An efficient local Chan–Vese model for image segmentation

Xiao-Feng Wang, De-Shuang Huang, Huan Xu

Abstract

In this paper, a new local Chan–Vese (LCV) model is proposed for image segmentation, which is built based on the techniques of curve evolution, local statistical function and level set method. The energy functional for the proposed model consists of three terms, i.e., global term, local term and regularization term. By incorporating the local image information into the proposed model, the images with intensity inhomogeneity can be efficiently segmented. In addition, the time-consuming re-initialization step widely adopted in traditional level set methods can be avoided by introducing a new penalizing energy. To avoid the long iteration process for level set evolution, an efficient termination criterion is presented which is based on the length change of evolving curve. Particularly, we proposed constructing an extended structure tensor (EST) by adding the intensity information into the classical structure tensor for texture image segmentation. It can be found that by combining the EST with our LCV model, the texture image can be efficiently segmented no matter whether it presents intensity inhomogeneity or not. Finally, experiments on some synthetic and real images have demonstrated the efficiency and robustness of our model. Moreover, comparisons with the well-known Chan–Vese (CV) model and recent popular local binary fitting (LBF) model also show that our LCV model can segment images with few iteration times and be less sensitive to the location of initial contour and the selection of governing parameters.

Keywords:
Extended structure tensor
Image segmentation
Intensity inhomogeneity
Level set method
Local Chan–Vese model

1. Introduction

Image segmentation has always been a fundamental problem and complex task in the field of image processing and computer vision. Its goal is to change the representation of an image into something that is more meaningful and easier to analyze [1]. In other words, it is used to partition a given image into several parts in each of which the intensity is homogeneous. Up to now, a wide variety of algorithms have been proposed to solve the image segmentation problem. Researchers have also done great efforts to improve the performance of the image segmentation algorithms.

Active contour model, or snakes, proposed by Kass et al. [2], has been proved to be an efficient framework for image segmentation. The fundamental idea of active contour model is to start with a curve around the object to be detected, and the curve moves toward its interior normal and stops on the true boundary of the object based on an energy-minimizing model. The main drawbacks of this method are its sensitivity to initial conditions and the difficulties associated with topological changes like the merging and splitting of the evolving curve. Since the active contour model was proposed, many methods have been proposed to improve it, in which level set method proposed by Osher and Sethian [3] is the most important and successful one.

Level set method is based on active contour model and particularly designed to handle the segmentation of deformable structures. Generally, the classical active contour model uses spline curves to model the boundary of an object. However, the level set method is to use a deformable curve front for approximating the boundary of an object. In the level set framework, the curve is represented by the zero level set of a smooth function, usually called the level set function. Moving the curves can be done by evolving the level set functions instead of directly moving the curves. Therefore, level set methods exhibit interesting elastic behaviors and can efficiently handle the topological changes which is also a main advantage compared with classical active contour model. Formally, the evolution of the curve is driven by a time-dependent partial differential equation (PDE) where the so-called velocity term reflects the image features characterizing the object to be segmented [4].

Generally, a classical level set framework consists of an implicit data representation of a hypersurface, a set of partial differential
equations (PDEs) that govern how the curve moves, and the corresponding numerical solution for implementing this method on computers [5].

It should be mentioned that the early edge-based level set methods [6–8] usually depend on the gradient of the given image for stopping the evolution of the curve. Therefore, these methods can only detect objects with edges defined by the gradient. However, the corresponding discrete gradients are generally bounded and the energy functional will hardly approach zero on the boundaries in practice. So, the evolving curve may pass through the true boundaries, especially for the models in [6–8].

Recently, region-based level set methods [10–13] have been proposed and applied to image segmentation filed by incorporating region-based information into the energy functional. Unlike edge-based level set methods using image gradient, region-based methods usually utilize the global region information to stabilize their responses to local variations (such as weak boundaries and noise). Thus, they can obtain a better performance of segmentation than edge-based level set methods, especially for images with weak object boundaries and noise. Among the region-based methods, Chan–Vese model [10] is a representative and popular one.

Based on the Mumford–Shah functional [9] for segmentation, Chan and Vese [10] proposed an easily handled model, or called Chan–Vese (CV) model, to detect objects whose boundaries are not necessarily detected by the gradient. Mumford–Shah model was firstly proposed as a general image segmentation model by Mumford and Shah in [9]. Using this model, the image is decomposed into some regions. Inside each region, the original image is approximated by a smooth function. The optimal partition of the image can be found by minimizing the Mumford–Shah functional. Chan and Vese successfully solved the minimization problem by using level set functions, which utilized the global image statistics inside and outside the evolving curve rather than the gradients on the boundaries.

CV model has achieved good performance in image segmentation task due to its ability of obtaining a larger convergence range and handling topological changes naturally. However, it still has some intrinsic limitations. First, CV model generally works for images with intensity homogeneity since it assumes that the intensities in each region always maintain constant. Thus, it often leads to poor segmentation results for images with intensity inhomogeneity due to wrong movement of evolving curves guided by global image information. Second, the segmentation of CV model is usually dependent on the placement of the initial contour, especially for the complicated images. Sometimes, the different results will be obtained on the same image by using different initial contours. Thus, the placement of initial contour is still an important issue for CV model to get successful segmentation in complicated images. Third, the CV model may become time-consuming if the periodical re-initialization step is adopted, which is a technique for periodically re-initializing the level set function to a signed distance function during the evolution. It has been regarded as a numerical remedy for maintaining stable curve evolution and ensuring precise results, which also leads to time-consumption as the side effect.

To solve the limitations of CV model, many efficient implementation schemes have been proposed [14–20]. For example, in [14], Vese and Chan extended their original model in [10] by using a multiphase level set formulation. However, the involved computation in this model is very expensive, which also limits its applications in practice. In addition, to reduce the computational cost, this method usually requires that the initial contour should be near to the object boundaries. In [15], an initialization scheme for the CV model was introduced, in which the initial curve is found by searching among the externals of the fidelity term in [10]. However, this one-dimensional search method is also time-consuming and fails to work when the gray difference between object and background is small. Xia et al. [16] proposed another initialization method which generates initial closed curves by iteratively connecting edge points obtained by canny detector and morphological filter. This method can efficiently work for some simple images. To reduce the computational load of curve evolution for CV model, the implementation schemes without solving the PDEs were proposed [17,18]. However, they are still sensitive to the selection of initial curves and sometime sensitive to noise. Li et al. [19] proposed a so-called penalizing energy which acts as a metric to characterize how close the level set function to a signed distance function. This metric can also be adopted by CV model to avoid the re-initialization step. In [20], the penalizing energy proposed in [19] and a discrimination function based on color information was combined into the CV model for segmenting the color images.

To efficiently perform the segmentation of images with intensity inhomogeneity, a new class of models has been proposed which not only utilize region-based techniques but also incorporate the benefits of local information. There have been several literatures [12,21–27] which are relevant to the existing works. Paragios et al. [12] presented a method in which edge-based energies and region-based energies were explicitly summed to create a joint energy which was then minimized. In [21], Sum et al. took a similar approach and minimized the sum of a global region-based energy and a local energy based on image contrast. Brox et al. [22] proposed the idea of incorporating localized statistics into a variational framework which shows that segmentation with local means is a first order approximation of the piecewise smooth simplification of the Mumford–Shah functional. Piovano et al. [23] employed convolutions to quickly compute localized statistics and yielded results similar to piecewise-smooth segmentation in a much more efficient manner. The work of An et al. [24] also noted the efficiency of localized approaches versus full piecewise smooth estimation and introduced a way in which localizations at two different scales can be combined to allow sensitivity to both coarse and fine image features. In [25], the authors proposed a similar flow based on computing geodesic curves in the space of localized means rather than approximating a piecewise-smooth model. In [26], a novel localization framework was proposed which allows the region-based energy to be localized in a fully variational way so that objects with heterogeneous statistics can be successfully segmented with the localized energies. Recently, Li et al. proposed an efficient region-based level set method by introducing a local binary fitting (LBF) energy with a kernel function [27]. The LBF model enables the extraction of accurate local image and can be used to segment the images with intensity inhomogeneity. It has attracted extensive attentions for its good segmentation performance in limited iterations. However, the LBF model usually needs to perform four convolution operations at each iteration, which greatly increases the computational complexity. In addition, it is also sensitive to the selection of governing parameters and the location of initial curve.

