the external force field. The resultant contour overflows the weak boundaries into the brain tissue, even if different initial locations and initial contour sizes could not change the outcome. In the level set method group, the original level set method, like the snake methods, yielded a resultant contour that marched over the weak boundaries. The geodesic and the area and length active contour methods could extract the appropriate boundary when the middle point was set as the starting location, and the results are shown in Fig. 5b and c, respectively. However, due to the very blur contour segments at the lower left corner and middle bottom region of the corpus callosum, only a partial extraction occurred when either left or right starting location was used for these methods. If further deformation is allowed, the contour will flow out from these gaps before it extracts the whole corpus callosum. The constrained optimization method could yield acceptable results for all three locations and a result is shown in Fig. 5d. Note that only three methods were able to obtain reasonable resultant contours with the constrained optimization method providing the best qualitative result. The corresponding parameters are listed in Table 7.

4.5. Ultrasound pig heart

The pig heart is an ultrasound image and the segmentation problems result from the complex shape with inhomogeneous interior and gaps. Three initial contour locations (Fig. 1e) were used. The sensitivity to the initial location affected all the methods. Due to the proximity of the gaps the results from the bottom initial location had the worst performance. The extracted contours of all the methods were able to represent most of the pig heart. Again, the GVF snake requires a much bigger initial contour (e.g., five times of the initial contour radius) than others in order to catch the attraction forces from edge points in all directions. The best qualitative result was the GVF snake[6] and the parameters of different methods are listed in Table 8.

Table 7
Parameter settings for Fig. 5 results

	Geodesic	Area/length	Constrained opt.
IF	13	10	–
AF	–	1	–
SM	–	–	6
DS	500	600	600

4.6. CT kidney salt and pepper noise

This test uses a CT cross-section of the kidney for the segmentation. Salt and pepper noise (10%) was added to the image to test how the deformable contour methods will handle noisy image. The topology snake was able to completely extract the kidney and was insensitive to the starting location. The result is displayed in Fig. 7c and it is close to the balloon snake result in Fig. 7b. The difference is there is a self-intersection in the balloon snake result, which is due to the approach incapability of contour topology change. The original level set method, geodesic active contour, and area and length active contour extract the majority of the kidney before marching out and are sensitive to the initial locations. The best results are displayed in Fig. 7d–f with different initial locations in Fig. 1f (e.g., Fig. 7d–f were using the right, top, and left initial locations, respectively). Fig. 7g displays the result of the constrained optimization method. This method is able to extract the kidney and is insensitive to the starting location. The topology snake and constrained optimization methods were determined to provide the best qualitative results. Note for the kidney image without noise, almost all the methods studied were able to extract the kidney except for distance snake and GVF snake. However, with large initial contours (i.e., the initial contour is big enough in the external force field to "feel" the attraction forces from remote edge points), these two methods could also extract the kidney with no noise present. The parameters of different methods to derive the results in Fig. 7 are listed in Table 9.

[6] According the quantitative error measurements in Table 10, the best result was produced by the constrained optimization method, with an irregular contour. However, because the regular shape produced by the GVF snake looks closer to the expert contour, the radiologist ranked it as the best by his visual observation, which is actually the second best with the quantitative error measurements.

Table 8
Parameter settings for Fig. 6 results

	Balloon/topology	GVF	Distance	Level set	Geodesic	Area/length	Constrained opt.
E	0.107	3.477	0.554	–	–	–	–
R	0.14	12	1.385	–	–	–	–
V	4.024	1.44	4.564	–	–	–	–
X	3.4	3.114	2.96	–	–	–	–
K_1	0.224	–	–	–	–	–	–
GVF N	–	0.196	–	–	–	–	–
GVF U	–	200	–	–	–	–	–
GT	–	–	0.156	–	–	–	–
IF	–	–	–	16	18	18	–
AF	–	–	–	–	–	1	–
SM	–	–	–	–	–	–	10
DS	1080	55	360	20,000	20,000	20,000	7000

Table 9
Parameter settings for Fig. 7 results

	Balloon/topology	Level set	Geodesic	Area/length	Constrained opt.
E	0.5	–	–	–	–
R	0	–	–	–	–
V	20	–	–	–	–
X	6	–	–	–	–
K_1	0.31	–	–	–	–
GVF N	–	–	–	–	–
GVF U	–	–	–	–	
GT	–	–	–	–	–
IF	–	6	13	20	–
AF	–	–	–	1	–
SM	–	–	–	–	4
DS	12,000	2700	1350	6000	2400

4.7. CT kidney Gaussian noise

The kidney CT cross-section with additive Gaussian noise (variance 3000) had the worst overall results. Except for the constrained optimization, all the other methods could only extract partial sections of the kidney. The constrained optimization method was able to obtain the kidney boundary and starting location had little effect on the resultant contour (Fig. 8b). The parameters for the result in Fig. 8b are: SM is 4 and DS is 2400. Even though most medical images today have little Gaussian or salt and pepper noise, the addition of these noises to the kidney image is used to show the robustness and extended applicability of the DCMs in the study.

