ACTIVE CONTOUR FOR OVERLAP RESOLUTION USING WATERSHED BASED INITIALIZATION (ACOReW): APPLICATIONS TO HISTOPATHOLOGY

Sahirzeeshan Ali

Rutgers University
Electrical And Computer Engineering
94 Brett Road
Piscataway, NJ 08854-8058

Anant Madabhushi *

Rutgers University
Biomedical Engineering
599 Taylor Road,
Piscataway, NJ 08854-8058

ABSTRACT

In recent years, shape based active contours have emerged as a natural solution to overlap resolution. However, most of these shape-based methods are limited to finding and resolving one object overlap per scene and require user intervention for model initialization. In this paper, we present a novel synergistic segmentation scheme called Active Contour for Overlap Resolution using Watershed (ACOReW). ACOReW combines shape priors with boundary and region-based active contours in a level set formulation with a watershed scheme for model initialization for identifying and resolving multiple object overlaps in an image scene. The energy functional for the variational active contour model is composed of three complimentary terms (a) a shape model which constrains the active contour to a pre-defined shape, (b) boundary based term which directs the active contour model to the image gradient, and (c) a third term driving the shape prior and the active contour towards a homogeneous intensity region. In this paper we show an application of ACOReW in the context of segmenting nuclear and glandular structures on prostate and breast cancer histopathology. The results of qualitative and quantitative evaluation on 100 prostate and 14 breast cancer histology images reveals that ACOReW outperforms two state of the art segmentation schemes (Geodesic Active Contour (GAC) and Rousson's shape based model) and resolves up to 92% of overlapping/occluded lymphocytes and nuclei on prostate and breast cancer histology images.

1. Introduction

A number of deformable segmentation schemes (Active Contours (AC)) have been developed to date; they can be roughly divided into boundary-based (first generation) and region-based (second generation) schemes [1, 2, 3]. These models are typically unable to handle object occlusion or scene clutter. Therefore, the integration of prior shape knowledge of the objects represents a natural way to overcome occlusion. Third generation AC models involve combining a shape prior with geometric/geodesic active contours that simultaneously achieves registration and segmentation [4]. In the methods proposed in [4, 5, 6], authors have incorporated various shape priors into active contour formulation to resolve overlaps. Rousson et al proposed a novel approach for introducing shape priors into level set representations targeting 2D closed structures. A limitation of these models is that they introduce shape priors into a level set framework so that usually only one pair of overlapping objects can be accurately resolved per scene. Further, all first 3 generation AC methods are sensitive to model initialization and typically require varying degrees of user intervention.

In the context of histopathology and microscopic imagery, being able to resolve overlapping objects into independent shapes has translational relevance in the context of a number of different diagnostic and prognostic applications [7]. In [7], Basavanhally quantified the extent of lymphocytic infiltration (LI) in HER2+ breast cancers using a nuclear detection and graph feature based approach. LI has been identified as an important prognostic marker of outcome in Her2+ breast cancer. Automated segmentation and quantification of nuclear and glandular structures is critical for classification and grading of cancer [7, 8]; cancer grade like LI being important prognostic markers in the context of several diseases. In [8], Fatakdwala combined an expectation maximization scheme with an explicit concavity based overlap resolution scheme to separate overlapping nuclei.

The main contribution of this work is a new variational active contour scheme, ACOReW, that requires minimal user intervention and segments all overlapping and non-overlapping objects simultaneously. ACOReW, extends the Geodesic Active Contour (GAC) by adding a shape prior and a region-based energy term based on the Mumford-Shah functional [3]

*This work was made possible via grants from the Wallace H. Coulter Foundation, National Cancer Institute (Grant Nos. R01CA136535-01, R01CA14077201, R21CA12718601, and R03CA143991-01) and The Cancer Institute of New Jersey. We would like to thank Dr. John E. Tomasezwski and Dr. Michael D. Feldman from the Pathology Department at the Hospital of the University of Pennsylvania for providing the prostate histology imagery and ground truth annotation.

with wastershed based initialization. Additionally, ACOReW is able to handle overlaps between multiple intersecting and adjacent objects.

