AUTOMATED PARAMETER SELECTION FOR SEGMENTATION OF TUBE-LIKE BIOLOGICAL STRUCTURES USING OPTIMIZATION ALGORITHM AND MDL

by

Muhammad Amri Abdul Karim

A Thesis Submitted to the Graduate
Faculty of Rensselaer Polytechnic Institute
in Partial Fulfillment of the
Requirements for the degree of
DOCTOR OF PHILOSOPHY
Major Subject: Computer and Systems Engineering

Approved by the
Examining Committee:

Badrinath Roysam, Thesis Adviser	Shivkumar Kalyanaraman, Member
Qiang Ji, Member	Alejandra Mercado, Member
William Shain, Member	Deanna Thompson, Member

Rensselaer Polytechnic Institute
Troy, New York

September 2005
(For Graduation December 2005)

Contents

List of Tables

List of Figures

Acknowledgment

I am obliged to dedicate the first non-automatically-generated page to all that makes the subsequent pages possible. Thank God.

Many thanks go to Prof. Badri Roysam for all guidance, mentoring, support, and especially patience throughout my entire career as a graduate student in his lab. Also thank you to the doctoral committee members, Prof. Shiv Kalyanaraman, Prof. Qiang Ji, Prof. Alejandra Mercado, Prof. William Shain, and Prof. Deanna Thompson for suggestions and constructive feedbacks that strengthen this work.

This work builds upon previous doctoral works by Khalid Al-Kofahi and Tao Ye, and software development efforts by MicroBrightField Inc. Thank you to Murat Yuksel for assistance in integrating the recursive random search algorithm. This work also would not be motivating without real-world applications and contributed image data by collaborators from The Wadsworth Center (William Shain, James Turner, Natalie Dowell-Mesfin, Chris Bjornsson, and Karen Smith), Baylor College of Medicine (Andreas Jeromin), David Geffen School of Medicine at UCLA (Joshua Trachtenberg), Harvard Medical School (Rakesh Jain and Edward Brown), Rensselaer Polytechnic Institute (Deanna Thompson), and Oregon Health and Science University (Gary Banker). Also thank you to Michael Goldbaum *et al.* and Meindert Niemeijer *et al.* for their publicly-available image and ground truth segmentation databases.

Funding for various parts of this work come from MARA of Malaysia, MicroBrightField Inc., the Center for Subsurface Sensing and Imaging Systems, Rensselaer Polytechnic Institute, and various NSF and NIH grants.

I would also like to thank Gang Lin, Omar Al-Kofahi, Jeff Sprenger, Jack Glaser, Alex Tyrrell, and Dirk Padfield for related discussions. Thank you to members of the lab, Nicolas Roussel, Hussein Sharafeddin, Yousef Al-Kofahi, Hari Iyer, Justin LaPre, Ying Chen, Andy Cohen, and Mathieu Malaterre for numerous help and assistance. Thank you to Laraine Michaelides and ECSE Dept. staff members for all assistance.

At the personal level, this work is dedicated to my parents A. Karim A. Aziz and Zaiton Shariff, and my wife Dr. Siti Mariah Hussain.

May peace be upon you.

Amri

Abstract

Segmentation of tube-like biological structures, e.g., neurons and blood vessels, allows extraction of structural measurements supporting quantitative studies ranging from understanding neural growth to cancer research. While automated segmentation algorithms minimize subjectivity associated with tedious manual segmentation, they generally have parameter settings to cope with high variability in image data across applications. Currently, these settings are chosen empirically, formulated heuristically, or by trial-and-error with no assurance towards optimality. This work is motivated by the need to automatically select parameter settings for segmentation algorithms since they directly affect segmentation accuracy.

An objective trade-off between a probabilistic measure of image-content coverage of a segmentation and its conciseness is based on the minimum description length principle (MDL). The recursive random search (RRS) optimization algorithm is used to efficiently explore combinations of segmentation algorithm parameter settings. For 3-D images, computation time is reduced by coordinated-optimizations on non-empty, representative subimages based on intensity and structural information.

The method is initially applied to 223 2-D images of human retinal vasculature and cultured neurons, from four different sources, using a single segmentation algorithm with 8 parameters. Relative to default settings, improvements in the proposed MDL-based segmentation quality metric are strongly correlated with improvements in agreement with ground truth ($r = 0.78$), ranging between 4.7 – 21% using 1000 function evaluations. Paired *t*-tests showed that improvements are statistically significant ($p < 0.0005$). Most of the improvement occurred in the first 44 function evaluations. For 223 images, RRS outperforms other optimization algorithms (controlled random search, multi-start pattern search, simulated annealing, and genetic algorithm) on average at all 1000 function evaluations. RRS with 1000 function evaluations is on average within 3.56% of the global optimum (6,804,000 function evaluations). The proposed coordinated subimage-optimization method results in average speedup of 11.2X for 22 3-D neurons images.

This enables non-expert users to use segmentation algorithms without knowledge of underlying algorithms, increases objectivity, and broadens applicability of the segmentation algorithm. It simplifies the user interface to just one optional parameter, creating a consistent interface that allows developers to freely modify the algorithms. The method allows adaptation of parameters across batches of images and delivers higher morphometric accuracy.

1. Introduction

A solution to the problem of automatically selecting segmentation algorithm parameters/settings for tube-like biological structures such as neurons and vasculature is presented using an optimization algorithm and the minimum description length (MDL) principle to construct the objective function.

The primary goal is to enable a non-expert user to select parameters effectively, and objectively, treating the segmentation software as a “black box.” The secondary goal is to enable the algorithm developer to modify the internal details while maintaining a consistent and simple external interface; and to minimize the cost of technical support.

