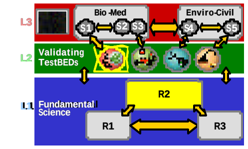
Automated Parameter Selection for Optimal Biological Image Segmentation

M. Amri Abdul-Karim[1], Edward Brown[2], Mark Booth[3], Andreas Jeromin[4], Jeff Sprenger[6], Murat Yuksel[1], Shivkumar Kalyanaraman[1], William Shain[5], Jack R. Glaser[6], Badrinath Roysam[1]*

[1]Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180
[2]Edwin L. Steele Laboratory, Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114
[3]Center for Behavioral Neuroscience, Georgia Tech/Emory BME, Atlanta, GA 30332
[4]Division of Neuroscience, Baylor College of Medicine, Houston, Texas 77030
[5]The Wadsworth Center, New York State Department of Health, Albany, NY 12201
[6]MicroBrightField, Inc., Williston, VT 05495

Rensselaer

why not change the world?



This work was supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation (Award Number EEC-9986821), Microbrightfield Inc., and MARA of Malaysia.

Abstract

The goal of this project is to automate and simplify the ordinarily arcane task of selecting the parameters ("settings") for biological image analysis systems using the common notions of coverage and complexity. This is essential for broadening the scope of applicability of these algorithms. The problem of automated parameter selection is formulated as a variational optimization based on mathematical descriptions of segmentation conciseness and coverage, a probabilistic vesselness measure, scale, and the Minimum Description Length (MDL) principle. Of specific interest to us is the analysis of tube-like biological objects such as tumor micro-vasculature and neurons. Next, the efficient Recursive Random Search (RRS) global optimization algorithm is utilized to explore the parameter space in an efficient and parallel manner. As an example, the resulting parameter set after 1000 RRS steps for all nine parameters of our automated tracing algorithm produces segmentation results that are within 5% of the globally optimal result, as compared to more than 300,000 steps that would be needed by an exhaustive search. Several examples representing progressively more complex search spaces are provided to demonstrate the power and broad applicability of this method. The parameter sets that produce optimal segmentation across different biological objects and imaging modalities, using different segmentation algorithms are presented to demonstrate the generality of this approach. Instead of manually hand-tuning empirical parameters that still inevitably contain human subjectivity, users can now objectively obtain optimal segmentation result automatically, across a broad domain of applications.

Technical Approach

The variational formulation for image segmentation is expressed mathematically using the Minimum Description Length (MDL) formulation [1]. This is analogous to the problem of fitting a polynomial to a set of points, where the goal is to find a trade-off between the degree of the polynomial (conciseness) and the corresponding fitting error (coverage).

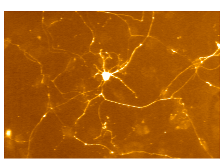
Let an image segmentation algorithm be a function $f:\{I,\xi\}\rightarrow M$ that maps the object(s) imaged in the image I to the segmented object(s) M using a parameter vector $\xi\in\Omega$ where Ω is the space containing all possible parameter vectors. Now, the goal becomes to obtain the optimal M, called $\hat{M}$, in terms of conciseness and coverage.

$$L_{2-p}(I\,|\,M)=\underbrace{-\log_2 P(I\,|\,M)}_{\text{coverage}}+\underbrace{\frac{k}{2}\log_2 n}_{\text{conciseness}}$$

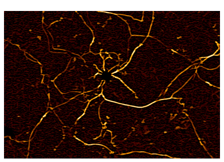
where

$$-\log_2 P(I\,|\,M)\cong m\left(\sum_{\mathbf{x}=\{\mathbf{x}\,|\,\mathbf{x}\in M^C,\,P(\mathbf{x})\geq a\}}-\log_2 P(\mathbf{x})\right)$$

Since the likelihood term is not defined by the MDL principle, we use the probabilistic Vesselness measure [2] for tube-like objects.



Input image



Vesselness measure (scaled for display)

Then, the image is segmented using a particular set of parameters, and the resulting segmented image is evaluated using the MDL-based metric.

