An active contour model for image segmentation based on elastic interaction

Yang Xiang [a],*, Albert C.S. Chung [b], Jian Ye [a]

[a] *Department of Mathematics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*
[b] *Department of Computer Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

Abstract

The task of image segmentation is to partition an image into non-overlapping regions based on intensity or textural information. The active contour methods provide an effective way for segmentation, in which the boundaries of the objects are detected by evolving curves. In this paper, we propose a new edge-based active contour method, which uses a long-range and orientation-dependent interaction between image boundaries and the moving curves while maintaining the edge fidelity. As a result, this method has a large capture range, and is able to detect sharp features of the images. The velocity field for the moving curves generated by this elastic interaction is calculated using the fast Fourier transform (FFT) method. Level set representation is used for the moving curves so that the topological changes during the evolution are handled automatically. This new method is derived based on the elastic interaction between line defects in solids (dislocations). Although it is derived originally for two dimensional segmentation, we also extend it to three dimensions. The features of the new method are examined by experiments on both synthetic images and medical images of blood vessels. Comparisons are made with the existing active contour methods.
© 2006 Elsevier Inc. All rights reserved.

Keywords: Active contour; Segmentation; Edge-based; Level set method; Elastic interaction

1. Introduction

The task of image segmentation is to partition an image into non-overlapping regions based on the intensity or textural information. For example, in medical image segmentations, people identify anatomical structures (e.g., blood vessels and brain tissues) and other regions of interest from the images acquired by different medical devices. The active contour method for image segmentation was introduced by Kass et al. in late 1980's [1]. In this method, a curve is evolved towards the object boundary under a force, until it stops at the boundary. More precisely, the curve moves to minimize the energy

* Corresponding author.
 E-mail address: maxiang@ust.hk (Y. Xiang).

$$E(l) = \int_l \left(\frac{1}{2}\alpha|l'(s)|^2 + \frac{1}{2}\beta|l''(s)|^2 - \lambda|\nabla I(l(s))|^2 \right) ds, \tag{1}$$

where $l(s)$ represents the parameterized curve, $I(x,y)$ is the image grey-level function, and constants α, β, $\lambda > 0$. The first two terms in the energy functional smooth the curve. The third term attracts the curve to the object boundary, where the value of image gradient is large. The dynamics equation of the curve to minimize the total energy is given by

$$l_t(s) = \alpha l''(s) - \beta l''''(s) + \lambda \nabla|\nabla I(l(s))|^2. \tag{2}$$

The image function $I(x,y)$ can be replaced by its smoothed version $G_\sigma(x,y) * I(x,y)$, where $G_\sigma(x,y)$ is a two dimensional Gaussian function with zero mean and standard deviation σ, and the operator $*$ is the convolution operator.

Many efforts have been made to improve this method. In [2,3], a constant force (balloon force) was added in the normal direction of the curve to accelerate its motion and increase the capture range. In [4,5], the level set framework [6] was used to handle the topological changes such as merging or splitting of the moving curve. Some other methods based on the level set formulation were proposed in which the velocity of the moving curve was written as the minimization of an energy [7–10], which gives some additional terms to attract the curve to the object boundary from its both sides. The dynamics equation in [7] is

$$\phi_t = |\nabla\phi|\nabla \cdot \left(g(\nabla I)\frac{\nabla\phi}{|\nabla\phi|} \right) + cg(\nabla I)|\nabla\phi|, \tag{3}$$

where $\phi(x,y)$ is the level set function whose zero level set represents the curve, $g(\nabla I(x,y))$ is a function whose value is very small at the boundary where the image gradient is large, so that the velocity of the curve is small and the curve will stop there. One choice of $g(\nabla I(x,y))$ is

$$g(\nabla I(x,y)) = 1/(1 + |\nabla G_\sigma(x,y) * I(x,y)|^p), \quad p \geqslant 1. \tag{4}$$

The first term in the dynamics equation Eq. (3) is related to the curvature of the curve, which smooths and shortens the curve. This term also contains a force attracting the curve to the boundary due to the presence of the function $g(\nabla I(x,y))$. The second term in Eq. (3) is proportional to a constant velocity c in the normal direction of the curve, which is similar to the balloon force mentioned above, making the curve expand or shrink depending on its sign ($c > 0$ was required in [7], but not in other methods, e.g., [5,8,10]). There are also some methods taking the direction of the image gradient into consideration [11,12]. However, for these methods, without the balloon force, the capture range is short and the curve cannot reach the narrow concave parts of the boundary. This is because the effect of $\nabla I(x,y)$, defined by $\nabla|\nabla I(x,y)|^2$ in Eq. (2) or $g(\nabla I(x,y))$ in Eq. (3), is localized near the boundary. While with the balloon force, there is a limitation that the balloon force cannot make some parts of curve shrink while other parts of the curve expand, so the initial curve must be placed entirely outside or inside the object to be detected.

To solve the problem of these methods, Xu and Prince [13] proposed the gradient vector flow (GVF) method. Some modifications on this method were made afterwards [14,15]. In this method, the attractive force near the object boundary $\nabla|\nabla I(l(s))|^2$ (see Eq. (2)) is extended to the whole computational region by diffusion. More precisely, they obtained a force field $\mathbf{v} = (u,v)$ from the image by minimizing the energy

$$E = \int\int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2|\mathbf{v} - \nabla f|^2 dx\,dy, \tag{5}$$

where $f = |\nabla I(x,y)|^2$, $\nabla f = \nabla|\nabla I(x,y)|^2$ represents the attractive force to the boundary, and μ is the diffusion constant. In this equation, near the object boundary, $|\nabla f|$ is large, the second term dominates and the minimization gives $\mathbf{v} = \nabla f$; while away from the boundary, $|\nabla f|$ is small and thus the second term is small, the energy is dominated by the diffusion term, which means the force $\mathbf{v}$ is extended smoothly from its value near the object boundary. Therefore, the capture range is larger than that of the previous methods and there is no need to place the initial curve entirely inside or outside the object. However, the initial curve

must still be placed near the object boundary to be detected, otherwise it might be attracted to a boundary of a wrong object or a wrong part of the boundary (see the image and the initial contour in Fig. 5(a) for an example).

There are also other region-based active contour methods, which rely on the homogeneity of the localized spatial features instead of the image gradient, such as the method proposed by Chan and Vese [16–19] based on the Mumford–Shah segmentation techniques [20] and the level set framework [6], and other region-based methods, e.g., [21–25], as well as the combinations of the edge-based and region-based methods, e.g., [26–28,12].

In this paper, we propose a new edge-based active contour method, which uses an orientation-dependent and long-range interaction between image boundaries and the moving curves. The incorporation of the orientations of the object boundaries and the moving curves enables the moving curves to be attracted to the correct object boundaries, thus enlarges the capture range significantly compared with the GVF method. In our new method, the attractive interaction between the moving curves and the object boundaries is very strong near the object boundaries, and decays relatively slowly away from them, which gives an extension of the effect of the image gradient to the whole image domain without smoothing the image gradient itself. The resultant velocity of the moving curve is an integral involving the image gradient over the whole image domain, and is calculated efficiently using the fast Fourier transform (FFT) method. While in the GVF method, the attractive force $\nabla|\nabla I(x,y)|^2$ near the object boundaries is extended by diffusion, which smears more or less the object boundaries. The long-range and orientation-dependent interaction between the object boundaries and moving curves in our new method is defined based on the elastic interaction between line defects (dislocations) in solids [29–32]. Although the method is derived originally for two dimensional segmentation, we also extend it to three dimensions. Finally, in our method, the level set representation [6] is used for the moving curves so that the topological changes during the evolution are handled automatically.

The rest of the paper is organized as follows. In Section 2, we define the elastic interaction between object boundaries and the moving curves, and then derive the expression of the resultant velocity. In Section 3, we analyze the properties of this interaction, and show how the image noise can be removed under this interaction. In Section 4, we present the numerical implementation for this method. In Section 5, we extend this method to image segmentation in three dimensions. In Section 6, we show the advantages of the new method by the experiments on both synthetic images and medical images of blood vessels. Finally, in Section 7, we present the conclusions.

2. Derivation

In this section, we present the new active contour model for image segmentation based on elastic interaction. We focus on two-dimensional images in this section. We first define an energy associate with lines, which represent the object boundaries or the moving contour. Then an evolution equation for the moving contour is derived to find the object boundaries by minimizing the total energy. The level set framework is used for the moving contour to handle the topological changes.

2.1. Energy of lines

In the dissipative dynamics of dislocations (line defects) in crystals [29–32] or vortices in superconductivity [33–36], there are energies associated with these lines, and the lines move and interact to minimize the total energies. In the two-dimensional segmentation problems, both the boundaries of the objects and the moving curves are lines. We now define an energy associated with them similar to those of dislocations or vortices. We put the plane containing the two-dimensional images (say the xy plane) into a three-dimensional space (the xyz space), as shown in Fig. 1. The energy associated with these lines in the image plane will be defined in the whole three-dimensional space.

Considering a parameterized curve $\gamma(s)$ in the three-dimensional space, we define an energy associated with $\gamma(s)$
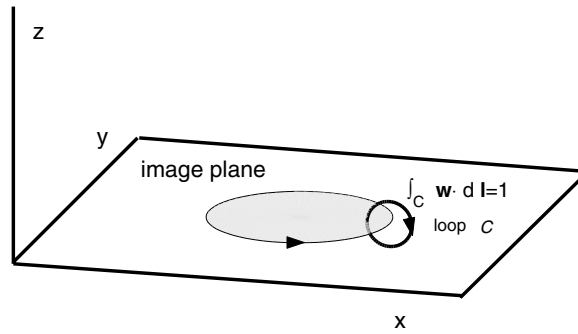
Fig. 1. The image plane is contained in a three dimensional space. The integral $\int_C \mathbf{w} \cdot d\mathbf{l} = 1$ along any curve C enclosing a line which represents an object boundary or the moving contour.

$$E(\gamma) = \min \int \frac{1}{2}(w_1^2 + w_2^2 + w_3^2)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{6}$$

subject to the constraint

$$\nabla \times \mathbf{w} = \delta(\gamma)\boldsymbol{\tau}, \tag{7}$$

where $\mathbf{w} = (w_1, w_2, w_3)$ is a function in the three dimensional space, $\boldsymbol{\tau}$ is the unit tangent vector of $\gamma(s)$, $\delta(\gamma)$ is the delta function of the curve $\gamma(s)$, which is the two dimensional Dirac delta function in the plane perpendicular to the tangent vector at any point on $\gamma(s)$, and $\nabla \times \mathbf{w}$ is the curl of $\mathbf{w}$ in three dimensions. The meaning of Eq. (7) is that the integral $\int_C \mathbf{w} \cdot d\mathbf{l} = 1$ along any curve C enclosing $\gamma(s)$, where $d\mathbf{l}$ is a line element of C, as shown in Fig. 1. This energy is a simplified version of the elastic energy associated with dislocations in solids [29–32].