1.1 Motivation

Segmentation partitions an image into its constituent elements or objects [1]. Segmentation is performed to obtain morphometrics, or structural measurements, of structures captured in the images. It is one of the main components of image analysis systems used in image-based biological studies. An image analysis system accepts images as input and produces morphometrics. It may be fully-automated or one that require user interactions. In this work, image analysis systems are limited in scope to those that segment the structures and report morphometrics based on the segmented structures (e.g., [2-6]).

The reported morphometrics allow image-based studies to answer biological questions in a quantitative manner. For example, Figure 1 displays an image from a study that investigates the guidance of neurite outgrowth by non-neural cells [7]. This is possible because images contain spatial information, an advantage absent in other quantitative study methods such as the rapid flow cytometry (up to 10,000 cells/sec [8, 9]) that requires the specimen tissue to be broken up. Furthermore, image-based studies allow locations of specific responses or activities or structures of interest to be pinpointed rather than quantified for an entire cell or a population of cells.

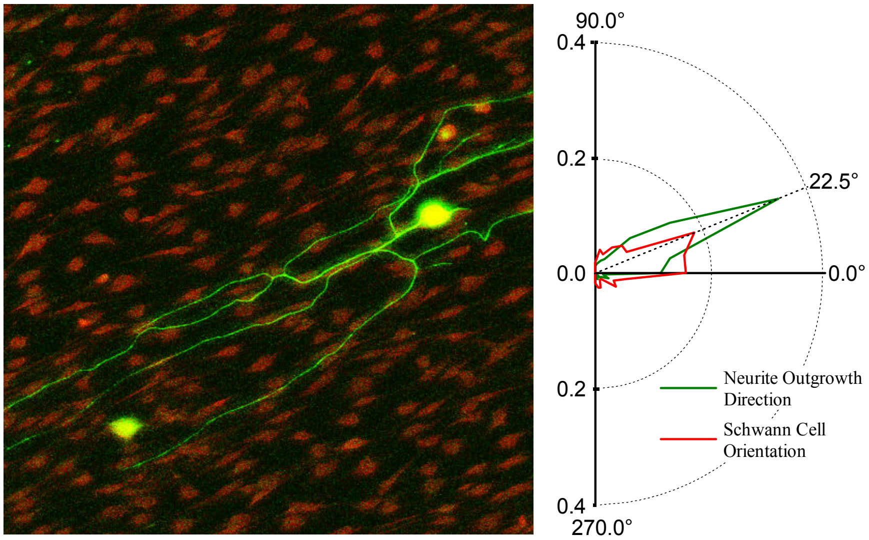
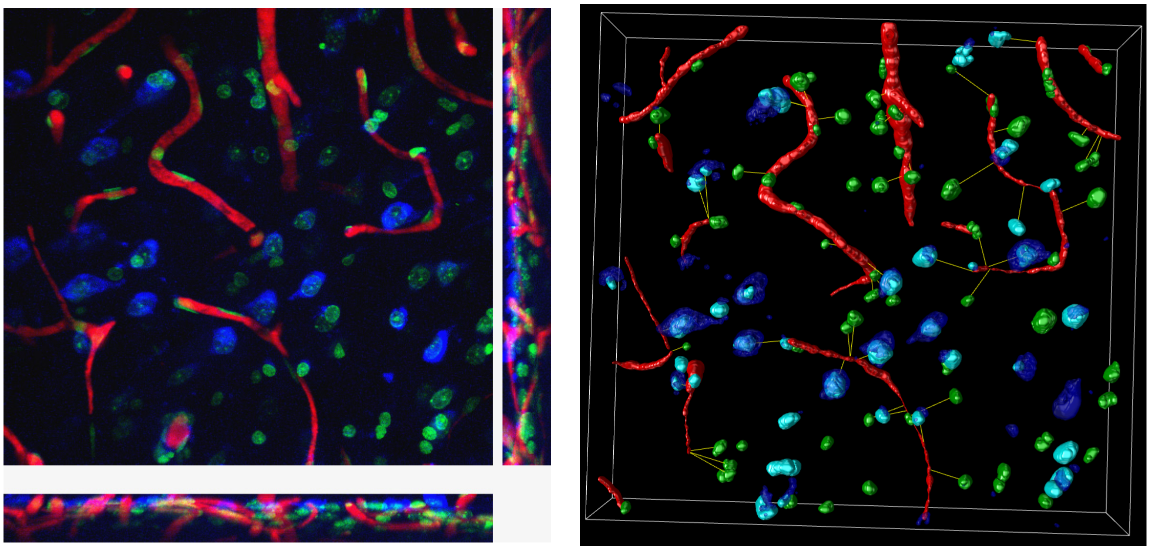


Figure 1. Illustrates the effect of non-neuronal cells (Schwann cells) in guiding neural outgrowth (Image source: D. Thompson, M. Cross, RPI). Schwann cells shown in red are grown on microlithographically-patterned substrates. The angular histogram (right) displays the extracted orientation of the Schwann cells and the outgrowth direction of the neurites shown in green (cell analysis: G. Lin).



(a) (b)

Figure 2. Image-based biological studies allow spatial structures to be investigated in association with each other. (a) A 512×480×51 3-channel image of the neurovascular unit containing vasculature (red), cell nuclei (green) and cytoplasm (blue) shown in *x-y*, *x-z* and *y-z* maximum intensity projections. (Image source: C. Bjornsson, K.L. Smith, W. Shain, J. Turner, Wadsworth Center). (b) A 3-D rendering of the segmented data, showing neuronal (cyan) and non-neuronal (green) nuclei after associating the segmented structures in all 3 channels.

