A HYBRID FRAMEWORK FOR IMAGE SEGMENTATION

Huiyu Zhou[1], *Tangwei Liu*[2], *Huosheng Hu*[1], *Yusheng Pang*[2], *Faquan Lin*[2], *and Ji Wu*[2]

[1]University of Essex, Colchester, CO4 3SQ, United Kingdom
[2]Guangxi Medical University, Nanning, 530027, P.R. of China

ABSTRACT

This paper presents a new approach of image segmentation by combining the classical gradient vector flow (GVF) algorithm with mean shift. Due to the dependence on the gradient vectors of an edge map, the classical GVF is sensitive to the shape irregularities, and hence the snake cannot be ideally located on the concave boundaries. We propose an improved representation of the internal energy force by reducing the Euclidean distance between the guessed centroid and the estimated one of the snake. Experimental work shows the performance of this approach in different tests.

1. INTRODUCTION

1.1. Background

Image segmentation is an important topic in computer vision, which is intended to distinguish objects from the background. The success of such a process may lead to adequate video surveillance, image compression and visual representation. An image segmentation process technically can be divided into two steps: (1) discontinuity measurement and (2) grouping and discrimination of similar pixels.

It is hence desirable to devise schemes which can be used to detect brightness or color discontinuity between objects and set up a distinguishing criterion. This can be dealt with by using a global threshold technique, e.g. [1]. However, this technique is not capable of handling the non-uniform regions. Consequently, a system of masking images with different thresholds is proposed in [2] to find the image edge. Evidence shows that this method is not trivial as the collected images may be inhomogeneous, where proper thresholding is too difficult to be generated. Alternatively, gradient-based approaches have consistently attracted the attention of academics due to the fast and prominent appearence of discontinuous areas. People have looked at a variety of perspectives, consisting of edge based, clustering, region growing, histogram thresholding, partial differential equations (PDE) and others. In this context, we expect to cope with the segmentation of black-and-white images. Therefore, gray-level is the unique factor to be concerned with in the segmentation.

Edge based approaches demand edge representation and smoothing. The extracted edges are then joined into segments while features are computed. It is acknowledged that edge based segmentation has not been completely successful due to small gaps that can merge dissimilar regions [3]. Clustering algorithms take into account pixel comparisons but they fail to identify local variation that possibly affects the cluster prototypes. Region growing based techniques try to merge the adjoining pixels similar to the selected seeds within an identical region. The main limitation of these techniques is their inability to segment sharp corners or concave regions. As a typical example of histogram thresholding framework, mean shift is a kernel type of measurement, where the kernel defines distance between pixels that encompass both spatial and temporal information. The significant drawback is that the mean shift cannot obtain globally optimal bandwidths when the local regions contain significantly varied pixels [4]. PDE methods, such as active contour methods, compute the energy function by considering elasticity and rigidity of contours [5]. To make the computation fast, level set and fast marching are then introduced [6], [7]. To challenge the problems associated with initialization and poor convergence to concave boundaries, [8] presents a new external force for active contours, which is called gradient vector flow (GVF) computed as a diffusion of the gradient vectors of a gray-level or binary edge map. These strategies, however, still suffer from instability and misleading convergence as well as the requirement of *a priori* kowledge of the region of interest.

1.2. Our contribution

In this paper, we propose a novel approach to improve the existing GVF performance on segmentation, incorporating a well established mean shift framework. This proposal comprises of significant contributions as follows: firstly, the mean shift technique is used to constrain the spatial diffusion of the gradient so that it is able to handle boundary cancavities or gaps that the original GVF cannot deal with efficiently. Secondly, the energy function representing the enhanced GVF snake is properly parameterized by adapting the coefficients of elasticity and rigidity along the potential contour. Thirdly, the proposed system is less sensitive to the initial contours than the classical GVF snake as the error between the estimated and the ideal density is dynamically rectified in the image domain, which prevents the snake missing the real boundary.

The remainder of this paper is organized as follows: in Section 2 we describe the classical snake and then GVF scheme, while their drawbacks are also pointed out. Section 3 introduces the application of mean shift in the improved GVF snake. Some examples will be shown in Section 4 to justify the performance of our approach. Finally, Section 5 provides conclusions and future work.

