



Fast Statistical Level Sets Image Segmentation for Biomedical Applications

Sophie Schüpp, Abderrahim Elmoataz, Jalal M. Fadili, Daniel Bloyet

► To cite this version:

Sophie Schüpp, Abderrahim Elmoataz, Jalal M. Fadili, Daniel Bloyet. Fast Statistical Level Sets Image Segmentation for Biomedical Applications. Third International Conference Scale-Space 2001, 2001, Vancouver, Canada. pp.380-388. hal-01016046

HAL Id: hal-01016046

<https://hal.science/hal-01016046>

Submitted on 27 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fast Statistical Level Sets Image Segmentation for Biomedical Applications

Sophie Schüpp, Abderrahim Elmoataz,
Mohamed-Jalal Fadili, and Daniel Bloyet

Groupe de Recherche en Informatique, Image et Instrumentation de Caen UMR 6072
GREYC-ISMRA 6, Bd Maréchal Juin, F-14050 Caen, France
sophie.schupp@greyc.ismra.fr

Abstract. In medical microscopy, image analysis offers to pathologist a modern tool, which can be applied to several problems in cancerology: quantification of DNA content, quantification of immunostaining, nuclear mitosis counting, characterization of tumor tissue architecture. However, these problems need an accurate and automatic segmentation. In most cases, the segmentation is concerned with the extraction of cell nuclei or cell clusters. In this paper, we address the problem of the fully automatic segmentation of grey level intensity or color images from medical microscopy. An automatic segmentation method combining fuzzy clustering and multiple active contour models is presented. Automatic and fast initialization algorithm based on fuzzy clustering and morphological tools are used to robustly identify and classify all possible seed regions in the color image. These seeds are propagated outward simultaneously to refine contours of all objects. A fast level set formulation is used to model the multiple contour evolution. Our method is illustrated through two representative problems in cytology and histology.

Keywords: Segmentation, active contour models, level set method, fuzzy clustering, medical microscopy.

1 Introduction

Image analysis offers a modern tool to pathologist, which can be applied to several problems in cancerology : quantification of DNA content, quantification of immunostaining, nuclear mitosis counting, characterization of tumor tissue architecture, etc... However, its introduction in clinical daily practice implies complete automation and standardization of procedures, together with the evaluation of the clinical interest of measured parameters. One of the bringing out steps is the segmentation process, which has to provide the interesting objects to be measured.

Segmenting medical images of soft tissues to form regions related to meaningful biological structures (such as cells, nuclei or organs) is a difficult problem, due to the wide variety of structures characteristics. Many strategies can be used; their performances depend largely on images to be processed and on a priori knowledges relative to the object features. Efforts have been made towards the unification of the contour and region based approaches, and level

set theories have been used in the formulation of the unification of these approaches [6][5][4][7]. To the best of our knowledge, the only work applying level set approach to medical microscopy images is reported by Sarti in [8]. In this work a partial differential equation based analysis is used as a methodology for computer-aided cytology. All the approaches using the level set approach for active contours can deal with gradient or regions information and can handle topological changes automatically. However, when an automatic segmentation or a quantitative segmentation such as studies in medical microscopy [9] is needed, robust and automatic specification of initial curves is required.

In this paper, we present a method for automatic segmentation of biomedical microscopic sections as a combination of fast level set approach and fuzzy clustering based on global color information. An initial automatic detection algorithm based on fuzzy clustering is used to robustly identify and classify all possible seed regions in the image. These seeds are propagated outward simultaneously to localize the final contours of all objects in the image.

The originality of the method is to classify markers obtained by morphological operators. The technique is fast, because the markers represent only 1 to 5% of the total number of pixels in the image. These markers resulting from this classification, are distributed symmetrically inside the objects of interest. They provide a good automatic initialization of the contours which allows to active contours and level set methods to operate in good conditions.