4.8. Different initializations

In general, initial contours for DCMs in medical image segmentations can be constructed by (a) placing a small contour within the object as the experiments in the previous sections, (b) drawing an approximate (polygonal) boundary across the desired object boundary, or (c) placing a large shape (polygon or circle) outside the desired boundary. The main experiments focused on the first condition, (a), because it requires the least user interaction for automated image segmentation. As described in Section 3.3.3, two other initializations, (b) and (c), were used to show the sensitivity of the DCMs. Due to the space limitation, the blood cell image was selected as an example for illustrating these different initial conditions. The constrained optimization method was also not utilized for these tests because it requires the initial contour being located within the object.

Fig. 9a and b shows the expert contour and the initial contour crossing the cell boundary. Fig. 9c–h shows the results obtained from the DCMs. For the initial contour segments within the object, the distance, and GVF snakes could correctly extract some segments of the cell boundary if there is no object overlapping at these segments. For the initial contour segments outside of the cell, they were attracted to edges of other cells if the initial contour segments were closer to the edges of the adjacent cells than those of the cell of interest. If the surrounding area of the cell is clean with no other objects, it can be predicted the distance and GVF snakes will be attracted to the desired boundary. The balloon snake and the level set methods failed to extract the object correctly. Actually in this initialization, they can not obtain the desired boundary even if the surrounding area of the cell of interest is clean because the balloon force will either grow or shrink the whole contour instead of growing the contour segments inside the object and shrinking the contour segments outside the object. All the methods had difficulties in handling the overlapping parts at the top and bottom of the cell.

Fig. 10a and b shows the expert contour and the initial contour surrounding the cell of interest. Fig. 10c–h shows the results obtained from the DCMs. The distance and GVF snakes extracted almost all the boundary except the top right contour segment that is being attracted to another cell, which is because the top right initial contour segment is closer to the edges of the adjacent cell than those of the cell of interest. The balloon snake and level set methods all had a negative inflation force value in order to shrink the initial contour to the cell boundary. It can be seen that the contour deflated into the cell from the overlapping regions of the top and bottom of the cell before the left contour segment was attracted to the correct edge.

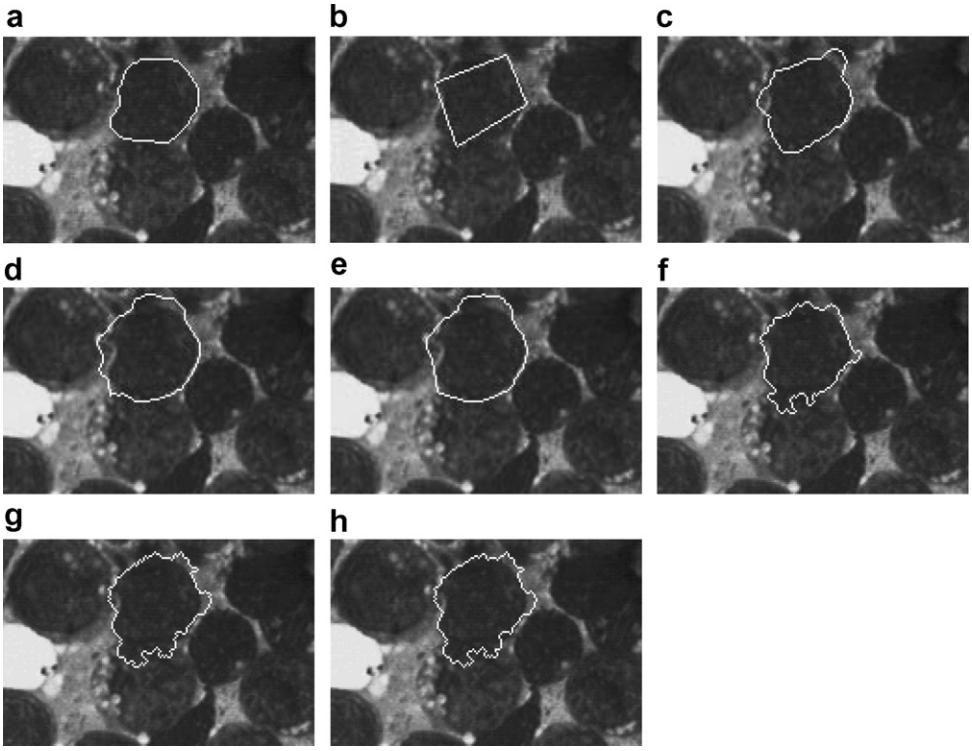

Fig. 9. (a) Cell expert contour. (b) Initial contour crossing object boundary. (c) Balloon/topology snake. (d) Distance snake. (e) GVF snake. (f) Original level set. (g) Geodesic active contour. (h) Area and length active contour.