In addition, imaged-based biological studies allow spatial structures to be investigated in association with each other, i.e., investigation of functional and structural relations among elements of a complex system [10]. For example, Figure 2a displays a 512×480×51 3-channel image of the neurovascular unit containing vasculature (red), cell nuclei (green) and cytoplasm (blue). Panel (b) displays a 3-D rendering of the segmented data, showing neuronal (cyan) and non-neuronal (green) nuclei after associating the segmented structures in all 3 channels [10].

For this work, the structures of interest are neurons and vasculature. Neurons and vasculature represents a broad range of quantitative studies in biology and medicine. As an example for neurons, the morphometric study of directing and promoting neuronal growth is important to understand the requirements to establish defined neural networks [11]. For vasculature, the study of tumor vasculature morphometrics reveals insights on controlling tumor growth and eventually shutting down the vasculature to starve the cancer cells [12].

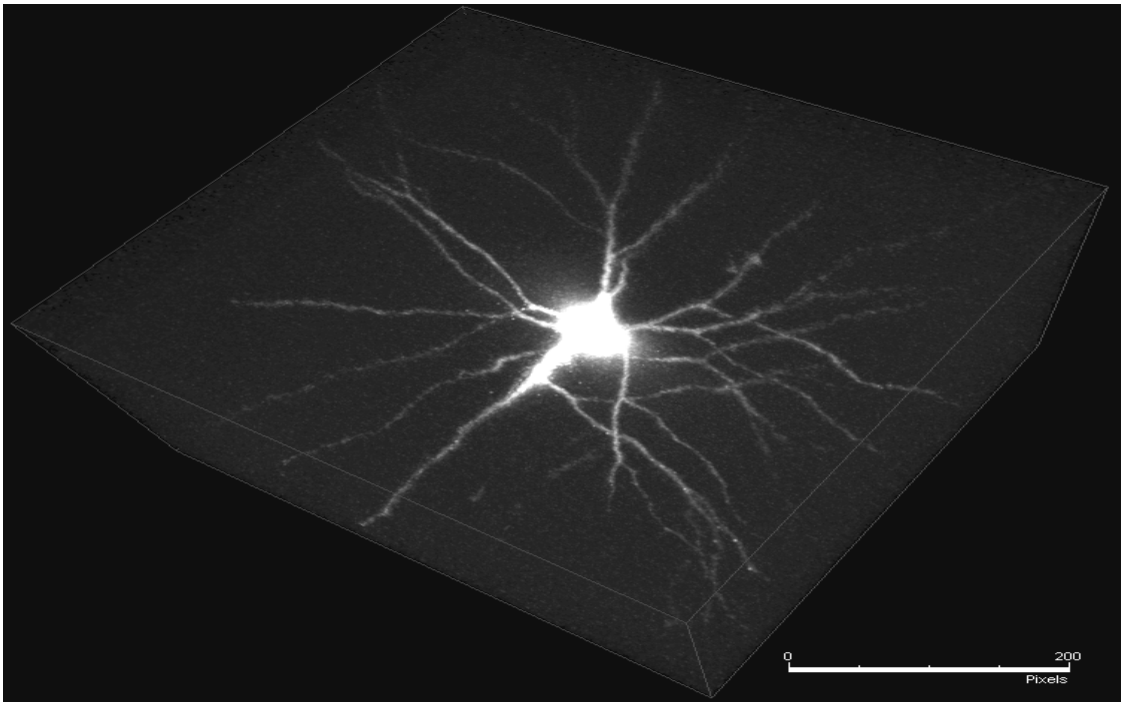


Figure 3. Displays a 3-D 512×480×301 image containing a dye-injected neuron in a thick brain slice. Over a short distance, the dendritic and axonal segments resemble tubes. (Image source: S. Lasek, D. Szarowski, J. Turner, Wadsworth Center).

Neurons and vasculature are categorized as tube-like structures because they can be locally[1] approximated as a tube[2] throughout the structure [13], with exception of the soma (cell body) of a neuron (Figure 3). In terms of representing these structures, they can be represented as centerlines[3] containing width information [13], or simply called as traces [14].

Segmentation of neurons and vasculature can be done manually by a human observer (e.g., [15-18]). Manual segmentation, not limited to neurons and vasculature, is tedious and subjective. In addition to inter-observer disagreements, the same observer may produce different results at different times, referred to as intra-observer variability. On the other hand, automated segmentation algorithms produces fast and objective results [3, 6]. Within the previously-defined scope of image analysis systems in this work, automated segmentation algorithms are the critical components of automated image analysis systems since they directly affect the accuracy of resulting morphometrics.