For CV model using the intensity average only, texture image segmentation is another difficult issue since intensity averages cannot represent the texture information inside and outside the target objects. In many texture images, due to the difference of the intensity averages of neighboring textures, the small textures in objects will be segmented while the desirable whole objects will be not separated. Therefore, other information should be introduced for texture image segmentation. Chan and Vese [10] suggested using texture information or features extracted from the original image, such as the curvature or the orientation of level sets, to overcome the difficulty. However, the proposed texture information in [10] cannot work well in many complicated texture images due to their simple properties. Note that the texture image segmentation greatly relies on the extraction of suitable texture information from the image. Recently, Gabor filters have been efficiently incorporated into level set
methods and CV model for the texture image segmentation [12,28,29]. Unfortunately, Gabor filters have the fatal drawback that they induce a lot of redundancies and thus lots of feature channels [30]. As another efficient texture representation, the structure tensor [31,32] is a kind of low dimensional feature computed from the spatial derivatives of the image. It is a common tool for local orientation estimation and image structure analysis which is formed as the outer product of the image gradient with itself. So far, the structure tensor has been applied in many image segmentation algorithms, most notably, in the early geodesic active contours framework [33–35].

In this paper, we proposed a so-called local Chan–Vese (LCV) model which utilizes both global image information and local image information for image segmentation. The energy functional for the proposed model consists of three parts: global term, local term and regularization term. By using the local image information, the images with intensity inhomogeneity can be efficiently segmented in limited iterations. Moreover, the time-consuming re-initialization step widely adopted in traditional level set methods can be avoided by introducing a new penalizing energy to the regularization term. As a result, the time-consumption is greatly decreased. Specially, the evolving curve in level set evolution process can automatically stop on true boundaries of objects according to a termination criterion which is based on the length change of evolving curve. Finally, we proposed constructing an extended structure tensor (EST) by adding the intensity information into the classical structure tensor for texture image segmentation. By incorporating the EST into the proposed LCV model, the texture image can be easily segmented no matter whether it presents intensity inhomogeneity or not. Moreover, the comparisons with the CV model and recent LBF model show that our LCV model can segment ordinary/texture images with or without intensity inhomogeneity in fewer iterations. Particularly, it can be found that our model is less sensitive to the location of initial contour and the selection of governing parameters.

The rest of this paper is organized as follows: In Section 2, we briefly review the Mumford–Shah model and Chan–Vese (CV) model. Our local Chan–Vese (LCV) model is presented in Section 3. In Section 4, the proposed model is validated by some experiments on synthetic and real images. Finally, some conclusive remarks are included in Section 5.

2. Previous works

2.1. Mumford–Shah model

The Chan–Vese model is the curve evolution implementation of a piecewise-constant case of the Mumford–Shah model [9]. The Mumford–Shah model is an energy-based model introduced by Mumford and Shah via an energy functional. The basic idea is to find a pair of \((u, C)\) for a given image \(u_0\), where \(u\) is a nearly piecewise smooth approximation of \(u_0\), and \(C\) denotes the smooth and closed segmenting curve. The general form for the Mumford–Shah energy functional can be written as

\[
E^{MS}(u, C) = \int_\Omega |u_0(x, y) - u(x, y)|^2 \, dx \, dy + \mu \int_{\partial C} [\nabla u(x, y)]^2 \, dx \, dy + \nu \cdot \text{Length}(C),
\]

where \(\mu\) and \(\nu\) are positive constants, \(\Omega\) denotes the image domain, the segmenting curve \(C \subset \Omega\). To solve the Mumford–Shah problem is to minimize the energy functional over \(u\) and \(C\). Note that the removal of any of the above three terms in (1) will result in trivial solutions for \(u\) and \(C\). However, with all three terms, it becomes a difficult problem to solve since \(u\) is a function in the \(N\)-dimensional space \(\{N = 2\} \) in 2D image segmentation, while \(C\) is an \((N-1)\)-dimensional data set.

2.2. Chan–Vese model

The Chan–Vese (CV) model is an alternative solution to the Mumford–Shah problem which solves the minimization of (1) by minimizing the following energy functional:

\[
E^{CV}(c_1, c_2, C) = \mu \cdot \text{Length}(C) + \lambda_1 \cdot \int_{\text{inside}(C)} |u_0(x, y) - c_1|^2 \, dx \, dy + \lambda_2 \cdot \int_{\text{outside}(C)} |u_0(x, y) - c_2|^2 \, dx \, dy, \tag{2}
\]

where \(\mu\), \(\lambda_1\) and \(\lambda_2\) are positive constants, usually fixing \(\lambda_1 = \lambda_2 = 1\). \(c_1\) and \(c_2\) are the intensity averages of \(u_0\) inside \(C\) and outside \(C\), respectively.

To solve this minimization problem, the level set method [3] is used which replaces the unknown curve \(C\) by the level-set function \(\phi(x, y)\), considering that \(\phi(x, y) > 0\) if the point \((x, y)\) is inside \(C\), \(\phi(x, y) < 0\) if \((x, y)\) is outside \((x, y)\), and \(\phi(x, y) = 0\) if \((x, y)\) is on \(C\). Thus, the energy functional \(E^{CV}(c_1, c_2, C)\) can be reformulated in terms of the level set function \(\phi(x, y)\) as follows:

\[
E^{CV}_\phi(c_1, c_2, \phi) = \mu \cdot \int_\Omega \delta_\phi(|\nabla \phi(x, y)|) \, dx \, dy + \lambda_1 \cdot \int_\Omega |\nabla \phi(x, y)| - \lambda_1 |u_0(x, y) - c_1|^2 H_\phi(x, y) \, dx \, dy + \lambda_2 \cdot \int_\Omega |\nabla \phi(x, y)| - \lambda_2 |u_0(x, y) - c_2|^2 (1 - H_\phi(x, y)) \, dx \, dy, \tag{3}
\]

where \(H_\phi(x, y)\) and \(\delta_\phi(z)\) are, respectively, the regularized approximation of Heaviside function \(H(z)\) and Dirac delta function \(\delta(z)\) as follows:

\[
H(z) = \begin{cases} 
1 & \text{if } z \geq 0, \\
0 & \text{if } z < 0,
\end{cases} \quad \delta(z) = \frac{d}{dz}H(z). \tag{4}
\]

This minimization problem is solved by taking the Euler–Lagrange equations and updating the level set function \(\phi(x, y)\) by the gradient descent method:

\[
\frac{\partial \phi}{\partial t} = \delta_\phi(|\nabla \phi|) \left[ \frac{\mu \div (\nabla \phi)}{|\nabla \phi|} - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right]. \tag{5}
\]

where \(c_1\) and \(c_2\) can be, respectively, updated at each iteration by

\[
c_1(\phi) = \frac{\int_\Omega u_0(x, y)H_\phi(x, y) \, dx \, dy}{\int_\Omega H_\phi(x, y) \, dx \, dy}, \quad c_2(\phi) = \frac{\int_\Omega u_0(x, y)(1 - H_\phi(x, y)) \, dx \, dy}{\int_\Omega (1 - H_\phi(x, y)) \, dx \, dy}. \tag{6}
\]

The main advantages of this model are: First, it can deal with the detection of objects whose boundaries may be either smooth or not necessarily defined by gradient. In such cases, the edge-based level set methods commonly fail and result in boundary leakage [36]. Second, it does not require image smoothing and thus can efficiently process the images with noise. Therefore, the true boundaries are preserved and could be accurately detected. Third, it can automatically detect interior contours with the choice of Dirac delta function \(\delta(z)\) that has non-compact support.

