PDE Based 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. PDE Based image segmentation for biomedical applications. International Conference on Image and Signal Processing, 2001, Agadir, Morocco. pp.204-215. hal-01016088

HAL Id: hal-01016088
https://hal.archives-ouvertes.fr/hal-01016088
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.
PDE Based Image Segmentation for biomedical Applications

S. Schüpp* A. Elmoataz J. Fadili D. Bloyet

Groupe de Recherche en Informatique, Image et Instrumentation de Caen UMR 6072,
GREYC-ISMRA 6, Bd Maréchal Juin, F-14050 Caen, France
*e-mail: 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 oncology: quantification of DNA content, quantification of immunostaining, counting of nuclear mitosis, characterisation of tumour tissue architecture. However, these problems need a quantitative and automatic segmentation. In most cases, the segmentation concerns the extraction of cell nuclei or cell clusters. In this paper we address the problem of automatically segmenting intensity or color images from medical microscopy. An automatic segmentation method combining fuzzy clustering multiple active contour models is presented. Automatic and fast initialization algorithm based on fuzzy clustering and morphological tools is used to robustly identify and classify all possible seed regions in the color image. These seeds are propagated outward simultaneously to localize the final contours of all objects. A fast level set formulation is used to model the multiple contour evolution. We illustrate our method by presenting two representatives problems in cytology and histology.

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

1 INTRODUCTION

Image analysis offers to the pathologist a modern tool, which can be applied to several problems in oncology: quantification of DNA content, quantification of immunostaining, counting of nuclear mitosis, characterisation of tumour tissue architecture, etc. However, its introduction in clinical practice implies a complete automation and standardisation of procedures and goes 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. In most cases, it concerns the extraction of cell nuclei or cell clusters with, for example, the following objectives:

– The analysis of isolated nuclei for DNA quantification.
– The in situ analysis of nuclei inside histological sections for immunostaining quantification.
The characterisation of tissue architecture or the determination of immunostaining topography [?].

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 large variations of the characteristics of the structures. There are several strategies for segmenting images; their performances depend largely on the type of images to be processed and on a priori knowledge relative to the object features. These methods can be roughly classified into several categories: contour-based methods [?][?], region-based methods, variational methods such as the global optimization approach minimizing an energy function or some bayesian criteria [?][?] model-based methods as active contours models [?][?] [?]. Efforts have also been made towards the unification of the contours and regions based approach, and level set theories have been used in the formulation of the unification of theses approach [?][?][?][?][?][?][?].

To the best of our knowledge, the only work applying level set approach to medical microscopy images has reported by Sarti in [?]. In this work a partial differential equation based analysis is used as methodology for computer-aided cytology. For extraction and classifying the shapes of nuclei from noisy confocal microscopy he build a chain that include edge-preserving image smoothing, a segmentation method, a geometry-driven scheme to regularize the shapes, and an interactive method to split clusters of cells and reclassify them.

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, for problems needing an automatic segmentation or a quantitative segmentation such as studies in medical microscopy [?] [?], robust and automatic specification of initial curves is required. A good edge localisation is only obtained if the initial curves are placed nearly symmetrically with respect to the object boundaries.

In this paper, we present a method for automatic segmentation 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 a given 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. With the result of this classification, these markers are distributed in a symmetrical way inside the interesting objects. They constitute thus a good automatic initialisation and suppress the principal disadvantage from active contours and level set methods.

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 [16]; 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 considers two representative problems of image color segmentation in quantitative medical image microscopy.

2 The original level set approach

After its introduction, the level set approach has been successfully applied to a wide collection of problems that arise in computer vision and image processing. We now describe the original level set idea of Osher and Sethian [5] for tracking the evolution of an initial front \( \Gamma_0 \) as it propagates in a direction normal to itself with speed function \( F \). The main idea is to match the one-parameter family of fronts \( \{ \Gamma_t \}_{t \geq 0} \), where \( \Gamma_t \) 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, we then need to find and solve a partial differential equation for the motion of the evolving surface. To be more precise, let \( \Gamma_0 \) be an initial front in \( \mathbb{R}^n \), \( n \geq 2 \) and assume that the so-called level set function \( u : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R} \) is such that at time \( t \geq 0 \) the zero level set \( u(x,t) = 0 \) is the front \( \Gamma_t \). We further assume that 
\[
\begin{align*}
    u(x,0) &= \pm d(x, \Gamma_0) \\
    d(x, \Gamma_0) &= \text{distance from } x \text{ to the curve } \Gamma_0.
\end{align*}
\]
We use sign + if \( x \) is inside \( \Gamma_0 \) and - if \( x \) is outside. Let each level set of \( u \) flow along its gradient field with speed \( F \). This speed function should match the desired speed function for the zero level set \( u \). Now consider the motion of, e.g., the level set

\[
\{ x \in \mathbb{R}^n : u(x,t) = 0 \}
\]