Now we look for the Euler–Lagrange equation of this energy. Using the Lagrange multiplier method, we have

$$E = \int \left[\frac{1}{2}(w_1^2 + w_2^2 + w_3^2) + \lambda \cdot (\nabla \times \mathbf{w} - \delta(\gamma)\boldsymbol{\tau}) \right] \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{8}$$

where the Lagrange multiplier $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ is a vector function. Taking variation with respect to $\mathbf{w}$ and λ, respectively, we have the equations for the energy minimum,

$$\begin{cases} \frac{\delta E}{\delta \mathbf{w}} = \mathbf{w} + \nabla \times \lambda = 0, \\ \frac{\delta E}{\delta \lambda} = \nabla \times \mathbf{w} - \delta(\gamma)\boldsymbol{\tau} = 0. \end{cases} \tag{9}$$

The first equation is equivalent to

$$\nabla \cdot \mathbf{w} = 0, \tag{10}$$

and the second one is the constraint equation (7).

This is a simplified version of the elasticity system associated with dislocations in solids [29–32]. This system is also similar to those for vortices in superconductivity [33–36] or fluid dynamics [37,38]. This system for $\mathbf{w}$ can be solved analytically [29,30,37,38], which is

$$\mathbf{w}(x,y,z) = -\frac{1}{4\pi} \int_\gamma \frac{\mathbf{r} \times d\mathbf{l}}{r^3}, \tag{11}$$

where $d\mathbf{l}$ is a line element of the curve γ, $\mathbf{r} = (x - x(s), y - y(s), z - z(s))$ is the vector between the point (x, y, z) and a point $(x(s), y(s), z(s))$ on $\gamma(s)$, and $r = \sqrt{(x - x(s))^2 + (y - y(s))^2 + (z - z(s))^2}$ is the distance between them. The total energy is

$$E = \frac{1}{8\pi} \int_{\gamma} \int_{\gamma'} \frac{d\mathbf{l} \cdot d\mathbf{l}'}{r}, \tag{12}$$

where γ' is the line γ using another parameter s'.

The solution $\mathbf{w}$ is singular on $\gamma(s)$ and as a result, the energy $E(\gamma)$ is also singular. The singularities can be smeared out if we use a regularized delta function in Eq. (7), as in the dynamics of dislocations or vortices. For image segmentation, we shall use the Gaussian function for the regularization of images, and the derivative of the regularized Heaviside function given by Eq. (29) for the regularization of the moving curves, see Section 2.2. Any other regularized delta functions can also be used and will not change the segmentation results, as long as they satisfy the symmetry and some other conditions described in Section 3.2. Using the relationship $d\mathbf{l} = \boldsymbol{\tau}\delta(\gamma) \, dx \, dy \, dz$ and a regularized delta function (for simplicity we still use the same notation), we have

$$\mathbf{w}(x,y,z) = -\frac{1}{4\pi} \int_{\mathbf{R}^3} \frac{\mathbf{r} \times \boldsymbol{\tau}}{r^3} \delta(\gamma') dx' dy' dz' \tag{13}$$

and

$$E = \frac{1}{8\pi} \int_{\mathbf{R}^3} \delta(\gamma) dx \, dy \, dz \int_{\mathbf{R}^3} \frac{\boldsymbol{\tau} \cdot \boldsymbol{\tau}'}{r} \delta(\gamma') dx' dy' dz', \tag{14}$$

where $\mathbf{r} = (x - x', y - y', z - z')$, $r = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}$, and $\boldsymbol{\tau}$ (or $\boldsymbol{\tau}'$) is the unit tangent vector of γ (or γ').

As an example for the interaction energy, consider two curves γ_1 and γ_2. For simplicity of notation, we still use the un-regularized energy expression in Eq. (12). The total energy is

$$E = \frac{1}{8\pi} \int_{\gamma_1 \cup \gamma_2} \int_{\gamma_1' \cup \gamma_2'} \frac{d\mathbf{l} \cdot d\mathbf{l}'}{r} = \frac{1}{8\pi} \int_{\gamma_1} \int_{\gamma_1'} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_1'}{r} + \frac{1}{8\pi} \int_{\gamma_2} \int_{\gamma_2'} \frac{d\mathbf{l}_2 \cdot d\mathbf{l}_2'}{r} + \frac{1}{4\pi} \int_{\gamma_1} \int_{\gamma_2} \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{r}, \tag{15}$$

where γ_1' and γ_2' are γ_1 and γ_2 using parameters with prime, $d\mathbf{l}_1$ and $d\mathbf{l}_2$ are line elements of γ_1 and γ_2, respectively. The first two terms are the self-energy of the two curves, respectively, which are proportional to their length. The third term is the interaction energy between them, which depends on their line directions, their relative position and their length. This interaction energy between any two points on the two curves, respectively, is proportional to $1/r$, where r is the distance between them, which decays very slowly as r goes to infinity, and might give a strong interaction between two long curves even if they are far away from each other. In this sense, this interaction has the long-range nature, as the interaction between dislocations [29–32]. This interaction energy between any two points on the two curves, respectively, also depends on their local directions. The interaction is strongly attractive when they have opposite directions (when the energy is minimum among all possible relative directions), and is strongly repulsive when they have the same direction (when the energy is maximum among all possible relative directions). The orientation-dependence nature of the interaction is also very important, which will enable the moving contour to find the correct part of the object boundary (see Fig. 5 for an example). Both the long-range and orientation-dependent natures give a large capture range of the moving contour, as will be shown by the analysis in Section 3 and the experimental results in Section 6.

If we allow $\gamma(s)$ to move in order to minimize $E(\gamma)$, the dynamics equation in the steepest descent direction using an artificial time t is

$$\gamma_t = \mathbf{w} \times \boldsymbol{\tau}. \tag{16}$$

Note that now γ, $\mathbf{w}$ and $\boldsymbol{\tau}$ depend on time t.

We now derive Eq. (16) using the exact delta function. Consider the variation of $E(\gamma)$ with respect to γ. Rewrite γ as $\gamma(q)$ where $q \in [0,1]$. Let $\gamma(u,q)$ be a family of closed curves with $u \in [-\epsilon, \epsilon]$ where ϵ is a small number, and $\gamma(0,q) = \gamma(q)$. Since the unit tangent vector at a point on the curve $\gamma(u,q)$ is $\boldsymbol{\tau} = \gamma_q(u,q)/|\gamma_q(u,q)|$, and $\delta(\gamma) \, dx \, dy \, dz = ds = |\gamma_q(u,q)| dq$, from Eq. (8), we have

$$
\begin{aligned}
\frac{\mathrm{d}E(\mathbf{w},\lambda,\gamma(u,q))}{\mathrm{d}u} &= -\frac{\mathrm{d}}{\mathrm{d}u}\int_0^1 \lambda(\gamma(u,q))\cdot\gamma_q(u,q)\mathrm{d}q \\
&= -\int_0^1 \frac{\mathrm{d}}{\mathrm{d}u}\lambda(\gamma(u,q))\cdot\gamma_q(u,q)\mathrm{d}q - \int_0^1 \lambda(\gamma(u,q))\cdot\gamma_{qu}(u,q)\mathrm{d}q \\
&= -\int_0^1 \frac{\mathrm{d}}{\mathrm{d}u}\lambda(\gamma(u,q))\cdot\gamma_q(u,q)\mathrm{d}q + \int_0^1 \frac{\mathrm{d}}{\mathrm{d}q}\lambda(\gamma(u,q))\cdot\gamma_u(u,q)\mathrm{d}q \\
&= -\int_0^1 [J(\lambda(\gamma(u,q)))\gamma_u(u,q)]\cdot\gamma_q(u,q)\mathrm{d}q + \int_0^1 [J(\lambda(\gamma(u,q)))\gamma_q(u,q)]\cdot\gamma_u(u,q)\mathrm{d}q \\
&= -\int_0^1 [J^{\mathrm{T}}(\lambda(\gamma(u,q)))\gamma_q(u,q)\cdot\gamma_u(u,q)]\mathrm{d}q + \int_0^1 [J(\lambda(\gamma(u,q)))\gamma_q(u,q)]\cdot\gamma_u(u,q)\mathrm{d}q \\
&= \int_0^1 [(J(\lambda(\gamma(u,q))) - J^{\mathrm{T}}(\lambda(\gamma(u,q))))\gamma_q(u,q)]\cdot\gamma_u(u,q)\mathrm{d}q,
\end{aligned}
\tag{17}
$$

where $J(\lambda)$ is the Jacobian matrix of the function λ, and $J^{\mathrm{T}}(\lambda)$ is its transpose.

Since $\gamma(0,q)=\gamma(q)$, we have

$$
\begin{aligned}
\left.\frac{\mathrm{d}E(\mathbf{w},\lambda,\gamma(u,q))}{\mathrm{d}u}\right|_{u=0} &= \int_0^1 [(J(\lambda(\gamma(q))) - J^{\mathrm{T}}(\lambda(\gamma(q))))\gamma_q(q)]\cdot\gamma_u(q)\mathrm{d}q. \\
&= \int_0^{L(\gamma(s))}\left[(J(\lambda(\gamma(s))) - J^{\mathrm{T}}(\lambda(\gamma(s))))\frac{\gamma_q(s)}{|\gamma_q(s)|}\right]\cdot\gamma_u(s)\mathrm{d}s,
\end{aligned}
\tag{18}
$$

where $L(\gamma(s))$ is the length of $\gamma(s)$.

Therefore the dynamics equation for the energy minimization in the steepest descent direction is

$$
\gamma_t = -[J(\lambda) - J^{\mathrm{T}}(\lambda)]\frac{\gamma_q}{|\gamma_q|} = -[J(\lambda) - J^{\mathrm{T}}(\lambda)]\boldsymbol{\tau}.
\tag{19}
$$

It is easy to verify that

$$
[J(\lambda) - J^{\mathrm{T}}(\lambda)]\boldsymbol{\tau} = (\nabla\times\lambda)\times\boldsymbol{\tau}.
\tag{20}
$$

Then using the first equation in (9), we get Eq. (16).

Under this velocity field, the energy is reduced by the shrinking of $\gamma(s)$, attraction and annihilation of segments of $\gamma(s)$ with opposite directions, and repulsion of segments with same direction, see the expression of the total energy of two curves given by Eq. (15) for an example. The properties of the velocity field will be analyzed in details in Section 3.

2.2. The new active contour model

In the active contour method for image segmentation [1], a curve is evolved to find the object boundaries. As in [4,5] and many other works, we use the level set framework [6] to represent and evolve the curve. The level set method handles the topological changes automatically during the evolution of the curve, and is implemented conveniently on a uniform grid.