However, CV model also has some drawbacks which have been described in Section 1, i.e., the unsuccessful segmentation of images with intensity inhomogeneity (as shown in Fig. 3(c)), the sensitivity to the placement of initial contour and the extraordinary time-consumption if re-initialization step is adopted for maintaining stable curve evolution and ensuring more precise results.
3. Local Chan–Vese model

In this section, we shall present and discuss the details of our proposed local Chan–Vese (LCV) model and its numerical implementation. What should be stressed is that our model is defined based on the techniques of curve evolution, local statistical function and level set methods. It is well-known that some traditional level set methods used either the image gradient [6–8] or the global information [10,11] to drive the evolving curve(s) towards the true boundaries. However, none of them can achieve success in segmenting images with intensity inhomogeneity. In the proposed model, we combined both global and local statistical information to overcome the inhomogeneous intensity distribution in some images and provided more satisfying segmentation result. Particularly, by incorporating a so-called extended structure tensor (EST) into the proposed LCV model, the texture image can also be segmented no matter whether it presents intensity inhomogeneity or not. The overall energy functional in our proposed LCV model consists of three parts: global term $E^G$, local term $E^L$ and regularization term $E^R$. Thus the overall energy functional can be described as

$$E^{LCV} = \alpha \cdot E^G + \beta \cdot E^L + E^R.$$  

(7)

3.1. Global term

The global term $E^G$ is directly derived from (2) in the Chan–Vese model, in which it is also called the fitting term. It can be seen that the global term is defined based on the global properties, i.e., the averages of $u_0$ inside $C$ and outside $C$, which is stated as follows:

$$E^G(c_1, c_2, C) = F_1(C) + F_2(C) = \int_{\text{inside}(C)} |u_0(x, y) - c_1|^2 \, dx \, dy + \int_{\text{outside}(C)} |u_0(x, y) - c_2|^2 \, dx \, dy,$$

(8)

Using the level set formulation, the boundary $C$ is represented by the zero level set of a Lipschitz function $\phi : \Omega \to \mathbb{R}$. $\phi(x, y) > 0$ if $(x, y)$ is inside $C$, $\phi(x, y) = 0$, $(x, y) \in C$, and $\phi(x, y) < 0$ if $(x, y)$ is outside $C$.

Accordingly, the global term in (8) can be rewritten so as to evaluate the level set function $\phi$ on the domain $\Omega$:

$$E^G(c_1, c_2, \phi) = \int_{\Omega} |u_0(x, y) - c_1|^2 H(\phi(x, y)) \, dx \, dy + \int_{\Omega} |u_0(x, y) - c_2|^2 (1 - H(\phi(x, y))) \, dx \, dy,$$

(10)

where $H(z)$ is the Heaviside function described in (4).

Usually, after (10) comes to a steady state, or approximately to be zero, the evolving curve $C$ (zero level set of $\phi$) will separate the object from the background. However, for the images with intensity inhomogeneity, the finally obtained curve $C$ can hardly divide the image into object region and background region even after a long iteration time. The reason is that the global term assumes that the image intensity is piecewise constant like the CV model. Thus, the averages $c_1$ and $c_2$ actually act as the global information and cannot represent the inhomogeneous intensities of object region and background region in the images with intensity inhomogeneity. So, to achieve a good performance in segmenting the images with intensity inhomogeneity, the local image information needs to be included.

3.2. Local term

Before introducing the local term, we should firstly discuss the intensity inhomogeneity problem. Intensity inhomogeneity usually arises from the imperfect factors of acquisition process for ordinary images or medical images, such as non-uniform daylight and artificial illumination, static field inhomogeneity, radio-frequency excitation field non-uniformity and inhomogeneity of reception coil sensitivity, etc. These inhomogeneities are known to appear in images as systematic changes in the local statistical characteristics of target object. Although the presence of intensity inhomogeneity is usually hardly noticeable to a human observer, many image processing methods, including image segmentation methods, are highly sensitive to the spurious variations of image intensities since they are based on the assumptions that the intensities in each region are constant.

The generally accepted assumption on intensity inhomogeneity is that it manifests itself as a smooth spatial varying function over the image [37]. The most common model in describing the acquired images $X$' with intensity inhomogeneity effect is

$$X' = BX + N,$$

(11)

where $X$ is the inhomogeneity-free image, $B$ denotes the intensity inhomogeneity field and $N$ is the noise. To simplify the computation, the noise is often ignored. Also, there have been theoretical modeling approaches to approximate the intensity inhomogeneity field. However, due to the complexity that causes the intensity inhomogeneity, it is difficult for ones to model the intensity inhomogeneity under a variety of image acquisition conditions [38].

Since the intensity inhomogeneity is slowly varying in the image domain, its spectrum in frequency domain will be concentrated in the low-frequency area. Thus, the intensity inhomogeneity effect mainly influences the non-contour pixels in the image, whereas for the pixels belonging to contour, this influence is less. Motivated by this observation, we proposed incorporating the local statistical information into the level set method for segmenting the images with intensity inhomogeneity effect. It should be noted that we do not try to eliminate the intensity inhomogeneity from the images which is still not a completely solved problem [38].

Here, the local term is introduced in (12) which uses the local statistical information as the key to improve the segmentation capability of our model on the images with intensity inhomogeneity.

$$E^L(d_1, d_2, C) = \int_{\text{inside}(C)} |g_k * u_0(x, y) - u_0(x, y)| - d_1|^2 \, dx \, dy + \int_{\text{outside}(C)} |g_k * u_0(x, y) - u_0(x, y)| - d_2|^2 \, dx \, dy,$$

(12)

where $g_k$ is a averaging convolution operator with $k \times k$ size window. $d_1$ and $d_2$ are the intensity averages of difference image $(g_k * u_0(x, y) - u_0(x, y))$ inside $C$ and outside $C$, respectively.

The assumption behind the proposed local term is that smaller image regions are more likely to have approximately homogeneous intensity and the intensity of the object is statistically different from the background. It is significant to statistically analyze each pixel with respect to its local neighborhood. The most simple and fast statistical information function is the average of the local intensity distribution, the rationale being that if the object pixels are brighter than the background, they should also be brighter than the average. It should be noticed that the size of the neighborhood has to be properly selected so as to cover sufficient object and background pixels, which may make the local term less sensitive to the existence of noise. However, at a larger neighborhood size, the local term will probably lose some fine detail of images. So, it needs to be combined with the global term in the image segmentation process.

By subtracting the original image from the averaging convolution image, the contrast between foreground intensities and background
where \( \text{dist}(\mathbf{x}, C_t) \) is the shortest Euclidean distance of \( \mathbf{x} \) to the points on the evolving curve \( C_t \) at time \( t \).