The derivations described in [5] yield the time-dependant level set equation:

\[
\begin{align*}
    \frac{\partial u}{\partial t} &= F|\nabla u| \\
    u(x, t = 0) &= \pm d(x, \Gamma_0)
\end{align*}
\]

An example of these models is the geodesic active contours, which have been introduced in by Kichenassamy [7] and Caselles [6].

If the non-regularised model given by the equation (2) is considered, this lead to a interesting and fast model being able to take 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 \( \frac{dT}{dx} \neq 0 \), \( T \) satisfy the stationary eikonal equation

\[
\begin{align*}
    |\nabla T| \cdot F &= 1 \\
    T_{d(x) = 0} &= 0
\end{align*}
\]

This equation states simply 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 we can solve it numerically by fast techniques. Sethian [19] combined heap sort algorithm with variant of Dijkstra algorithm to
solve the equation \((\theta)\), and has become known as the fast marching method \((FMM)\). The fast marching method uses heap sort algorithm to update \(T\) at specified pixel in an increasing order. Because of heap sort, if \(N\) is the number of image pixels, the complexity is \(O(N\log(N))\). Lately an alternative sweeping strategy was suggested and used by Kim [16] to derive fast algorithm known as group marching methods \((GMM)\). The cost is \(O(N)\). This is precisely what we will do in this paper.

3 The group marching algorithm

We begin by finding a discrete version of the Eikonal equation \((\theta)\). The easiest way to obtain such discretisation is to replace the gradient by the first-order approximation [20]:

\[
\sqrt{\left( \max (D_{ij}^{-x}T, -D_{ij}^{+x}T, 0) \right)^2 + \left( \max (D_{ij}^{-y}T, -D_{ij}^{+y}T, 0) \right)^2} = \frac{1}{T_{ij}}
\]

where the standard finite differences are given by: \(D_{ij}^{-x} = T_{ij} - T_{i-1,j}\) and \(D_{ij}^{+x} = T_{i+1,j} - T_{ij}\). \(T_{ij}\) is the value of \(T\) for each pixel \((i, j)\). The backward and forward operators \(D_{ij}^{-y}\) and \(D_{ij}^{+y}\) in other coordinate directions are similar.

Consider a neighborhood \(\Gamma\) of the front \(\Gamma_i\), in the current stage of \(GMM\), a group of points \(G\) are selected from \(\Gamma\). The solution is recomputed at neighboring points of \(G\) that are not completed. The neighboring points are registered as members of \(\Gamma\) if they are not, and finally tag "completed" for the points in \(G\).

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

\[
F_{\Gamma,\min} = \min \left\{ \frac{1}{F_{i,j}} : (i, j) \in \Gamma \right\}
\]

\[
\delta \tau = \frac{1}{\sqrt{2}} F_{\Gamma,\min}
\]

\[
\tau_{\Gamma,\min} = \min \{ \tau_{ij} : (i, j) \in \Gamma \}
\]

and select \(G\) as follows:

\[
G = \{ p = (i, j) \in \Gamma, \tau_p < \tau_{\Gamma,\min} + \delta \tau \}
\]

Proceeding as in [18], 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 \(\text{id}\) \(T(i, j) = 2\)
  - **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 \(\text{id}\) \(T(i, j) = 1\)
    - * save those point indices to the interface indicator array \(\Gamma(i, j)\) set \(T_{M}\) to be the minimum of \(T\) on those points
- **Unprocessed pixels**: Pixels away from the markers $T(i, j) = ∞$; label them as $idT(i, j) = 0$.