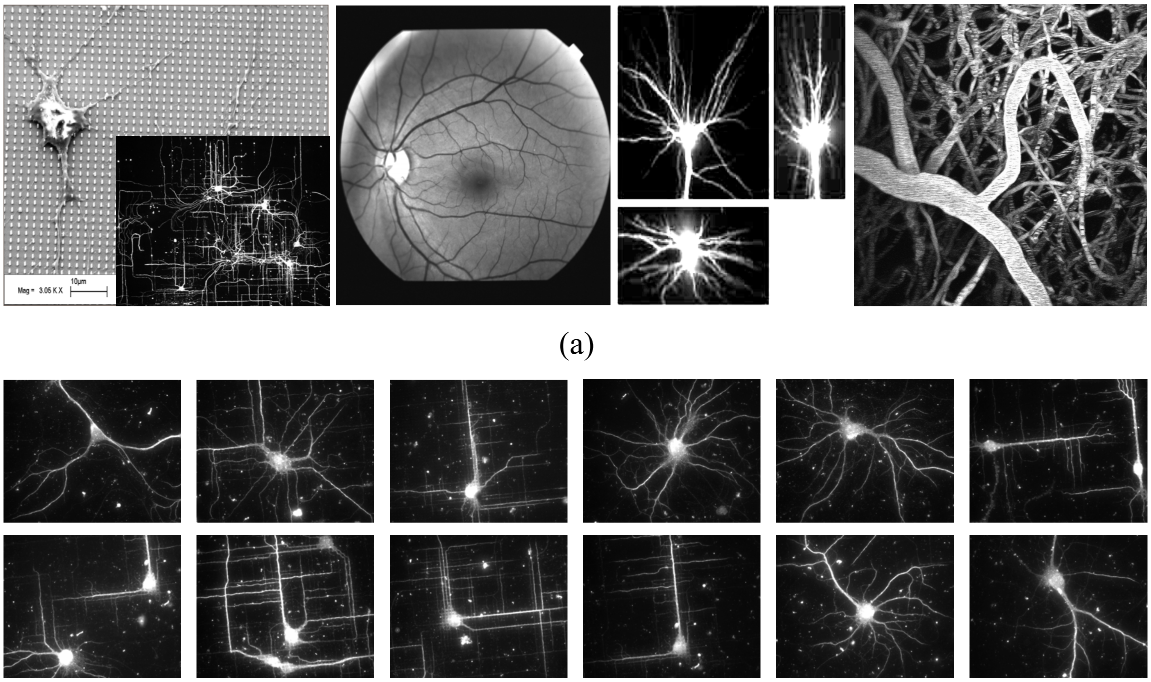
Despite the advantages offered, one practical barrier to more widespread adoption of automated image analysis systems in quantitative biomedicine is the need to adapt/customize them to cope with biological variability (illustrated in Figure 4). To achieve this, algorithm designers are forced to incorporate user-settable parameters. As a result, users are faced with the difficulty of selecting these parameters without sufficient knowledge of the internal mechanisms. Time-consuming manual trial-and-error, as well as extensive developer support is often necessary to properly configure the software for a given application. Even then, these settings are subjective, and there is no assurance of optimality. Currently, settings for these algorithms are chosen empirically (e.g., [2, 3, 19, 20]) or formulated heuristically (e.g., [21]).

---

[1] The term “local” is defined as “within a short distance”, e.g., 2-3 pixels.

[2] The term “generalized cylinder” used by Al-Kofahi *et al.* may be more appropriate for structures that vary in width such as neuronal dendrites and vasculature.

[3] Also called as “medial axes” in the literature.



(a)

(b)

Figure 4. Inter- and intra-application variability that exists in neuron/vasculature images, in panel (a) and (b) respectively. From left, panel (a) displays images of: (i) neurons grown on topographically-modified semiconductor surface; (ii) human retinal vasculature; (iii) fluorescently-labeled neuron in 3-D; (iv) brain vasculature. Panel (b) displays 12 neuron images captured within the same study.

Accuracy of extracted measurements may also be affected by different algorithm settings. For example, Figure 5a shows a phase contrast image of cultured neurites grown on an imprinted surface with known orientations of 45° and 90°. Panel (b) displays the automatically-generated traces using default settings. Panel (c) displays the traces obtained with automatically selected parameters using the method presented in this dissertation. The normalized angular histogram of measured segment orientations extracted from the automatically-generated traces is displayed in panel (d). Note the correct peak at 45° obtained using automatically-selected settings vs. 34° using default settings.