In the level set method, two dimensional curves in a plane are represented by the zero level set of a level set function $\phi(x,y)$ defined everywhere in the plane. The evolution of the curves is implicitly determined by the evolution of the level set function, which is

$$
\phi_t + \mathbf{v}\cdot\nabla\phi = 0,
\tag{21}
$$

where $\mathbf{v}(x,y)$ is the velocity of the curves extended to the whole plane. In our new active contour model, the velocity comes from the elastic interaction between the boundaries of objects and the moving curve defined in the previous subsection. The object boundaries are fixed while the curve moves to find the object boundaries by minimizing the total energy (see Fig. 2(a) for an example).
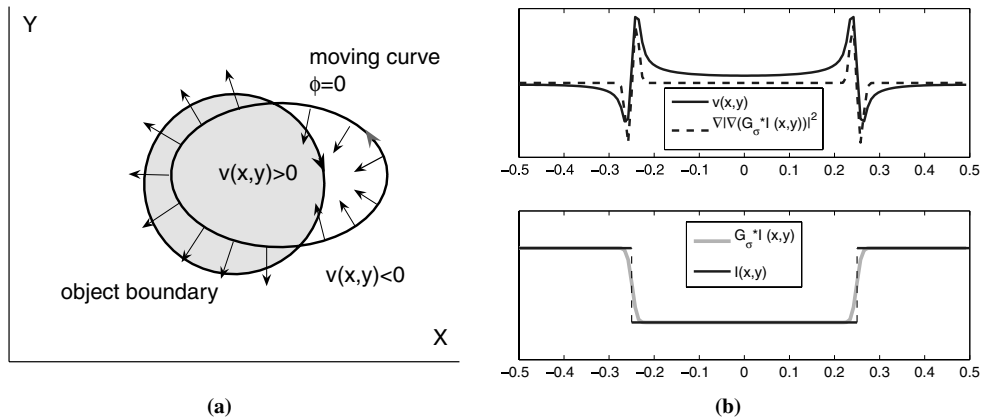
Fig. 2. An example of our normal velocity $v(x, y)$. (a) The object, the moving curve, and the direction of the velocity of the moving curve (shown by the arrows along the curve). (b) The normal velocity field $v(x, y)$ in our method and comparison with the outer radial velocity generated by $\nabla|\nabla(G_\sigma * I)|^2$ (top image), and the image function $I(x, y)$ and its smoothed version $G_\sigma * I(x, y)$ (bottom image) along a line passing through the center of the disc object shown in (a). The standard deviation σ used for the smoothed image is $1/32$ of the radius of the disc.

To get the velocity field, both the objects and the moving curve are placed in the $z = 0$ plane as described in the previous subsection. The energy associated with the object boundaries and the moving curve is defined by Eqs. (6) and (7), where γ is the collection of all the object boundaries and the moving curve. We first consider the object boundaries. If $I(x, y)$ is the image grey-level function, the direction of the object boundary is defined as

$$\tau = \frac{\nabla(G_\sigma * I)}{|\nabla(G_\sigma * I)|} \times \mathbf{k}, \tag{22}$$

where $\mathbf{k} = (0, 0, 1)$ is the unit vector in z direction, G_σ is the two dimensional Gaussian function with mean zero and standard deviation σ, which is used to smooth the object boundaries, and the delta function in Eq. (7) is given by

$$\delta(\gamma) = |\nabla(G_\sigma * I)|\delta(z), \tag{23}$$

where $\delta(z)$ is the one dimensional delta function of z. We use the convention that the range of the intensity function $I(x, y)$ is $[0, 1]$, and $I(x, y) = 0$ when it is black and $I(x, y) = 1$ when it is white. Under this convention, if the image of an object is darker than the background, the direction of the object boundary is clockwise; and is counterclockwise, vice versa. As such, for the contribution of the object boundaries in Eq. (7),

$$\delta(\gamma)\tau = \nabla(G_\sigma * I) \times \mathbf{k}\delta(z). \tag{24}$$

Note that this equation holds when the difference between the intensity outside and inside the object is 1. For other values of the difference, the right-hand side of the above equation is a delta function multiplied by a constant, but it does not change the nature of the interaction. So we simply use Eq. (24) for the right-hand side of Eq. (7) generated by the image.

Similarly, the tangent vector of the moving curve represented by the zero level set of ϕ is

$$\tau = \frac{\nabla\phi}{|\nabla\phi|} \times \mathbf{k}, \tag{25}$$

and the delta function of the moving curve is

$$\delta(\gamma) = \delta(\phi(x, y))|\nabla\phi(x, y)|\delta(z). \tag{26}$$

The direction of the moving curve is chosen such that the moving curve has an opposite direction with the object boundary, so that when they stay together, their effects cancel out and the energy is minimized. As such, for the contribution of the moving curve in Eq. (7),

$$\delta(\gamma)\boldsymbol{\tau} = \delta(\phi)\nabla\phi \times \mathbf{k}\delta(z) = \nabla H(\phi) \times \mathbf{k}\delta(z), \tag{27}$$

where $H(x)$ is a one dimensional (regularized) Heaviside function. Using a linear combination of Eqs. (24) and (27) for the right-hand side of Eq. (7), we have

$$\nabla \times \mathbf{w} = \nabla(G_\sigma * I + \alpha H(\phi)) \times \mathbf{k}\delta(z), \tag{28}$$

where $\alpha > 0$ is a constant to match the strength of the effects of the object boundary and the moving curve (note that the difference between the intensities inside and outside the object may not be *1* exactly; while the Heaviside function always gives a difference *1* between inside and outside the moving curve). We shall explain in the next section that α is also an important parameter for noise removal. We use the following regularized Heaviside function which has the same regularization parameter σ as that in image smoothing

$$H(\phi) = \begin{cases} 0 & \text{if } \phi \leqslant -\sigma, \\ \frac{1}{2}\left(\sin\left(\frac{\pi\phi}{2\sigma}\right) + 1\right) & \text{if } -\sigma < \phi < \sigma, \\ 1 & \text{if } \phi \geqslant \sigma. \end{cases} \tag{29}$$

The regularized delta and Heaviside functions have been used in many applications of the level set method (e.g., [17,39]).

Using Eq. (11), it can be verified that $w_1(x,y,0) = w_2(x,y,0) = 0$. As such, by Eq. (16), the velocity of the moving curve is

$$\mathbf{v}(x,y,0) = w_3\mathbf{k} \times \left(\frac{\nabla\phi}{|\nabla\phi|} \times \mathbf{k}\right) = w_3\frac{\nabla\phi}{|\nabla\phi|} = \mathbf{w} \cdot \mathbf{k}\frac{\nabla\phi}{|\nabla\phi|}. \tag{30}$$

Using Eqs. (13) and (28), we have

$$\begin{aligned} \mathbf{v}(x,y) &= -\left(\frac{1}{4\pi}\int_{\mathbf{R}^2}\frac{\mathbf{r} \times (\nabla(G_\sigma * I + \alpha H(\phi)) \times \mathbf{k}) \cdot \mathbf{k}}{r^3}dx'dy'\right)\frac{\nabla\phi}{|\nabla\phi|} \\ &= \left(\frac{1}{4\pi}\int_{\mathbf{R}^2}\frac{\mathbf{r} \cdot \nabla(G_\sigma * I + \alpha H(\phi))}{r^3}dx'dy'\right)\frac{\nabla\phi}{|\nabla\phi|}, \end{aligned} \tag{31}$$

where $\mathbf{r} = (x - x', y - y')$ and $r = \sqrt{(x-x')^2 + (y-y')^2}$. Note that although we derive this formula in three dimensions, the formula itself is still in two dimensions.

It is convenient to write the evolution equation for the level set function ϕ as

$$\phi_t = v|\nabla\phi|, \tag{32}$$

where the velocity in the normal direction of the moving curve v is

$$v(x,y) = -\mathbf{v} \cdot \frac{\nabla\phi}{|\nabla\phi|} = -\frac{1}{4\pi}\int_{\mathbf{R}^2}\frac{\mathbf{r} \cdot \nabla(G_\sigma * I + \alpha H(\phi))}{r^3}dx'dy'. \tag{33}$$

Note that the direction $\nabla\phi/|\nabla\phi|$ is perpendicular to the moving curve, which is the zero level contour of ϕ. For convenience, we call $-\nabla\phi/|\nabla\phi|$ the normal direction of the moving curve.

Under this velocity field, the curve will move to find the object boundaries by minimizing the total energy, which is

$$E = \frac{1}{8\pi}\int_{\mathbf{R}^2}dx\,dy\int_{\mathbf{R}^2}\frac{\nabla(G_\sigma * I + \alpha H(\phi))(x,y) \cdot \nabla(G_\sigma * I + \alpha H(\phi))(x',y')}{r}dx'dy'. \tag{34}$$

If the domain of an image is Ω, then the velocity in the normal direction and the total energy are

$$v(x,y) = -\frac{1}{4\pi}\int_\Omega\frac{\mathbf{r} \cdot \nabla(G_\sigma * I + \alpha H(\phi))}{r^3}dx'dy', \tag{35}$$

and

$$E = \frac{1}{8\pi}\int_\Omega dx\,dy\int_\Omega\frac{\nabla(G_\sigma * I + \alpha H(\phi))(x,y) \cdot \nabla(G_\sigma * I + \alpha H(\phi))(x',y')}{r}dx'dy', \tag{36}$$

respectively, where $\mathbf{r} = (x - x', y - y')$ and $r = \sqrt{(x-x')^2 + (y-y')^2}$.

We can add a small curvature term to smooth the moving curve, then the evolution equation can be written as

$$\phi_t = \left(\mu \nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + v \right) |\nabla \phi|, \tag{37}$$

where μ is a small positive constant, and v is given by Eq. (35). This is our new active contour model for image segmentation. Its properties will be analyzed and illustrated in details in the next section.

3. Properties of the interaction

In this section, we analyze the properties of the interaction, which gives the normal velocity v of the moving curve in Eq. (33). We first consider the direction and magnitude of the velocity generated by the object alone, i.e., $\alpha = 0$ in Eq. (33), which is

$$v = -\frac{1}{4\pi} \int_{\mathbf{R}^2} \frac{\mathbf{r} \cdot \nabla(G_\sigma * I)}{r^3} \, dx' dy'. \tag{38}$$

We then discuss how to remove noise by using a nonzero α.

3.1. Long-range attractive interaction

We first show that the moving curve is attracted to the object boundary for a single object under the normal velocity given by Eq. (38). Consider an object which is darker than the background, as shown in Fig. 2(a). For any point (x,y) in the image plane, it is easy to obtain using Green's formula that

$$v(x,y) = \frac{1}{4\pi} \int_{\mathbf{R}^2} \frac{(G_\sigma * I - c_b)}{r^3} \, dx' dy', \quad \text{when } (x,y) \text{ is outside the object}, \tag{39}$$

and

$$v(x,y) = \frac{1}{4\pi} \int_{\mathbf{R}^2} \frac{(G_\sigma * I - c_o)}{r^3} \, dx' dy', \quad \text{when } (x,y) \text{ is inside the object}, \tag{40}$$

where $r = \sqrt{(x - x')^2 + (y - y')^2}$, and c_o and c_b are the intensities of the object and background, respectively. Note that to remove the singularity when $(x',y') = (x,y)$, c_0 and c_b appear in the above equations.