- Set $δτ = \frac{1}{\sqrt{a}} \min \left( \frac{1}{T(i, j)} \right)$

Marching Forward:

(M1) $Set TM = TM + δτ$

(M2) For each $(i, j)$ in $Γ(i, j)$, in the reverse order, if $T(i, j) ≤ TM$, update the solution at neighboring points $(l, m)$ where $idT(l, m) ≤ 1$;

(M3) For each $(i, j)$ in $Γ$, in the forward order, if $T(i, j) > TM$,

(a) update the solution at neighboring points $(l, m)$ where $idT(l, m) ≤ 1$;

(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 $(i, j)$ out $Γ$; set $idT(i, j) = 0$;

(M4) if $Γ ≠ ∅$, go to (M1);

The GMM is in fact an iterative update procedure, converging in two iterations. One may want to select $G$ with a larger $δτ$. In this case, the number of iterations becomes larger. Rouy and Tourin [21] 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 FMM ($δτ → 0$) and the purely iterative algorithm of Rouy and Tourin ($δτ → ∞$).

This algorithm can be easily extended to deal with evolution and labeling of multiple curves. Let us assume we have $C$ seeds regions $G_i$. To deal with the evolution any of independent contours propagating with possibly 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:

\[
\frac{|∇Γ|}{T_{/G_i}} = 0
\]  

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

The implementation of this algorithm is at the root of a certain number of interesting image processing 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 to different labels and in determining the initial seed regions.

Our method consists of two steps:

1) A detection of a set of germs, placed in a symmetrical way inside all the interesting objects in the image. Mathematical morphology operators mainly extract these germs.
2) All or a part of these germs is gathered in classes of germs according to their color, by using a method of fuzzy classification. This latter can be supervised or unsupervised using the available a priori information of considered images.

4.1 Feature extraction

A color image is a function where each pixel, whose coordinates are \((x_1, x_2)\), corresponds to three values: the grey level of each of the three planes of the color space RGB.

\[
I : \mathbb{R}^2 \rightarrow \mathbb{R}^3
\]

\[
(x_1, x_2) \rightarrow (I_1(x_1, x_2), I_2(x_1, x_2), I_3(x_1, x_2))
\]

The gradient amplitude is obtained by the contour information \(f\) defined as follows:

\[
|\nabla I| = \sqrt{\lambda_+ + \lambda_-}
\]

\(\lambda_+\), \(\lambda_-\) 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 seeds regions placed nearly symmetrically with respect to the object boundaries.

The h-minima are the minima of a given height \(h\) for an image \(U_0\). They can be formulated by

\[
h_{\text{min}}(U_0) = \left\{ P \mid \left( U_0(P) - \gamma^{(\text{rec})}(U_0, U_0 + h)(P) \right) < 0 \right\}
\]

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

These seeds are classified according to their color and characterised by region information that is given by \(\mu_i, \rho_i\): the mean and variance of each class \(i\).

4.2 Fuzzy classification of seeds regions

For classification, a modified fuzzy c-mean algorithm [2][3] is applied to classify all seed pixels in a given image into \(C\) classes by minimizing the following objective function:

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

\(u_{ij}\) is the membership value at pixel \(j\) in the class \(i\) such that \(\sum_{i=1}^{C} u_{ij} = 1 \quad \forall \; j \in [0, N]\). \(p_i = \frac{1}{N} \sum_{j=1}^{N} u_{ij}\) 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 Euclidean distance and $m$ is a weighting coefficient on each fuzzy membership, we take $m=2$.

In the algorithm the number of classes $C$ can be known or automatically determined by choosing a high value of $C$ and eliminating the class $i$ with the smallest probability $p_i$. This is the main difference with the classical algorithm for which the number of cluster is fixed.

5 Localization

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

$$F^i(I)(x,y) = 1 - e^{\frac{1}{2} \left( \sum_{k=1}^{n} \left| I_k(x,y) - \mu^i_k \right|^2 + \left| \nabla_c I(x,y) \right|^2 \right)}$$

$I_k$ is the $k^{th}$ channel of color image $\mu^i_k$ is the mean of classes $i$ on channel $k$ and $\nabla_c I$ is the color gradient amplitude.

6 Biomedical applications

To illustrate the robustness of the proposed method, two segmentation problems in the field of quantitative microscopy analysis are presented and detailed.

