Informed Segmentation: A Framework for Using Context to Select an Algorithm and a Case Study Using Humans in the Loop

Danna Gurari, Diane Theriault, and Margrit Betke

Boston University Computer Science Department

Abstract. Over the past 30 years, many image segmentation algorithms have been proposed to cope with challenges arising from varying object appearance, environmental conditions, and image quality. Which algorithm will work best for any particular image set is not known a priori, and the search for a single algorithm that will work well in all circumstances seems futile. In this paper, we propose to abandon the search for a single algorithm that works well in general and instead propose to select a segmentation algorithm most appropriate for the image context. We implemented this framework by linking segmentation algorithms with domain-expert-provided classifications to find the boundaries of fibroblast cells, which exhibit large appearance variability, in phase-contrast images. A case study reveals that our system yields higher quality segmentations than nine popular freely-available standalone algorithms.

1 Introduction

By analyzing cell shape over time, researchers hope to gain an understanding of fundamental biological processes and use this knowledge in turn to diagnose diseases and engineer biomaterials. They often rely on single-cell analysis to discover the relationship between cell shape and function [1, 2] and how the environment influences cell phenotype [3]. Such quantitative analysis can depend on detecting subtle appearance variation which motivates the need for a laboratory protocol that consistently produces high quality segmentations.

The problem of segmentation of cell images has been extensively studied over the last few decades [4]. While many algorithms are reported to work well for particular image sets, it is not clear which single algorithm will work best in general for the large variety of shapes and appearances, wide range of environmental conditions, and large number of image acquisition methods. Among the older methods, thresholding, region growing, and feature-based algorithms are still actively used due to their widespread availability in bioimage analysis systems and ease of use. *Thresholding* methods are based on the assumption that cells ("foreground") have different intensity values than the background. *Otsu thresholding* sets the threshold to the value that minimizes the average variance between all foreground and background pixels, respectively [5]. *Adaptive thresholding* adjusts the threshold applied throughout the image, such that it is set to

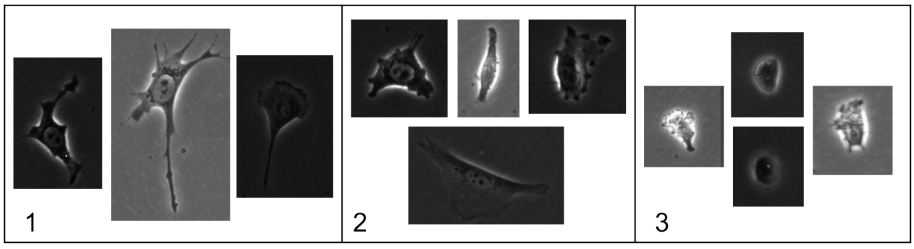

Fig. 1. Phase-contrast microscopy images of fibroblasts with relative size preserved. Context is defined as one of three states of the cell: State 1: Cells that are larger with respect to the rest of the population and typically have weak edges and thin branch-like protrusions. State 2: Cells that have average size and exhibit some some protrusions. State 3: Cells that are smaller with respect to the rest of the population and commonly have a circular shape, surrounded by a bright-intensity "halo," with few or no protrusions.

the mean of the local neighborhood for each pixel. *Seeded watershed* is a popular *region-growing* method that clusters pixels based on spatial proximity and intensity homogeneity, assigning every pixel to the cluster of exactly one of the seed points. *Hough transform with circles* is a popular *feature-based* method that uses as the foreground the combination of circles that have at least a pre-specified number of pixels on their boundary in the edge map of the image [6].

Level set based methods have become popular with bioimage analysis systems [7]. In general, these methods deform an initial contour to a final contour, separating image foreground from background so that some method-specific image partition condition is enforced. *Geodesic active contours* evolves the initial contour to end up in regions with strong edges (high contrast) [8]. *Active contours without edges* evolves the initial contour to try to separate the image into two homogeneous regions [9]. *Lankton region-based active contours* evolves the initial contour by using the local neighborhood statistics for each pixel in order to adjust how to separate the sub-region into two homogeneous regions [10]. The *Shi approximation method* computationally speeds up the evolution process by replacing slow real-valued calculations with faster integer-based calculations [11]. The method by *Bernard et al.* uses a linear combination of B-spline basis functions for process speedup [12]. Recent work suggests that the success of *level set* based methods for phase contrast images depends on specialized contour initialization methods to avoid common curve evolution failures [13, 14].

