

MULTIPHASE LEVEL SET FOR AUTOMATED DELINEATION OF MEMBRANE-BOUND MACROMOLECULES

Hang Chang and Bahram Parvin

Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720

ABSTRACT

Membrane-bound macromolecules play an important role in tissue architecture and cell-cell communication, and is regulated by almost one-third of the genome. At the optical scale, one group of membrane proteins expresses themselves as linear structures along the cell surface boundaries, while others are sequestered. This paper targets the former group, whose intensity distributions are often heterogeneous and may lack specificity. Segmentation of the membrane protein enables the quantitative assessment of localization for comparative analysis. We introduce a three-step process to (i) regularize the membrane signal through iterative tangential voting, (ii) constrain the location of surface proteins by nuclear features, and (iii) assign membrane proteins to individual cells through an application of multi-phase geodesic level-set. We have validated our method against a dataset of 200 images, and demonstrated that multiphase level set has a superior performance compared to gradient vector flow snake.

Index Terms— Segmentation, multiphase levelset, membrane proteins

1. INTRODUCTION AND MOTIVATION

Different types of cell surface proteins control and regulate cell-cell interactions and the physical properties of tissue structure. This paper focuses on the segmentation of membrane proteins that demonstrate a diffused signal along the cell surface boundaries when imaged by fluorescence microscopy. For example, the cadherin family of membrane proteins have a diffused signature, while the connexin family of proteins are sequestered between neighboring cells. In this paper, we use samples that are labeled for E-cadherin as a proxy for validating computational steps for a wider class of cell surface proteins. The E-, R-, and N-cadherin families of adhesion molecules are known to regulate the dynamic properties of cell-cell adhesion. For example, E-cadherin is a calcium-dependent cell adhesion molecule that influences differentiation and tissue structure; it is a class of an adherent junction between epithelial cells with access to the actin cytoskeleton through cadherin attachment proteins. As an endpoint, E-cadherin has been studied extensively, since it appears to function as a barrier to cancer. Loss of E-cadherin has been associated with (i) increased motility, (ii) potential cancer progression and metastasis, and (iii) increased resistance to normal cell death. Since down-regulation of E-cadherin is a critical endpoint for quantitative

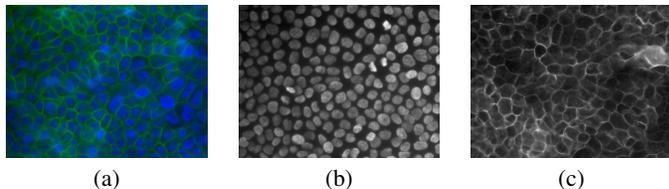


Fig. 1. An example of an E-cadherin signal: (a) A composite image of nuclei and membrane signals; (b) The corresponding nuclear channel; and (c) Membrane channel.

systems cell biology, detailed segmentation of the E-cadherin signal provides important clues for understanding biological processes under different sets of experimental factors and perturbation. An example of the localization of E-cadherin adhesion molecules for a monolayer system is shown in Figure 1.

Even though cell-cell adhesion molecules express themselves as linear structures, these signals suffer from perceptual gaps, nonuniformity in intensity and scale (e.g., thickness), and noise. The expressions of these molecules are often examined in the context of nuclear substructures with both the nuclear and adhesion molecules labeled with different fluorescent probes. Current literature is rich in terms of the delineation of nuclear morphology and its shape features [1, 2, 3, 4]; however, the segmentation of cell-cell adhesion molecules or cell surface proteins remains largely unexplored. In this paper, we introduce a novel method for segmentation of adhesion molecules that are bound to the basal-lateral region of the cell. These signals correspond to local linear features that delineate cell boundaries as shown in Figure 1. Our approach is to group and enhance these linear structures based on continuity, and then apply an evolving front for detailed segmentation. Our method enables the inference of saliency from incomplete boundary information through voting, perceptual grouping, and multi-phase geodesic level-set.

The organization of this paper is as follows: Section 2 provides a brief review of the previous research; Section 3 describes our approach and details our implementation of tangential voting; Section 4 demonstrates the experimental results; and Section 5 concludes the paper.

2. REVIEW OF PREVIOUS WORK

The difficulties in the segmentation of surface protein localization are often due to variations in scale, noise, and topology. Other complexities originate from missing data and perceptual boundaries that lead to diffusion and dispersion of the spatial grouping in the image space. Techniques for grouping local image features into globally salient structures have incorporated clustering and graph theoretic

THIS RESEARCH WAS SUPPORTED IN PART BY A GRANT FROM NATIONAL CANCER INSTITUTE (1P01107584-01A1) AND THE U.S. DEPARTMENT OF ENERGY, OFFICE OF BIOLOGICAL AND ENVIRONMENTAL RESEARCH, LOW DOSE RESEARCH PROGRAM UNDER CONTRACT NUMBER DE-AC02-05CH1131.