It is crucial to keep the evolving level set function as an approximate signed distance function during the evolution, especially in the neighborhood around the zero level set [5]. The most straightforward way of implementing the re-initialization operation is to extract the zero level set and then explicitly compute the distance function from it. However, this method is generally time-consuming. To overcome this difficulty, a now widely accepted method has been proposed [39] in order to re-initialize the level set function by solving the following partial differential equation:

\[
\frac{\partial \phi}{\partial t} = \text{sign}(\phi_0)(1 - |\nabla \phi|),
\]

where \( \phi_0 \) is the function to be re-initialized, and \( \text{sign}(\phi_0) \) is the sign function. When the steady state of Eq. (18) is reached, \( \phi \) will be a distance function with the same zero level set as \( \phi_0 \) despite \( \phi_0 \) is a distance function or not. This is commonly known as the standard re-initialization procedure.

Another equivalent approach is to solve the following eikonal equation:

\[
|\nabla \phi| = 1,
\]

with the boundary condition \( \phi = 0 \) on \( \{\phi_0 = 0\} \).

Re-initialization has been extensively used as a numerical remedy for maintaining stable curve evolution and ensuring desirable results in the level set methods. Unfortunately, it is obviously a disagreement between the theory of level set method and its implementation, since it has an undesirable side effect of moving the zero level set away from its original location. Moreover, it is quite complicated and time-consuming, and when and how to apply it is still a serious problem [40]. Accordingly, some fast techniques [41,42] were proposed for finding a solution to these problems. Among these methods, fast marching method [42] is a representative and popular one. It is the optimal technique for solving the Eikonal equation \( F(\nabla \phi) = 1 \), where \( F \) denotes the speed of interface. For more detailed technical description about fast marching method, readers can refer to literature [42]. Though fast marching method is more efficient than traditional approach, the computational time is still large.

In this paper, we did not directly use the re-initialization step to keep the level set function as a signed distance function but add to the regularization term a penalty term as follows:

\[
P(\phi) = \int_{\Omega} \frac{1}{2} (|\nabla \phi(x,y)| - 1)^2 \, dx \, dy
\]

which can force the level set function to be close to a signed distance function.

Actually, this penalty term is more like a metric which characterizes how close a function \( \phi \) is to a signed distance function. The metric plays a key role in the elimination of re-initialization in our method. To explain the effect of the penalty term \( P(\phi) \), we give its gradient flow as follows:

\[
\nabla \phi - \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) = \text{div} \left[ \left( 1 - \frac{1}{|\nabla \phi|} \right) \nabla \phi \right].
\]

Notice that the above gradient flow has the factor \( 1 - (1/|\nabla \phi|) \) which can act as the diffusion rate. If \( |\nabla \phi| > 1 \), the diffusion rate is positive and the effect of the penalty term is the usual diffusion, i.e. making \( \phi \) more even and therefore reduce the gradient \( |\nabla \phi| \). If \( |\nabla \phi| < 1 \), the penalty term has effect of reverse diffusion and therefore increase the gradient. It can be seen that the penalty term continually adjusts the deviation of level set function \( \phi \) from the signed distance function in the evolution process. Therefore, it can naturally and automatically force the level set function to be an approximate signed distance function during the evolution.
energy functional in (26) with respect to two pairs of constants: $c_1$ and $c_2$, $d_1$ and $d_2$. By calculus of variations, it can be shown that the constant functions $c_1(\phi)$, $c_2(\phi)$, $d_1(\phi)$ and $d_2(\phi)$ that minimize $E^{CV}(c_1, c_2, d_1, d_2, \phi)$ for a fixed function $\phi$ are given by

\[\begin{align*}
&c_1(\phi) = \frac{\int_\Omega u_0(x,y)H_1(\phi(x,y))\,dx\,dy}{\int_\Omega H_1(\phi(x,y))\,dx\,dy}, \\
&c_2(\phi) = \frac{\int_\Omega u_0(x,y)(1 - H_1(\phi(x,y)))\,dx\,dy}{\int_\Omega (1 - H_1(\phi(x,y)))\,dx\,dy}, \\
d_1(\phi) = \frac{\int_\Omega g_k u_0(x,y) - u_0(x,y)H_1(\phi(x,y))\,dx\,dy}{\int_\Omega H_1(\phi(x,y))\,dx\,dy}, \\
d_2(\phi) = \frac{\int_\Omega g_k u_0(x,y) - u_0(x,y)(1 - H_1(\phi(x,y)))\,dx\,dy}{\int_\Omega (1 - H_1(\phi(x,y)))\,dx\,dy}.
\end{align*}\]

Keeping $c_1$, $c_2$, $d_1$ and $d_2$ fixed, and minimizing the overall energy function $E^{CV}$ in (26) with respect to $\phi$, we can deduce the associated Euler–Lagrange equation for $\phi$. The minimization of (26) can be done by introducing an artificial time variable $\tau \approx 0$, and moving $\phi$ in the steepest descent direction to a steady state with the initial condition defined in (28b) and boundary condition defined in (28c):

\[\frac{\partial \phi}{\partial \tau} = \phi_0(x,y) \text{ in } \Omega,\]

where $\overline{\Omega}$ denotes the exterior normal to the boundary $\partial \Omega$.

In the above partial differential equation, the Neumann boundary condition in (28c) is chosen as the boundary condition. Usually, Neumann boundary condition has many advantages. First, it is easy to implement since there are no values to assign for $\phi$ at the boundary. Second, it implies that the solution of (26) satisfies a maximum principle. Moreover, the gaps of the curve $C$ may appear only when advancing the zero level set which would change its topology, and cannot come from outside of $\Omega$ because of the spurious values created by the Neumann boundary condition.

### 3.5. Termination criterion for curve evolution

Evolving curve $C$, or zero level set function, will gradually split according to the topological structure of objects and evolves towards the true boundaries of objects in images. When evolving curves finally arrive at the position of true boundary $C^*$, the curves should stop evolving. Now, a problem of when and how the curves automatically stop evolving or what termination criterion is for curv​es evolution is arising.

It can be obviously seen that the global term and local term keep decreasing while the curve evolves towards the true boundaries of objects and the true boundary $C^*$ is the minimizer of the global term and local term, which can be written as follows:

\[\inf \{\zeta \cdot E^G(C) + \beta \cdot E^L(C) \} \approx 0 \approx \zeta \cdot E^G(C^*) + \beta \cdot E^L(C^*).\]

Therefore, we can approximately obtain the energy minimization value by

\[\inf_{c_1, c_2, d_1, d_2, C} E^{CV}(c_1, c_2, d_1, d_2, C) \approx \mu \cdot L(C^*).\]
Actually, (30) can be the termination criterion for curves evolution. But, it cannot be directly implemented in practice because of its approximate equal sign. In our numerical solution, we shall use another form of this termination criterion by determining the length of evolving curve \( L(C(t)) \) at each iteration. When curve reaches the true boundary, it is obvious that the variable \( L(C(t)) \) will remain almost constant. Here, we shall give the termination criterion for the curve evolution in an image as follows:

Termination criterion: If the absolute value of the change of the curve length \( |L(C(t))−L(C(t−1))| \) remains smaller than a given threshold \( ε_{\text{length}} \) over a fixed threshold of iterations \( T_l \), the evolution of curves will be stopped.