The key computer vision challenge addressed in this paper is how to automatically obtain accurate contours of objects that can exhibit large variability in shape and appearance. Instead of searching for a single algorithm that will work well in all circumstances, we propose a strategy for using context to determine which algorithm to apply. In our study, we work with phase-contrast microscopy images showing populations of fibroblasts that undergo significant changes in

appearance over time, as exemplified in Figure 1. We define context as one of three cell states used by biologists to characterize fibroblasts.

Our work relates to efforts for combining human and computer resources to address the segmentation problem. Proposed methods include reducing human involvement from drawing the boundary of an object to either drawing a cross [15] or clicking on points both inside and outside of the true boundary [16] to guide a segmentation algorithm. Our work examines how to link a biologically-defined taxonomy used to distinguish between different fibroblast appearances with computer vision algorithms. Specifically, we examine whether we can bridge the gap of terminology between two communities such that the state of a fibroblast classified by a biologist can be used to choose the best-suited segmentation algorithm. The key contributions of this paper are:

1. A framework for identifying, based on the image context, the optimal segmentation source.
2. A context-based system that incorporate humans in the loop to generate segmentations and a case study that demonstrates the system yields higher quality segmentations than nine popular, freely-available cell segmentation algorithms.
3. A contour initialization approach for level set based algorithms, called *Variance Maps*, that led to high-quality segmentations in phase contrast images.

2 Methods

We propose a framework for identifying, based on image context, which among many segmentation algorithm options will lead to the highest quality segmentation (**Fig. 2**). We then describe a system implementing this framework that combines human and computer expertise to segment highly deformable cells in phase contrast images. We finally propose an algorithm to automatically create initial contours needed by level set based algorithms.

2.1 Proposed Context-based Segmentation Framework

Our proposed framework uses image context to identify the best-suited segmentation algorithm to apply. It includes an off-line phase where one defines the contexts, then establishes a method to identify each context, and finally learns the optimal segmentation algorithm for each context. In the online phase, a context label is first assigned to the specified image (or image region) and then the optimal segmentation algorithm for that context is applied. This two-step process produces the final output segmentation.

In the off-line phase, contexts may be defined manually or automatically. They could be defined by domain experts who already have built a taxonomy to differentiate between object or image appearances. One could also use machine learning algorithms to create classifiers that distinguish between different contexts based on the image information. Mapping segmentation algorithms to

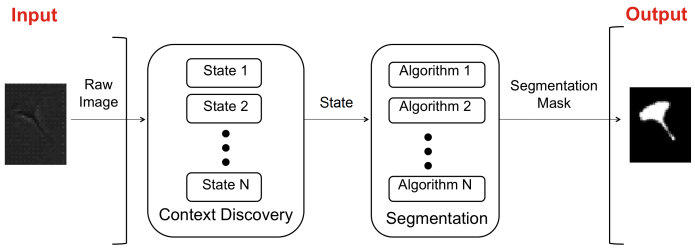

Fig. 2. Example of a system that applies the online phase of the proposed context-based segmentation framework. For each image, the state of the fibroblast cell is discovered and then an appropriate final segmentation algorithm is applied.

context may similarly be achieved manually or automatically. Manually, a human could inspect segmentations created by numerous algorithms for all images in a particular context and then recommend a best-performing algorithm. This process would be repeated for each context. Automatically, one would instead quantitatively validate a set of algorithms and then apply the best performing one for each context. To do so, one could choose a collection of images for each context, collect gold standards representing the true boundaries of each object in every image, define performance metrics, and then choose the algorithm for each context that returns the highest performance scores.

2.2 Proposed Context-based Segmentation System

We applied the proposed framework to consistently collect high-quality segmentations for fibroblasts in phase contrast images by linking each classified fibroblast state from a biologically-defined taxonomy with a best-performing segmentation algorithm for that state. In the off-line phase, a mapping between fibroblast states and segmentation algorithms was established. In the online phase, the context-based segmentation system used these learned pairings to select the appropriate segmentation algorithm on the fly.