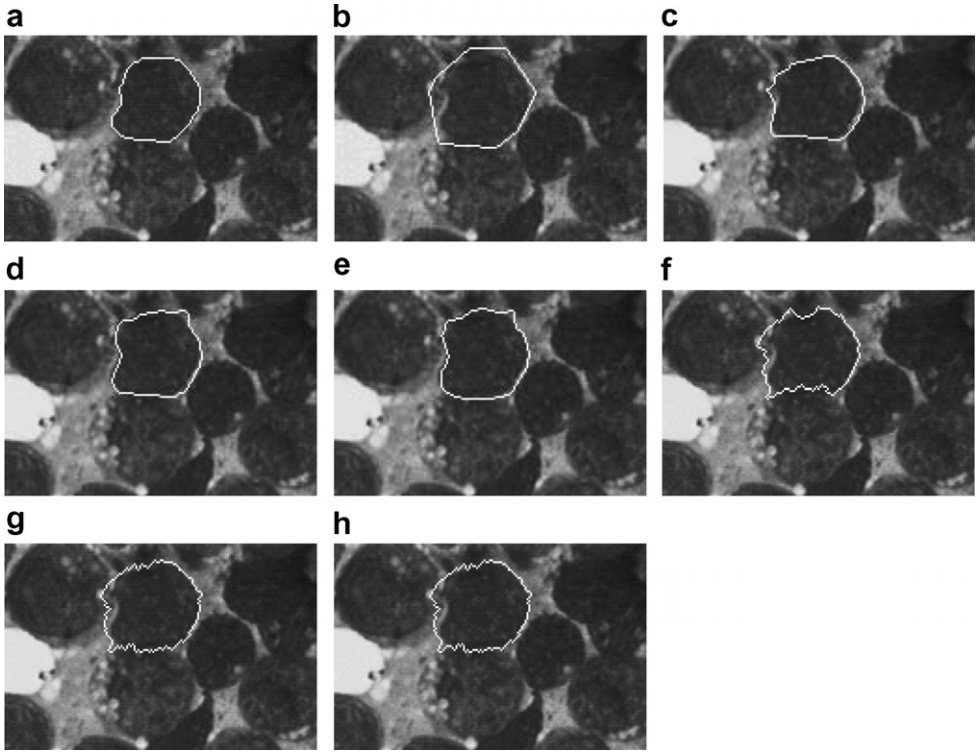

Fig. 10. (a) Cell expert contour. (b) Initial contour surrounding object. (c) Balloon/topology snake. (d) Distance snake. (e) GVF snake. (f) Original level set. (g) Geodesic active contour. (h) Area and length active contour.

4.9. Error measure results

Table 10 displays the results of both quantitative and qualitative evaluation of the experiments of the test set of images based on the experiments from Sections 4.1–4.7. As indicated in Section 3.1, the qualitative evaluations are implemented by the radiologist. The quantitative evaluations are done based on the two error measures defined

Table 10
Error measures: the bolded values represent the best quantitative result and the values marked by asterisk represent the best qualitative result

Measure	Medical image	Methods						
		Balloon/topology	Distance	GVF	Level set	Geodesic	Area/length	Constrained opt.
ε_1	Knee	0.20*	0.21*	**0.17***	0.34	0.28	0.26	0.30
	Cell	0.13	0.10*	**0.09***	0.19	0.21	0.20	0.18
	Brain (sulci)	0.51	0.97	0.91	0.23	0.26	0.35	**0.19***
	Brain (corpus callosum)	0.71	0.72	0.85	0.69	0.34	0.4	**0.18***
	Heart	0.16	0.17	0.12	0.14	0.13	0.13	**0.07***
	Kidney (SP)	0.17*	0.75	0.89	0.18	0.42	0.26	**0.12***
	Kidney (G)	0.57	0.80	0.87	0.57	0.54	0.76	**0.12***
ε_2	Knee	5.0*	5.0*	**3.0***	7.0	6.0	5.8	9.0
	Cell	6.1	**3.0***	3.6*	5.8	6.7	6.0	7.1
	Brain (sulci)	12.8	76.7	72.0	10.8	12.6	14.2	**7.1***
	Brain (corpus callosum)	42.2	59.3	57.1	57.3	7.6	9.8	**2.2***
	Heart	11.7	16.3	8.6	14.9	14.9	14.8	**6.4***
	Kidney (SP)	6.0*	48.4	51.0	7.8	12.2	7.1	**3.6***
	Kidney (G)	35.2	47.0	48.3	29.4	26.7	42.4	**4.0***

in Section 3.2 and the expert contours in Fig. 1. In Sections 4.1–4.7, the balloon snake and the topology snake have obtained the same results of the test images except the brain sulci segmentation, which involves contour topology change. Therefore, their results are merged together in Table 10 and the brain sulci contour is generated by the topology snake. The boldfaced values in the table represent the best result for the given test. The values marked by asterisk represent the best qualitative result. Note that in most of the cases, the qualitative and quantitative results are the same. Due to the inclusion of both edge and region features in deformation, the constrained optimization approach obtains better results than others in most cases.