2. ACTIVE CONTOUR MODEL

2.1. Traditional snake

Active contour (or snake) model has been broadly used in image-related research due to explicit description of boundary evolving. This is the consequence of using a combinatorial energy function to represent the dynamic characteristics of the boundary propagation. The traditional snake in 2-D is delineated as follows

$$E = \int_0^1 \frac{1}{2}(\alpha \parallel \mathbf{c}'(s) \parallel^2 + \beta \parallel \mathbf{c}''(s) \parallel^2) + E_{ext}(\mathbf{c}(s))ds, \tag{1}$$

where $\mathbf{c}(s) = (x(s), y(s))$, and $s \in [0, 1]$. The first two terms in Eq. 1 stand for the internal forces, inlcuding elasticity and bending, where $\mathbf{c}'(s)$ and $\mathbf{c}''(s)$ are the first and second derivative respectively, and the coefficients α and β control the tension and rigidity of the snake. The last term in the equation consists of potential forces that drag the snake towards the edges.

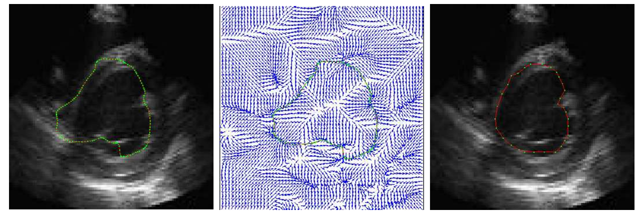
To overcome the limitation imposed by the traditional snake, e.g. inefficiency at concave boundaries, [8] have proposed the GVF snake based on the gradient vectors.

2.2. GVF snake

The GVF snake yields an external force field called *GVF field* in the continuous gradient domain. Technically, a binary edge map is demanding, which forces the snake to effectively approach the edge-like areas. This smoothing process is described by the following function [9] $g(s) = \frac{1}{2\pi\sqrt{\sigma}} \exp^{-\frac{|\nabla(G_I(s))|^2}{2\sigma^2}}$, and $f(x, y) = 1 - g(s)$, where $G_I(s)$ represents the convolution of the image I with a Gaussian kernel, $f(x, y)$ is the edge map and σ is the variance of the kernel.

A GVF field is defined as a 2-D vector $\mathbf{v} = (u(s), v(s))$ that minimizes the energy function [8]

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



(a) GVF. (b) Gradient map. (c) Manual.

Fig. 1. Comparison between GVF and manual segmentation.

where μ is a blending parameter.

When the GVF snake is finally settled, where the internal and external forces are balanced, we shall have the relationship as $\alpha\mathbf{c}''(s) - \beta\mathbf{c}''''(s) + \gamma\mathbf{v} = 0$, where γ is a proportional coefficient.

2.3. Defects of GVF snake

It has been recognized that the GVF snake primarily relies on the detected edges and the corresponding strength. A cluttered or noisy scene, e.g. a medical image, possibly contain multiple structures, betraying the snake to incorrect contours. Furthermore, the GVF snake is discovered to be sensitive to the shape irregularities. As a result, the GVF force field cannot push or pull the snake to the ideal region (see Fig. 1). To enhance the behavior of the classical GVF snake, people have focused on modifying the external force function so as to regulate the evolution of the snake in general circumstances [9], [10], [11]. Unfortunately, this is an image-specific operation so a generalized framework is being deveploed.

It is evident that Eq. 2 becomes valid after a number of iterations is undertaken. People expect a consistent minimization in a real image by adjusting the parameterization of the external energy. For example, Yu and Bajaj [11] propose to compute the GVF force field by using a polar coordinate form rather than cartesian coordinates. This approach performs better than the classical GVF snake in areas of long thin concave boundaries and boundary gaps. However, due to less capture range the snake possesses slow convergence.

3. IMPROVED GVF SNAKE

3.1. Probabilistic model

In contrast to the approaches related to the external force, we assign two weight functions to the relationship described in Section 2.2 so that it looks like

$$h_\alpha\mathbf{c}''(s) - h_\beta\mathbf{c}''''(s) + \gamma\mathbf{v} = 0, \tag{3}$$

where h_α and h_β denote the weight functions of elasticity and bending forces, respectively. Taking such an important change is based on the fact that given a fixed external force, the instability between the elasticity and rigidity leads to an ill-determined boundary, e.g. excess or insufficiency. In other words, the snake may tend to shrink (or expand) when the elastic property dominates the energy field. Otherwise, the snake will have to keep the original shape due to the rigid materials. In practice, people usually adjust the elasticity and rigidity coefficients in an attempt to arrive to a compromise on the ideal contour. However, this adjusting does not hold any theoretical basis and therefore requires a large number of attempts in different image environments.