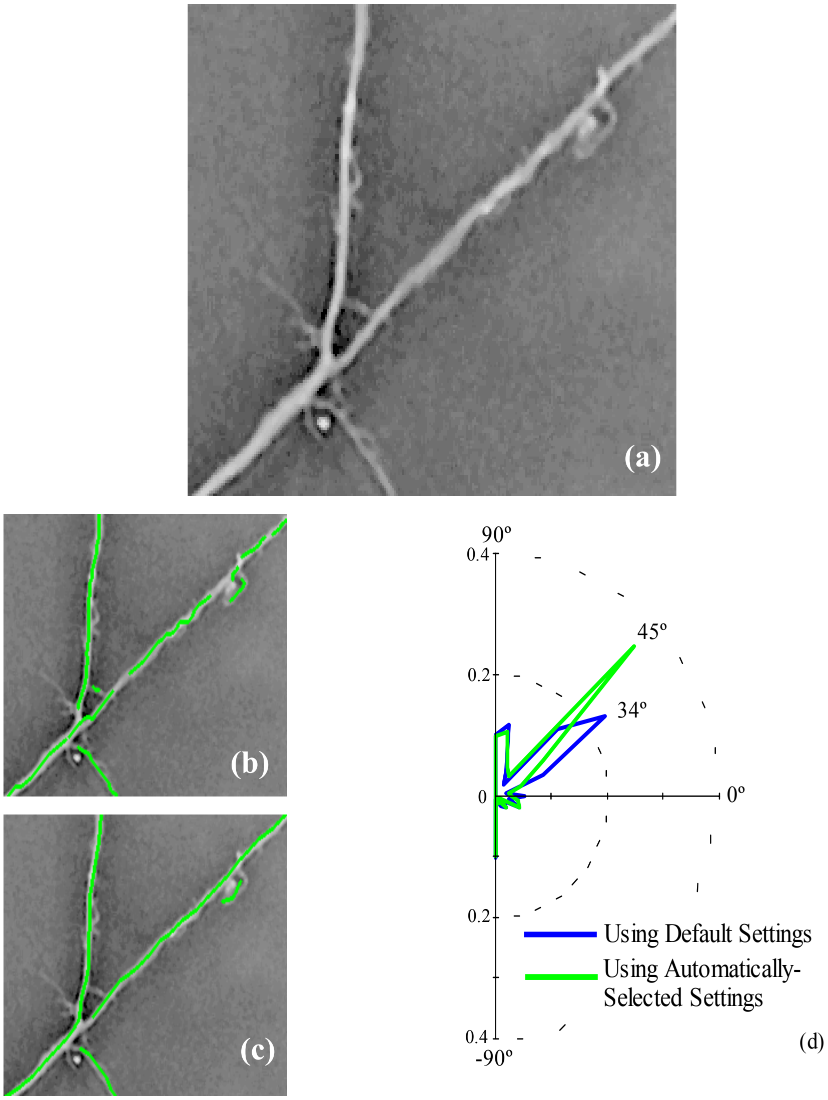


Figure 5. Improved segmentation settings lead to more accurate measurements. (a) A phase contrast image of cultured neurites grown on an imprinted surface with known orientations of 45° and 90°. (b) Automatically-generated traces using default settings. (c) Traces obtained with automatically selected parameters using the method presented in this dissertation. (d) The normalized angular histogram of measured segment orientations extracted from the automatically-generated traces. Note the correct peak at 45° obtained using automatically-selected settings vs. 34° using default settings. (Image source: G. Banker, OHSU).

1.2 Groundwork

To date, this group has designed and implemented a segmentation algorithm to trace tube-like structures, with a wide range of successful application areas including neurobiology [3, 19], ophthalmology [14], and angiogenic tumor vasculature studies [2]. In principle, this algorithm is broadly applicable to the segmentation of tube-like structures. In practice, the segmentation performance depends upon the characteristics of the image data. For instance, different levels of noise, structural irregularities, and imaging artifacts are encountered in different applications.

To cope with the challenges posed by a novel application, the simplest approach to generalize the algorithm is to vary certain constant values such as tracing step size and sizes of correlation kernels [19]. This simple-minded approach was often adequate when performed by an experienced algorithm designer, who had strong intuitions as well as specific knowledge of the segmentation algorithm. However, it proved unacceptable in the hands of an end user, especially a user from a different discipline who did not have knowledge of the internal workings of the segmentation algorithm. A practical trade-off to assist the user is to provide “default” settings that are known to produce acceptable results over a broad range of applications. Even these are quite limiting. They are, in essence, merely a non-systematic accumulation of the empirical temporary solutions for previously-studied applications.

A somewhat more expensive approach to adapt an algorithm to a novel application is to modify the algorithm itself. This proved necessary in some cases. For instance, the tracing algorithm of Can *et al.* [14] was extended by Al-Kofahi *et al.* [19] to incorporate variable-length correlation kernels in order to cope with the discontinuities encountered in images of neurons grown on topographically-modified surfaces [11]. Subsequently, the 3-D tracing algorithm of Al-Kofahi *et al.* [3] was extended by Abdul-Karim *et al.* [2] to incorporate a median-based kernel instead of the standard correlation kernel (e.g., as in [3]) in order to cope with the irregularity of tumor microvasculature. While such extensions are inevitable and desirable, they pose a practical difficulty to the non-specialist user. The newer algorithms are more complex and introduction of additional parameters to handle diverse cases is unavoidable. The problem of selecting these parameters has now become even more challenging to the user.

Interestingly, the end user is not the only affected party. This occurs even when the number of parameters goes down due to advancements in the algorithm design. The designer of graphical user-interfaces for the segmentation software is faced with the challenge of encapsulating and hiding the algorithm details from the external user interface. This is harder to do with changing number and type of parameters required by the algorithm.

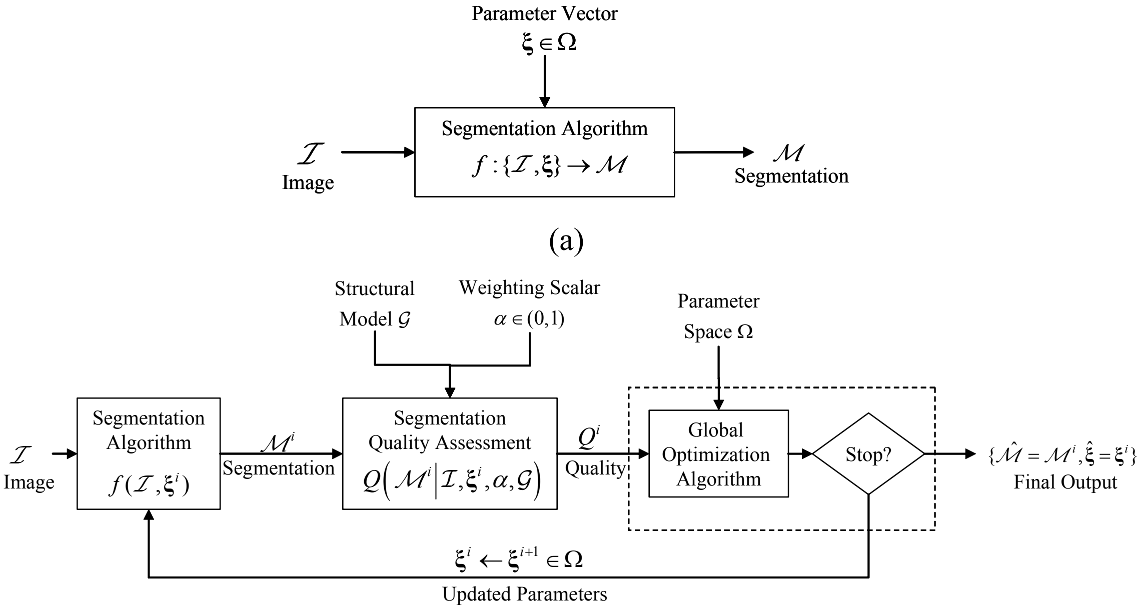
With the above considerations in mind, it is helpful to identify key application-specific aspects of the segmentation algorithm. In this work, four such elements are identified: (i) mathematical model describing the local anatomic structure of the objects of interest $\mathcal{G}$; (ii) mathematical model describing the local intensity profile of the objects of interest $\mathcal{T}$; (iii) parameters related to the fitting of the local model (items 1 and 2 above) to the image data; (iv) parameters related to global (image-wide) model fitting.