This paper is organized as follows. In section 2, the level set algorithm is reviewed. Section 3 presents a fast level set algorithm called the Group Marching Algorithm [10]; it describes how we extend the later to deal with multiple active contour evolution. In section 4, we consider the problem of automatic initialization of level set, and propose automatic fuzzy clustering combined with local morphology tools as an automatic initialization algorithm. As a conclusion, section 5 illustrates the robustness of the proposed method with two representative problems of color quantitative segmentation in medical image microscopy.

2 The Original Level Set Approach

Since its introduction, the level set approach has been successfully applied to a wide collection of problems that arise in computer vision and image processing. Let us describe the original level set idea of Osher and Sethian [11] for tracking the evolution of an initial front Γ_0 as it propagates in a direction normal to itself with a speed function F . The main idea is to match the one-parameter family of fronts $\{\Gamma_t\}_{t \geq 0}$, where Γ_t is the position of the front at time t , with a one-parameter family of moving surfaces in such a way that the zero level set of the surface always yields the moving front. To determine the front propagation, it is necessary to solve a partial differential equation for the motion of the evolving surface. Assume that the so-called *level set function* $u : \mathbb{R}^n \times \mathbb{R}^+ \rightarrow \mathbb{R}$ is such that at time $t \geq 0$ the zero level set $u(x, t)$ is the front Γ_t . The derivations

described in [12] yield the time-dependant level set equation:

$$\begin{aligned} \frac{\partial u}{\partial t} &= F |\nabla u| \\ u(x, t = 0) &= \pm d(x, \Gamma_0) \end{aligned} \quad (1)$$

where $d(x, \Gamma_0)$ is the distance from x to the curve Γ_0 . The distance is positive if x is inside Γ_0 and negative if x is outside. If the non-regularized model given by equation (1) is considered, this leads to an interesting and fast model being able to take into account the simultaneous evolution of several contours. In this model, the speed function F is either always positive or always negative. We can introduce a new variable (the arrival time function) $T(x)$ defined by $u(x, T(x)) = 0$. In other words, $T(x)$ is the time when $u(x, t) = 0$. If $\frac{dx}{dt} \neq 0$, T satisfy the stationary eikonal equation

$$\begin{aligned} |\nabla T| \cdot F &= 1 \\ T_{d(x)=0} &= 0 \end{aligned} \quad (2)$$

This equation states that the gradient of the arrival time function is inversely proportional to the velocity of the contour at any given point. The advantage of this formulation is that it can be numerically solved by fast techniques. Sethian [12] combined heap sort algorithm with variant of Dijkstra algorithm to solve equation (2), this method is known as the fast marching method (*FMM*). The heap sort algorithm is used to update T at any specified pixel in an increasing order. If N is the number of image pixels, the complexity goes as $O(N \log(N))$. Lately, an alternative sweeping strategy was suggested and used by Kim [10] to derive fast algorithm known as Group Marching Methods (*GMM*); its cost is $O(N)$. The latter is used in this paper.

3 The Group Marching Algorithm

Now let us derive a discrete version of the Eikonal equation (2). The easiest way to obtain such discretization is to replace the gradient by the first-order approximation [12] :

$$\sqrt{\left(\max(D_{ij}^{-x}T, -D_{ij}^{+x}T, 0)\right)^2 + \left(\max(D_{ij}^{-y}T, -D_{ij}^{+y}T, 0)\right)^2} = \frac{1}{F_{ij}} \quad (3)$$

where the standard finite differences are given by: $D_{ij}^{-x} = T_{ij} - T_{i-1j}$ and $D_{ij}^{+x} = T_{i+1j} - T_{ij}$, where T_{ij} is the value of T for each pixel (i, j) . Expressions of D_{ij}^{-y} and D_{ij}^{+y} in the other direction are similar.

Consider a neighborhood Γ of the front Γ_t , in the current stage of *GMM*, a group of points G is selected from Γ . Some pixels of Γ are included in G and are already labeled "completed". For the other pixels of Γ the equation (3) is solved.