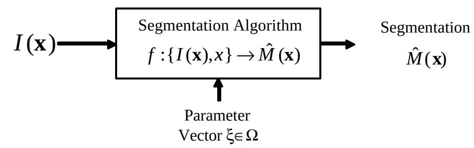
An efficient global optimization algorithm, the Recursive Random Search [3], is used to explore the parameter space using the metric as defined using the MDL principle.

Conclusions

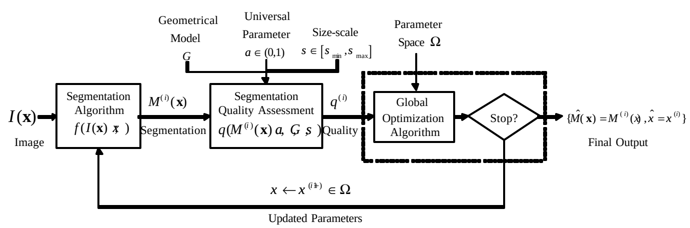
The presented approach removes the human subjectivity factor when selecting parameters/settings for an image analysis system. It also reveals the practical limitations of a particular image analysis system, especially when the system is used to analyze images of kinds that it is not designed for.

The broad applicability of this method is demonstrated by the experimental results. Essentially, all the parameters of an image analysis system are mapped onto user-friendlier terms.

The Traditional Approach to Image Segmentation



The Automated Parameter Selection Approach to Image Segmentation



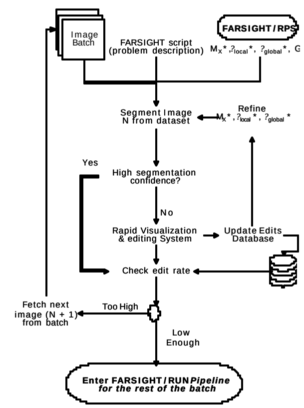
State of the Art

Metrics for segmentation evaluation [4-10] can be either goal-oriented, i.e. evaluation based on the performance of post-segmentation steps such as pattern classification [9], based on other application-defined criteria, such as the probability of false detection [6, 7, 10], or based on visual inspections [4]. Here, we restrict ourselves to the latter category, since we believe that optimality of the segmentation result should be compatible with and independent of any post-segmentation image analysis steps.

The *Minimum Description Length* (MDL) principle [1] is chosen since it offers a systematic way to obtain an objective balance between segmentation conciseness and coverage. We present a different MDL-based formulation, i.e. compared to [11-12], where the formulation now contains a term to account for the probability of structural presence at each pixel. This search of optimal parameters is analogous to the variational formulation for image segmentation [12].

The FARSIGHT* Framework

*Fluorescent Association Rules for multi-dimensional Insight



References

[1] Barron, A., J. Rissanen, and B. Yu, *The minimum description length principle in coding and modeling.* Ieee Transactions on Information Theory, 1998. **44**(6): p. 2743-2760.

[2] Frangi, A.F., et al., *Multiscale Vessel Enhancement Filtering.* Lecture notes in computer science, 1998(1496): p. 130.

[3] Ye, T. and S. Kalyanaraman, *Internet Traffic Engineering - A Recursive Random Search Algorithm for Large-Scale Network Parameter Configuration.* ACM SIGMETRICS Performance evaluation review, 2003. **31**(1): p. 10.

[4] Heath, M.D., et al., *A robust visual method for assessing the relative performance of edge-detection algorithms.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997. **19**(12): p. 1338-1359.

[5] Heath, M., et al., *Comparison of Edge Detectors: A Methodology and Initial Study.* Computer Vision and Image Understanding, 1998. **69**(1): p. 38-54.

[6] Bowyer, K., C. Kranenburg, and S. Dougherty, *Edge Detector Evaluation Using Empirical ROC Curves.* Computer Vision and Image Understanding, 2001. **84**(1): p. 77-103.

[7] Kanungo, T., et al., *A methodology for quantitative performance evaluation of detection algorithms.* IEEE Transactions on Image Processing, 1995. **4**(12): p. 1667-1674.

[8] Shin, M.C., D.B. Goldgof, and K.W. Bowyer, *Comparison of Edge Detector Performance through Use in an Object Recognition Task.* Computer Vision and Image Understanding, 2001. **84**(1): p. 160-178.

[9] Trier, O.D. and A.K. Jain, *Goal-directed evaluation of binarization methods.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995. **17**(12): p. 1191-1201.