Models, in the context of this work, are assumptions in precise mathematical forms specific to the biological structure to be segmented. Locally, models are associated with the structural/geometrical/anatomical characteristics of the biological object and/or based on observable image features such as intensity profile and texture. Then the local models are fitted to the image using strategies such as matched-filtering, generalized likelihood-ratio test, regression analysis, or by formulations of goodness-of-fit measures between the models and the image content. Then the structures are segmented from the image by a global fitting strategy, which can be a pixel-wise operation [20, 22, 23], exploratory [2, 3, 13, 14, 21, 24], or by minimum-cost path estimations [25-27]. These concepts will be further clarified with respect to the current literature in Chapter 2.

1.3 Systematic Approach to Select Parameters

Borrowing the feedback-loop concept from signal processing, a segmentation-optimization framework is proposed, coupling a segmentation algorithm with an optimization algorithm. To guide the optimization algorithm, the framework incorporates an automatic evaluation/assessment of the quality of a segmentation resulting from a given combination of the algorithm parameter settings. Figure 6 illustrates the traditional approach to segmentation of biological images in contrast to this framework. In the latter, the automated segmentation quality assessment drives the parameter selection process. Ideally, the optimal segmentation algorithm parameters—a vector $\hat{\xi} \in \Omega$ in the space of all possible algorithm parameter combinations—

corresponds to an optimal segmentation and is the one that we look for. In practice, the search for the optimal parameter settings $\hat{\xi}$ may be limited by certain computational budget. In either case, the notion of segmentation optimality is made precise in the next section. In other words, given 2 segmentations, it tells how to quantitatively conclude one is better than the other.



(b)

Figure 6. Illustrates the proposed self-optimizing image segmentation approach. Panel (a) shows a traditional segmentation algorithm in which the parameter settings ξ are set empirically for each input image $\mathcal{I}$. Panel (b) illustrates the proposed method in which a global optimization algorithm efficiently explores the parameter space Ω driven by a segmentation quality assessment value based on trading off conciseness of the segmentation versus its image-content coverage. The user optionally specifies a single universal parameter α to override the trade-off.

1.4 Segmentation Optimality

We are in effect deciding which segmentation $\mathcal{M}_i$ to be optimal among the set of all *computable* segmentations $\{\mathcal{M}_i\}$—obtained by varying the algorithm parameters. The traditional *maximum-a-posteriori* (MAP) criterion that minimizes the probability of error [28] can be opted if prior probabilities on the segmentations can be specified. In this work, error is defined as the disagreement between a segmentation and the ground truth[4] at each image pixel. The MAP criterion is to choose the segmentation $\hat{\mathcal{M}}_i$ that maximizes the posterior probability $P(\mathcal{M}_i|\mathcal{I})$:

$$P(\mathcal{M}_i|\mathcal{I}) = \frac{P(\mathcal{I}|\mathcal{M}_i)P(\mathcal{M}_i)}{P(\mathcal{I})}, \tag{1}$$

such that:

$$\hat{\mathcal{M}}_i = \arg\max_{\mathcal{M}_i} P(\mathcal{I}|\mathcal{M}_i)P(\mathcal{M}_i), \tag{2}$$

for the input image $\mathcal{I}$, dropping the constant $P(\mathcal{I})$ term. Unfortunately, the prior probability $P(\mathcal{M}_i)$ is not well defined [29]. Furthermore, if a uniform probability distribution for the prior probability $P(\mathcal{M}_i)$ is assumed, then the MAP criterion reduces to the *maximum-likelihood* (ML) criterion that measures only the goodness-of-fit and has a tendency to overfit models to the data.

It turns out that the MAP criterion can still be used if optimal descriptive languages for a segmentation and the image given the segmentation are specified. Then the minimum description length (MDL) [30, 31] criterion can be adopted to choose the optimal segmentation. In other words, the MDL principle is chosen in this work because it has been shown to be equivalent to the MAP method that minimizes the probability of error [29].

MDL is a principle for statistical model selection and statistical inference based on the simple idea that the best way to capture regular features in data is to construct a model in a certain class which permits the shortest description of the data and the model itself [32]. It guides the trade-off between model overfit and model simplicity in an objective manner. For the purpose of this work, it guides the trade-off between segmentation coverage of image structural content and conciseness of the segmentation itself. As an illustrative example for fitting a polynomial to a set of n points (reproduced

---

[4] To avoid introducing further subjectivity in this work, ground truth is obtained from public-domain databases of manual segmentations and by consensus of manual observers when provided by these databases.

from [33]), Figure 7 displays fit with a line (1^{st}-degree polynomial), a complex fit with $(n-1)^{\text{th}}$ degree polynomial[5], and a trade-off fit with a 3^{rd} degree polynomial. Figure 8 illustrates an analogous example to image segmentation. Given the amount of structure shown in panel (a), panel (b) displays a complex segmentation, and panel (c) displays the trade-off segmentation. Both segmentations were obtained using different parameter settings for a segmentation algorithm [19].