To test the comparison values' invariability to the inter-individual and intra-individual variations of the manual segmentations, the radiologist drew two more sets of expert contours and three non-experts drew three sets of the non-expert contours based on the atlas [40–44] and guidance from the radiologist. Figs. 11 and 12 illustrate two examples of the segmentation error of ε_1 and ε_2 on the blood cell image (Fig. 1b) and the pig heart image (Fig. 1e), which are obtained by comparing the DCM results with the six sets of expert and non-expert contours. Ideally, if the boundaries drawn by different people (or by one people with multiple times) are identical, the error bars should have the same height. Therefore, a small variation of an error indicates a good invariability of the result to the inter- and intra-individual variations. As an example in Fig. 11, the mean and the standard deviation of ε_1 of the GVF result are 0.102 and 0.009 for the three expert contours, 0.098 and 0.005 for the three non-expert contours; the mean and the standard deviation of ε_2 are 2.997 and 0.167 for the three expert contours, 3.202 and 0.35 for the three non-expert contours. For both expert and non-expert model contours, the standard deviations are generally small compared with the mean values for the two measurements, thus this result is about invariable to human segmentation variations. The bars in Figs. 11 and 12 all have small variations and the method obtaining the best result in Table 10 in most cases also has the best results when compared with other models. From other testing images, we can also observe the similar results, which show that the quantitative comparison values in Table 10 are about invariable to the inter-individual and intra-individual variations.

4.10. Summary

As the emphasis was on obtaining correctly segmented contours and not on how efficiently implementing the algorithms, the computational complexity of the algorithms was not broached in this comparative study. Besides the topology changing operations in the topology snake and the external force fields construction in the distance and GVF snakes, the snake computation complexity in the coarse detection is $O(mP)$ where P is the number of snaxels (snake pixels) in each iteration and m is the iteration number before the snake reaches equilibrium. Level set methods implemented by the original level set formulation generally have the computational complexity $O(N^3)$ for 2D image segmentation, with N being the number of points in each coordinate direction. When narrow-band scheme is applied instead of level set formulation, the computational complexity will be reduced to $O(wN^2)$, where w is width of the narrow-band around the deforming contour. When the contour deformation velocity F is non-negative, fast marching scheme can be applied to deform the contour. The computational expense reduces from $O(N^3)$ of the level set formulation to $O(N^2 \log N)$ per time step in fast marching approach. In the cases of objects with simple shape, like the cell and knee, narrow-band method and fast marching method are obviously faster than the original level set method. On the other hand, the narrow-band method could be very slow in the cases of complex objects because the overhead to maintain a valid band is very large.

In summary, the snake methods have an overall smoother extracted contour than the level set methods. The smoothness is a result of the internal force parameters used in the snake definition and is why snakes have a more