The pseudocode of this termination criterion is listed as follows:

\[
\begin{align*}
&k = 1; \\
&\text{if} \ |\text{Length}(C(t))−\text{Length}(C(t−1))|≤ \varepsilon_{\text{length}} \text{ then} \\
&\{ \\
&\text{if} \ k = T_l \ \text{then} \ \text{stop evolution of curve;} \\
&\quad \ k = k+1; \\
&\quad \} \ \text{else} \ k = 1; \\
&\text{end}
\end{align*}
\]

If the condition \( |L(C(t))−L(C(t−1))| ≤ \varepsilon_{\text{length}} \) is firstly satisfied at a later stage, it is unnecessary to stop the evolution of curve immediately. Instead, we shall examine whether or not the condition is maintained over a fixed threshold of iterations \( T_l \) before stopping the evolution of curve. This examination must be performed because the evolution process often slows down temporarily even before true boundary is reached. In practice, the choice of the two thresholds is flexible. Generally, we always set \( T_l = 10 \) and \( \varepsilon_{\text{length}} = 5 \).

Remark 1. Note that the energy functional (26) of proposed model is non-convex and sometimes it may have multiple local minima, these being properties inherited from the Mumford–Shah model. To improve the ability of curve to approach the true boundary, the local information is incorporated into the energy functional. The global term is also employed to restrain the noise and maintain the boundary details. In practice, we do not guarantee that our model can always converge to a global minimizer. Sometimes, a local minimizer may be obtained, but close to a global minimizer. In this case, the energy functional will become stationary and the proposed termination criterion can also be applied.

3.6. Numerical implementation of the model

The partial differential equation in the continuous domain defined in (28a) can be solved by a finite difference method in numerical scheme. All the spatial partial derivatives are approximated by the central difference and the temporal partial derivatives are approximated by the forward difference.

Then, (28a) can be discretized using the forward difference as follows:

\[
\frac{\phi_{ij}^{n+1}−\phi_{ij}^n}{\Delta t} = L(\phi_{ij}^n),
\]

where \( \Delta t \) is the time-step and \( L(\phi_{ij}^n) \) is the numerical approximation of the right-hand side in (28a).

The corresponding curvature \( \kappa = \text{div}(\nabla \phi / \| \nabla \phi \|) \) in the \( L(\phi_{ij}^n) \) can be discretized using a second-order central differencing scheme:

\[
\kappa = \text{div} \left( \frac{\nabla \phi}{\| \nabla \phi \|} \right) = \frac{\phi_{xx} \phi_y^2−2\phi_x \phi_y \phi_{xy}+\phi_{yy} \phi_x^2}{(\phi_x^2+\phi_y^2)^{3/2}}.
\]

where \( \phi_x, \phi_y, \phi_{xx}, \phi_{yy} \) and \( \phi_{xy} \) are computed as follows:

\[
\begin{align*}
\phi_x &= \frac{1}{2h}(\phi_{i+1,j}−\phi_{i−1,j}), \quad \phi_y = \frac{1}{2h}(\phi_{i,j+1}−\phi_{i,j−1}), \\
\phi_{xx} &= \frac{1}{h^2}(\phi_{i+1,j}+\phi_{i−1,j}−2\phi_{i,j}), \quad \phi_{yy} = \frac{1}{h^2}(\phi_{i,j+1}+\phi_{i,j−1}−2\phi_{i,j}), \\
\phi_{xy} &= \frac{1}{h^2}(\phi_{i+1,j+1}−\phi_{i+1,j−1}−\phi_{i−1,j+1}+\phi_{i−1,j−1}).
\end{align*}
\]

where \( h \) is the grid spacing. Eq. (28a) is then implemented as follows:

\[
\frac{\phi_{ij}^{n+1}−\phi_{ij}^n}{\Delta t} = \delta_i(\phi_{ij}^n)−(\alpha(u_{ij}−c_1(\phi^n)))+\beta(g_k∗u_{ij}−u_{ij}−d_k(\phi^n)^2) \\
+ (\alpha(u_{ij}−c_2(\phi^n)))+\beta(g_k∗u_{ij}−u_{ij}−d_k(\phi^n)^2)) \\
+ [\mu \cdot \psi_{ij}^\phi + (\phi_{i+1,j}+\phi_{i−1,j}+\phi_{i,j+1}+\phi_{i,j−1}−4\phi_{ij}−\kappa)],
\]

where \( \delta_i \) and \( \kappa \) are computed according to (25) and (32), respectively.

3.7. Texture image segmentation

Usually, the texture image segmentation algorithms include two steps: First, a texture representation is selected and corresponding texture features are extracted from initial images. Second, some objective function can be defined using the texture features, and the segmentation is formulated as an optimization or minimization problem. In this paper, the extended version of classical structure tensor was proposed to act as the texture feature and further incorporated into the LCV model for texture image segmentation.

For a scalar image \( I \), the classical structure tensor \( t \) is obtained by Gaussian smoothing of the tensor product of the image gradient, i.e.:

\[
J_t = K_\sigma ∗ (\nabla I \cdot \nabla I^T) = \left( \begin{array}{cc} K_\sigma ∗ I_x^2 & K_\sigma ∗ I_x I_y \\
K_\sigma ∗ I_x I_y & K_\sigma ∗ I_y^2 \end{array} \right).
\]

where \( K_\sigma \) is a Gaussian kernel with standard deviation \( \sigma \), and subscripts \( x \) and \( y \) denote the partial derivatives. For vector-valued images, the following expression is employed:

\[
J_t = K_\sigma ∗ \left( \sum_{i=1}^{n} \nabla I_i \nabla I_i^T \right).
\]

It can be seen from (35) that the classical structure tensor yields three feature channels for each scale. The matrix representation using the image gradient allows the integration of information from a local neighborhood without considering cancellation effects. Such effects would appear if gradients with opposite orientation were included directly [43]. Moreover, the usage of Gaussian smoothing will cause the classical structure tensor robust to noise and make the orientation estimation at certain pixel to be performed. Comparing the number of features obtained by the classical structure tensor to that of Gabor filters reveals that the degree of freedom for the orientation known from Gabor filters is replaced by the smoothed versions of the image derivatives [30]. Since the image derivatives hold the whole orientation information, the components of the structure tensor are as powerful for the discrimination of textures as a whole set of Gabor filters with a fixed scale. Due to the mixed matrix component, texture discrimination of classical structure tensor is also fully rotation invariant.

However, the classical structure tensor has the disadvantage of not using any intensity information (or color information in the case
of vector-valued images) at all. When the intensity (color) of texture object is different from that of background in a texture image, the introduction of intensity information will be of benefit to the texture object segmentation. Moreover, if texture images present the distinct intensity inhomogeneity as we discussed in Section 3.2, the segmentation using the classical structure tensor will also fail even if the LCV model with local term is adopted. The reason is due to the fact that the intensity information is missed of which the LCV model is strongly dependent. Therefore, we consider constructing an extended structure tensor (EST) by incorporating the intensity information into the classical structure tensor for texture image segmentation.

For a scalar image $I$, the extended structure tensor (EST) $f_E^I$ can be defined as follows:

$$f_E^I = K_I \ast (v v^T) = \begin{pmatrix} K_x \ast I_x^2 & K_x \ast I_x I_y & K_x \ast I_y I_x & K_x \ast I_y^2 \\ K_y \ast I_x I_y & K_y \ast I_y^2 & K_y \ast I_x I_y^2 & K_y \ast I_y^2 \\ K_x \ast I_x I_y & K_y \ast I_y I_x & K_{xy} \ast I_x I_y & K_{xy} \ast I_y I_x \\ K_x \ast I_y I_x & K_y \ast I_x I_y & K_{xy} \ast I_x I_y & K_{xy} \ast I_y I_x \end{pmatrix},$$  

(37)

where $v = [I_x, I_y, I]^T$.

For vector-valued images, the following expression can be used:

$$f_E^I = K_I \ast \left( \sum_{i=1}^{N} v_i v_i^T \right),$$  

(38)

where $v_i = [I_x, I_y, I]^T$.