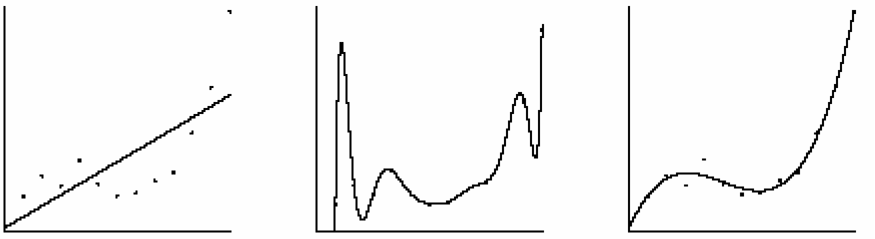
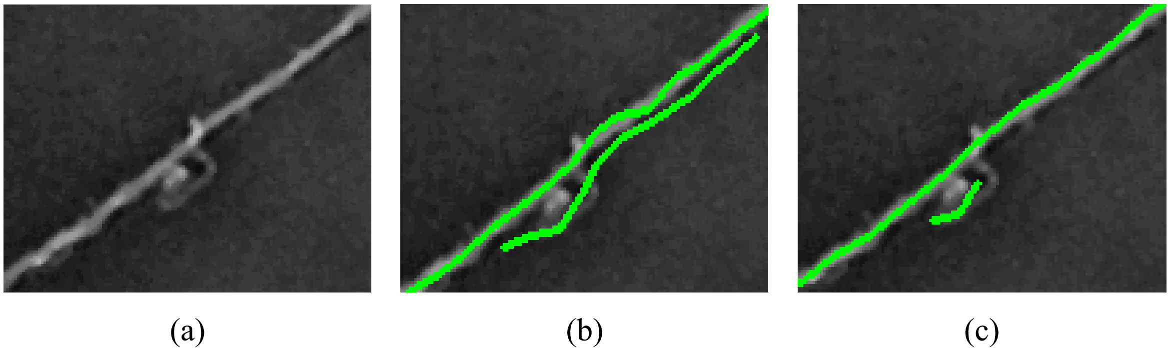


Figure 7. Demonstrates the MDL principle in guiding model selection, with an example to polynomial fitting to a set of n points. The plots show a simple, a complex, and a trade-off (3rd degree) polynomial. (Figure reproduced from [33]).



(a) (b) (c)

Figure 8. Given an image displayed in panel (a), panel (b) displays a complex segmentation, and panel (c) displays a trade-off segmentation.

Once the descriptive languages are specified, the MDL criterion is equivalent to the MAP criterion. This was first presented in the context of image partitioning by Leclerc

---

[5] For every n points, there exist a polynomial of the $(n-1)^{\text{th}}$ degree that goes through all n points.

[29]. By taking the negative log of the probabilities in (1) the MDL criterion becomes to choose the segmentation $\hat{\mathcal{M}}_i$ that minimizes:

$$-\log_2 P\left(\mathcal{M}_i \middle| \mathcal{I}\right) = -\log_2 P\left(\mathcal{I} \middle| \mathcal{M}_i\right) - \log_2 P\left(\mathcal{M}_i\right). \tag{3}$$

Now, let $|\cdot|$ to be the description length in bits. Given an optimal descriptive language $\mathcal{L}_m$ to describe the segmentation $\mathcal{M}_i$, we need

$$\left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right| = -\log_2 P\left(\mathcal{M}_i\right) \text{ bits} \tag{4}$$

to describe the segmentation $\mathcal{M}_i$ [29]. Likewise, given the optimal descriptive language $\mathcal{L}_d$ to describe the image $\mathcal{I}$ given the segmentation $\mathcal{M}_i$, the description length becomes:

$$\left|\mathcal{L}_d\left(\mathcal{I} \middle| \mathcal{M}_i\right)\right| = -\log_2 P\left(\mathcal{I} \middle| \mathcal{M}_i\right) \text{bits.} \tag{5}$$

Then we can see that the MDL criterion is equivalent to the MAP criterion, which is to choose the segmentation $\hat{\mathcal{M}}_i$ such that

$$\hat{\mathcal{M}}_i = \underset{\mathcal{M}_i}{\arg\min} \left|\mathcal{L}_d\left(\mathcal{I} \middle| \mathcal{M}_i\right)\right| + \left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right|. \tag{6}$$

In Chapter 3 we will describe the methods to exactly compute the description length. Experimental results for segmentations that minimize the description length are presented in Chapter 4.

1.5 Summary of Contributions

The primary contribution of this work is to the field of automated segmentation of biological imagery—specifically of tube-like structures such as neurons and vasculature. The related literature is summarized in Chapter 2. The detailed contributions of this work are summarized below along with brief mentions of the state of the art:

1. **Automated parameter selection for segmentation of tube-like biological structures.**
    a. **Minimizes subjectivity associated with hand-tunings** of parameter settings for automated image segmentation algorithms for tube-like structures. Currently, parameters for these algorithms are chosen empirically (e.g., [2, 3, 14, 19, 20, 27]) or

formulated heuristically (e.g., [21, 34]). Unlike related works in computer vision applications by Min *et al.* [35] for range image segmentation and the closed-loop reinforcement learning framework by Peng and Bhanu [36], the proposed parameter selection method is fully unsupervised.

b. The segmentation computed with the automatically-obtained optimal algorithm parameters **reveals limitations of the segmentation algorithm**. This can be used as an objective measure when choosing a segmentation algorithm to use for the specific application at hand. Currently, segmentation accuracy reports in the literature were of segmentations computed with empirical parameters, with no notion of optimality in terms of the parameters and the resulting segmentation (e.g., [2-4, 11, 13, 14, 19, 20]).

c. Segmentation algorithms in their optimal configurations can be used in a high-throughput **batch execution system** for large amount of similar images—acquired within the same study and may be of different specimens. To date, other than for time-series images [4, 37], batch segmentation for images containing tube-like biological structures has not been explicitly presented in the surveyed literature (Chapter 2).