For the image shown in Fig. 2(a), $c_o < c_b$. So the direction of the object boundary is clockwise from Eq. (22). The level set function ϕ is chosen such that $\nabla \phi$ is in the inner normal direction of the moving curve, which makes the direction of the moving curve counterclockwise from Eq. (25). From Eqs. (39) and (40), $v(x,y) < 0$ when (x,y) is outside the object, and $v(x,y) > 0$ when (x,y) is inside the object. It is easy to see that the part of the moving contour outside the object shrinks, and the part of the moving contour inside the object expands: both parts are attracted to the object boundary, as shown in Fig. 2(a). Note that we use the convention that a positive v means a velocity in the $-\nabla \phi/|\nabla \phi|$ direction (pointing outward in this example).

Now we consider the magnitude of the velocity. The top image in Fig. 2(b) shows the normal velocity $v(x,y)$ given by Eq. (38) along a line passing through the center of the disc object in Fig. 2(a). The image function $I(x,y)$ and its smoothed version $G_\sigma * I(x,y)$ are plotted in the bottom image in Fig. 2(b). For convenience, we have normalized the velocity using its maximum magnitude. We can see that as described above, the velocity is negative outside the object, positive inside the object, and zero on the object boundary. The magnitude of the velocity is large near the object boundary, which reflects the strong attractive interaction to it. Away from the object boundary, the velocity decays but never becomes zero, which gives an extension of the strong attractive interaction near the object boundary. These features of the velocity field are highly desired in the edge-based active contour methods.

Fig. 2(b) also shows the comparison between our velocity and $\nabla |\nabla(G_\sigma * I)|^2$ in the outer radial direction of the disc, which has also been normalized for convenience. The latter was used in the traditional active contour methods (see Eqs. (2) and (4)). We can see that both velocities have strong attractive effects near the object boundary, but our velocity has an influence over the whole region. While for the $\nabla |\nabla(G_\sigma * I)|^2$, its influence

is only localized near the object boundary, which requires that initially the moving curve must be placed very close to the object boundary.

Using the asymptotic expansion method, we can obtain the quantitative behavior of the velocity in our method when (x,y) approaches the object boundary or ∞. The velocity field Eq. (11) or (38) is the same as the outer solution of the Ginzburg–Landau equation in superconductivity [33–36]. Using the results for the solution of the Ginzburg–Landau equation [35], the limit of the velocity as the point (x,y) approaches the object boundary, but remains outside the region in which the delta function of the boundary is regularized, is

$$v = (c_\mathrm{b} - c_\mathrm{o})\left[\pm\frac{1}{2\pi\sigma} \pm \frac{\kappa}{4\pi}\log\sigma\right] + \mathrm{O}(1), \tag{41}$$

where $\sigma \ll 1$ and 2σ is the width of the region in which the delta function of the object boundary is regularized, κ is the curvature of the object boundary, the positive or negative sign before the first term depends on whether (x,y) approaches the object boundary from inside or outside the object, and the positive or negative sign before the second term depends on whether the object is concave or convex at the nearest point on the object boundary. This limit is approximately the maximum magnitude of the velocity near the object boundary. For the asymptotic property at ∞, since the integrand in the integral expression of the velocity given by Eq. (39) decays at $\mathrm{O}(1/R^3)$, where $R = \sqrt{x^2 + y^2}$, the normal velocity v as (x,y) approaches ∞ is approximately of an order between $O(1/R^3)$ and $\mathrm{O}(1/R)$, depending on the area of the object (Note that the integrand is zero outside the boundary in Eq. (39)). This also shows that the interaction is long-range, and by which the moving contour can "see" the far away object boundary.

3.2. Velocity near the object boundary

We have shown by an example that the velocity given by Eq. (38) has a very strong attractive effect to the object boundary. Now we consider this behavior quantitatively.

We consider a point P' on an object boundary γ. Let $(\mathbf{l}, \mathbf{n})$ be the Frenet coordinates at P', where $\mathbf{l}$ and $\mathbf{n}$ are unit vectors in the tangent and normal directions of γ, respectively. We rotate the coordinate system so that it coincides with $(\mathbf{l}, \mathbf{n})$ at P'. We assume that the delta function of the object boundary γ is regularized within a narrow band along γ with width 2σ. If we use a Gaussian function to smooth the image, then σ is its standard deviation and the contribution is small outside this narrow band, so the results obtained from a compact-support regularized delta function still hold. We assume that the regularized delta function of the object boundary $\delta(x) = \mathrm{O}(1/\sigma)$, $\delta'(x) < 0$ for $x > 0$, $\delta'(x) > 0$ for $x < 0$, and $\delta(x)$ is symmetric with respect to $x = 0$. We consider the normal velocity v at the point $P(0,\eta)$ within this narrow band, where $\eta \ll \sigma$, see Fig. 3.

Let γ be parameterized by a shifted arclength s such that $P' = \gamma(0)$. If κ is the curvature of γ at P' which is assumed to be O(1), a point Q' on γ and near P' has the representation

$$Q' = (\alpha(s), \beta(s)), \tag{42}$$

where $\alpha(s) = s - \frac{1}{6}\kappa^2 s^3 + \cdots$ and $\beta(s) = \frac{1}{2}\kappa s^2 + \frac{1}{6}\kappa' s^3 + \cdots$, and a point Q inside this narrow band and near P' can be written as
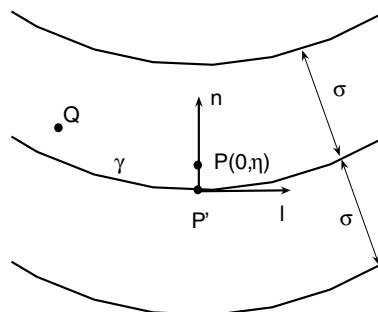

Fig. 3. The Frenet coordinates at a point P' on the object boundary γ.

$$Q = (\alpha(s), \beta(s)) + u(-\beta'(s), \alpha'(s))/\sqrt{\alpha'(s)^2 + \beta'(s)^2}, \tag{43}$$

where u is a parameter in the interval $[-\sigma, \sigma]$. Therefore for the integral expression Eq. (38) of the normal velocity v at P, we have

$$\mathbf{r} = QP = \left(-s + \left(\kappa s + \frac{1}{2}\kappa' s^2\right)u, \eta - \frac{1}{2}\kappa s^2 - \left(1 - \frac{1}{2}\kappa^2 s^2\right)u\right) + O(s^3), \tag{44}$$

and

$$\nabla(G_\sigma * I) = (c_{\mathrm{b}} - c_{\mathrm{o}})\delta(u)\left(\kappa s + \frac{1}{2}\kappa' s^2, -1 + \frac{1}{2}\kappa^2 s^2\right) + O(s^3/\sigma). \tag{45}$$

Note that $\delta(u) = O(1/\sigma)$. We have assumed that the object is convex at the point P', otherwise there is a minus sign in Eq. (45). Let $v(P) = v_1(P) + v_2(P)$, where $v_1(P)$ is the integral on a segment of γ where $s \in [-1, 1]$, and $v_2(P)$ is the integral on the other part of γ. We have

$$v_1(P) = \frac{c_{\mathrm{b}} - c_{\mathrm{o}}}{4\pi}\int_{-1}^{1}\mathrm{d}s\int_{-\sigma}^{\sigma}\mathrm{d}u\,\frac{\delta(u)\left(\eta - u + \frac{1}{2}\kappa s^2\right) + O(s^2) + O(s^3/\sigma)}{[s^2 + (\eta - u)^2]^{3/2}(1 + O(s))}. \tag{46}$$

Now we look for the leading order terms in $v_1(P)$. We have

$$\int_{-1}^{1}\mathrm{d}s\int_{-\sigma}^{\sigma}\mathrm{d}u\,\frac{\delta(u)(\eta - u)}{[s^2 + (\eta - u)^2]^{3/2}} = \int_{-\sigma}^{\sigma}\frac{2\delta(u)}{(\eta - u)\sqrt{1 + (\eta - u)^2}}\,\mathrm{d}u = 2\pi\mathbf{H}[\delta](\eta) + O(\sigma), \tag{47}$$

where $\mathbf{H}[\delta](x) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\delta(u)}{x - u}\,\mathrm{d}u$ is the Hilbert transform of the function $\delta(x)$, and

$$
\begin{aligned}
\int_{-1}^{1}\mathrm{d}s\int_{-\sigma}^{\sigma}\mathrm{d}u\,\frac{\frac{1}{2}\kappa s^2\delta(u)}{[s^2 + (\eta - u)^2]^{3/2}} &= \int_{-\sigma}^{\sigma}\kappa\delta(u)\left[\log\left(\sqrt{1 + (\eta - u)^2} + 1\right) - \log|\eta - u| - \frac{1}{\sqrt{1 + (\eta - u)^2}}\right]\mathrm{d}u \\
&= -\kappa\int_{-\sigma}^{\sigma}\delta(u)\log|\eta - u|\,\mathrm{d}u + O(1) \\
&= -\kappa\int_{-\sigma}^{\sigma}\log|\eta - u|\,\mathrm{d}(H(u) - H(\eta)) + O(1) \\
&= -\kappa\left[(1 - H(\eta))\log|\sigma - \eta| + H(\eta)\log|\sigma + \eta| - \int_{-\sigma}^{\sigma}\frac{H(u) - H(\eta)}{u - \eta}\,\mathrm{d}u\right] + O(1) \\
&= -\kappa\log\sigma + O(1),
\end{aligned} \tag{48}
$$

where $H(x)$ is a regularized Heaviside function whose derivative is the regularized delta function $\delta(x)$ used above. We have used the mean value theorem and the fact that $\delta(x) = O(1/\sigma)$ in the above equation.

It is not difficult to verify that the contribution to $v_1(P)$ from all the error terms in the integrand is $O(1)$. Therefore, we have

$$v_1(P) = (c_{\mathrm{b}} - c_{\mathrm{o}})\left(\frac{1}{2}\mathbf{H}[\delta](\eta) - \frac{\kappa}{4\pi}\log\sigma\right) + O(1). \tag{49}$$

Since we have assumed that the regularized delta function $\delta(x)$ is symmetric with respect to $x = 0$, we have $\mathbf{H}[\delta](0) = 0$ and

$$v_1(P) = (c_{\mathrm{b}} - c_{\mathrm{o}})\left(\frac{1}{2}\mathbf{H}[\delta]'(0)\eta - \frac{\kappa}{4\pi}\log\sigma\right) + O(1), \tag{50}$$

where

$$\mathbf{H}[\delta]'(0) = O(1/\sigma^2) \tag{51}$$

is the derivative of $\mathbf{H}[\delta](u)$ at $u = 0$.

Since $v_2(P) = O(1)$, we have

$$v(P) = (c_b - c_o)\left(\frac{1}{2}\mathbf{H}[\delta]'(0)\eta - \frac{\kappa}{4\pi}\log\sigma\right) + O(1) \tag{52}$$

when $\eta \ll \sigma \ll 1$. In the original coordinate system, η is the signed distance to the object boundary γ (positive inside). If the object boundary is concave at P', the sign of the second term in the above equation is changed to positive.