The evolution of the set of active pixels is done by choosing, at the initial time, a subset G of Γ which corresponds to all the points that have to be processed. The formal definition of this principle is given by:

$$F_{\Gamma, \min} = \min \left\{ \frac{1}{F_{i,j}} : (i,j) \in \Gamma \right\} \quad (4)$$

$\delta\tau = \frac{1}{\sqrt{2}} F_{\Gamma, \min}$, $\tau_{\Gamma, \min} = \min \{ \tau_{ij} : (i,j) \in \Gamma \}$ and select G as follows:

$G = \{p = (i,j) \in \Gamma, \tau_p < \tau_{\Gamma, \min} + \delta\tau\}$. Proceeding as in [10], the Group Marching Method goes as follows:

– **Initialization**

- **Processed pixels:** All pixels under markers; assign a distance transform value of zero to them $T(i,j)=0$ and label them $idT(i,j)=2$
- **Active pixels:** Pixels at the outside boundary of the markers; their distance transform is known $T(i,j)=1/F(i,j)$
 - * label them as $idT(i,j)=1$
 - * save those point indices to the interface indicator array $\Gamma(i,j)$ set TM to be the minimum of T on those points
- **Unprocessed pixels:** Pixels away from the markers $T(i,j) = \infty$; label them as $idT(i,j) = 0$,
- Set $\delta\tau = \frac{1}{\sqrt{2}} \cdot \min_{i,j} \left(\frac{1}{F_{i,j}} \right)$

– **Marching Forward:**

- (M1) Set $TM = TM + \delta\tau$
- (M2) For each (i,j) in $\Gamma(i,j)$, in the reverse order, if $T(i,j) \leq TM$, update the solution at neighboring points (l,m) where $idT(l,m) \leq 1$;
- (M3) For each (i,j) in Γ , in the forward order, if $T(i,j) \leq TM$,
 - (a) update the solution at neighboring points (l,m) where $idT(l,m) \leq 1$;
 - (b) if $idT(l,m) = 0$ at a neighboring point (l,m) , set $idT(i,j) = 1$ and save (l,m) into Γ ;
 - (c) remove the index (i,j) from Γ ; set $idT(i,j) = 0$;
- (M4) if $\Gamma \neq \emptyset$, go to (M1);

The *GMM* is in fact an iterative update procedure, converging in two iterations. Rouy and Tourin [13] have chosen all the grid points as one group and carried out iterations up to convergence. *GMM* can be viewed as an intermediate algorithm between *FSM* ($\delta\tau \rightarrow 0$) and the purely iterative algorithm of Rouy and Tourin ($\delta\tau \rightarrow \infty$).

This algorithm can be easily extended to deal with evolution and labeling of multiple curves. Let us assume we have C seed regions G_i , to deal with the evolution. Any of the independent contours possibly propagates with different speeds, we label all seeds with C labels according to the results of fuzzy classification, and then we propagate these labels while computing G_i , by solving the equation:

$$\begin{aligned} |\nabla T| &= \frac{1}{F(T)} \\ T_{/G_i} &= 0 \end{aligned} \quad (5)$$

For each pixel, two properties are calculated: the arrival time and the region label that reached that pixel first. All curves are thus allowed to evolve simultaneously and no limiting evolution time is necessary.

The implementation of this algorithm is at the root of a wide range of image processing applications to do various image analysis tasks that one typically encounters in the study of medical microscopic images.

4 Automatic Initialization

An essential step of the whole framework consists in estimating features associated with different labels and in determining the initial seed regions.

Our method consists of two steps:

1) A detection of a set of germs, located in a symmetrical way inside all the interesting objects in the image. Mathematical morphology operators are mainly used to extract these germs.

2) All or a part of these germs is gathered in classes of germs according to their color, by using a fuzzy classification. These seeds are classified according to their color and characterized by region information : the mean and variance of each class i . The latter can be supervised or unsupervised using the available a priori information on the images considered.