6.1 Color Cytology

For this first biomedical application, images from serous cytology are considered. The cell preparations are obtained by centrifugation and slides are stained by the Papanicolaou international standard of coloration. Images were taken with a 20x objective using a 3-CCD color camera. The images are from a database of digitised cells images, collected from pleural and peritoneal effusions with different pathologies. In this class of images, both cytoplasm and nuclei have to be segmented, the cytoplasm to obtain the context information (to characterise isolated or clustered cells) and the nucleus for grading malignancy. Once segmented, the cell can be classified among cellular types (ranging from normal to abnormal). The segmentation has to be incorporated in a system we develop called ARCTIC which aim is to assist the screening of serous cytology slides by an automatic recognition of the cells [17]. The 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 (respectively the figure 1 (c), (d), (e)). The figure (f) presents the final segmentation obtained.
Fig. 1. (a) A serous cytology color image. (b) Gradient amplitude minima. (c) Nuclei markers. (d) Cytoplasm markers. (e) Background markers. (f) Final segmentation.
6.2 Color Histology

In the second image class, acquisitions were performed on 5µm sections of immunohistochemically stained tissues. The analysis aimed at quantifying some proteins (estrogen or progesterone receptors, proliferation markers) by revealing a staining bound to the associated antibody. A double diaminobenzidine-peroxidase and hematoxylin staining is processed onto sections after they have been placed onto slides. Diaminobenzidine and peroxidase are both combined to reveal immunohistochemical markers and involve a brown coloration for positive nuclear locations. Hematoxylin is a DNA specific staining revealing a blue coloration for unmarked nuclei (negative locations).

Images of this class are more complex than in the previous case. One has to distinguish here many categories of objects: clusters of tumoral cells (called lobules in carcinoma) and nuclear profiles presenting specific characteristics inside the clusters. So a greater magnification (×33) is used to complete this problem, it allows observing tumoral lobules (figure 2a). The goal of the analysis is to evaluate the immunostaining ratio defined as the positive nuclear area divided by the whole nuclear area to limit this measure inside the lobules. The automation of this procedure requires a precise segmentation at the tissue level (identification of lobules) as well as at the cellular level (identification of positive and negative profiles).

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 problem is that fibrovascular and lymphocytes nuclei are often gathered so that the distance criterion is not sufficient. Other features such as size or intensity are then to be considered. To extract lobules in these images, the color image is regularized and a morphological closing is applied. Then, our fuzzy clustering algorithm is applied to extract the corresponding markers splitted among two classes. The plan retained is therefore the following:

a) Image simplification is used to remove lymphocytes and to make 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 \( \delta_B \) and \( \varepsilon_B \) are the dilation and erosion of the \( i^{th} \) plane of the color image \( I \) by a plan 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 2b).

Individualization of nuclei inside the lobules The retained plan is twofold:

a) Extraction of 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 \( I_a \). This step corresponds to the following operation:
Original histological breast cancer image (x33). (b) Segmentation result: binary mask of lobules.

\[ I_R = I_L * G(x, \sigma_1) - I_L * G(x, \sigma_2) \]
where \( G(x,\sigma) \) is a gaussian function of standard deviation \( \sigma \).

b) Splitting of touching nuclei. An inverted image of distance \( I_d \) is computed from \( I_a \) and the watershed transformation is applied onto. The distance and the watershed transformations are computing by setting in equation (??) respectively \( F=1 \) and \( F = |\nabla I_d| \)

In order to limit the process inside the sole lobules previously extracted, a logical intersection between the image of lobules and the the image of all nuclei is computed giving a new image (figure 3b).

**Immunostaining Characterization** In order to characterise the positive nuclei, we classify only nuclei pixels inside the lobules (figure 3c) according to their color in two classes: the marked nuclei (appearing in brown) are easily discriminated from the counterstained nuclei (appearing in blue). A simple binary thresholding on image, which represent 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 in this paper. This method is based on the integration of two attractive techniques: fuzzy clustering and level set active contours. It can both take into account some local information, such as the gradient modulus, and some statistical information, such as the mean colour levels in an object. According to their properties, the initialisation and localization process by means of morphological tools, classification and the level set approach of active contours can be easily extended from
2D images to 3D images. This allows segmenting, for example, images provided by a confocal microscope.

![Segmentation images](image)

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

**References**