Addressing this problem, we explore a mean shift based framework for iteratively determing h_α and h_β. Let us first look at a classical Bayes formulation within the narrow band of the target contour:

$$P(C|D_\alpha, D_\beta) = \frac{P(C|D_\beta)P(D_\alpha|C, D_\beta)}{P(D_\alpha|D_\beta)}, \quad (4)$$

where D_α and D_β are the contributions of the elasticity and rigidity components respectively, C is the location of the target contour, $P(C|D_\beta)$ is the prior probability while $P(D_\alpha|C, D_\beta)$ denotes the likelihood of the boundary. To obtain an ideal contour, $P(C|D_\alpha, D_\beta)$ needs to be maximized. The probability $P(C|D_\beta)$ can be determined by specifying an initial contour, or just assuming a uniform distribution in the image. We seek a solution for maximizing $P(D_\alpha|C, D_\beta)$ by using the mean shift algorithm, which will be detailed next.

3.2. Mean shift algorithm

Mean shift is employed to search for a contour candidate that has the most similar characteristics to that of the target contour. In this context, we regard the centroid of a snake as an energy descriptor, showing the approximate position of the region outlined by the snake. To efficiently constrain the contour propogation, the CAMSHIFT algorithm by [12] is used due to its preference of accounting for dynamically changing distributions during the evolution. The CAMSHIFT algorithm is the variation of the mean shift algorithm. So, we briefly review the principle of mean shift.

Given an image point sequence $\mathbf{s}_i$ $(i = 1, 2, ..., n)$ in the m-dimensional space R^m, then the multivariate kernel density estimate with kernel $K(\mathbf{s})$ and window radius r is given as

$$F(\mathbf{s}) = \frac{1}{nr^m}\sum_{i=1}^{n} K((\mathbf{s} - \mathbf{s}_i)/r). \quad (5)$$

The multivariate *Epanechnikov* kernel can be estimated by

$$K_E(\mathbf{s}) = \begin{cases} \frac{(m+2)(1-\|\mathbf{s}\|^2)}{2c_m}, & \| \mathbf{s} \| < 1 \\ 0, & otherwise \end{cases} \quad (6)$$

where c_m is the volume of the unit m-dimensional sphere.

Assuming a kernel $\Psi(\mathbf{s}) = c_0\psi(\| \mathbf{s} \|^2)$, where c_0 is a normalization constant, the mean shift vector is expressed as

$$MS(\mathbf{s}) \equiv \frac{\sum_{i=1}^{n} \mathbf{s}_i \psi(\| (\mathbf{s} - \mathbf{s}_i)/r \|^2)}{\sum_{i=1}^{n} \psi(\| (\mathbf{s} - \mathbf{s}_i)/r \|^2)} - \mathbf{s}, \quad (7)$$

where $\psi(\cdot)$ is an intermediate function [13]. The mean shift procedure, in fact, is a recursive evolution by computing the mean shift vector $MS(\mathbf{s})$ and adjusting the centroid of kernel Ψ by $MS(\mathbf{s})$.

To exploit the CAMSHIFT algorithm, we first calculate the zeroth moment M_{00}, moment M_{10} for x-coordinates, and moment M_{01} for y-coordinates of image points on the contour candidate. This requires an estimate of the Euclidean distance between the origin (0,0) and individual points before the moment calculation is conducted. The centroid (x_c, y_c) of the contour is then calculated by $x_c = M_{10}/M_{00}$ and $y_c = M_{01}/M_{00}$. The Euclidean distance d_e between the initial centroid and the estimated one is consistently calculated so that h_α and h_β can be estimated by

$$\begin{cases} h_\alpha = \alpha d_e, \\ h_\beta = \frac{\beta}{d_e}. \end{cases} \quad (8)$$

If $d_e \leq 1$ and Eq. 3 holds, then the evolution will stop as the convergence to the target contour has been reached; otherwise, the search has to continue. Meanwhile, the likelihood of the boundary $P(D_\alpha|C, D_\beta)$ also approaches its maxima.

4. EXPERIMENTAL RESULTS

We have tested the proposed mean shift based algorithm using a couple of images. The first one is a short-axis cardiac ultrasonic image. Obviously, this human heart image contains a number of structures besides the endocardial border, referred to Fig. 2. Before the segmentation is conducted, we need to initiate the snake and the centroid of the target contour, both of which have been illustrated in Fig. 2. According to Fig. 1 and 2, our approach has better performance on the border segmentation than the classical GVF algorithm. The second image is a hand placed in a cluttered scene (Fig. 3). The challenge for the traditional GVF is that it cannot effectively locate the concave areas in the scene. The rigidity component is dominant so it limits the diffusion of the snake to the real concaves. Fig. 3 demonstrates that the proposed mean shift based strategy allows the snake to be located at the ideal boundary. However, it also shows that the settled contour has less efficiency to avoid a couple of edge "bumps", indicating necessary improvements have to be launched later.

