Computers in Biology and Medicine 41 (2011) 292–301

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Homogeneity- and density distance-driven active contours for medical image segmentation

Phan Tran Ho Truc [a], Tae-Seong Kim [b], Sungyoung Lee [a], Young-Koo Lee [a,*]

[a] *Department of Computer Engineering, Kyung Hee University, Repuplic of Korea*
[b] *Department of Biomedical Engineering, Kyung Hee University, Republic of Korea*

ARTICLE INFO

Article history:
Received 26 January 2010
Accepted 18 March 2011

Keywords:
Active contours
Image segmentation
Level set methods
Medical images

ABSTRACT

In this paper, we present a novel active contour (AC) model for medical image segmentation that is based on a convex combination of two energy functionals to both minimize the inhomogeneity within an object and maximize the distance between the object and the background. This combination is necessary because objects in medical images, e.g., bones, are usually highly inhomogeneous while distinct organs should generate distinct image configurations. Compared with the conventional Chan–Vese AC, the proposed model yields similar performance in a set of CT images but performs better in an MRI data set, which is generally in lower contrast.

1. Introduction

Medical image analysis has played a more and more important role in many clinical procedures due to the advancements in medical imaging modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound [1]. Medical image analysis deals with enhancement, segmentation, registration, and visualization, among which segmentation is a very crucial task because it provides the organ of interest, such as bone [2,3], brain [4], or heart [5], that is necessary for clinical diagnosis and/or treatment [6]. Image segmentation is to partition the image into its constituent parts which correspond to separated objects. One then may think of the extraction of object boundaries, where simple edge detectors like the gradient-based and the second-order derivative-based operators [7] or a more elaborated approach like Canny edge detector [8] are widely used. However, an edge detector is usually not suitable for extracting object boundaries due to many reasons. Firstly, extracted edges do not always correspond to object boundaries. For example, one may think of texture. Secondly, edge detectors usually yield discontinued edges, whereas objects are necessarily separated by closed contours. So post-processing tasks are needed to link discontinued edges, which are complex and prone to be erroneous. Finally, edge detectors depend on the local information in a neighborhood of a pixel. Being local is sometimes advantageous, yet in many cases, the global view of the object appearance is of significant clues. Therefore, image segmentation in general is different from edge detection. The former is to provide regions, represented by closed boundaries, not edges.

Region growing [7,9] is one of the simple techniques that provide regions. Starting with a set of seed points, the algorithm successively appends to each seed point its neighboring pixels that share similar image features such as intensity, texture, or color to form larger regions. This is an iterative process that stops when all pixels are processed. The algorithm can be regarded as a heuristic minimization for the Mumford-Shah functional [10] where the energy decreases while the regions are growing. Therefore, just like its energy optimization counterpart, region growing suffers from the sensitivity to seed selection as the initial condition, which can lead to under- or over-growing.

Another region-providing method is *snake* or *active contour* (AC). An AC model is the description of contours in 2D or surfaces in 3D which evolve under an appropriate energy to move toward desired features, such as object boundaries. Because contours are always closed, object boundaries extracted are continuous, making post-processing tasks to connect discontinued edges no longer necessary. Since it was first introduced by Kass, Witkin, and Terzopoulos [11], active contour has attracted a large amount of researches: many AC models have been proposed, which can be categorized into parametric-type and geometric-type ACs. In parametric ACs [11–15], the curve (contour) C is explicitly represented using its parameterization: $C(p) = [x(p), y(p)]$, where $p (0 \leq p \leq 1)$ parameterizes the curve. This makes the parametric ACs non-intrinsic because their energy functional depends on the parameterizations but not on the geometry of the contour. As a result, these models cannot naturally handle topological changes to simultaneously detect multiple objects. Many special (usually heuristic) procedures have been proposed in detecting possible splitting and merging [16–18] but prior knowledge about the number of objects needs to be given.