Off-line Phase: First, we define context by adopting three fibroblast states used by domain experts to distinguish between the spectrum of appearances fibroblasts exhibit (**Fig. 1**). We then evaluated the following nine popular freely-available cell segmentation algorithms [7] to establish optimal pairings between each fibroblast state and a segmentation algorithm: adaptive thresholding, Otsu thresholding [5], Hough transform for circles [6], seeded watershed, geodesic active contours [8], active contours without edges [9], Lankton region-based active contours [10], Shi approximation level set method [11], and Bernard level set method [12]. Since the five level-set based algorithms are sensitive to initialization and we observed that using basic geometric shapes such as bounding rectangles as initial contours did not work well for our image sets, we instead used the output of our proposed segmentation algorithm *Variance Maps* (Section 2.3). It consistently undersegments the fibroblasts while closely hugging

their boundaries. The "best segmentation algorithm" for each of the three cell states is determined using the algorithm that returns the highest performance score for each context (section 3).

On-line Phase: The context-based segmentation system links fibroblast classifications indicating appearance with computer vision algorithms to create accurate segmentations. The system first collects from the domain expert the fibroblast location (i.e., pixel coordinate) labeled with the fibroblast state and then applies the optimal segmentation algorithm for that state. The resulting segmentation is post-processed by filling any holes in the output object(s) and, if there is more than one object, filtering out all pixels except those in the object closest to the centroid of the detection region. This results in a binary mask identifying the outline of the fibroblast in the image.

2.3 Proposed Contour Initialization Algorithm: Variance Maps

Variance Maps takes advantage of the observation that in our phase contrast image sets, the variability of intensity values inside cells is often larger than the variability of intensity values in the background. The algorithm runs in linear time with respect to the number of pixels in the image and consists of two steps: 1) for every pixel in the image, compute the image texture and 2) threshold the resulting texture image to distinguish between foreground and background pixels. Our key design decisions were the image texture measure, neighborhood type, and threshold value on the texture image.

The texture measure we ultimately chose builds off the standard deviation transform proposed by Bradhurst et al. [17], which computes the variance $\sigma^2(N)$ of the intensity values $I(x, y)$ of pixels in a circular neighborhood N with radius r as $\sigma^2(N) = 1/|N| \sum_{(x,y)\in N} I(x, y)^2 - (1/|N| \sum_{(x,y)\in N} I(x, y))^2$. To improve the computational efficiency of the variance equation, we used square instead of circular neighborhoods. We used a width and height of $2r + 1$ pixels, resulting in a neighborhood size of $|N| = (2r + 1)^2$ and the following variance equation:

$$\sigma^2(N) = \frac{1}{|N|} \sum_{x-r}^{x+r} \sum_{y-r}^{y+r} I(x, y)^2 - \left(\frac{1}{|N|} \sum_{x-r}^{x+r} \sum_{y-r}^{y+r} I(x, y) \right)^2, \tag{1}$$

which can then be computed in time constant in the number of pixels with the integral image S. The integral image S assigns each pixel a value equal to the sum of all of the image intensity values above and to the left of it: $S(X, Y) = \sum_{x=1}^{X} \sum_{y=1}^{Y} I(x, y)$. Then, the sum of the intensity values in a square neighborhood of size $(2r + 1)^2$ surrounding a pixel located at (x, y) is computed by $\sum_{x-r}^{x+r} \sum_{y-r}^{y+r} I(x, y) = S(x+r, y+r) - S(x-r-1, y+r) - S(x+r, y-r-1) + S(x-r-1, y-r-1)$. A similar integral image can be computed for the squared image values in order to compute $1/(2r+1)^2 \sum_{x-r}^{x+r} \sum_{y-r}^{y+r} I(x, y)^2$. We freely share this code (`http://www.cs.bu.edu/~betke/BiomedicalImageSegmentation`).

Fig. 3. Representative segmentation results. For a fibroblast in each state (rows 1–3), the figure shows the raw image (col. 1), the gold standard (col. 2), the algorithm-drawn segmentations (cols. 3–11), and the segmentation created by *Variance Maps* used to initialize the level set based methods (col. 12).

3 Experiments and Results

Our experiments involved an evaluation of the proposed context-based segmentation system in comparison to the nine stand-alone cell segmentation methods. The image library shows fibroblasts of the Balb /c 3T3 mouse strain cultured at 37°C in 5% CO_2. It was generated by a domain expert to represent the variability of fibroblast appearances when captured using various image acquisition parameters under different environmental conditions. It contains 125 images showing 361 fibroblasts.