In conclusion of this subsection, using Eq. (52), when the moving curve is close to the object boundary, the full normal velocity Eq. (33) generated by both the object boundary and the moving curve is

$$v = (c_b - c_o)\frac{1}{2}\mathbf{H}[\delta]'(0)\eta - (c_b - c_o)\frac{\kappa}{4\pi}\log\sigma + \alpha\frac{\kappa}{4\pi}\log\sigma + O(1), \tag{53}$$

where η is the signed distance to the object boundary, and $\eta \ll \sigma \ll 1$. Since $\mathbf{H}[\delta]'(0) = O(1/\sigma^2)$, the first term, which represents the strong attractive interaction to the object boundary, dominates near the object boundary. If we choose $\alpha = c_b - c_o$, then the leading order terms in the normal velocity cancel out when the moving curve stays exactly on the object boundary, i.e., $\eta = 0$. However, since the first term dominates in the above equation, we do not need to choose α to be $c_b - c_o$ exactly. This will allow us to use one α for multiple objects, and will also help to remove noise, see next subsection.

3.3. Noise removal

Using the normal velocity in Eq. (33), the image noise also generates a normal velocity for the moving curve, which attracts the moving curve and results in spurious contours. However, either (a) the intensity difference $|c_b - c_o|$ between outside and inside the noise is relatively small compared with that across the object boundary, so that the term $\alpha\frac{\kappa}{4\pi}\log\sigma$ in the local expansion of the velocity near noise (Eq. (53)) dominates (because α is chosen to be of the same order of the intensity difference across the object boundary), or (b) the size of the noise is small so that curvature term controlled by the parameter μ dominates in the evolution equation (37). Both will force the spurious contours caused by noise to shrink and disappear.

4. Numerical formulation

4.1. Numerical formulation for the velocity field

The normal velocity field v (Eq. (35)) is an integral over the whole image domain. We use the fast Fourier transform (FFT) method to compute it efficiently.

Eq. (35) is the convolution of the two vector functions $\nabla(G_\sigma * I(x, y) + \alpha H(\phi(x, y)))$ and $\nabla\frac{1}{R}$ with a factor $1/4\pi$, where $R = \sqrt{x^2 + y^2}$. Assuming that the Fourier transform of $G_\sigma * I(x, y) + \alpha H(\phi(x, y))$ is d_{mn}, where m and n are frequencies in the Fourier space, we have

$$\nabla(\widehat{G_\sigma * I + \alpha H(\phi)}) = (im, in)d_{mn}. \tag{54}$$

Since

$$\widehat{\frac{1}{R}} = \frac{1}{2\pi\sqrt{m^2 + n^2}}, \tag{55}$$

we have

$$\widehat{\nabla\frac{1}{R}} = \frac{(im, in)}{2\pi\sqrt{m^2 + n^2}}. \tag{56}$$

Therefore the Fourier transform of the normal velocity is

$$\widehat{v} = \frac{1}{4\pi}(2\pi)^2\frac{(im, in)}{2\pi\sqrt{m^2 + n^2}} \cdot (im, in)d_{mn} = -\frac{\sqrt{m^2 + n^2}}{2}d_{mn}, \tag{57}$$

where d_{mn} is the Fourier transform of $G_\sigma * I(x, y) + \alpha H(\phi(x, y))$. A calculation for the Fourier transform of $1/R$ (Eq. (55)) can be found in Appendix.

4.2. Implementation of the evolution equation

Regarding the numerical implementation for the evolution equation of the level set function ϕ (Eq. (37)), we use the central difference for the curvature term, Godunov's scheme [40,41] combined with the WENO derivative [42] for the term $v|\nabla\phi|$, and the forward Euler method in time. To reduce the numerical errors, re-initialization [43] is used to maintain an accurate signed distance function near the zero level set of the level set function ϕ during its evolution.

4.3. Initial curve

The object boundaries alone generate a velocity field

$$v_0(x,y) = -\frac{1}{4\pi} \int_\Omega \frac{\mathbf{r} \cdot \nabla(G_\sigma * I)}{r} \mathrm{d}x'\mathrm{d}y', \tag{58}$$

which has different signs inside and outside the objects, as discussed in Section 3. We can let

$$\phi(x,y) = v_0/|v_0|, \tag{59}$$

for the initial value of the level set function ϕ. This gives an initial contour very close to the object boundary, but in the mean time, the noise is also kept. The object boundary can then be found and the noise be removed after short-time evolution of ϕ.

4.4. Normalized velocity

A normalized velocity field $v(x,y)/|v(x,y)|$ can be used to accelerate the evolution in the early stage. When the motion of the curve converges using the normalized velocity, the moving curve should get close to the object boundary and most of the noise disappears. Then we can use the un-normalized velocity field to evolve the curve for a few steps to find the object boundary accurately. Such process greatly reduces the computation time and does not change the final result.

5. Three dimensional model

In this section, we extend our new active contour model for segmentation to three dimensions, in which the object boundaries are surfaces and a surface moves under some velocity field to find the object boundaries. We have not found a physically meaningful energy for surfaces that has the similar properties to that of the dislocations in solids. It is also difficult to define an energy for surfaces similar to that for lines given by Eqs. (6) and (7). Instead, we directly extend the normal velocity field in two dimensions for the moving curves Eq. (33) to three dimensions for the moving surfaces, and show that it has the same properties as the elastic interaction in two dimensions.

Using the level set method, the moving surface is represented by the zero level set of a function $\phi(x, y, z)$, and the evolution equation of ϕ is also given by Eq. (37), in which now the gradient is taken in three dimensions.

We define the normal velocity field for the moving surface as

$$v(x,y,z) = -\int_{\mathbf{R}^3} \frac{\mathbf{r} \cdot \nabla(G_\sigma * I + \alpha H(\phi))}{r^4} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z', \tag{60}$$

where $\mathbf{r} = (x - x', y - y', z - z')$ and $r = \sqrt{(x - x')^2 + (y - y')^2 + (z - z')^2}$.

It is easy to verify that this velocity field minimizes an energy similar to that in two dimensions Eq. (34). The energy in three dimensions is

$$E = \int_{\mathbf{R}^3} \mathrm{d}x\mathrm{d}y\mathrm{d}z \int_{\mathbf{R}^3} \frac{\nabla(G_\sigma * I + \alpha H(\phi))(x,y,z) \cdot \nabla(G_\sigma * I + \alpha H(\phi))(x',y',z')}{r^2} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z'. \tag{61}$$

Similar to that in two dimensions, the normal velocity generated by an object alone ($\alpha = 0$ in Eq. (60)) can be written as

$$v(x,y,z) = \int_{\mathbf{R}^3} \frac{(G_\sigma * I - c_\mathrm{b})}{r^4} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z', \quad \text{when}(x,y,z) \text{ is outside the object}, \tag{62}$$

and

$$v(x,y,z) = \int_{\mathbf{R}^3} \frac{(G_\sigma * I - c_\mathrm{o})}{r^4} \mathrm{d}x'\mathrm{d}y'\mathrm{d}z', \quad \text{when}(x,y,z) \text{ is inside the object}, \tag{63}$$

where c_o and c_b are the intensities of the object and background, respectively. So this interaction is long-range and the moving surface is attracted to the object boundary under this normal velocity, as demonstrated in the two dimensional case.

When the moving surface is very close to the object boundary, the leading order terms in the normal velocity generated by a single object is

$$v = (c_\mathrm{b} - c_\mathrm{o})\left(2\pi^2 \mathbf{H}[\delta]'(0)\eta - \frac{\pi}{2}\kappa \log \sigma\right) + \mathrm{O}(1), \tag{64}$$

where η is the signed distance to the object boundary (positive inside), κ is the mean curvature at the nearest point on the surface, and $\eta \ll \sigma \ll 1$. The calculation is similar to that in two dimensions. Having this expansion, the choice of the parameter α and the mechanism for noise removal are the same as those in the two dimensional model.

The normal velocity Eq. (60) is the convolution of the two vector functions $\nabla(G_\sigma * I(x,y,z) + \alpha H(\phi(x,y,z)))$ and $\nabla\frac{1}{R}$, where $R = \sqrt{x^2 + y^2 + z^2}$, which can also be calculated efficiently using FFT, as the velocity in two dimensions. The formulation is

$$\widehat{v} = -2\pi^2 \sqrt{m^2 + n^2 + l^2} d_{mnl}, \tag{65}$$

where d_{mnl} is the Fourier transform of $G_\sigma * I(x,y,z) + \alpha H(\phi(x,y,z))$ with frequencies m, n, and l in directions x, y, and z, respectively. Here we have used

$$\widehat{\frac{1}{\sqrt{x^2 + y^2 + z^2}}} = \frac{1}{4\pi\sqrt{m^2 + n^2 + l^2}}, \tag{66}$$

which can be derived similarly as that shown in the appendix for the Fourier transform of $1/\sqrt{x^2 + y^2}$ in two dimensions.

6. Experimental results

We have applied our method on the synthetic images and clinical images of blood vessels. The results were obtained on 128×128 pixel two dimensional images or $128 \times 128 \times 128$ pixel three dimensional images. The size of an image was 2×2 or $2 \times 2 \times 2$ units. As such, the pixel width was $\mathrm{d}x = \mathrm{d}y = \mathrm{d}z = 2/128$. Unless specified otherwise, we chose $\sigma = 0.8 \sim 2\,\mathrm{d}x$, $\alpha = 0.3$, and $\mu = 0.002$ in the evolution equation Eq. (37) and the normal velocity field Eq. (33) or (60).

Fig. 4 shows the experimental result on an image of an object which has very thin convex feature on its right and concave feature on its left, at which these features are pointed at by the arrows, see Fig. 4(a). Both features are only one pixel width. The zero level contour of the normalized velocity equation (59) gives a good initial result, except that at the tip of the one-pixel convex, the contour width is a little bit bigger than one pixel. After a few steps of evolution from this initial contour, the object is found accurately, see Fig. 4(b). This example shows that our method can find sharp features of the objects accurately. The reason is that in our method, the
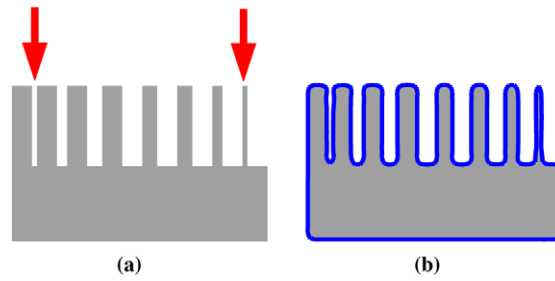
Fig. 4. Thin concave and convex. (a) The image. (b) Contour found by our method.

velocity of the moving curve is defined over the whole image domain due to the long-range interaction, which gives an extension of the effect of the image gradient to the whole image domain without smoothing the image gradient itself.