* Corresponding author. Fax: +82 31 202 2520.
E-mail address: yklee@khu.ac.kr (Y.-K. Lee).
URL: http://uclab.khu.ac.kr (Y.-K. Lee).

0010-4825/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compbiomed.2011.03.006

Geometric-type ACs [19–24], on the other hand, can handle topological changes without additional efforts because they are implemented using level-set framework [25,26]. In this framework, the curve C is implicitly represented by the zero level set of a function $\phi(\mathbf{x},t):\mathbb{R}^n\times[0,\infty)\rightarrow\mathbb{R}$, $n=\{2,3\}$ such that

$$C=\{\mathbf{x}\in\mathbb{R}^n:\phi(\mathbf{x},t)=0\}. \quad (1)$$

The function ϕ is then evolved using the following general equation

$$\frac{\partial\phi}{\partial t}=\underbrace{b\kappa|\nabla\phi|}_{\text{Curvature Based}}+\underbrace{V_N|\nabla\phi|}_{\text{Normal Direction}}+\underbrace{\vec{S}\cdot\nabla\phi}_{\text{Vector Field Based}}, \quad (2)$$

where κ is the Euclidean curvature and (b, V_N, and $\vec{S}$) are three parameters determining the velocity and direction of the evolution. The curvature-based force is to smooth the curve; the normal direction force shrinks or expands the curve along its normal direction; and the external vector field-based force acts as a translation operator. Although the function ϕ itself moves up and down on a fixed coordinate system without changing its topology, its zero level set (or the curve C) may automatically split or merge.

The first geometric-type AC is the (original) *Geometric AC* which was introduced independently by Caselles et al. [19] and Malladi et al. [27]. The main idea is to move the curve using the curvature and the normal direction forces and stop the motion at the object boundaries using an edge-based function $g(\mathbf{x})=g(|\nabla I|^2)$ (where ∇I is the gradient of the input image I) which approaches 0 on the edges and 1 otherwise, e.g., $g(\mathbf{x})=e^{-(1/\sigma_e^2)|\nabla G_\sigma * I(\mathbf{x})|^2}$ with σ_e a scaling factor and G_σ the smoothing Gaussian kernel of scale σ. Caselles et al. [20] proposed another geometric-type AC, called *Geodesic AC*, using an energy functional to search for a curve of minimal edge-weighted length (geodesic curve). This model is similar to the Geometric AC, except that ∇g is used as a vector field force to increase the curve attraction towards weak edges. Then, Paragios et al. [24] proposed to replace the vector field force ∇g with the well-known gradient vector flow (GVF) introduced by Xu and Prince [15] to increase the capture range, leading to the *GVF Fast Geometric AC* (shortly, GVF-Geo AC). Differently, Chan and Vese [22] proposed a new model, which we call the *CV AC*, using an approximation of the Mumford-Shah functional. All these geometric-type ACs are considered as classical models in the research field and their level-set parameters are summarized in Table 1. Here, we do not consider other approaches that incorporate prior knowledge about object shape [28–30] or texture [31] since they require a training stage which is generally application specific.

We can see from Table 1 that the first three models depend heavily on the edge function $g(\mathbf{x})$, making themselves prone to be trapped in false edges caused by noises. This can be alleviated by performing smoothing with larger σ, yet it in turn leads to inexact results because edges are smoothed as well. The CV AC, on the other hand, does not depend on g (this gives it the name "without-edge AC") but on the homogeneity assumption, i.e., image features within a segment should be similar. In this case, the image I is assumed to be consisted of two segments with approximately piecewise-constant intensities. The CV energy functional $F(C)$ is defined as

$$F(C)=F_1(C)+F_2(C)=\int_{inside(C)}|I(\mathbf{x})-c_{in}|^2\mathrm{d}\mathbf{x}+\int_{outside(C)}|I(\mathbf{x})-c_{out}|^2\mathrm{d}\mathbf{x}, \quad (3)$$

where c_{in} and c_{out} are, respectively, the average intensities inside and outside the variable curve C. Compared to the other three AC models, the CV AC can detect the objects more exactly since it does not need to smooth the initial image (via the edge function $g(|\nabla I_\sigma|^2)$ where $I_\sigma=G_\sigma * I$), even if it is very noisy. In other words, this model is more robust to noise and thus suitable for medical images since they are often noisy and low contrast. Also, it was shown to provide a relaxed initial position requirement [22].