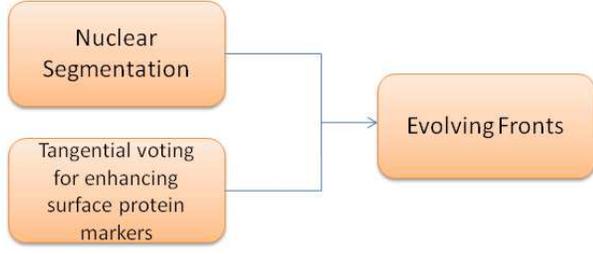


Fig. 2. Computational steps in segmenting cell surface proteins.

methods [5], Bayesian models for combining tangential representations of sparse contours [6], tensor voting [7] for grouping or interpolating distant features, etc. While these techniques differ in their concepts, they are model-free and share a common thread of continuity and proximity along the minimum energy path to infer global saliency. More recently, prior shape models have also been incorporated for boundary closure [8].

In [9], an interactive method with evolving fronts was proposed for cell segmentation. In [10], we proposed a three step strategy based on (i) tangential voting strategy for membrane signal enhancement and grouping, (ii) delineating of nuclear regions through a combination of zero-crossing and gradient filter, and (iii) snake-based for delineating membrane-bound signal. The main limitations are that segmentation of clumps of nuclei was exhaustive, and the snake model of evolving front did not incorporate interactions between neighboring nuclei. These issues are being addressed in this paper.

3. APPROACH

Figure 2 shows the computational steps in segmenting cell surface proteins, Figure 3 gives an example for each step. Nuclear segmentation provides one set of constraints for membrane segmentation. The second set of constraints is provided through the enhancement and regularization of cell surface protein markers. The application of multi-phase geodesic levelset provides the final assignment of surface proteins for each cell.

3.1. Delineation and quantification of the membrane-bound macromolecules

A typical assay often provides a label for the nuclear regions (e.g., a counter-stain). Therefore, segmentation of nuclei provides the context for characterizing membrane-bound proteins. The two-step process consists of:

1. *Nuclear segmentation*: Among different strategies [11, 12, 13, 14], we have adopted the method described in [11]. The advantage of this method is that geometric constraints are directly applied to delineate clumps of nuclear regions. Thus, facilitating analysis on a cell-by-cell basis, where the key geometric constraint is expressed in term of the convexity of the nuclear geometry. Accordingly, when nearby nuclei overlap, they form folds (e.g., positive curvature maxima). By identifying folds and applying Delaunay triangulation between folds, a set of hypothesis can be generated for partitioning a clump of cells. These hypotheses are then processed through

a constraint satisfaction network for revealing a correct partitioning for each nucleus. The method is superior to an earlier method [2], which was based on exhaustive search and did not use Delaunay triangulation for limiting the search space.

2. *Membrane signal grouping*: Membrane-bound signals can be noisy and contain perceptual gaps, as shown in Figure 1. Therefore, the signal is regularized through tangential voting, which has been described in our earlier paper [15]. The net result is that evolving contour is more smooth along the cell boundaries.

3.2. Grouping of the Membrane Signal

Our intent is to quantify membrane-bound protein localization on a cell-by-cell basis, and the nuclear channel is used as a context in which the segmentation problem is constrained. Here, we propose to use multi-phase level-set framework [16, 17, 4]. Within this framework, we have two principles:

1. The evolving contour (zero level-set) represents the membrane for each nucleus, which means the contour must attach to the membrane signal.
2. Each phase represents a unique cell, and different phase regions (cell regions) do not overlap.