In Fig. 5, we show that our method has less sensitivity to the initialization of the moving contour. We start from a small circle intersecting with the right-hand side of a U shape object, as shown in Fig. 5(a), and evolve it using our evolution equation Eq. (37). The part of the contour outside the object is attracted to the object boundary and the part of the contour inside the object is pushed far away to approach the other side of the object boundary, see Fig. 5(d) for the direction of the velocity of the initial contour. An intermediate result after several steps of evolution is shown in Fig. 5(b). Finally, the contour converges to the object boundary, as shown in Fig. 5(c). In this example, our method has found the object with all detailed features from an initial contour that is not close to the object boundary. As a comparison, we have also applied the gradient vector flow (GVF) method [13–15] on the same U shape object with the same initial contour as shown in Fig. 5(a). The contour fails to find the object; it shrinks and finally disappears. This is because both sides of the contour
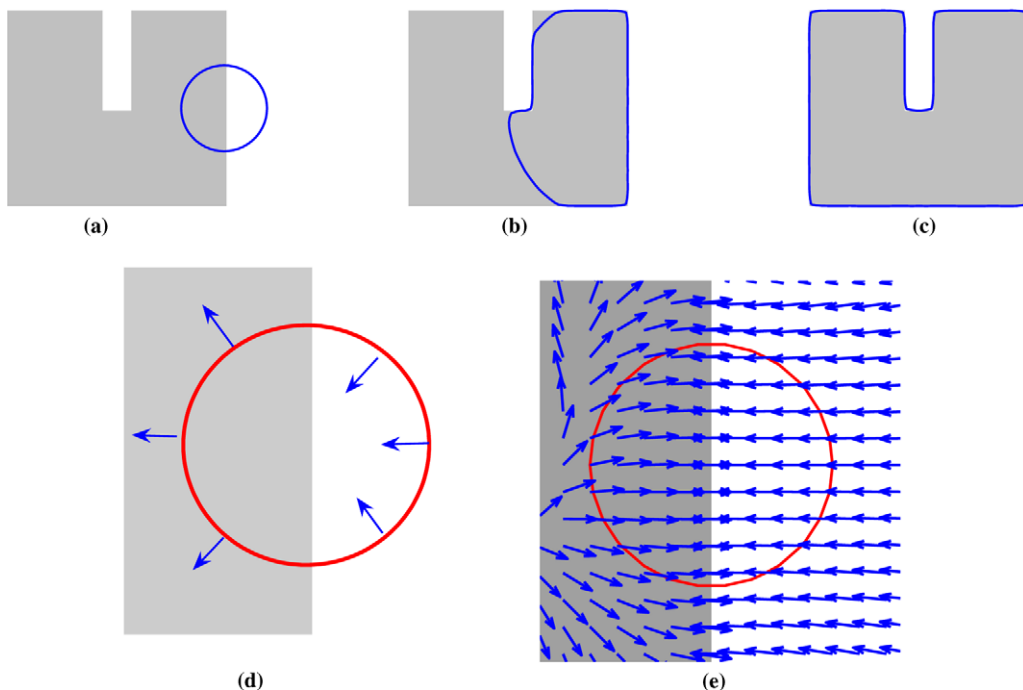

Fig. 5. Contour found from an initial contour not close to the object boundary using our method and comparison with the result of the GVF method. (a) Initial contour and object. (b) An intermediate result using our method. (c) Final result using our method. (d) Direction of the velocity of the initial contour in our method. (e) The force field of the GVF method near the object boundary and the same initial contour. The initial contour is driven to shrink and disappear under the GVF force, thus cannot find the object.

are attracted to the object boundary inside the initial contour under the GVF force, which can be seen from the force field of the GVF method near the object boundary and the initial contour plotted in Fig. 5(e). This example illustrates that our method has less sensitivity to the initialization of the moving contour than the GVF method does. Note that in this example we start from such an initial contour in our method only for the purpose of illustrating the advantage of our method. For this kind of images without noise, the objects can be found more efficiently using our method by choosing the zero level contour of the normalized velocity field Eq. (59) as the initial contour, as in the previous example.

Fig. 6 shows the results using our method and the GVF method on an image with multiple objects which have different intensity and topology. Given that the intensity values ranging between 0 (black) and 1 (white), the intensity values of the four objects are 3/9, 4/9, 5/9 and 6/9, respectively, see Fig. 6(a). For this image without noise, the zero level contour of the normal velocity field in our method Eq. (59) gives the object boundary directly and no evolution is needed, see Fig. 6(b). The normal velocity field after normalization $v(x,y)/|v(x,y)|$ is shown in Fig. 6(c), which is positive inside the objects and negative outside. The objects can also be found if we start from a large contour, as shown in Fig. 6(d), which has an opposite direction with that of the object
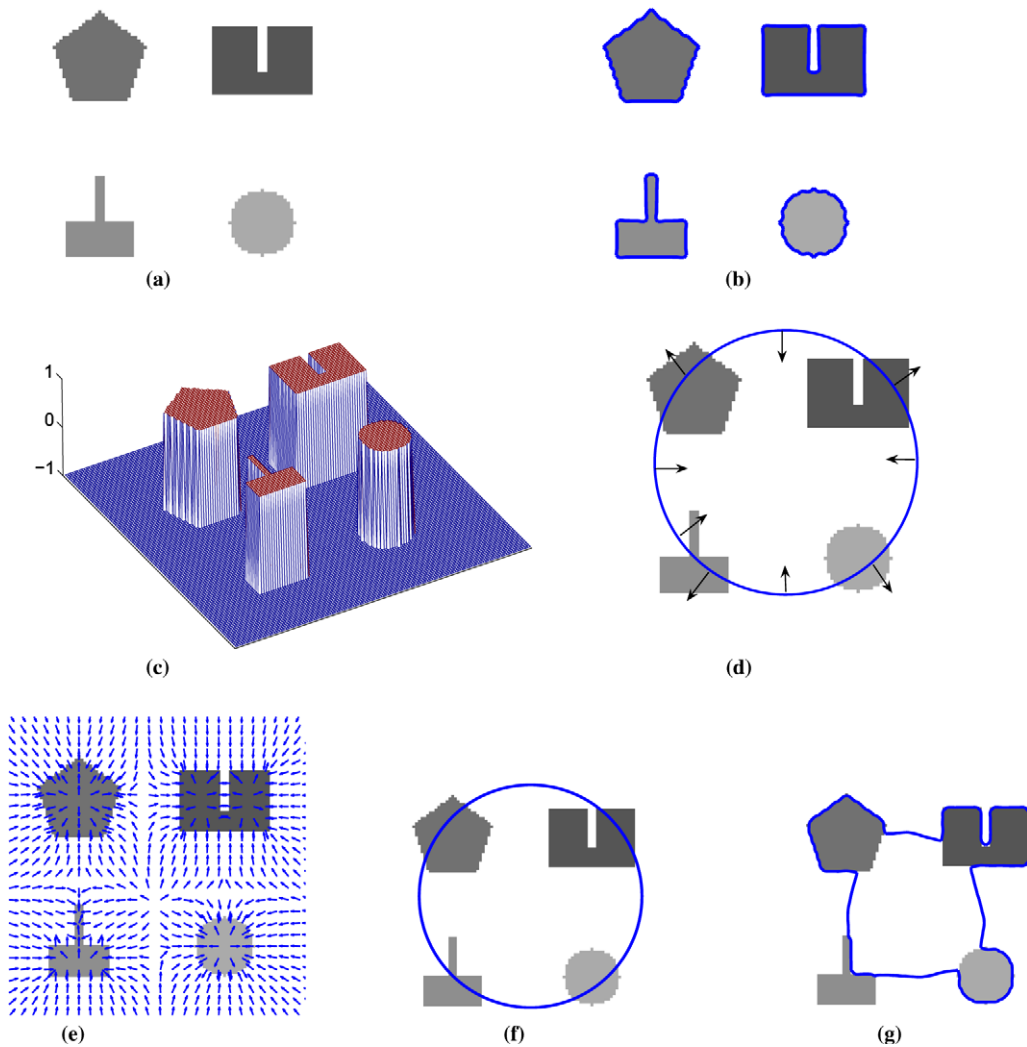

Fig. 6. (a) An input image with multiple objects having different intensity and topology. (b) Result obtained using our method. (c) The normal velocity field $v(x,y)/|v(x,y)|$ in our method. (d) The moving direction of an initial contour under the velocity field shown in (c). (e) The GVF force field. (f) Initial contour in the GVF method. (g) Result obtained using the GVF method.

boundary. (See the definition of these directions in Section 2.2. Here the image boundary is clockwise, hence the contour is counterclockwise). The positive value of the normal velocity field makes this contour move in its outer normal direction, and the negative value makes it move in its inner normal direction, as shown in Fig. 6(d). The contour moves under this normal velocity field and finally finds the objects. The result is the same as that using the zero contour of the normal velocity directly and shown in Fig. 6(b).

We have compared our result with that obtained using the GVF method on this image. Note that using the original GVF method [13] on an image with multiple objects like this example, some parts of the moving contour may be parallel to the force field, thus cannot move further to find the objects. A balloon force was added in a modified GVF method in [15] when this deadlock happens. However, this modification brings back the problem of the balloon force method, which was tried to solve by the original GVF method [13], namely the direction of the balloon force (e.g., expanding or shrinking) cannot be determined automatically within the method; it has to be prescribed thus requires careful initialization. The force field of the GVF method for this image is shown in Fig. 6(e). We have used the modified GVF method [15] on this image, with parameters $\mu = 0.02$, $\sigma_E = 0.32dx$, $\beta = 0.002$, and $\epsilon = 0.1$ (See [15] for definitions of the parameters.), and an initial contour as shown in Fig. 6(f), which is the same as that in our method. As shown by the final result in Fig. 6(g), the GVF method cannot provide a satisfactory result. A part of the contour is attracted to wrong boundary of the objects (see the bottom left object in Fig. 6(g)), and the contour cannot move further in the middle between two adjacent objects. The former has been explained in the previous example that in the GVF method, the moving curve is always attracted to the nearest object boundary. The latter can be understood by the corresponding force field of the GVF method shown in Fig. 6(e). Recall that in the GVF method, the force field is obtained by extending the local attractive force near the object boundaries to the whole image domain by diffusion. From Fig. 6(e), it is observed that in some regions near the central line between two adjacent objects, the GVF force is pointing outside (See, e.g., the region between the two objects on the right in Fig. 6(e)), which prevents the contour from moving further inside to find the objects. The balloon force added in [15] does not help because the GVF force that prevents the contour from moving inside is not parallel to the contour (Actually, the angle between them is large, see, e.g., the region between the two objects on the right in Fig. 6(e)), therefore the GVF force dominates there from their method. On the other hand, in our method, the normal velocity is always negative outside the objects, which can push the segments of the contour between two objects all the way inward to find the objects, see Fig. 6(c) and (d). This example shows that our method is less sensitive to the initial contour for images with multiple objects. It also shows that the objects can be found by the zero level contour of the normal velocity field for images without noise. For images with noise, the evolution of the contour is needed, and the zero level contour of the normal velocity field provides an initial value for the moving contour, see next example.