However, the convergence of the CV AC depends on the homogeneity of the segmented objects. When the inhomogeneity becomes large like in carpal bones or knee bones, the CV AC provides unsatisfactory results. To address this, let us consider a synthetic image (size 128×128) with an inhomogeneous object having five different parts over the bright background as shown in Fig. 1. The image intensity is scaled on the range [0, 1], with 1 the brightest. The CV fitting term $F(C)$ is calculated at each iteration during the evolution and plotted in Fig. 2.[1] As expected, the curve moves in the direction of decreasing $F(C)$ and stops when $F(C)$ reaches a minimum value, which is $F(C^*)=34$ (at iteration number 16) in this case. Nevertheless, this is not the "desirable" result, whose minimum fitting term is $F(C^{des})=197$. Clearly, the desirable minimum here is larger (more local) than the practically resulting minimum $F(C^*)$.

From the above example, we can see that the global minimum of the CV energy functional does not always guarantee the desirable results, especially when a segment is highly inhomogeneous. To provide flexibility in searching for the desirable minimum (which is often neither the most local nor global), Li and Yezzi [32] proposed a *dual-front AC* model with the active region's width as a controlling factor. The model is an iterative process consisting of the active region relocation and the dual front evolution which is another iterative process, demanding a high computational cost.

Vese and Chan [23] and Tsai et al. [33] independently and contemporaneously proposed to use the original Mumford-Shah functional [10] to segment inhomogeneous objects. Because the minimizer of the Mumford-Shah functional is difficult to get (and remains an issue) due to the term of discontinuities, the authors in [23,33] presented the set of discontinuities in form of a curve evolution problem. The resulting optimization process involves both evolving a level-set function and solving Poison partial differential equations. Although it can generate a piecewise smooth approximation of the input image that well represents the objects of interest, this process is very complicated and computationally expensive and requires a good initialization. Another piecewise smooth approach was presented in [34]. The authors elegantly generalized the mean intensities c_{in} and c_{out} in (3) to the local weighted averaging using a Gaussian kernel convolution. This leads to a model that approximates the original Mumford-Shah functional but has a complexity close to that of the CV model. When the variance of the Gaussian kernel approaches infinity, this model becomes the CV model. On one hand, this variance parameter

Table 1
Level-set parameters of the classical AC models.

	b	V_N	$\vec{S}$
Geometric	$g(\mathbf{x})$	$\eta g(\mathbf{x})$	$\vec{0}$
Geodesic	$g(\mathbf{x})$	$\eta g(\mathbf{x})$	$\nabla g(\mathbf{x})$
GVF-Geo	$g(\mathbf{x})$	$\eta K(\mathbf{x})g(\mathbf{x})$	$g(\mathbf{x})(1-\lvert K(\mathbf{x})\rvert)[\hat{u},\hat{v}]$
CV	γ	$\eta+(I-c_{in})^2-(I-c_{out})^2$	$\vec{0}$

γ and η are constants, $[\hat{u},\hat{v}]$ the GVF [15], $K(\mathbf{x})$ a function depending on the curve normal and the GVF, $I=I(\mathbf{x})$ the image intensity, and c_{in} and c_{out}, respectively, the average intensity values inside and outside the variable curve.

[1] Note that Fig. 2 plots only the term $F(C)$ in (3), whereas the result in Fig. 1 was obtained from the original CV model [22] that includes both $F(C)$ and a length term to guarantee the smoothness of the contour.