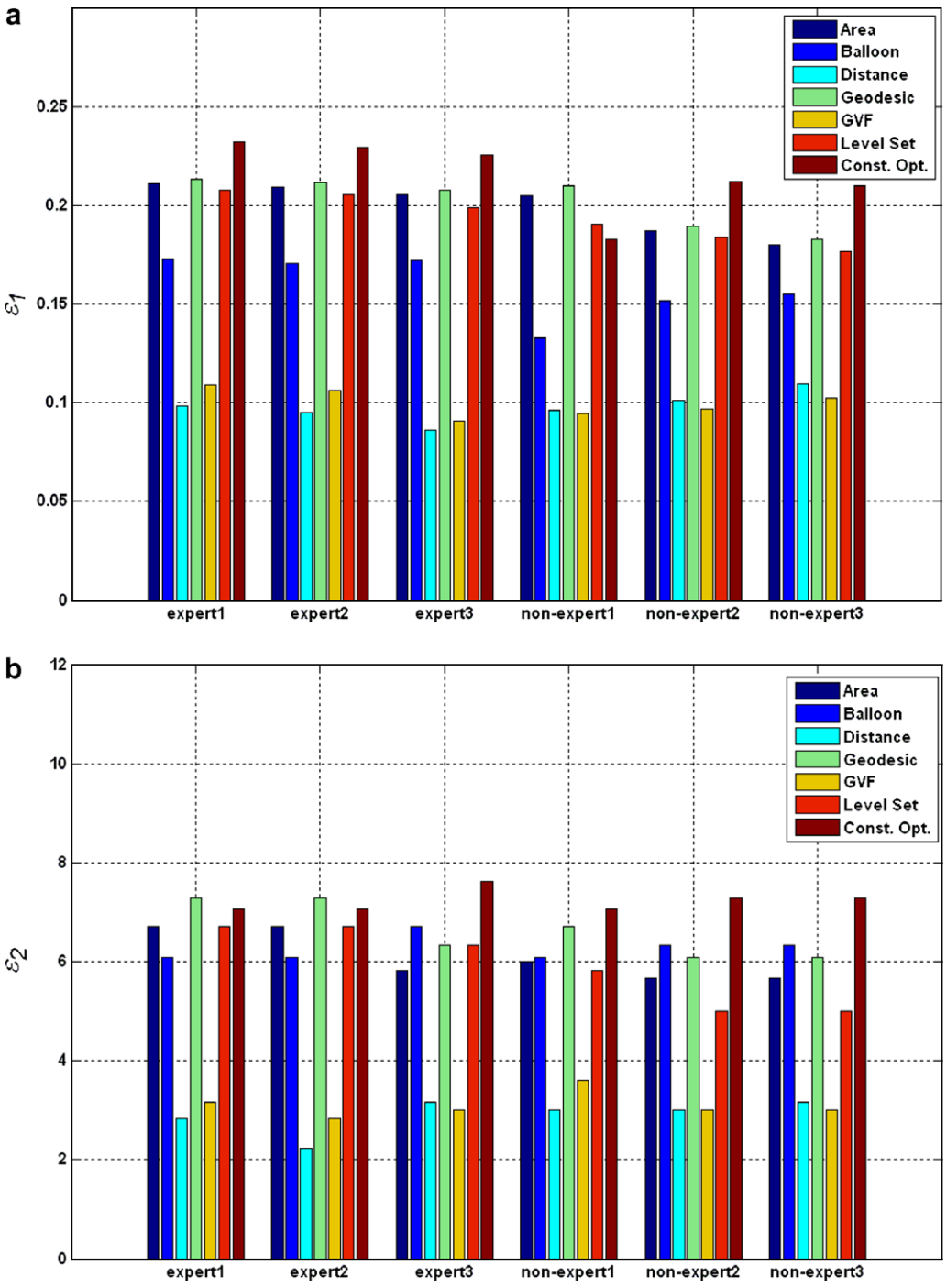

Fig. 11. (a) Blood cell error curves of ε_1. (b) Blood cell error curves of ε_2.

difficult time extracting protrusions in complex images. After the GVF snake, a generalized GVF (GGVF) [39] snake was proposed to handle the sharp protrusions with some improvements. In our testing images, it had similar results with the GVF snake. However, it still cannot handle the complex case due to the topology change limitation. Additionally, the GGVF snake performs like the distance and GVF snake in the case of weak edges, especially when a weak boundary is close to a strong spurious one, the snake readily steps through the weak edge and stops at the strong one, such as the MRI knee example in this study. On the other hand, the internal force enables snakes perform better than level set methods in gaps, such as blood cells. Furthermore, this smoothness is also reflected in the local error measure ε_2, which on average are larger than level set methods in images that have complex boundaries.[7]

[7] This is because ε_2 measures the maximum Euclidean distance between the extracted contour and the target boundary, for a complex boundary with a large curvature variation, the smoother the contour (i.e., the smaller the curvature variation), the larger the ε_2.