4.1 Feature Extraction

A 2D color image I is a function where each pixel (x_1, x_2) , by three grey level values in the RGB color space. The gradient amplitude is obtained by the contour information I defined by : $|\nabla I| = \sqrt{\lambda_+ + \lambda_-}$. λ_+, λ_- are the largest, (resp. smallest) eigenvalues of the quadratic form associated to f . The local minima or the h-minima of this contrast image give a set of seed regions placed nearly symmetrically with respect to the object boundaries. The h-minima of the image can be formulated by :

$$h_{min}(U_0) = \left\{ P / \left(U_0(P) - \gamma^{(rec)}(U_0, U_0 + h)(P) \right) < 0 \right\}$$

where $\gamma^{(rec)}(U_0, U_0 + h)$ denotes the morphological reconstruction by erosion of the $U_0 + h$ image with U_0 .

4.2 Fuzzy Classification of Seed Regions

For classification, a modified fuzzy c -mean algorithm [14] is applied to classify all seed pixels in a given image into C classes by minimizing the following objective function :

$$J = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d^2(x_j, c_i) - \alpha \sum_{i=1}^C p_i \log(p_i)$$

where u_{ij} is the membership value at pixel j in the class i such that $\sum_{i=1}^C u_{ij} = 1$

$\forall j \in [0, N]$. $p_i = \frac{1}{N} \sum_{j=1}^N u_{ij}$ is interpreted as "the probability" of all the pixels

j to belong to the class i . c_i is the centroid of class i , N is the total number of pixels in image, $d^2(x_j, c_i)$ is the standard Euclidian distance and the fuzziness index m is a weighting coefficient on each fuzzy membership, $m=2$ is an usual value.

In the algorithm, the number of classes C can be known or automatically determined by choosing initially a high value of C and eliminating the class i with the smallest probability p_i .

5 Localization

In order to take into account the information about regions and contours obtained in the classification step, we consider an adaptive speed function F . This function is defined in each point by the following equation:

$$F^i(I)(x, y) = 1 - e^{-\frac{1}{2} \left(\sum_{k=1}^3 |I_k(x, y) - \mu_i^k|^2 + |\nabla_C I(x, y)|^2 \right)} \quad (6)$$

where I_k is the k^{th} channel of color image μ_i^k is the mean of classes i on channel k and ∇I_C is the color gradient amplitude.

6 Biomedical Applications

6.1 Color Cytology

For this first biomedical application, images from serous cytology are considered. The images are from a database of digitized cells images, collected from pleural and peritoneal effusions with different pathologies. In this class of images, both cytoplasm and nuclei have to be segmented. Once segmented, the cells can be classified among cellular types (ranging from normal to abnormal). Figure 1(b) gives the set of minima extracted from the amplitude of the gradient. From these ones, markers are obtained for each class : nuclei, cytoplasm and background (figure 1 (c), (d), (e), respectively). Figure 1(f) presents the final result.

6.2 Color Histology

In the second example, acquisitions were performed on sections of immunohistochemically stained tissues (figure 2 (a)). The markers involve a brown coloration for positive nuclear locations and a blue coloration for unmarked nuclei (negative locations). Images of this class are more complex than in the previous case. One could be interested by many categories of objects : the clusters of tumoral cells (called lobules in carcinoma), the marked and unmarked tumor nuclei presenting specific characteristics inside the clusters. The goal of this analysis is to evaluate the immunostaining ratio defined as the positive nuclear area to the whole nuclear area within the lobules limits.