2. ACTIVE CONTOUR FOR OVERLAP RESOLUTION USING WATERSHED (ACOReW)

2.1. Shape Term - F_{shape}

Each shape in the training samples is embedded as the zero level set of a higher dimensional surface. The Signed Distance Function (SDF), used to encode the distance to the F_{shape} is a functional that depends on the AC providing the boundaries. This functional evaluates the shape difference between the level set ϕ and the zero level set of the shape function ψ at each iteration. It should be noted that PCA applied on aligned SDFs of a training set produces shape functions very similiar to SDFs [9]. The level set formulation of the shape functional is expressed as:

$$F_{shape} = \int_{\Omega} (\phi(\mathbf{x}) - \psi(\mathbf{x}))^2 |\nabla\phi| \delta(\phi) d\mathbf{x} \tag{1}$$

where $\{\phi\}$ is a level set function, ψ is the shape prior, $\delta(.)$ is the Dirac function, and $\delta(\phi)$ is the contour measure on $\{\phi = 0\}$. Since ϕ undergoes a similarity transformation to adjust the pose and scale, we can also write F_{shape} in terms of rotation, translation and scaling factor using standard linear transformations (not shown).

The model described in Equation 1 introduces a shape prior in such a way that only objects of interest similar to the shape prior can be recovered, and all unfamiliar image structures are suppressed. However, this formulation only solves for a single level set consistent with the shape prior. If there are several objects of the same shape in the scene, this model finds at most *one*, and may not find all shapes of interest in the image. Therefore we incorporate a method to deal with overlap between multiple objects of similar shape (Section 2.4).

2.2. Region Homogeneity Term

We define a functional to drive the shape model towards a homogeneous intensity region corresponding to the shape of interest. If the objects of interest are supposed to have a smooth intensity surface, then the Mumford-Shah (MS) model is the most appropriate model to segment these objects [3]. Since the MS method applied on the AC will extract globally homogeneous regions and our objective is to capture an object corresponding to a specific shape, the best solution is to apply the MS-based force on the shape prior [6] to drive it towards a homogeneous intensity region. The functional F_{region} can be written with the shape function ψ and statistics of partitioned foreground and background regions, u_{in}, u_{out}:

$$F_{region}(\psi, u_{in}, u_{out}) = \int_{\Omega} \Theta_{in} H_{\psi} d\mathbf{x} + \int_{\Omega} \Theta_{out} H_{-\psi} d\mathbf{x}, \tag{2}$$

where ψ is shape function, $H(.)$ is the Heaviside function, Θ_r = $|I - u_r|^2 + \mu|\nabla u_r|^2$ and $r \in \{in, out\}$.

2.3. Combining Shape, Boundary and Region-based Functionals

ACOReW integrates a geometric shape prior with local and global intensity information within a variational framework [9]: $F = F_1 + F_{region}(\psi, u_{in}, u_{out})$, where $F_1 = \beta_1 F_{boundary}(C) + \beta_2 F_{shape}(\phi, \psi)$ with $F_{boundary} = \int_0^1 g(|\nabla I(C(q))|)|C'(q)|dq$, where C is the active contour, ψ is the shape function of the object of interest given by the PCA, g is an edge detecting function and $\beta_1, \beta_2 > 0$, are constants that balance the contributions of the boundary, shape and region terms. This is an extension of the work of Chen et al in [6]. ACOReW extends [6] by incorporating a shape term.