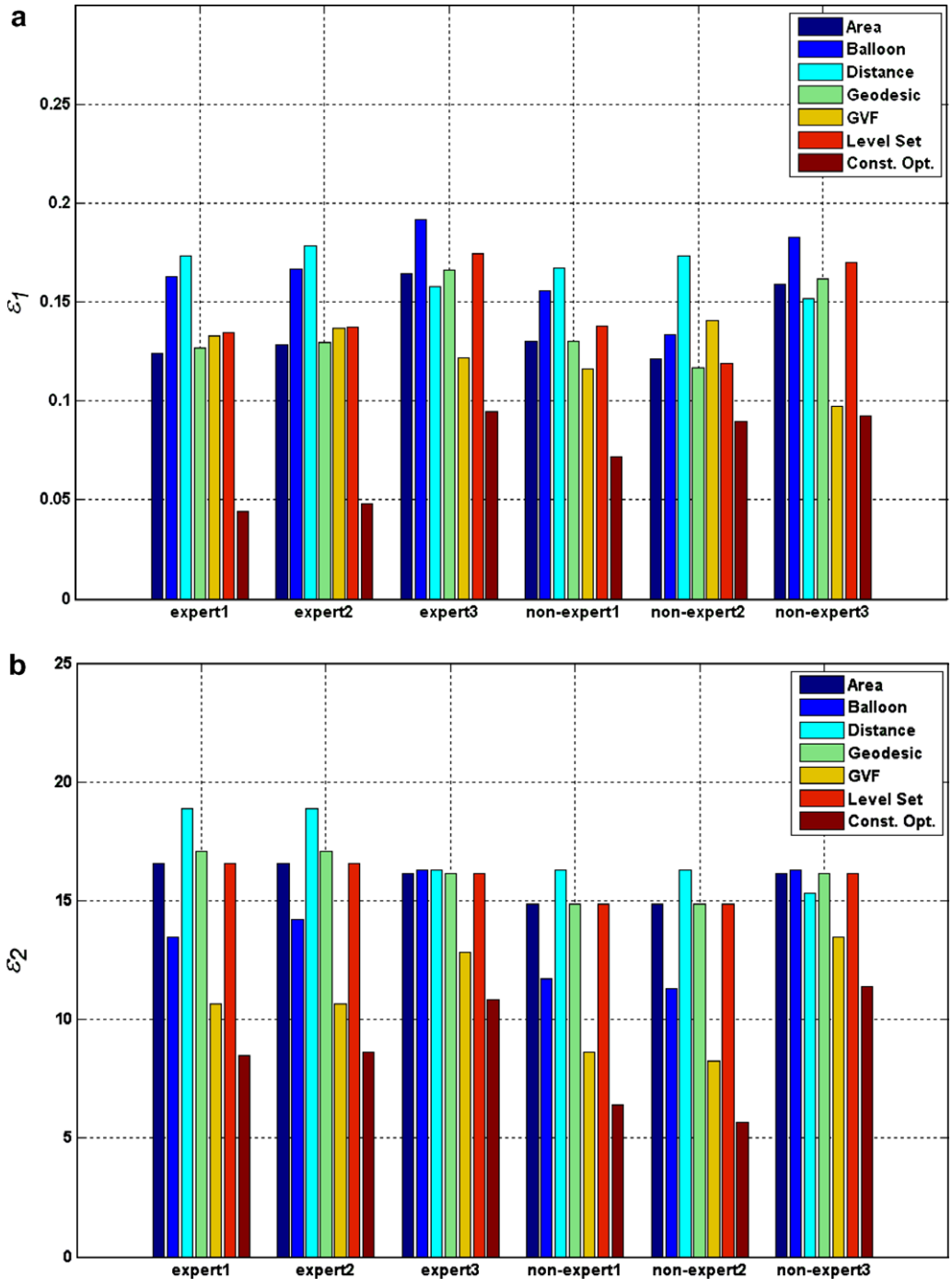

Fig. 12. (a) Pig heart error curves of ε_1. (b) Pig heart error curves of ε_2.

The balloon-type methods (balloon snake and level set methods) are more insensitive to initial contour locations than the methods do not have pressure forces like the distance snake and GVF snake. This is due to the shape and edge strength of the desired boundary. The distance snake and GVF snake usually require the initial contour to be placed in the center of object of interest, in order to avoid the whole contour being attracted by a part of boundary. As indicated in Sections 1 and 2, the object shape (edge map) generates the image external force field, which attracts the deformable contour to the desired boundary, thus the object shape plays an important role in the determination of the final result. For example, the cell image and knee image have shapes that are more "circular". These types of images, for the distance snake and GVF snake, are less sensitive to initial contour location and have a more uniform attraction to the boundary edge if the initial contour is large enough. In contrast, the corpus callosum in brain image has a narrow elongated shape and thus the position of the initial location causes drastic differ-

Table 11
Deformable contour method selection

	Multiple contours	Gap or blurry edges	Complex shapes with a large curvature variation	Circular shape[a]	Elongated shape	Noise interior	Initial contour outside or cross boundary
Balloon		√		√	√	√	
Distance		√		√		√	√
GVF		√		√		√	√
Topology	√	√		√	√	√	
Level set	√		√		√		
Geodesic	√		√		√		√
Area and length	√		√		√		√
Const. opt.	√		√		√	√	

[a] As indicated in Sections 4.1 and 4.5, a big distance or GVF snake with center initialization in the object is desired to obtain a good result in this condition.

ences on the extracted contours. GGVF is a good improvement of distance and GVF snakes on the robustness to the object shape variations. For the methods with ''balloon force'' to inflate or deflate the deformable contour, in ideal case (with only object and background) their results on the same image should be invariable to different initial contour locations and object shapes. However, in practice the medical images are rather complex and noisy. Therefore, the initial contour locations matter and the result will be good if the initial contour has a good location, like far away from gaps and close to clear edges.

The level set methods have the ability for handling topological change. This ability makes the level set methods preferable for segmenting complex shapes. Also, level set methods have fewer parameters than their snake methods counterparts and are more insensitive to the parameter variations making parameter tuning easier. However, in practice, the stopping criteria of the level set methods are mostly based on the number of iterations. This means that the number of iterations required for extracting the desired contour must be approximately known or the final contour will be either smaller or march over the expected boundary. For both snakes and level set methods, the segmentation results are insensitive to the parameter variations if their values are within the acceptable ranges, as described in Section 3.3.2 and shown in Tables 2 and 3. As a brief summary for above observations, Table 11 provides a general reference on DCM selection under different conditions. Note that practical problems are always complicated as combinations of several challenges, thus DCM selection is usually a tradeoff among different candidates.