We quantitatively analyzed the quality of each segmentation by computing how closely it matched each gold standard region, using the *Jaccard similarity index*, a standard evaluation metric. This metric computes the number of pixels common to both the algorithm-generated and gold-standard regions that is also in the combination of their regions (i.e., $\frac{|A \cap B|}{|A \cup B|}$ where A represents the set of pixels in the gold standard segmentation and B represents the set of pixels in the segmentation to analyze). Because of known limitations with the traditional approach for using manual annotation to establish gold standard segmentations[18], we instead established the gold standard segmentation by having a domain expert choose the segmentation best representing each fibroblast from the collection of the manual annotation, nine algorithm-generated segmentations, and the *Variance Maps* segmentation. The domain expert chose the best segmentation from the 11 segmentations shown simultaneously. To prevent that the experts may learn the algorithmic or manual source of the segmentations, the user interface randomized the order of presented segmentations.

To analyze the standalone segmentation algorithms, we created nine sets of segmentations for each of 361 fibroblasts using the nine algorithms supported in the proposed context-based segmentation system (**Fig. 3**). We then computed scores indicating the quality for all 3,610 generated segmentations and report average scores for the whole image library (**Table 1, row 2**) as well as based on each of the three cell states (**Table 1, rows 3–5**).

To analyze our context-based segmentation system, we partitioned the data into three sets and then conducted three experiments. In each experiment, we

Table 1. Mean precision score for each algorithm evaluated for all 361 cells (row 2) and based on fibroblast state (rows 3–5).

Dataset	AdTh	Otsu	HoTr	Wate	Case	ChVe	Lank	Shi	Bern
All	0.36	0.44	0.61	**0.76**	0.65	0.51	0.65	0.37	0.32
State 1	0.22	0.25	0.55	**0.80**	0.66	0.41	0.67	0.14	0.12
State 2	0.33	0.38	0.51	**0.85**	0.69	0.47	0.72	0.29	0.19
State 3	0.49	0.64	**0.76**	0.65	0.61	0.63	0.56	0.63	0.60

trained on one third of the data to learn the optimal segmentation algorithm for each cell state (offline phase) and then tested the system on the remaining two thirds of the data with the optimal algorithm (online phase). We averaged scores from the three experiments. We found that in each experiment, the context-based segmentation system selected the Hough Transform for fibroblasts in state 3 and the seeded watershed method for fibroblasts in states 1 and 2. The context-based segmentation algorithm gives an average score for the three experiments of 0.80, which exceeds the average score for the best standalone segmentation algorithm (Watershed) by 4 percent points.

To analyze our proposed *Variance Maps* algorithm for creating the initial contour for the level set based methods, we used two measures: *accuracy* to calculate the fraction of the true cell region captured by the segmented region (i.e., $\frac{|A \cap B|}{|A|}$ where A represents the set of pixels in the gold standard segmentation and B represents the set of pixels in the segmentation to analyze) and the aforementioned *Jaccard similarity index*. We found the mean accuracy score was 0.97 and the mean Jaccard similarity index was 0.51 for all 361 cells.

4 Discussion

We found that linking domain-expertise from the computer vision and biology communities led to benefits for both communities. Biology experts could obtain high quality annotations efficiently by relying on their domain-specific taxonomy of images rather than by learning about segmentation algorithms. Similarly, computer vision experts could leverage the biology-defined taxonomies distinguishing between different image contexts in order to achieve improved segmentation performance over nine stand-alone methods.

We infer from our experiments that we can learn more about the strengths and weaknesses of various segmentation algorithms and the simple versus difficult images to segment by analyzing algorithm performance on a variety of image sets. The algorithms generally performed the best with cells in state 3 and the worst for fibroblasts in state 1. This suggests the commonly smooth, circular shape of state-3 fibroblasts surrounded by a bright "halo" is easier to capture automatically than state-1 fibroblasts that commonly have weak edges and thin, branch-like protrusions. Furthermore, the results demonstrate that the variation between the best and worst performing algorithms for fibroblasts in state 1

is nearly twice the variation of the best and worst performing algorithms for fibroblasts in state 3, indicating algorithm choice matters more for fibroblasts with ill-defined edges and protrusions. We hypothesize that similar algorithm performance will be observed for other cell types that also can often range in appearance from round with halos to thin with protrusions.