Compared with the classical structure tensor, the EST yields six feature channels for each scale, with three of which containing the intensity information. After the EST is computed, the LCV model can be adopted to address texture image segmentation by replacing the original image $u_0$ in (28) with the average of all the channels $f_{x,j}^I (i = 1, 2, ..., 9)$ belonging to the EST $f_E^I$ in (37), i.e.:

$$u^E_I = \frac{1}{9} \sum_{i=1}^{9} f_{x,j}^E, $$  

(39)

where $\sigma$ is the standard deviation of Gaussian kernel.

### 3.8. Description of algorithm steps

Now, we can describe the steps of our LCV model as follows:

Step 1: Input the original image $u_0$. If the image is texture image then compute the corresponding extended structure tensor $f_E^I$ according to (37) and replace $u_0$ with $u^E_I$ in (39) (the average of nine channels of the extended structure tensor $f_E^I$).

Step 2: Set the initial curve $C_0$ in $u_0$. Set the value of time-step $\Delta t$, the grid spacing $h$ and $\varepsilon$ in (34). In practice, $\Delta t = 0.1$ and $h = \varepsilon = 1$. Set the window size $k$ of averaging convolution operator in local term.

Step 3: Evolve level set function $\phi$ according to (28) and its numerical solution scheme described in (34).

Step 4: Extract the evolving curve $C$ from the zero level set function.

Step 5: Judge whether the termination criterion described in Section 3.5 is satisfied or not. If yes, the algorithm is stopped; otherwise, go to Step 3.

### 4. Experimental results

In this Section, we shall present the experimental results of our local Chan–Vese (LCV) model on some synthetic and real images. The proposed model was implemented by Matlab 7 on a computer with Intel Core 2 Duo 2.2GHz CPU, 2G RAM, and Windows XP operating system. The processing time referred later in this section starts after choosing the initial contour. We used the same parameters of the time-step $\Delta t = 0.1$, the grid spacing $h = 1$, $\varepsilon = 1$ for $H(z)$ and $D(z)$, the window size of averaging convolution operator $k = 15$, $T_0 = 10$ and $\omega_{\text{length}} = 5$ (to be used in the proposed termination criterion), $\sigma=2$ (the standard deviation of Gaussian kernel in extended structure tensor) for all the experiments in this section. It can be found that we were able to get good segmentation results on a wide range of images with these parameters.

The couple of controlling parameters of global term and local term $\alpha, \beta$ should be set according to the image intensity property as described in Section 3.8. Generally, we fixed $\beta = 1$ and dynamically adjusted the value of $\alpha$. In our experiments, $\alpha$ has two corresponding values: 0.1 and 1 for images with/without intensity inhomogeneity. The length controlling parameter $\alpha$ also has a scaling role like $\beta$. If we have to detect all or as many objects as possible and the objects of any size, then $\mu$ should be small. Otherwise, $\mu$ should be larger. In our experiments, two corresponding values of $\mu, 0.01 \times 255^2$ and $0.1 \times 255^2$, are adopted. In conclusion, there are totally two parameters whose values need to be dynamically adjusted in our experiments, i.e., $\alpha$ and $\mu$.

We firstly considered the simplest case: segmentation of images with the intensity homogeneity. Fig. 1(b)–(d) show the segmentation process of a noisy synthetic image using the proposed LCV model. The robustness property of our model is due to the usage of global term and length penalty term. It can be seen from Fig. 1(a) that there are four separated shapes within the original image. To show the high efficiency of our model, the initial contour was placed at the bottom left corner unlike other methods in which the initial contour was usually placed in the center of images or touches the target objects. Afterwards, the initial contour evolves and further splits into several parts according to the topological structure of target shapes, as shown in Fig. 1(c). Finally, the evolving curves stop on the true boundary of each shape after 40 iterations, as can be seen in Fig. 1(d). Since the parameter of evolution termination criterion $T_T = 10$ is predefined before the evolution process, it is obvious that the evolving contour has already arrived at the true boundary at the 30th iteration. It should be noticed that this conclusion is suitable for all experiments using our model in this paper.

Fig. 2 shows the segmentation process of a real image of DNA channel using the LCV model. The objective is to extract the nuclei which appear much brighter than the background in the DNA channel, as shown in Fig. 2(a). Some nuclei are very close to each other, and traditional thresholding methods may fail to segment them. The initial contour was still placed at the bottom left corner, as shown in Fig. 2(b). After 4 iterations, the evolving curves successfully approach and surround all the nuclei, as shown in Fig. 2(c). It can be seen from Fig. 2(d) that the majority of nuclei have been segmented after 10 iterations while two pairs of nuclei are still connected. The evolving curves continue to split and all nuclei are successfully segmented at the 25th iteration (accomplished at the 15th iteration due to $T_T = 10$, as shown in Fig. 2(e)). Here, $\mu$ is set to $0.01 \times 255^2$ since the objective is to segment those small nuclei.

In the next experiments (Figs. 3–5), we shall illustrate the ability of the proposed LCV model to segmenting images with the intensity...
Fig. 1. Noisy synthetic image segmentation using the proposed LCV model: (a) original noisy image; (b) initial contour; (c) intermediate segmentation result at the 20th iteration; and (d) final segmentation result at the 40th iteration. Size = 114×112, $\alpha = \beta = 1$, $\mu = 0.1 \times 255^2$, and processing time = 3.1 s.

Fig. 2. Segmentation of a real image of DNA channel using the LCV model: (a) original image; (b) initial contour; (c) intermediate segmentation result at the 4th iteration; (d) intermediate segmentation result at the 10th iteration; (e) intermediate segmentation result at the 15th iteration; and (f) final segmentation result at the 25th iteration. Size = 217×160, $\alpha = \beta = 1$, $\mu = 0.01 \times 255^2$, and processing time = 5.2 s.

Fig. 3. The comparisons of the CV model and the proposed LCV model on segmenting a synthetic image with the intensity inhomogeneity: (a) original image; (b) initial contour; (c) final segmentation result using the CV model; and (d) final segmentation result using the proposed LCV model. Size = 88×85, $\alpha = \beta = 1$ (LCV), $\mu = 0.1 \times 255^2$ (CV/LCV), and processing time = 5.3 s (CV), 1.1 s (LCV).
inhomogeneity. Fig. 3 shows the segmentation results for the well-known synthetic image using both the CV model and the proposed LCV model. It can be seen from Fig. 3(a) that the intensity decreases gradually from the left to the right. The initial contour was placed at the intersectional region of high intensity area and low intensity area, as shown in Fig. 3(b). The failing segmentation of this synthetic image using the CV model is illustrated in Fig. 3(c), which shows that the evolving curve of the CV model cannot pass through the intersectional region of high intensity area and low intensity area even after 100 iterations. Fig. 3(d) shows the segmentation result of the LCV model where the evolving curve of the LCV model can successfully stop on the boundary of object after 19 iterations. Here, the value of the controlling parameter $\alpha = 0.1$ is less than that of $\beta = 1$ to maintain a restriction effect on the intensity inhomogeneity.

Fig. 4 shows the segmentation results for three real blood vessel images with the intensity inhomogeneity using both the CV model and the proposed LCV model. To achieve a fair comparison, we added the penalty term described in (20) into the CV model and the proposed LCV model. To achieve a fair comparison, we added the penalty term described in (20) into the CV model to avoid the re-initialization step. In addition, the proposed termination criterion has also been applied to the CV model. It can be seen from the third column of Fig. 4 that the CV model failed to segment all three images with the intensity inhomogeneity as we anticipated. The reason is due to the inherent disadvantage of not using the local information. The fourth column of Fig. 4 shows that the proposed LCV model can successfully segment the images in the first column. It should be noticed that the iteration times will be efficiently decreased if the initial contours are placed on certain part of the objects, as shown in Fig. 5.