[10] Yitzhaky, Y. and E. Peli, *A method for objective edge detection evaluation and detector parameter selection.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003. **25**(8): p. 1027-1033.

[11] Zhu, S.C. and A. Yuille, *Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996. **18**(9): p. 884-900.

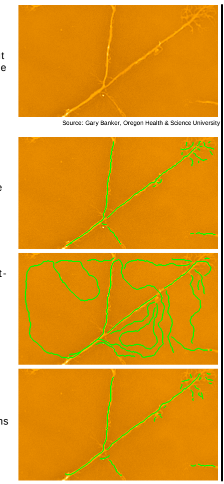
[12] Morel, J.-M. and S. Solimini, *Variational methods in image segmentation : with seven image processing experiments*. Progress in nonlinear differential equations and their applications ; v. 14. 1995, Boston: Birkhèauser. xvi, 245 p.

[13] Al-Kofahi, K.A., et al., *Median-based robust algorithms for tracing neurons from noisy confocal microscope images.* Information Technology in Biomedicine, IEEE Transactions on, 2003. **7**(4): p. 302-317.

[14] Abdul-Karim, M.A., et al., *Automated tracing and change analysis of angiogenic vasculature from in vivo multiphoton confocal image time series.* Microvasc Res, 2003. **66**(2): p. 113-25.

Experimental Results: Automated Tracing Algorithm [13,14]

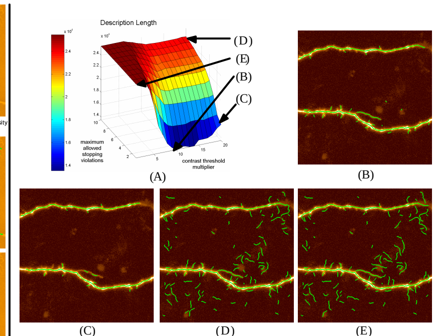
Input Image



Global Optimal: Exhaustive Search, 300000 iterations

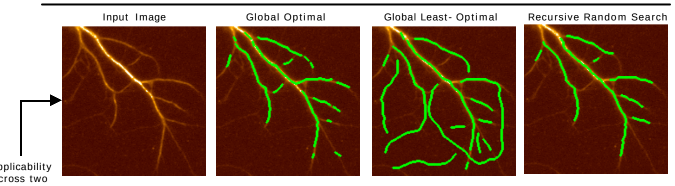
Global Least-Optimal: Exhaustive Search, 300000 iterations

Recursive Random Search: 1000 iterations

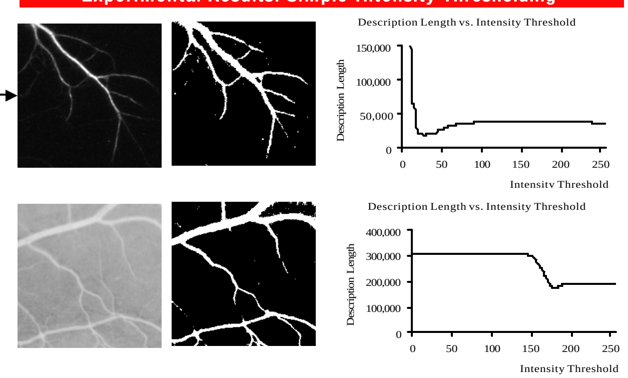


The conciseness and coverage *Description Length* metric, is calculated for traces that result from a full-mesh combination of two tracing parameters, *maximum allowed stopping violations* and *contrast threshold multiplier.* The traces are shown in (B), (C), (D), and (E), and the corresponding metric value is pointed in (A), where (B) is the optimal trace. *(Image courtesy of Andreas Jeromin, Baylor College of Medicine, Texas)*

Applicability across two different applications



Experimental Results: Simple Intensity Thresholding



Illustrates the segmentation result using a global intensity threshold value for tube-like images; upper row displays the result for a neuron image (source: Natalie Dowell, Wadsworth Center, New York), and the lower row displays the result for a retinal blood vessels image (source: The STARE Project). The left-most column contains the input image, the middle column contains the optimal segmentation image, and the right-most column shows plots of the description length (the conciseness and coverage metric) versus the intensity threshold value.