5. Conclusions

This paper presents a study of eight deformable contour methods applied to segmentation of medical imagery. These test images provide a small sample of some typical cross-sectional views of anatomic parts as well as the four different imaging technologies widely used in the biomedical image processing domain: CT, MRI, ultrasound, and optical. Due to the various challenges posed by the medical images, in terms of extracting the actual boundary of the target object within a given image, the various deformable contour methods studied in this paper showed their versatility and shortcomings. It is concluded that even though there are still many challenges to be faced, it is with a better understanding of both the problems and the features provided by the various methods that a successful solution could be devised. One important point to be noted here is these methods are not mutually exclusive. Thus in practice, a new deformable contour method can be proposed to incorporate features from other methods in order to handle specific applications, e.g., the user could has a balloon snake that uses distance forces or GVF. In addition, the more information from image or object integrated in the evolution framework, the better segmentation results can be obtained. For example, the deformable contour method (e.g., the constrained optimization approach) considering both edge and region features can obtain better results than those based on only edge information. We hope this paper provides a better insight to both the biomedical image segmentation problem and the powerful features of the DCMs.

Acknowledgements

The authors thank Chenyang Xu, Vicent Caselles, Kaleem Siddiqi, and Xiao Han for the suggestions on the algorithm development and parameter settings. The authors also thank the reviewers for valuable comments to improve this paper.

References

[1] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, Int. J. Comput. Vis. 1 (4) (1988) 321–331.
[2] L. Cohen, On active contour models and balloons, CVGIP Image Understand. 52 (2) (1991) 211–218.
[3] L. Cohen, I. Cohen, Finite element methods for active contour models and balloons for 2-D and 3-D images, IEEE Trans. Pattern Anal. Machine Intell. 15 (11) (1993) 1131–1147.
[4] C. Xu, J. Prince, Snakes, shapes, and gradient vector flow, IEEE Trans. Image Process. 7 (3) (1998) 359–369.
[5] R. Malladi, J. Sethian, B. Vemuri, Shape modeling with front propagation, IEEE Trans. Pattern Anal. Machine Intell. 17 (2) (1995) 158–171.