Here, we also used the recent popular local binary fitting (LBF) model [27] to segment the images in Fig. 4. As we discussed in Section 1, the LBF model usually needs to perform four convolution operations at each iteration and is sensitive to the selection of governing parameters and the location of initial contour. We tried many times and selected the best governing parameters $\mu = 0.001 \times 255^2$ (the length controlling parameter), $\sigma = 3.5$ (the standard deviation of Gaussian kernel for three images) and the initial contour for the LBF model. In contrast, the parameters of the LCV model are the same as the ones in Fig. 4. Fig. 5 illustrates the segmentation results of the LBF model and the proposed LCV model with the same initial contours. Both models have succeeded in the segmentation task. Their iteration numbers and processing time for segmenting images in Fig. 5 are presented in Table 1. It can be seen from Table 1 that the iteration number and processing time for the LCV model are both less than that of the LBF model for all three image segmentation. Note that the LCV model only needs to perform one convolution operation before level set evolution. Considering that the parameters and initial contours of the LBF model are selected elaborately, so the LCV model is proved to be more efficient in segmenting the images with the intensity inhomogeneity.

The next three experiments (Figs. 6–8) are focused on illustrating the ability of segmenting texture images of the LCV model combined with the extended structure tensor (EST). Fig. 6 shows the
Fig. 5. The comparisons of the LBF model and the proposed LCV model on segmenting the images in Fig. 4. The first column: initial contours. The second column: final segmentation results using the LBF model. The third column: final segmentation results using the LCV model.

Table 1
Iteration number and processing time for the LBF model and proposed LCV model in segmenting the images in Fig. 5.

<table>
<thead>
<tr>
<th>Image from the First Row</th>
<th>Iteration Number</th>
<th>Processing Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBF</td>
<td>28</td>
<td>4.9</td>
</tr>
<tr>
<td>LCV</td>
<td>27</td>
<td>2.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Image from the Second Row</th>
<th>Iteration Number</th>
<th>Processing Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBF</td>
<td>50</td>
<td>6.8</td>
</tr>
<tr>
<td>LCV</td>
<td>29</td>
<td>2.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Image from the Third Row</th>
<th>Iteration Number</th>
<th>Processing Time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBF</td>
<td>59</td>
<td>7.6</td>
</tr>
<tr>
<td>LCV</td>
<td>25</td>
<td>2.9</td>
</tr>
</tbody>
</table>

segmentation process of three kinds of synthetic texture images. The first kind of image contains one texture object (the first row), the second kind of image contains three objects with the same texture (the second row), and the third kind of image contains two objects with different texture (the third row). According to the algorithm steps described in Section 3.8, the average of nine channels of the extended structure tensor (EST) is utilized for level set evolution instead of original image. Since the texture objects are usually located in the center parts of original images, the center parts of the computed averages are always in high-frequency area. Correspondingly, the boundary parts of the computed averages are in low-frequency area. Thus, the boundary parts can be firstly captured by the initial contours at the image corner. The second column shows the intermediate segmentation results for three texture images after one iteration. It can be seen that the small initial contours in the first column have quickly expanded and involve the majority of the images. After 4 or 5 iterations, the evolving curves successfully approach the true boundaries of texture objects, as shown in the third column. It can be seen from the fourth column that all texture objects are successfully segmented. Fig. 7 shows the segmentation result for a real giraffe image which contains nature texture produced from the upper body of the giraffe. It can be seen from Fig. 7(c) that the desired upper body of giraffe was successfully segmented.

Fig. 8 shows the segmentation results for a texture image with the intensity inhomogeneity using the LCV model combined with the classical structure tensor and the EST. First, the classical structure tensor was incorporated into the LCV model to segment this texture image. It can be seen from Fig. 8(c) that the LCV model
Fig. 6. Segmentation of three synthetic texture images using the LCV model combined with the EST. The first column: initial contours. The second column: intermediate segmentation result. The third column: intermediate segmentation result. The fourth column: final segmentation results. Size = 140 × 130, 128 × 128, 128 × 128, \( x = \beta = 1 \), and \( \mu = 0.01 \times 255^2 \).

Fig. 7. Segmentation of a real giraffe image using the LCV model combined with the EST: (a) original image; (b) initial contour; and (c) final segmentation result. Size = 279 × 205, \( x = \beta = 1 \), and \( \mu = 0.01 \times 255^2 \).

combined with the classical structure tensor failed to segment the image. The reason is due to the fact that the intensity inhomogeneity effect can be regarded as a smooth spatial varying function over the original image and the classical structure tensor does not contain any original intensity information. Since the intensity information is incorporated into the EST, the segmentation process would not be influenced by the intensity inhomogeneity any further. The segmentation result illustrated in Fig. 8(d) has also proved the efficiency of the EST.

As we discussed in Section 3.3, re-initialization can be regarded as a numerical remedy for maintaining stable curve evolution in the level set methods. However, it is usually quite complicated and...
The performance comparisons of proposed LCV model using the fast marching method and the penalty term in segmenting the images in Figs. 2 and 3.

<table>
<thead>
<tr>
<th></th>
<th>Fig. 2(a), size = 217×160</th>
<th></th>
<th>Fig. 3(a), size = 288×285</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average computational</td>
<td>Iteration number</td>
<td>Average computational</td>
</tr>
<tr>
<td></td>
<td>time (s)</td>
<td></td>
<td>time (s)</td>
</tr>
<tr>
<td>Fast marching method</td>
<td>0.0938</td>
<td>80</td>
<td>0.0351</td>
</tr>
<tr>
<td>Penalty term</td>
<td>0.0313</td>
<td>25</td>
<td>0.0156</td>
</tr>
</tbody>
</table>

time-consuming. Therefore, the fast marching method [42] was proposed for finding a solution to these problems. In this paper, we added a penalty term into the regularization term to completely avoid the re-initialization step. To further show the advantage of this penalty term, the fast marching method was also used and compared with our non-re-initialization solution scheme. Accordingly, the penalty term was removed from the energy functional and the fast marching method was adopted to re-initialize level set function to a signed distance function at every iteration. Table 2 shows the performance comparisons of proposed LCV model using the fast marching method and the penalty term in segmenting the images in Figs. 2 and 3. Both images were successfully segmented by the LCV model using two solution schemes. It can be seen from Table 2 that the average computational time of fast marching method is greater than that of penalty term. In addition, the iteration number for the LCV model using fast marching method are greatly increased, especially in segmenting the larger image in Fig. 2. As a result, the total processing time for the LCV model using fast marching method is obviously greater than that for the LCV model using penalty term. It can be concluded that the penalty term is a good substitute for fast marching method while solving the re-initialization problem in the proposed LCV model.