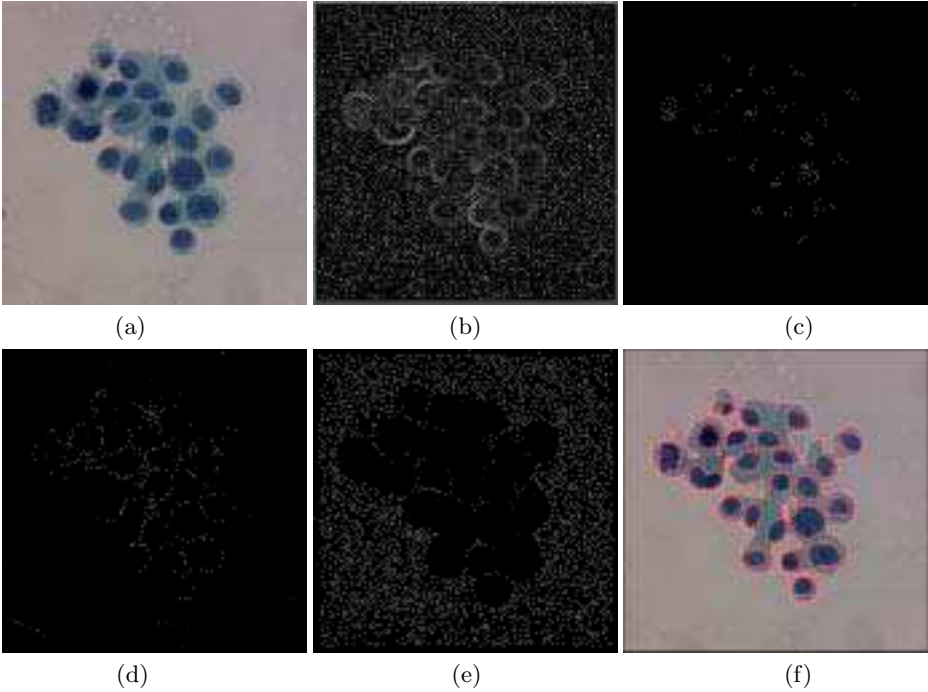


Fig. 1. (a) A serous cytology color image ($\times 20$). (b) Gradient amplitude minima. (c) Nuclei markers. (d) Cytoplasm markers. (e) Background markers. (f) Final segmentation.

Segmentation of Lobules. The tumoral lobules are made of clusters which can be characterized by a small inter-cellular distance and whose nuclei have a greater size than the other cell categories (lymphocytes or stroma cells for example). The processing goes according to the following: a) Image simplification is used to remove lymphocytes and to make the clustering of other cells easier. This step uses morphological closing performed on each color plane (i ranging from 1 to 3): $\gamma_B(I_i) = \varepsilon_B \circ \delta_B(I_i)$, where δ_B and ε_B are the dilation and erosion of the i^{th} plane of the color image I by a flat structuring element B .

b) The fuzzy clustering algorithm provides reliable markers for the two different classes of pixels to be used in the localization. The result is a binary mask I_b displaying the lobules (figure 2 (b)).

Detection of the Nuclei Inside the Lobules. This process is twofold:

- a) Extraction of the nuclei by residual analysis on the luminance component (I_L) provides a monochromatic image I_R whose positive and negative values form a binary image of nuclei I_a .
- b) An inverted image of distance I_d is computed from I_a and the watershed transformation is applied to split nuclei initially merged. The distance and

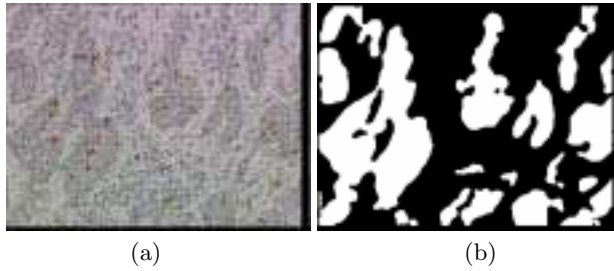


Fig. 2. (a) Original histological breast cancer image ($\times 33$). (b) Segmentation result : binary mask of lobules.

the watershed transformations are computed by setting $F=1$ and $F = |\nabla I_d|$ respectively, in equation (5).

The process is limited to the lobule area, by the means of a logical intersection between the image of lobules and that of all nuclei (figure 3 b).

Immunostaining Characterization. A simple binary thresholding, which represents the degree of membership to the class of marked nuclei, allows to detect the positive pixels (brown pixels). To extract the marked nuclei, the segmented objects are reconstructed from the positive pixels in order to assess the total area of positive profiles.