$$F = \underbrace{\int_{\Omega} (\phi(\mathbf{x}) - \psi(\mathbf{x}))^2 |\nabla\phi| \delta(\phi) d\mathbf{x} +}_{Shape+boundary force} \underbrace{\beta_r \int_{\Omega} \left(\Theta_{in} H_{\psi} + \int_{\Omega} \Theta_{out} H_{-\psi}\right) d\mathbf{x}}_{Region force} \tag{3}$$

2.4. Segment multiple objects under mutual occlusion

The level set formulation in Equation (3) is limited in that it allows for segmentation of only a single object at a time. In this work, we incorporate the method presented in [10] into Equation 3. Consider a given image consisting of multiple objects $\{O_1, O_2, \cdots, O_n\}$ of the same shape. For the problems considered in this work (nuclei segmentation on histopathology images), all nuclei are assumed to be roughly elliptical in shape. Instead of partitioning the image domain into mutually exclusive regions, we allow each pixel to be associated with multiple objects or the background. Specifically, we try to find a set of characteristic functions χ_i such that: $\chi_i(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in O_i \\ 0 & \text{otherwise.} \end{cases}$ We associate one level set per object in such a way that any $O_a, O_b, a, b \in \{1, 2, \cdots, n\}$ are allowed to overlap with each other within the image. These level set components may both be positive within the area of overlap, and enforce the prior on the shapes of objects extracted from the image. We consider a case of segmenting two objects within an input image, which is then generalized to N independent familiar objects.

Given an image with two similarly shaped objects $O_a, O_b, a, b \in \{1, \cdots, n\}$, and for simplicity, assume that

they are consistent with the shape prior ψ. Then simultaneous segmentation of O_a, O_b with respect to ψ is solved by minimizing the following modified version of Equation(3):

$$F(\Phi, \Psi, u_{in}, u_{out}) = \sum_{i=1}^{2} \int_{\Omega} (\phi_i(\mathbf{x}) - \psi(\mathbf{x}))^2 |\nabla \phi_i| \delta(\phi_i) d\mathbf{x} + \beta_r \int_{\Omega} \Theta_{in} H_{\chi_1 \vee \chi_2} d\mathbf{x} + \int_{\Omega} \Theta_{out} - H_{\chi_1 \vee \chi_2} d\mathbf{x} + \omega \int_{\Omega} H_{\chi_1 \wedge \chi_2} d\mathbf{x} + \sum_{i=1}^{2} \int_{\Omega} (\phi_i - \psi_i)^2 d\mathbf{x} \quad (4)$$

with $H_{\chi_1 \vee \chi_2} = H_{\psi_1} + H_{\psi_2} - H_{\psi_1} H_{\psi_2}, H_{\chi_1 \wedge \chi_2} = H_{\psi_1} H_{\psi_2}$ where $\Phi = (\phi_1, \phi_2)$ and $\Psi = (\psi_1, \psi_2)$. The fourth term penalizes the overlapping area between the two segmenting regions, and it prevents the two evolving level set functions from becoming identical. Minimizing Equation 4 iteratively with respective to dynamic variables, yields the associated Euler-Lagrange equations. The above model can be adapted for N objects (proof not shown).

2.5. Watershed Based Initialization

To address the issue of initialization, we use the popular watershed transformation to get the initial delineation and initialize the level set accordingly. The watershed transform can be classified as a region-based segmentation approach and has been widely applied to segment touching objects. The intuitive idea is that of a landscape or topographic relief which is flooded by water, watersheds being the divide lines of the domains of attraction of rain falling over the region [11].

Since our method requires N level sets for N objects, in regions where there are overlapping or adjacent objects, we empirically analyze the size of the initial delineated areas provided by watershed. For those regions smaller than a predetermined threshold, we only place one level set, in regions larger than the threshold, we place multiple level sets.

3. Experimental Results and Discussion

3.1. Model parameters & Data Description

In this paper, for the shape model we generate a training set of 30 ellipses (nuclei and lymphocyte being elliptical in shape) by changing the size of a principle axis with a gaussian probability function.