Based on the principles listed above, and using the Heaviside function H , and the one-dimensional Dirac measure δ , defined by

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0 \end{cases}, \delta(z) = \frac{dH(z)}{dz}$$

The energy form can be written as:

$$E = \mu \sum_i^M \int_{\Omega} g(I) |\nabla \Phi_i(x, y)| \delta_i(x, y) dx dy + \lambda \sum_{i=1}^M \sum_{j=1, j \neq i}^M \int_{\Omega} H(\Phi_i) H(\Phi_j) dx dy \quad (1)$$

in which, Ω is the image domain; M is the number of nuclei; I is the enhanced membrane image; Φ is the Lipschitz function, whose zero level-set is designated to attach to the membrane signal; μ and λ are constant coefficients, weighting different terms; and

$$g(I) = \frac{1}{1 + |\nabla I|^p} \quad (2)$$

The first term is the geodesic length of the zero level-set, enforcing the zero level-set to attach to the membrane signal; the second term is the penalty for overlapping. The minimization of the objective function is achieved by gradient descent based on the corresponding Euler-Lagrange equation:

$$\frac{\partial \Phi_i}{\partial t} = \delta(\Phi_i) \cdot \left(\mu \nabla g \frac{\nabla \Phi_i}{|\nabla \Phi_i|} + \mu g \text{div} \left(\frac{\nabla \Phi_i}{|\nabla \Phi_i|} \right) \right) - \delta(\Phi_i) \cdot \lambda \cdot \sum_{j=1, j \neq i}^M H(\Phi_j) \quad (3)$$

To discretize the equation: let h be the space step, Δt be the time step, and $(x_p, y_q) = (ph, qh)$ be the grid points. The finite differences are:

$$\begin{aligned} \Delta_+^x \Phi(p, q) &= \Phi(p+1, q) - \Phi(p, q) \\ \Delta_-^x \Phi(p, q) &= \Phi(p, q) - \Phi(p-1, q) \\ \Delta_+^y \Phi(p, q) &= \Phi(p, q+1) - \Phi(p, q) \\ \Delta_-^y \Phi(p, q) &= \Phi(p, q) - \Phi(p, q-1) \end{aligned}$$

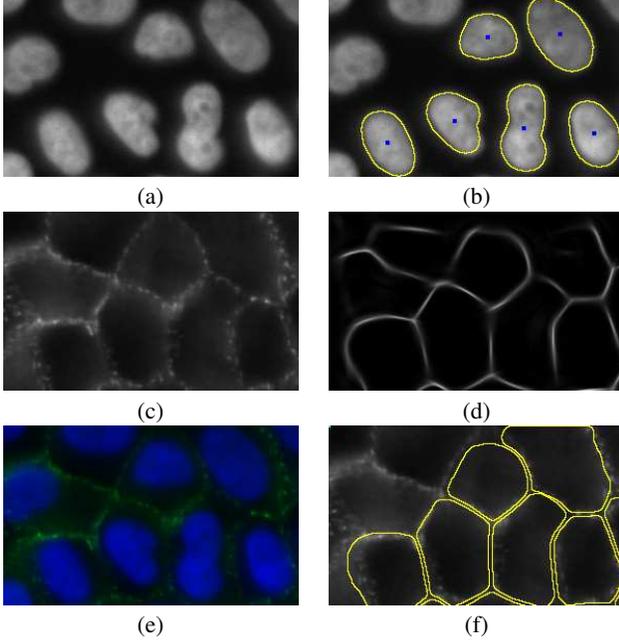


Fig. 3. Computational steps in the assignment of cell surface protein markers to each nucleus: (a) original nuclear regions; (b) segmented nuclear features; (c) original cell surface marker; (d) voted cell surface proteins; (e) composite image of both nuclear channel and membrane channel; (f) segmentation of cell surface protein.

We compute Φ^{n+1} by the following discretization:

$$\begin{aligned}
\frac{\Phi_i^{n+1} - \Phi_i^n}{\Delta t} &= \delta_h(\Phi_i^n) \left[\frac{\mu}{h^2} \frac{\Delta_+^x g \Delta_+^x \Phi_i^n}{\sqrt{\frac{(\Delta_+^x \Phi_i^n)^2}{h^2} + \frac{(\Delta_+^y \Phi_i^n)^2}{h^2}}} \right] \\
+ \delta_h(\Phi_i^n) &\left[\frac{\mu}{h^2} \frac{\Delta_+^y g \Delta_+^y \Phi_i^n}{\sqrt{\frac{(\Delta_+^x \Phi_i^n)^2}{h^2} + \frac{(\Delta_+^y \Phi_i^n)^2}{h^2}}} \right] \\
+ \delta_h(\Phi_i^n) &\left[\frac{\mu g}{h^2} \Delta_+^x \left(\frac{\Delta_+^x \Phi_i^n}{\sqrt{\frac{(\Delta_+^x \Phi_i^n)^2}{h^2} + \frac{(\Delta_+^y \Phi_i^n)^2}{h^2}}} \right) \right] \\
+ \delta_h(\Phi_i^n) &\left[\frac{\mu g}{h^2} \Delta_+^y \left(\frac{\Delta_+^y \Phi_i^n}{\sqrt{\frac{(\Delta_+^x \Phi_i^n)^2}{h^2} + \frac{(\Delta_+^y \Phi_i^n)^2}{h^2}}} \right) \right] \\
- \delta_h(\Phi_i^n) \cdot \lambda \cdot &\sum_{j=1, j \neq i}^M H(\Phi_j^n) \quad (4)
\end{aligned}$$