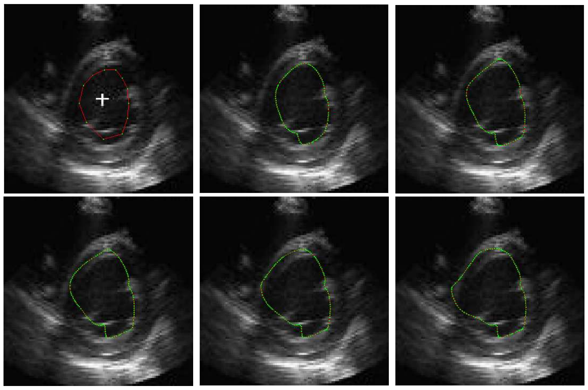


Fig. 2. Segmentation of a short-axis cardiac image using the improved GVF, where '+' shows the position of the initial centroid ($\alpha = 0.35; \beta = 0.4; \gamma = 1; \mu = 0.1$).

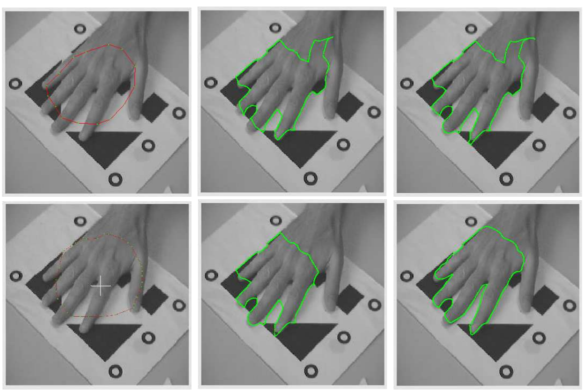


Fig. 3. Comparison between the classical GVF and our approach, where the upper row is by GVF and the lower one by our approach ($\alpha = 0.15; \beta = 0.2; \gamma = 1; \mu = 0.1$): image courtesy of Teo de Campos, University of Oxford.

5. CONCLUSIONS AND FUTURE WORK

We have presented a novel algorithm for image segmentation, integrating the classical GVF algorithm and the mean shift technique. The experimental results demonstrate that the snake is favourably performed in concave boundaries as well as in the cluttered scenes. However, we have not addressed the problem of computational costs. Future work will be mainly focused on reducing the computational cost of the optimization of the energy function.

6. ACKNOWLEDGEMENTS

This work is supported by Guangxi Science Foundation, P. R. of China (GKJ03429-3), and the UK EPSRC under Grant GR/S29089/01. Thanks for the comments from P. Jones.

7. REFERENCES

[1] F.F. Yin, M.L. Giger, K. Doi, C.E. Metz, C.J. Vyborny, and R.A. Schmidt, "Computerized detection of masses in digital mammograms: Analysis of bilateral substraction images," *Med. Phy.*, vol. 18, pp. 955–963, 1991.

[2] M. Abdel-Mottaleb, C.E. Carman, C.R. Hill, and S. Vafai, "Locating the boundary between the breast skin edge and the background in digitized mammograms," in *Proc. of 3rd International Workshop on Digital Mammography*, 1996, pp. 467–470.

[3] W.A. Perkins, "Area segmentation of images using edge points," *IEEE Tran. PAMI*, vol. 1, pp. 8–15, 1980.

[4] J. Wang, B. Thiesson, Y. Xu, and M. Cohen, "Image and video segmentation by anisotropic kernel mean shift," in *Proc. of ECCV*, 2004.

[5] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 4, pp. 321–331, 1987.

[6] R. Malladi, J.A. Sethian, and B.C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Tran. PAMI*, vol. 17, pp. 158–175, 1995.

[7] J.A. Sethian, "A fast marching level set method for monotonically advancing fronts," *Proc. Nat. Acad. Sci.*, vol. 93, pp. 1591–1595, 1996.

[8] C. Xu and J.L. Prince, "Gradient vector flow: A new external force for snakes," in *Proc. of CVPR*, 1997, pp. 66–71.

[9] N. Paragios, O. Mellina-Gottardo, and V. Ramesh, "Gradient vector flow fast geodesic active contour," in *Proc. of ICCV*, 2001, pp. 67–73.

[10] D. Gil and P. Radeva, "Curvature vector flow to assure convergent deformable models for shape modelling," in *Proc. of EMMCVPR, Portugal*, 2003.

[11] Z. Yu and C. Bajaj, "Image segmentation using gradient vector diffusion and region merging," in *Proc. of ICPR*, 2002, pp. 941–944.

[12] G.R. Bradski, "Computer vision face tracking for use in a perceptual user interface," *Intel Technology Journal*, , no. 2nd Quarter, 1998.

[13] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proc. of CVPR*, 2000, pp. 142–149.