We evaluate ACOReW on two different histopathology datasets: prostate cancer and breast cancer cohorts comprising 100 and 14 images respectively. A total of 70 nuclei from 14 images for prostate and 504 lymphocytes from 52 images for breast cancer were manually delineated by an expert pathologist (serving as the ground truth annotation for quantitative evaluation).

3.2. Comparative Strategies

We qualitatively and quantitatively compared the segmentation performance of ACOReW with the GAC (Geodesic Active Contour) [2] and the Rousson shape based model [5]. We also compared ACOReW against ACORe (ACOReW with random initialization).

3.3. Evaluating Detection, Segmentation Performance

Figure 1 showcases the performance of the ACOReW model for the problem of segmenting ((a), (k)) nuclei, and (f) lymphocytes, on prostate and breast histopathology. Watershed provides the initial object initialization ((a), (f) and (k)) and ACROReW is able to ((c), (h) and (m)) outperform a GAC model ((b), (g) and (l)) for both datasets. We also compared the number of true objects segmented via the ACOReW, GAC, and ACORe by comparing the foreground regions formed by level sets (where $\phi > 0$). For the prostate dataset, ACOReW segmented 92% of the true lymphocytes as compared to 65% by GAC and 82% by ACORe. ACOReW achieved a corresponding 90% lymphocyte detection performance on the breast cancer dataset. We also computed sensitivity and positive predictive value for segmentation performance of each of the 4 schemes considered in this work. Quantitative results for the different schemes are shown in Table 1.

3.4. Evaluating Overlap Resolution

We define a measure for evaluating overlap resolution, OR, as follows: $OR = \frac{\#\text{ overlaps resolved}}{\text{Total }\#\text{ of overlaps}}$ The OR values for each of the 4 models are reported in Table 1 and they reflect the superior performance of ACOReW over the GAC, ACORe, and Rousson models for both the prostate and breast cancer datasets. ACOReW was able to resolve 92.5% of overlaps of nuclei and lymphocytes across both datasets.

Table 1. Quantitative evaluation of segmentation and overlap resolution results for the 4 models and across both the prostate and breast cancer datasets.

	SN	PPV	OR
GAC	0.20	0.58	0.022
Rousson	0.59	0.63	0.64
ACORe	0.73	0.64	0.86
ACOReW	0.82	0.66	0.91