4. EXPERIMENTAL RESULTS

An experiment has been designed and images have been collected to validate the performance of the method. In this experiment, samples are labeled with a counter-stain to visualize nuclear features, and with an antibody against the E-cadherin for visualizing cell surface proteins. A total of 201 images were collected, nuclear regions were segmented, and an E-cadherin signal was assigned to each nuclear region. The parameters are fixed for the entire data set at $\mu = 100$, $\lambda = 5$, $\Delta t = 0.1$, and *iteration* = 100. Figures 4 shows one

example of our experimental result, various processing steps, and comparison with an earlier method based on GVF snake [18]. The GVF snake has the property that its evolution can stop due to absence of attracting forces. Several cells in Figure 4 have this property. In addition, Figure 4 shows intermediate results of tangential voting for smoothing the membrane-bound signal [15]. Tangential voting regularizes the membrane-bound signal through enhancement, noise reduction, and gap filling of small regions. In our experimental data set, the quality of segmentation for the membrane-bound protein is nearly perfect, and any erroneous segmentation is due to incorrect nuclear segmentation. Sometimes, during the sample preparation and fixation, nuclear regions overlap and form large clumps; this is a quality control issue associated with sample preparation, and in these cases, even with correct segmentation of the nuclear regions, membrane-bound signals are meaningless, since it is not clear which nuclei they are associated with.

5. CONCLUSION AND FUTURE WORK

A series of computational steps are proposed and tested on real data to delineate cell surface proteins. Cell surface proteins are heterogeneous in width and suffer from perceptual gaps. A multi-step process is proposed to segment membrane proteins and assign them to individual cells. First iterative tangential voting is applied to enhance and regularize membrane proteins while diffusing noise. Next, the nuclear segmentation provides context and the necessary reference for initializing the evolving fronts for quantifying membrane-bound signal on the cell-by-cell basis. Finally, the GVF snake is replaced with the multi-phase geodesic level-set model for improved performance and geometric stability. Our main contribution is to build the robust pipeline for delineating and capturing membrane-bound signal on cell-by-cell basis. Our current plan is to utilize this technology for understanding a number of biological processes.

6. REFERENCES

- [1] J. Byun, B. Sumengen, P. Lewis, B. Manjunath, S. Fisher, Automated tool for the detection of cell nuclei in digital microscopic images, *Molecular Vision* 12 (2006) 949–960.
- [2] S. Raman, C. Maxwell, M. Barcellos-Hoff, B. Parvin, Geometric approach segmentation and protein localization in cell cultured assays, *Journal of Microscopy* (2007) 427–436.
- [3] H. Chang, Q. Yang, B. Parvin, Segmentation of heterogeneous blob objects through voting and level set formulation, *Pattern Recognition Letters* 28 (13) (2007) 1781–1787.
- [4] S. Nath, K. Palanispan, F. Bunyak, Cell segmentation using coupled level-set and graph vertex coloring, *Medical image computing and computed-assisted intervention-Miccai* 4190 (2006) 101–108.
- [5] Z. Wu, R. Leahy, An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (11) (1993) 1101–1113.
- [6] J. H. Elder, S. W. Zucker, Computing contour closure, in: *ECCV* (1), 1996, pp. 399–412.
- [7] M. L. G. Medioni, C. Tang, *A Computational Framework for Segmentation and Grouping*, Elsevier, 2000.
- [8] T. Chan, W. Zhu, Level set based shape prior segmentation, in: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2005, pp. II: 1164–1170.