7 Conclusion

A fast statistical level set method for color image segmentation was presented. This method is based on the integration of two attractive techniques : the fuzzy clustering and the level set active contours. They can both take into account local information, such as the gradient modulus, and statistical information, such as the mean color levels in an object. According to their properties, the initialization and localization can be easily extended from 2D images to 3D images, provided by a confocal microscope.

Acknowledgments

Mrs. Herlin from Baclesse Center (Caen) and Mr Helie (Hospital of Cherbourg) have kindly provide applications and images.

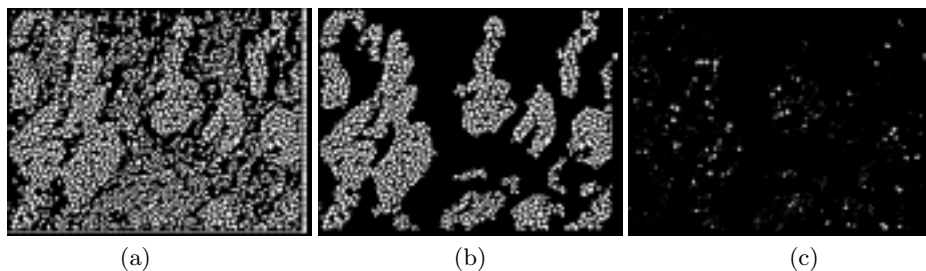


Fig. 3. Segmentation of nuclei inside the lobules. (a) Residual analysis of the luminance. (b) Intersection image. (c) Marked nuclei inside the lobules.

References

1. J. Morel, S. Soleimini, *Variational. Methods in Image Segmentation*. Progress in Nonlinear Differential Equations Equations and their application, Birkhäuser 1995.
2. V. Caselles, F. Catte, T. Coll , F. Dibos, A geometric model of active contours", *Numerische Mathematik*, 66 (1993) 1-3.
3. S. Kichenassamy, A. Kumar, P.J. Olver, A. Tannenbaum, A. Yezzi, Gradient flows and geometric active contours, *Proceeding of Fifth International Conference on Computer Vision, Cambridge*, (1995) 810-815.
4. C. Samson, L. Blanc-Feraud, G. Aubert, J. Zerubia, A level set model for image classification", *rapport de recherche, INRIA*, France, RR-3662 (1999).
5. T. Chan, B. Y. Sandberg, L. Vese, Active Contours without Edges for Vector-Valued Images, *J. of Visual Communication and Image Representation*, 11 (2000) 130-141.
6. T. Chan, B. L. Vese, Active Contours and Segmentations Models Using Geometric PDE's for Medical Imaging, *UCLA CAM Report* (2000).
7. R. Malladi, B. J.A Sethian, , A Real-time Algorithm for Meduical Shape Recovery, *Proceeding of International Conference on Computer Vision, Mubai, India* (1998) 304-310.
8. A. Sarti, C. Ortiz, S. Lockett, R. Malladi, A Geometric Model for 3D Confocal Microscope Image Analysis , *Preprint, Lurence Berkeley National Laboratory, LBLNL-41740*, 1999.
9. A. Elmoataz, S. Schüpp, R. Clouard, P. Herlin, D. Bloyet, Using active contours and mathematical morphology tools for quantification of immunohistochemical images, *Signal Processing*, 71 (1998) 215-226.
10. S. Kim, O(N) Level set , *Math-Report, University of Kentuckey*, (2000).
11. S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, *Journal of computational physics*, 79 (1988) 12-49,.
12. J. Sethian, *Level Set Methods: Evolving interfaces in geometry, fluid mechanics, computer vision, and material science*. Cambridge University Press, (1996).
13. E. Rouy, A. Tourin, " A viscosity solutions approach to shape-from-shading, *SIAM J. Numer. Anal.*, (1992) 29 867-884.
14. A. Lorette, X. Descombes, J. Zerubia, Urban areas extraction based on texture analysis through a markovian modelling, *International Journal of Computer Vision* 36 (2000) 219-234.