[6] J. Sethian, A fast marching level set method for monotonically advancing fronts, Proc. Natl. Acad. Sci. USA 93 (4) (1996) 1591–1595.
[7] R. Malladi, J. Sethian, A real time algorithm for medical shape recovery, Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Bombay, India, 1998, pp. 304–310.
[8] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vis. 22 (1) (1997) 61–79.
[9] X. Wang, L. He, W.G. Wee, Deformable contour method: a constrained optimization approach, Int. J. Comput. Vis. 59 (1) (2004) 87–108.
[10] T. McInerney, D. Terzopoulos, T-snakes: topologically adaptive snakes, Med. Image Anal. 4 (2) (2000) 73–91.
[11] K. Siddiqi, Y.B. Lauzière, A. Tannenbaum, S.W. Zucker, Area and length minimizing flows for shape segmentation, IEEE Trans. Image Process. 7 (3) (1998) 433–443.
[12] G. Stovik, A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing, IEEE Trans. Pattern Anal. Machine Intell. 16 (1994) 976–986.
[13] A. Lundervold, G. Storvik, Segmentation of brain parenchyma and cerebrospinal fluid in multispectral magnetic resonance images, IEEE Trans. Med. Imaging 14 (1995) 339–349.
[14] R. Grzeszczuk, D. Levin, "Brownian strings": segmenting images with stochastically deformable contours, IEEE Trans. Anal. Machine Intell. 19 (1997) 1100–1114.
[15] L.H. Staib, J.S. Duncan, Boundary finding with parametrically deformable models, IEEE Trans. Pattern Anal. Machine Intell. 14 (1992) 1061–1075.
[16] T.F. Cootes, A. Hill, C.J. Taylor, J. Haslam, Use of active shape models for locating structures in medical images, Image Vis. Comput. 12 (1994) 355–366.
[17] A.K. Jain, Y. Zhong, S. Lakshmanan, Object matching using deformable templates, IEEE Trans. Pattern Anal. Machine Intell. 18 (1996) 267–278.
[18] A. Chakraborty, J.S. Duncan, Game-theoretic integration for image segmentation, IEEE Trans. Pattern Anal. Machine Intell. 21 (1) (1999) 12–30.
[19] M. Rousson, N. Paragios, Shape priors for level set representations, in: Proceedings of the Seventh European Conference on Computer Vision (ECCV), 2002, pp. 78–92.
[20] R. Ronfard, Region-based strategies for active contour models, Int. J. Comput. Vis. 13 (2) (1994) 229–251.
[21] A. Chakraborty, L.H. Staib, J.S. Duncan, Deformable boundary finding in medical images by integrating gradient and region information, IEEE Trans. Med. Imaging 15 (1996) 859–870.
[22] S.C. Zhu, A. Yuille, Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation, IEEE Trans. Pattern Anal. Machine Intell. 18 (9) (1996) 884–900.
[23] T.F. Chan, L.A. Vese, Active contour without edges, IEEE Trans. Image Process. 10 (2001) 266–277.
[24] N. Paragios, R. Deriche, Geodesic active region and level set methods for supervised texture segmentation, Int. J. Comput. Vis. 46 (3) (2002) 223–247.
[25] B. Ostlad, A. Tonp, Encoding of a priori information in active contour models, IEEE Trans. Pattern Anal. Machine Intell. 18 (9) (1996) 863–872.
[26] S.D. Fenster, J.R. Kender, Sectored snakes: evaluating learned-energy segmentations, IEEE Trans. Pattern Anal. Machine Intell. 23 (9) (2001) 1028–1034.
[27] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, Geometric snake model for segmentation of medical imagery, IEEE Trans. Med. Imaging 16 (2) (1997) 199–209.
[28] W.J. Niessen, B. Romeny, M.A. Viergever, Geodesic deformable models for medical image analysis, IEEE Trans. Med. Imaging 17 (4) (1998) 634–641.
[29] L.D. Cohen, R. Kimmel, Global minimum for active contour models: a minimal path approach, Int. J. Comput. Vis. 24 (1) (1997) 57–78.
[30] A.K. Jain, Y. Zhong, M. Dubuisson-Jolly, Deformable template models: a review, Signal Process. 71 (2) (1998) 109–129.
[31] T. McInerney, D. Terzopoulos, Deformable models in medical image analysis: a survey, Med. Image Anal. 1 (2) (1996) 91–108.
[32] Y. Chen, S. Thiruvenkadam, F. Huang, H.D. Tagare, D. Wilson, E.A. Geiser, On the incorporation of shape priors into geometric active contours, in: Proceedings of the IEEE Workshop on Variational and Level Set Methods, 2001, pp. 145–152.
[33] M. Leventon, E. Grimson, O. Faugeras, Statistical shape influence in geodesic active contour, Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR) (2000) 316–322.
[34] D. Adalsteinsson, J.A. Sethian, A fast level set method for propagating interfaces, J. Comput. Phys. 118 (2) (1995) 269–277.
[35] J. Tohka, Surface extraction from volumetric images using deformable meshes: a comparative study, in: Proceedings of the Seventh European Conference on Computer Vision (ECCV), 2002, pp. 350–364.
[36] A.F. Frangi, W.J. Niessen, M.A. Viergever, Three-dimensional modelling for functional analysis of cardiac images: a review, IEEE Trans. Med. Imaging 20 (1) (2001) 2–25.
[37] H. Delingette, J. Montagnat, Shape and topology constraints on parametric active contours, Comput. Vis. Image Understand. 83 (2001) 140–171.
[38] C. Xu, D.L. Pham, J.L. Prince, Medical image segmentation using deformable models, in: SPIE Handbook on Medical Imaging – vol. III: Medical Image Analysis, May 2000, pp. 129–174.
[39] C. Xu, J.L. Prince, Generalized gradient vector flow external forces for active contours, Signal Process. 71 (2) (1998) 131–139.
[40] P. Fleckenstein, J. Tranum-Jensen, Anatomy in Diagnostic Imaging, second ed., W.B. Saunders, Philadelphia, 2001.
[41] A.V. Hoffbrand, J.E. Pettit, Color Atlas of Clinical Haematology, third ed., Mosby, London, New York, 2000.
[42] C.M. Otto, The Practice of Clinical Echocardiography, second ed., W.B. Saunders, Philadelphia, 2002.
[43] J.R. Haaga, C.F. Lanzieri, R.C. Gilkeson, CT and MR Imaging of the Whole Body, fourth ed., Mosby, St. Louis, MO, 2003.
[44] J.K. Mai, J. Assheuer, G. Paxinos, Atlas of the Human Brain, Academic Press, San Diego, CA, 1997.
[45] A.A. Amini, T.E. Weymouth, R.C. Jain, Using dynamic programming for solving variational problems in vision, IEEE Trans. Anal. Machine Intell. 12 (9) (1990) 855–867.
[46] D.J. Williams, M. Shah, A fast algorithm for active contours and curvature estimation, CVGIP Image Understand. 55 (1) (1992) 14–26.
[47] G.A. Giraldi, E. Strauss, A.A. Oliveira, A boundary extraction method based on dual-T-snakes and dynamic programming, Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR) (2000) 44–49.
[48] A.X. Falcao, J.K. Udupa, F.K. Miyazawa, An ultra-fast user-steered image segmentation paradigm: live wire on the fly, IEEE Trans. Med. Imaging 19 (1) (2000) 55–62.
[49] L. O'Donnell, C.F. Westin, W.E.L. Grimson, J. Ruiz-Alzola, M.E. Shenton, R. Kikinis, Phase-based user-steered image segmentation, in: Proceedings of the Fourth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Utrecht, Netherlands, 2001, pp. 1022–1030.