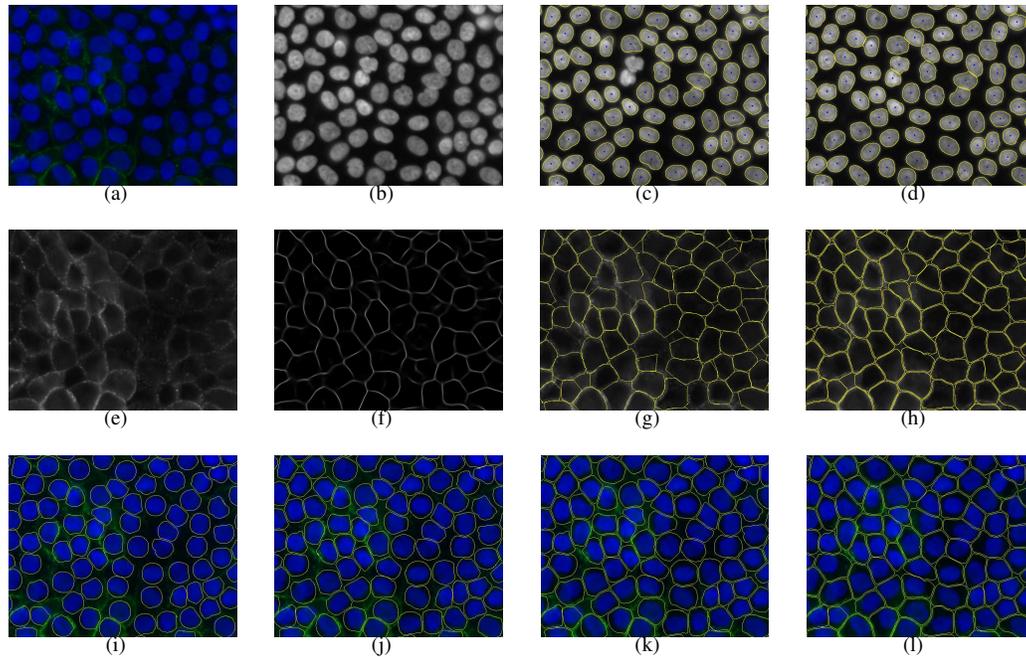


Fig. 4. Localization of membrane bound protein for a 2D cell culture model: (a) composite image of both nuclear channel and membrane channel; (b) original nuclear region; (c) segmented nuclear features with method in [2]; (d) segmented nuclear features with our approach; (e) original E-cadherin signal; (f) result of tangential voting along the membrane-bound signal; (g) segmentation of cell surface protein via GVF snake [18]; and (h) segmentation of cell surface protein with our approach (multi-phase geodesic level-set); (i-l) intermediate results in level-set evolution.

- [9] L. Coulot, H. Kirschner, A. Chebira, J. Moura, J. Kovacevic, E. Osuna, R. Murphy, Topology preserving stacs segmentation of protein subcellular location images, in: ISBI, 2006, pp. 566–569.
- [10] J. Han, H. Chang, K. L. Andarawewa, P. Yaswen, M. H. Barcellos-Hoff, B. Parvin, Multidimensional profiling of cell surface proteins and nuclear markers, *IEEE/AMC Transaction on Computational Biology and Bioinformatics* 7 (1) (2010) 80–90.
- [11] Q. Wen, H. Chang, B. Parvin, A delaunay triangulation approach for segmenting clumps of nuclei, in: ISBI, 2009, pp. 9–12.
- [12] T. R. Jones, A. Carpenter, P. Golland, Voronoi-based segmentation of cells on image manifolds, *LECTURE NOTES IN COMPUTER SCIENCE* (3765) (2005) 535–543.
- [13] S. K. Nath, K. Palaniappan, F. Bunyak, Accurate spatial neighborhood relationships for arbitrarily shaped objects using hamilton-jacobi gvd, *LECTURE NOTES COMPUTER SCIENCE* (4522) (2007) 421–431.
- [14] G. Srinivasa, M. Fickus, M. Gonzalez-Rivero, S. Hsieh, Y. Guo, A. Linstedt, J. Kovacevic, Active mask segmentation for the cell-volume computation and golgi-body segmentation of hela cell images, in: ISBI, 2008, pp. 348–351.
- [15] H. Chang, K. L. Andarawewa, J. Han, M. H. Barcellos-Hoff, B. Parvin, Perceptual grouping of membrane signals in cell-based assays, in: ISBI, 2007, pp. 532–535.
- [16] V. Caselles, R. Kimmel, G. Sapiro, Geodesic active contours, in: *Proceedings of the IEEE International Conference on Computer Vision*, 1995, pp. 694–699.
- [17] L. Vese, T. Chan, A multiphase level set framework for image segmentation using the mumford and shah model, *International Journal of Computer Vision* 50 (3) (2002) 271–293.
- [18] J. Han, H. Chang, K. Andarawewa, P. Yaswen, M. H. Barcellos-Hoff, B. Parvin, Integrated profiling of cell surface protein and nuclear marker for discriminant analysis, in: ISBI, 2007, pp. 532–535.