Our method has been applied to the same image in Fig. 6(a) but with noise, see Fig. 7(a). The standard deviation of the background noise is $\sigma_B = 0.1$, which is about 1.5 times of that obtained from the real 3D-RA images to be described in the last two examples in this section. Fig. 7(b) shows the initial contour given by the normalized velocity field Eq. (59). An intermediate stage during the evolution is shown in Fig. 7(c), and the final result is shown in Fig. 7(d). We can see that the initial contour gives a good guess of the object
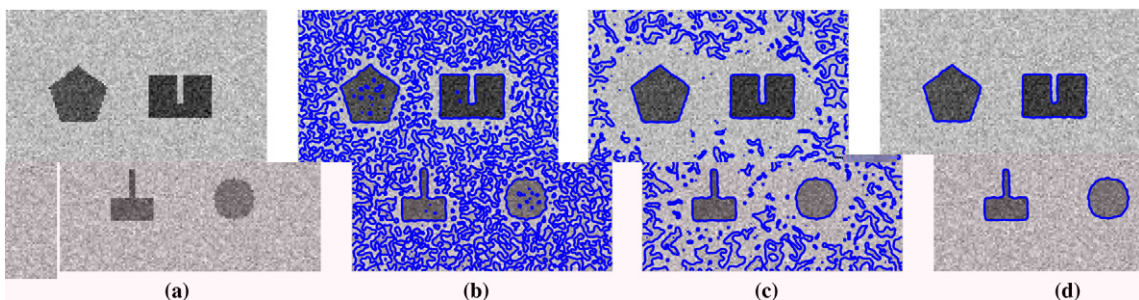

Fig. 7. Application of our method to an image with noise. (a) An image with noise. (b) Initial contour given by the normalized velocity field Eq. (59). (c) An intermediate stage during the evolution. (d) Final result.
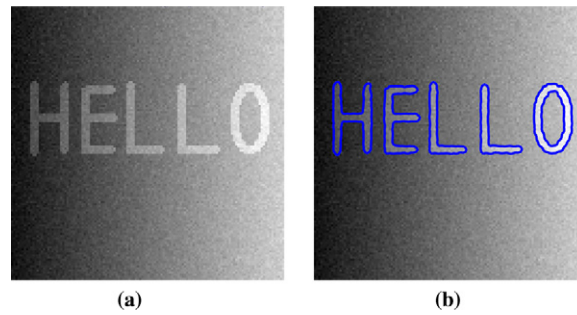
Fig. 8. Application of our method to an image with linearly changed background intensity. (a) The image. (b) The result using our method.

boundaries, but also contains a lot of spurious contours generated by the noise. During the evolution, these spurious contours disappear gradually, and finally the objects are found with high accuracy. This example shows that our method works well on images with noise.

Fig. 8 shows the result using our method on an image whose background intensity changes linearly. The background intensity is 0.05 on the left and is 0.8 on the right, the difference between the intensities of the letters and the local background is 0.2, and the standard deviation of the background noise is $\sigma_B = 0.03$, see Fig. 8(a). Our method finds the letters accurately, see Fig. 8(b). Actually, a constant gradient in the whole
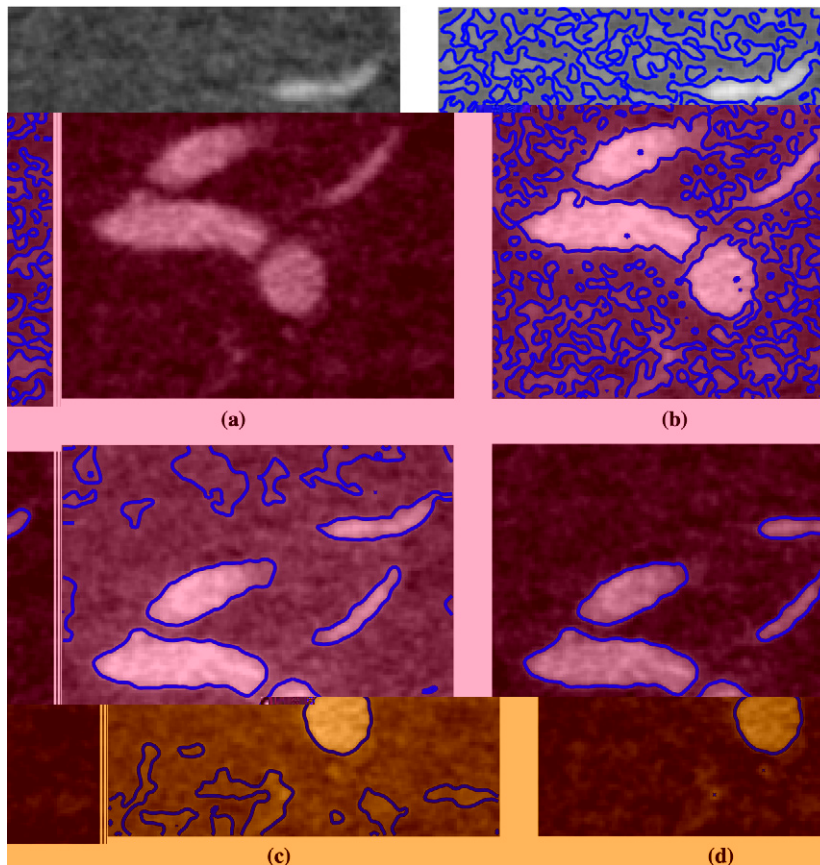

Fig. 9. Application of our method to an image slice of blood vessels. (a) Original image slice. (b) Initial contour given by the normalized velocity field Eq. (59). (c) An intermediate stage during the evolution of the contour. (d) The final result.

domain does not contribute to the normal velocity Eq. (33). Since the intensity is not periodic in this example, we need to use a finite difference method in the physical space to calculate the gradient of the image, then use FFT to calculate the normal velocity. Note that the region-based methods mentioned in the introduction do not work on this kind of images.
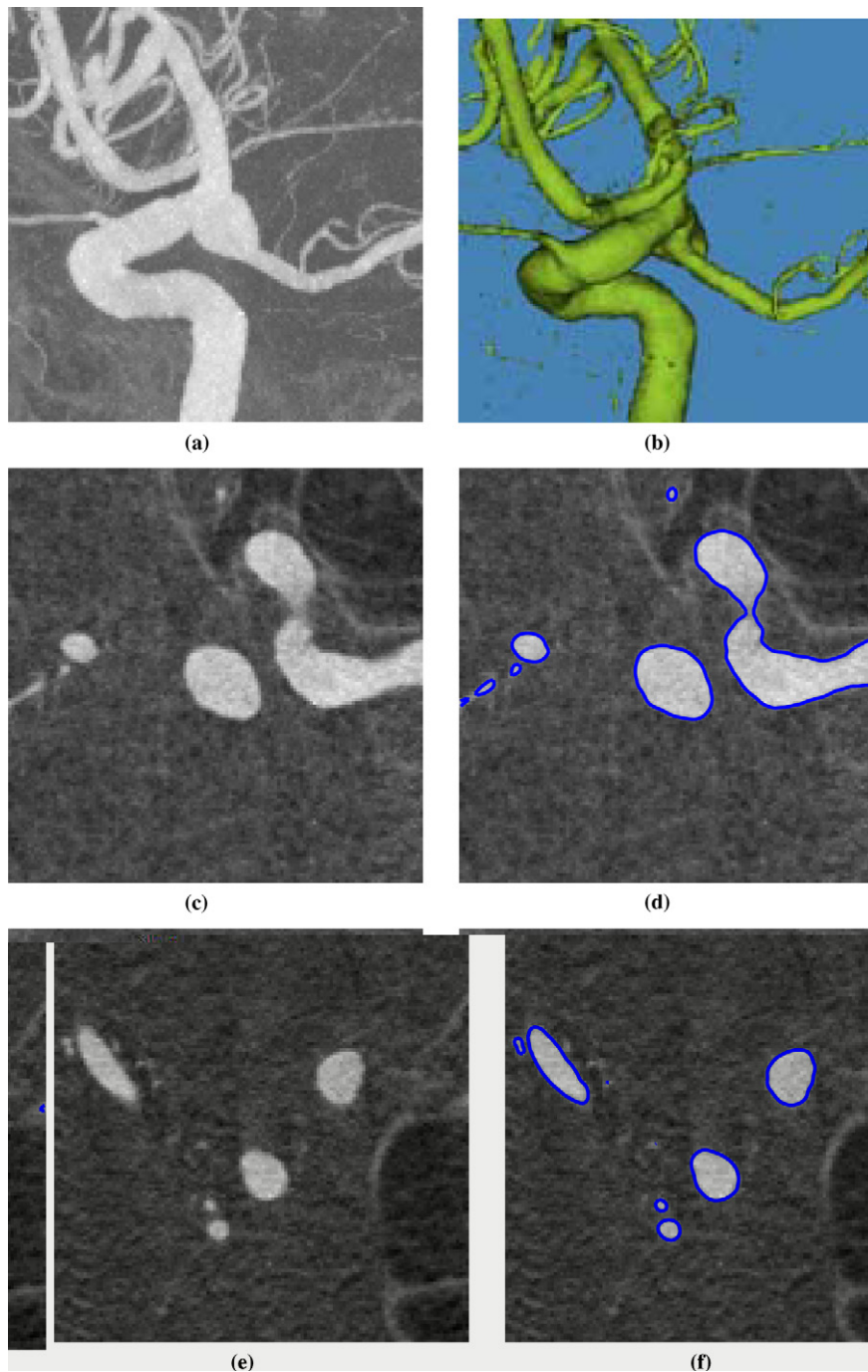

Fig. 10. Three dimensional segmentation results using our method on images of blood vessels. (a) The MIP image of the blood vessels. (b) The three dimensional result obtained using our method. (c) and (e) Two selected slices. (d) and (f) The segmentation results on slices (c) and (e).

We have also applied our segmentation method on three dimensional rotational angiographic (3D-RA) clinical datasets acquired by a Philips Integris Imager medical system at the Department of Diagnostic Radiology and Organ Imaging, Prince of Wales Hospital, Hong Kong. 3D-RA allows for visualization of vasculature and detection of vascular diseases, e.g. stenoses and aneurysms. The data volume was around $100 \times 100 \times 100$ voxels with a voxel size of $0.75 \times 0.75 \times 0.75 \text{ mm}^3$. Fig. 9 shows the result on one selected image slice of the data volume. The original image is shown in Fig. 9(a). Fig. 9(b) shows the initial contour given by the normalized velocity field Eq. (59). An intermediate stage during the evolution is shown in Fig. 9(c), and the final result is shown in Fig. 9(d). We can see that the initial contour identifies the blood vessels approximately, but the boundaries of them are irregular and a lot of spurious contours generated by the noise are present. During the evolution, the boundaries of the blood vessels become more and more regular and the spurious contours disappear gradually, and finally the blood vessels are found.