Here, we also provided more experiments on four different types of real images to further demonstrate the performance of the proposed LCV model. It should be noted that all the initial contours were placed in the background area, as shown in the second column of Fig. 9. The first row of Fig. 9 shows the segmentation of a wrist X-ray image with obvious intensity inhomogeneity. The true boundary of the wrist is covered by certain blurry light, which will result in wrong segmentation result if only the global information is used. The successful segmentation result using the proposed LCV model is illustrated in the third image of the first row of Fig. 9. The second row shows the segmentation of a real starfish image with nature texture. Thus, the LCV model combined with the EST was adopted and the successful segmentation result is shown in the third image of the second row of Fig. 9. The third row shows the segmentation of a real garden image in which segmenting the fire hydrant is our objective. Obviously, the segmentation process will be influenced by the existences of wall, gate and grass if the traditional solutions are used. Although this image does not contain obvious texture, the EST containing intensity information can still be adopted for eliminating the influences of wall, gate and grass. The segmentation result in the third image of the third row shows that the LCV model combined with the EST successfully segmented the fire hydrant while the wall, the gate and the grass were not included. As mentioned above, the LCV model combined with the EST can also segment the texture image with the intensity inhomogeneity, as can be seen from the synthetic texture image segmentation in Fig. 8. Next, we shall provide a real case to further validate this conclusion. The fourth row shows the segmentation of a real corpus callosum MR image with intensity inhomogeneity. The target is to segment the blurry corpus callosum region which is surrounded by the dim brain tissues. To eliminate the influences caused by both surrounding brain tissue and the intensity inhomogeneity effect, the EST was incorporated into LCV model for segmentation with the controlling parameters $x = 0.1$ and $\beta = 1$. It can be seen from the third image of the fourth row that the blurry corpus callosum region was successfully segmented.

Remark 2. In general, the Courant–Friedrichs–Levy (CFL) rule will restrict the allowable step size of the partial differential equation (PDE) when the explicit numerical scheme in Section 3.6 is applied. That is, for any given grid spacing $h$ and the velocity $V$, we require $F \cdot \Delta t \leq h$, where $\Delta t$ is the time-step. Thus, the CFL condition places an upper bound on the time-step and hence on the speed of curve evolution. In our experiments, we used the same evolving parameters of the time-step $\Delta t = 0.1$ and the grid spacing $h = 1$. It seems that the proposed LCV model can sometimes segment the target objects in very few iterations, especially in Figs. 2 and 6. As we discussed in Section 3.2, the utilization of local term can significantly increase the contrast between foreground intensities and background intensities. Thus, the ability of evolving curve to approach the true boundary has also been greatly improved while compared with
Fig. 9. Segmentation of four different types of real images using the LCV model. The first column: original images. The second column: initial contours. The third column: final segmentation results using the proposed LCV model with different schemes. Size = 164×323, 300×294, 255×230, 238×159, α = 0.1, β = 1 (the first and the fourth row), α = β = 1 (the second and the third row), and μ = 0.01 × 255².

traditional level set methods. Moreover, a larger window size of convolution operator k = 15 was adopted which also implicitly updates the grid spacing in the numerical solutions. As a result, the small initial contour can quickly expand and approach the true boundary even starting from the image corner, as can be seen from Fig. 2(c). For texture image segmentation, the average of nine channels of the extended structure tensor (EST) is utilized for level set evolution instead of original image. Since most texture objects are located in the center part of the original image, the center part of the computed average is in high-frequency area. Correspondingly, the boundary part of the computed average is in low-frequency area which can be firstly captured by the initial contour at the image corner. Hence, the evolving curves will expand to surround the center part of image and quickly approach the true boundaries of texture objects under the dual effect of the EST and the local term. The second to the fourth columns of Fig. 6 properly demonstrate this process. By incorporating the local term and the EST, our solution scheme adds some flexibility to the CFL constraint, which makes the curve evolution...
maintain stable and converge in fewer iterations while segmenting some ordinary images and texture images. However, it should be emphasized that the CFL rule still takes effect in our numerical scheme. It can be seen from the third and the fourth rows of Fig. 9 that the iteration times for two segmentations are relatively large due to the complicated background.

5. Conclusions and future works

In this paper, we propose a new local Chan–Vese (LCV) model for image segmentation, which is based on the techniques of curve evolution, local statistical function and level set theory. The energy functional for the proposed model consists of global term, local term and regularization term. By incorporating the local image information into our model, the images with intensity inhomogeneity can be efficiently segmented. To avoid the time-consuming re-initialization step, a new penalizing energy is introduced into the regularization term. Moreover, a termination criterion based on the length of the evolving curve is proposed to ensure that the evolving curve can automatically stop on the true boundaries of objects. Particularly, by adding the intensity information into the classical structure tensor, an extended structure tensor (EST) is constructed for texture image segmentation. Combining the EST with the proposed LCV model, the texture image can be efficiently segmented no matter whether it presents intensity inhomogeneity or not. Finally, experiments on some synthetic and real images have demonstrated the desired segmentation performance of our proposed model for the (texture) images with or without intensity inhomogeneity. The comparisons with the CV model and the LBF model also show that the proposed LCV model has a much faster convergence speed and less sensitivity to the location of initial contour and the selection of governing parameters.

It should be noted that the proposed LCV model is efficient for the two-modal (phase) images, which usually generate two segments, i.e., foreground and background. As a bimodal model, it cannot simultaneously detect multiple objects in different intensities and the triple junctions. In our future work, we will extend the current LCV model by using the multi-layer idea (transformation of single phase to layer) or multiple level set functions. Further extension can be made to obtain a multimodal LCV model, so that multi-modal (phase) images and regions with triple junctions can be segmented.

Acknowledgements

This work was supported by the grants of the National Science Foundation of China, Nos. 60873012, 60805021, 60705007 & 30700161, the grant from the National Basic Research Program of China (973 Program), No.2007CB311002, the grants from the National High Technology Research and Development Program of China (863 Program), Nos. 2007AA01Z167, the grant of the Guide Project of Innovative Base of Chinese Academy of Sciences (CAS), No.KSCX1-YW-R-30, the grant of the Graduate Students’ Scientific Innovative Project Foundation of CAS (X.-F. Wang).

Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.patcog.2009.08.002.

References


About the Author—XIAO-FENG WANG received the B.Sc. degree in Computer and Science Technology from Anhui University, Hefei, China, in 1999, the M.Sc. degree in Pattern Recognition and Intelligent System from Institute of Intelligent Machines (IIM), Chinese Academy of Sciences, Hefei, China, in 2005. He is now in pursuit for Ph.D. degree in Pattern Recognition and Intelligent System in University of Science and Technology of China (USTC). His research interests include pattern recognition, image processing and data mining.

About the Author—DE-SHUANG HUANG received the B.Sc., M.Sc. and Ph.D. degrees all in Electronic Engineering from Institute of Electronic Engineering, Hefei, China, National Defense University of Science and Technology, Changsha, China, and Xi’an University, Xian, China, in 1986, 1989 and 1993, respectively. During the 1993–1997 period he was a postdoctoral student, respectively, in Beijing Institute of Technology and in National Key Laboratory of Pattern Recognition, Chinese Academy of Sciences Beijing, China. In September, 2000, he joined the Institute of Intelligent Machines, Chinese Academy of Sciences as the Recipient of “Hundred Talents Program of CAS”. From September 2000 to March 2001, he worked as Research Associate in Hong Kong Polytechnic University. From April 2002 to June 2003, he worked as Research Fellow in City University of Hong Kong. From August to September 2003, he visited the George Washington University as visiting professor, Washington DC, USA. From October to December 2003, he worked as Research Fellow in Hong Kong Polytechnic University. From July to December 2004, he worked as the University Fellow in Hong Kong Baptist University. Dr. Huang is currently a senior member of the IEEE.

About the Author—HUAN XU received the B.Sc. degree in Central China Normal University, majored in Computer Software Science and minored in Mathematics, in 2006, the M.Sc. degree in Pattern Recognition and Intelligent System from University of Science and Technology of China (USTC), in 2009. She is now in pursuit for Ph.D. degree in the Electrical and Computer Engineering Department of Louisiana State University. Her research interests include image analysis, visual and geometric computing.