We observed that the proposed *Variance Maps* algorithm used to create the initial contours for the level set based algorithms consistently captured the entire true fibroblast region, albeit without capturing the fine boundary details (last column, Fig. 3). We found that this algorithm yielded a good initial contour estimate to pair with level set based methods for a wide range of cell appearances observed in phase contrast images.

5 Conclusions

Identifying a segmentation algorithm that works well in general can be challenging, especially when the data exhibit significant variability. We proposed a framework for selecting the best-suited segmentation algorithm to apply based on context. We described an implementation of this approach that combines domain experts with fully-automated algorithms to use a combination of segmentation algorithms instead of applying any single method. We found this method yielded improved segmentation results over nine popular, freely available cell segmentation algorithms on phase-contrast images of highly deformable cells. Future work will explore how to remove human involvement by building a classifier to automatically predict from the raw image the cell state. Possible future research directions will also include generalizing this work by linking other expert-based image taxonomies with segmentation algorithms and running a larger-scale study with more algorithms and image sets to analyze the factors that influence the successes and failures of algorithms.

Acknowledgments

The authors gratefully acknowledge funding from the National Science Foundation (IIS-0910908) and thank Matthew Walker and Joyce Y. Wong for the images and annotations they provided.

References

1. K. Wada, K. Itoga, T. Okano, S. Yonemura, and H. Sasaki. Hippo pathway regulation by cell morphology and stress fibers. *Development*, 138:3907–3914, 2011.
2. S. F. Tavazoie et al. Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nature Neuroscience*, 8(12):818–829, 2008.
3. T. Yeung et al. Effects of Substrate Stiffness on Cell Morphology, Cytoskeletal Structure, and Adhesion. *Cell Motility and the Cytoskeleton*, 60(1):24–34, 2005.
4. E. Meijering. Cell segmentation: 50 years down the road. *IEEE Signal Processing Magazine*, 29(5):140–145, 2012.

5. N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.
6. D. Ballard. Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, 13(2):111–122, 1981.
7. T. Dietenbeck, M. Alessandrini, D. Friboulet, and O. Bernard. Creaseg: A free software for the evaluation of image segmentation algorithms based on level-set. *IEEE International Conference on Image Processing (ICIP)*, pages 665–668, 2010.
8. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours *IEEE Transactions on Image Processing*, 22(1):61–79, 1997.
9. T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
10. S. Lankton and A. Tannenbaum. Localizing region-based active contours. *IEEE Transactions on Image Processing*, 17(11):2029–2039, 2008.
11. Y. Shi and W. C. Karl. A real-time algorithm for the approximation of level-set based curve evolution. *IEEE Transactions on Image Processing*, 17(5):645–656, 2008.
12. O. Bernard, D. Friboulet, P. Thevenaz, and M. Unser. Variational b-spline level-set: A linear filtering approach for fast deformable model evolution. *IEEE Transactions on Image Processing*, 18(6):1179–1191, 2010.
13. K. Li et al. Cell population tracking and lineage construction with spatiotemporal context. *Medical Image Analysis*, 12(5):546–566, 2008.
14. I. Ersoy et al. Cell spreading analysis with directed edge profile-guided level set active contours. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 376–383, 2008.
15. N. Ben-Zadok, T. Riklin-Raviv, and N. Kiryati. Interactive level set segmentation for image-guided therapy. *IEEE International Symposium on Biomedical Imaging*, pages 1079 - 1082, 2009.
16. A. Kronman and L. Joskowicz. Image Segmentation Errors Correction by Mesh Segmentation and Deformation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 206-213, 2013.
17. C. J. Bradhurst, W. Boles, and Y. Xioa. Segmentation of bone marrow stromal cells in phase contrast microscopy images. In *23rd IEEE International Conference on Image and Vision Computing New Zealand (IVCNZ)*, 2008. 6 pages.
18. D. Gurari et al. SAGE: An Approach and Implementation Empowering Quick and Reliable Quantitative Analysis of Segmentation Quality. *Proceedings of the IEEE Workshop on Applications in Computer Vision (WACV)* pages 475 - 481, 2013.