Finally, we present one of the three dimensional (3D) segmentation results using our method on these clinical datasets, see Fig. 10. Using the results of 3D segmentation, vascular surfaces can be accurately obtained so that quantitative 3D morphological information of vasculature can be provided for diagnosis. The 3D surface of the segmented image volume using our method is shown in Fig. 10(b). For visual comparison, the corresponding maximum intensity projection (MIP) is shown in Fig. 10(a). MIP is used widely for the visualization of brain blood vessels but cannot give accurate 3D morphological information. It is only a 2D projection of a 3D volume and is generated by taking the maximum intensity along the projection direction. From Fig. 10(a) and (b), we can see that the vessels, which are thin and elongated structures, are detected effectively using our method. The results on two selected slices of this three dimensional volume are shown in Fig. 10(d) and (f), and their corresponding original image slices are shown in Fig. 10(c) and (e). This example shows that our three dimensional method also works very well for real images.

7. Conclusions and discussion

We proposed a new edge-based active contour method, and applied the method to the synthetic and medical images. The new method is based on a long-range and orientation-dependent elastic interaction between the object boundary and the moving contour. The resultant velocity of the moving curve is an integral involving the image gradient over the whole image domain, which gives an extension of the effect of the image gradient to the whole image domain without smoothing the image gradient itself. The noise is removed by the self elastic interaction and the curvature effect of the moving contour. The velocity field can be calculated by the fast Fourier transform (FFT) method. Level set framework is used for the moving curve to handle the topological changes that might occur during its evolution. An initial contour very close to the object boundary is generated within this method, and then the noise is removed and the final result is achieved by solving an evolution equation of the level set function. Numerical experiments show that the new method is very effective, both for synthetic images and medical images of blood vessels. Compared with the existing edge-based methods, e.g., the methods based on the local image gradient and/or with the balloon force, and the gradient vector flow (GVF) method, our method is much less sensitive to the initialization of the moving contour, and moreover, an initial contour close to the object boundaries is provided within our method. Finally, the elastic interaction defined in our method is general and can be used in other frameworks for the representation of the moving contours such as the front-tracking methods.

When multiple junctions are present, the velocity in our method might be zero on some lines (or surfaces in three dimensions) that are not object boundaries, so the global convergence to the object boundaries no longer hold. We need to give an initial contour in a region near an object to find its boundary, as other active contour methods do. However, our method is still much less sensitive to the initialization of the moving contour in this case, see, e.g. Fig. 5. Besides, more level set functions are needed to model the boundaries of the multiple junctions, as in [19]. A direct extension of our method to color or texture image segmentation is to replace the image functions used here by the characteristic functions of the objects (1 inside and 0 outside). The features of the objects can be obtained first, e.g., by some statistical methods, then the characteristic function can be used to describe the image based on the identified features, and finally our active contour method can be used to find the object boundaries. Similar idea has been used by Paragios and Deriche in [28] for texture segmentation, in which the active contour stage was implemented using condition probabilities. These extensions and

more straightforward edge-based methods for color and texture segmentation are being considered. In our method presented in this paper, an initial contour close to the object boundaries is provided, but it may contain many spurious loops generated by noise (e.g., Fig. 7(b)), so it is not efficient to use the narrow band or the local level set methods ([5,43]). An alternative is to use a simple contour, say a large loop, as the initial value of the active contour. In this case, the initial contour is no longer close to the object boundaries, but the contour is simple and the narrow band or local level set methods can be used. The velocity generated by the image can be calculated only once at the beginning, and that generated by the moving contour can be calculated using a summation along itself instead of the FFT over the whole image domain. Which method is more efficient depends on the images, and will be examined in the future work.

Acknowledgements

Appendix. The Fourier transform of $1/\sqrt{x^2 + y^2}$

It is well known that in three dimensional space, $-1/4\pi\sqrt{x^2 + y^2 + z^2}$ is the solution of the three dimensional Poisson equation

$$-\Delta \frac{1}{4\pi\sqrt{x^2 + y^2 + z^2}} = \delta, \tag{67}$$

where δ is the three dimensional delta function. Taking Fourier transform, we have

$$\widehat{\frac{1}{\sqrt{x^2 + y^2 + z^2}}} = \frac{1}{2\pi^2(m^2 + n^2 + l^2)}, \tag{68}$$

where m, n and l are frequencies in the Fourier space. Therefore when $z = 0$, we have

$$\frac{1}{\sqrt{x^2 + y^2}} = \int_{\mathbf{R}^3} \frac{1}{2\pi^2(m^2 + n^2 + l^2)} e^{imx} e^{iny} \mathrm{d}m\, \mathrm{d}n\, \mathrm{d}l = \int_{\mathbf{R}^2} \frac{1}{2\pi\sqrt{m^2 + n^2}} e^{imx} e^{iny} \mathrm{d}m\, \mathrm{d}n, \tag{69}$$

which means

$$\widehat{\frac{1}{\sqrt{x^2 + y^2}}} = \frac{1}{2\pi\sqrt{m^2 + n^2}}. \tag{70}$$

References

[1] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, Int. J. Comput. Vision 1 (1988) 321–331.
[2] L.D. Cohen, On active contour models and balloons, Computer Vision, Graphics, and Image Processing: Image Understanding 53 (1991) 211–218.
[3] L.D. 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 (1993) 1131–1147.
[4] V. Caselles, F. Catte, T. Coll, F. Dibos, A geometric model for active contours in image processing, Numer. Math. 66 (1993) 1–31.
[5] R. Malladi, J.A. Sethian, B.C. Vemuri, Shape modeling with front propagation: a level set approach, IEEE Trans. Pattern Anal. Machine Intell. 17 (1995) 158–175.
[6] S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, J. Comput. Phys. 79 (1988) 12–49.
[7] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, in: Proc. Int. Conf. Computer Vision, Boston, USA, 1995, pp. 694–699.
[8] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, A. Yezzi, Gradient flows and geometric active contour models, in: Proc. Int. Conf. Computer Vision, Boston, USA, 1995, pp. 810–815.

[9] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, Int. J. Comput. Vision 22 (1997) 61–79.

[10] K. Siddiqi, Y.B. Lauziere, A. Tannenbaum, S.W. Zucker, Area and length minimizing flows for shape segmentation, IEEE Trans. Image Processing 7 (1998) 433–443.

[11] H. Park, T. Schoepflin, Y. Kim, Active contour model with gradient directional information: directional snake, IEEE Trans. Circuits Syst. Video Technol. 11 (2001) 252–256.

[12] M. Jacob, T. Blu, M. Unser, Efficient energies and algorithms for parametric snakes, IEEE Trans. Image Processing 13 (2004) 1231–1244.

[13] C. Xu, J.L. Prince, Snakes, shapes, and gradient vector flow, IEEE Trans. Image Processing 7 (1998) 359–369.

[14] C. Xu, J.L. Prince, Generalized gradient vector flow external forces for active contours, Signal Processing 71 (1998) 131–139.

[15] N. Paragios, O. Mellina-Gottardo, V. Ramesh, Gradient vector flow fast geometric active contours, IEEE Trans. Pattern Anal. Machine Intell. 26 (2004) 402–407.

[16] T.F. Chan, L.A. Vese, An active contour model without edges, in: M. Nielsen, P. Johansen, O.F. Olsen, J. Weickert (Eds.), Scale-Space Theories in Computer Vision, 1682, 1999, pp. 141–151.

[17] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Trans. Image Processing 10 (2001) 266–277.

[18] T.F. Chan, B.Y. Sandberg, L.A. Vese, Active contours without edges for vector-valued images, Journal of Visual Communication and Image Representation 11 (2000) 130–141.

[19] L.A. Vese, T.F. Chan, A multiphase level set framework for image segmentation using the Mumford and Shah model, Int. J. Comput. Vision 50 (2002) 271–293.

[20] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, Comm. Pure Appl. Math. 42 (1989) 577–685.

[21] R. Ronfard, Region-based strategies for active contour models, Int. J. Comput. Vision 13 (1994) 229–251.

[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 (1996) 884–900.

[23] O. Amadieu, E. Debreuve, M. Barlaud, G. Aubert, Inward and outward curve evolution using level set method, in: Proceedings of 1999 International Conference on Image Processing, 1999, pp. 188–192.

[24] A. Yezzi, A. Tsai, A. Willsky, Medical image segmentation via coupled curve evolution equations with global constraints, IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 2000, pp. 12–19.

[25] J. Lie, M. Lysaker, X.C. Tai, A variant of the level set method and applications to image segmentation, Math. Comp., in press.

[26] J. Ivins and J. Porrill, Statistical snakes: active region models, in: Fifth British Machine Vision Conference, 1994, pp. 377–386.

[27] A. Chakraborty, L.H. Staib, J.S. Duncan, Deformable boundary finding in medical images by integrating gradient and region information, IEEE Trans. Med. Imag. 15 (1996) 859–870.

[28] N. Paragios and R. Deriche, Geodesic active regions for supervised texture segmentation, in: Int. Conf. Computer Vision, 1999, 926–932.

[29] J.P. Hirth, J. Lothe, Theory of Dislocations, second ed., John Wiley, New York, 1982.

[30] R.W. Lardner, Mathematical Theory of Dislocations and Fracture, Univ. of Toronto Press, Toronto and Buffalo, 1974.

[31] Y. Xiang, L.T. Cheng, D.J. Srolovitz, W. E, A level set method for dislocation dynamics, Acta Mater. 51 (2003) 5499–5518.

[32] Y. Xiang, D.J. Srolovitz, L.T. Cheng, W. E, Level set simulations of dislocation-particle bypass mechanisms, Acta Mater. 52 (2004) 1745–1760.

[33] L.D. Landau, E.M. Lifshitz, Statistical Physics, Pergamon Press, London, 1980.

[34] L.M. Pismen, Vortices in Nonlinear Fields: from Liquid Crystals to Superfluids, from Non-equilibrium Patterns to Cosmic Strings, Oxford University Press, New York, 1999.

[35] L.M. Pismen, J. Rubinstein, Motion of vortex lines in the Ginzburg–Landau model, Physica D 47 (1991) 353–360.

[36] W. E, Dynamics of vortices in Ginzburg–Landau theories with applications to superconductivity, Physica D 77 (1994) 383–404.

[37] G.K. Batchelor, Introduction to Fluid Mechanics, Cambridge Univ. Press, Cambridge, 1967.

[38] A.J. Chorin, Vorticity and Turbulence, Springer, New York, 1994.

[39] Y.C. Chang, T.Y. Hou, B. Merriman, S. Osher, A level set formulation of Eulerian interface capturing methods for incompressible fluid flows, J. Comput. Phys. 124 (1996) 449–464.

[40] M. Bardi, S. Osher, The nonconvex multi-dimensional Riemann problem for Hamilton–Jacobi equations, SIAM J. Math. Anal. 22 (1991) 344–351.

[41] S. Osher, C.W. Shu, High-order essentially nonoscillatory schemes for Hamilton–Jacobi equations, SIAM J. Numer. Anal. 28 (1991) 907–922.

[42] G.S. Jiang, D. Peng, Weighted ENO schemes for Hamilton–Jacobi equations, SIAM J. Sci. Comput. 21 (2000) 2126–2143.

[43] D. Peng, B. Merriman, S. Osher, H.K. Zhao, M. Kang, A PDE-based fast local level set method, J. Comput. Phys. 155 (1999) 410–438.