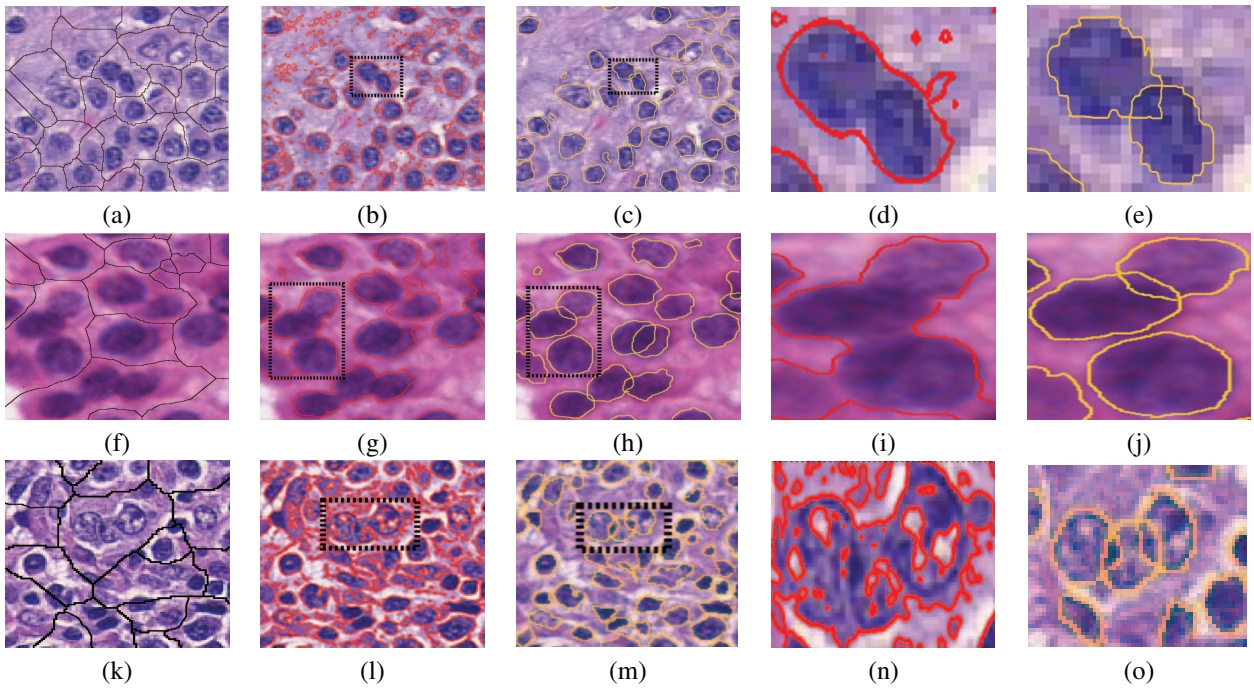

Fig. 1. Watershed initialization of nuclei and lymphocytes on prostate and breast cancer histopathology with corresponding segmentation results obtained via the GAC ((b), (g), (l)) and ACOReW ((c), (h), (m)) schemes. Magnified regions within (c), (h), (m) reveal the superior ability of ACOReW in both segmentation and overlap resolution compared to the GAC model.

4. Concluding Remarks

We presented a novel segmentation scheme Active Contour for Overlap Resolution using Watershed based Initialization (ACOReW) that employs minimal user intervention and uses boundary and region based active contours with a statistical shape model. Furthermore, we presented our model in a multiple level set formulation to segment multiple objects under mutual occlusion. We presented an application of ACOReW in context of segmenting nuclei and lymphocytes on prostate and breast histopathology. Our test results show that our model is more accurate compared to two state of the art active contour schemes. Our model was able to detect overlapping and non-overlapping lymphocytes and nuclei with 92% accuracy.

5. References

[1] Kass et al., "Snakes: Active contour models," *IJCV*, vol. 1, no. 4, pp. 321–331, 1988.

[2] Caselles et al., "Geodesic active contours," *IJCV*, vol. 22, no. 1, pp. 61–79, 1997.

[3] Chan et al., "Active contours without edges," *IEEE TIP*, vol. 10, no. 2, pp. 266 –277, Feb. 2001.

[4] Leventon et al., "Statistical shape influence in geodesic active contours," in *CVPR*, 2000, pp. 316–323.

[5] Rousson et al., "Shape priors for level set representations," in *ECCV*. 2002, pp. 78–92, Springer.

[6] Chan et al., "Level set based shape prior segmentation," in *CVPR*, 2005, vol. 2, pp. 1164 – 1170 vol. 2.

[7] Basavanhally et al., "Computer-aided prognosis of er+ breast cancer histopathology and correlating survival outcome with oncotype dx assay," in *IEEE TBE*, 282009-july1 2010, pp. 851 –854.

[8] Fatakdawala et al., "Em driven geodesic active contour with overlap resolution (emagacor): Application to breast cancer histopathology," in *IEEE TBE*, 2010, pp. 69 –76.

[9] Paragios et al., "Unifying boundary and region-based information for geodesic active tracking," in *CVPR*, 1999, vol. 2, p. 305 Vol. 2.

[10] Zhang et al., "Segmenting multiple familiar objects under mutual occlusion," in *ICIP*, 2006.

[11] Kiran et al., "Watersnake: integrating the watershed and the active contour algorithms," in *TENCON*, 2003, vol. 2, pp. 868 – 871 Vol.2.