2. **Formulation of a segmentation quality metric, or optimality measure, associated with each segmentation originating from segmentation algorithm parameter settings.** Specifically, the MDL-optimality criterion [29] is specialized for biological image segmentation using a local structure-indicator function that measures the goodness-of-fit of local models and effectively captures the pixel correlation under the modeling assumptions. Currently, segmentation quality metric and optimality measures has been devised for low-level computer vision tasks such as edge-detection [38-40] and image-partitioning [29, 41-44]. Unlike the published methods for biomedical images (e.g., [44-46]), the presented MDL-optimality criterion for tube-like biological structures does not require manual (or estimated [46]) ground truth segmentation.
3. **Systematic speedup method using image subvolumes.** A method to minimize computations is introduced by performing parameter selection on a **representative** subvolume, or a set of subvolumes containing a representative subvolume in a

coordinated manner. This is inspired by related work by Shen *et al.* [47] and Lin *et al*. [48] that prioritizes segmentation to achieve real-time performance for subsequent tasks. The proposed measure for being representative is based on both intensity and structural information. Representative measures presented in the literature for content-based image retrieval systems (e.g., [49-52]) are mainly based only on image intensity such as gradient and texture (e.g. [52]) with future research heading towards using structural information [51].

4. **Systematic reduction of the many parameters of image segmentation algorithms into one parameter that trades off segmentation coverage and segmentation conciseness.** This minimizes the guesswork on the behalf of the non-expert user in choosing segmentation algorithm parameters. In practical terms, all segmentation algorithm parameters are mapped to a single parameter that adjusts the "aggressiveness" of the algorithm, coherent to the users' intuitive sense. This creates a consistent user interface that allows algorithm designers to freely modify the underlying algorithms. This has not been reported or proposed in the surveyed literature (Chapter 2) for tube segmentation algorithms.
5. **Breaking down an image segmentation algorithm into modular components:** (i) local structure model; (ii) local intensity model; (iii) local model fitting; and (iv) global model fitting. This allows these modular components to be interchanged for specific applications at hand. Modularity enables algorithm designers to modify and improve each component independently. To date, segmentation algorithms reported in the literature (summarized in Chapter 2) are often application-specific with no explicit proposition for modularity in terms of these components.

All of these contributions are incorporated within a modular segmentation-optimization framework to ensure wide applicability with current image analysis systems and also current optimization algorithms. For applications of interest, modularity ensures that existing algorithm implementations can be interchanged and interfaced (Figure 6b) to fulfill their specific requirements such as computation time constraints and robustness to imaging artifacts. This is the key idea that will further generalize an image analysis system to a wider range of applications.

To summarize, this work does not propose a new image analysis system or new image segmentation algorithms within existing systems, but a *systematic* way to obtain the optimal result from such existing systems. The resulting segmentations are not only visually better than using default settings, we actually know for sure that they are as good as they can get.

2. Summary of Related Literature

This work draws upon four main bodies of literature:

1. Automated segmentation algorithms for tube-like/curvilinear biological structures.
2. Automated/quantitative/objective evaluation of segmentation results.
3. The minimum description length (MDL) principle.
4. Global optimization.

The primary contribution of this work is to the field of automated segmentation of biological imagery - specifically of tube-like structures such as neurons and vasculature. Therefore, this field is given great emphasis in this chapter. As the surveyed literature is vast, usage of synonyms are unavoidable, and are stated in the text separated by forward slashes. These synonyms also may be interchanged throughout the text.

2.1 Segmentation of Tube-like Structures

Segmentation of tube-like/curvilinear structures in biological images is one specific use of computer vision algorithms. Tube-like structures are defined as those of which medial-axes/centerline/center-points can be locally approximated as a line [13]. These structures include neurons, vasculature, pulmonary airways (e.g., [53]), and even certain markers on endografts (e.g., [54]). Computer vision techniques and concepts applicable to natural scene imagery are adapted to segment tube-like structures. They include edge detection [55], image partitioning, region growing [56, 57], model fitting, surface estimation, intensity ridge/valley [58, 59], and energy minimization techniques [42]. Segmentation algorithms may use the computer vision techniques in their general form or include further constraints specific to tube-like structures—perhaps the imaging modality/condition/environment as well. Currently, settings for these algorithms are chosen empirically (e.g., [2, 3, 19, 20]) or formulated heuristically (e.g., [21]).

In a broad sense, tube-like segmentation algorithms can be classified based on the modeling assumptions made about the presence or absence of tubes at each pixel in the image, and how these models are fitted locally and fitted globally to the entire image content. Locally, the assumptions are based on a *structure model*, and/or an *intensity model* (referred collectively as a "*local model*"). For example, the simple intensity thresholding algorithms (e.g., [60]) and region growing algorithms (e.g., [56]) can be thought of as fitting an intensity model at each pixel, but both usually have no explicit structure model. Using the local models, tube-like structures are segmented/extracted/delineated from an image using various global/image-wide model-fitting strategies in varying complexity from simple pixel-by-pixel operation to iterative exploratory techniques.

At this point, the distinction between local models and the global fitting strategies should be further distinguished. Two algorithms using the same local model may differ in the global fitting strategy. For example, ridge-based [58] local models, described later, can be applied either to all pixels (e.g., [61]), or iteratively in an exploratory manner (e.g., [13]). Similarly, two algorithms using the same exploratory strategy, such as by using the exploration path estimate from local model fittings, may differ in the underlying local models (e.g., the generalized cylinder structure model in [3] vs. the toroidal structure model in [27]).

2.1.1 Local Model: Structure Model and Intensity Model

For local models, the distinction between structure models and intensity models also needs further clarification. Both of these models are often tightly coupled (e.g., a Gaussian intensity model within a structure model that of a cylinder [13]), although either one may not be explicitly specified/used. They are dichotomized here solely for the purpose of classifying segmentation algorithms for tube-like structures in the literature.

Structure models correspond to the geometrical shape of the tube-like structure. Examples of structure models are parallel-edges in 2-D images (e.g., [14, 22]), generalized cylinder in 3-D images (e.g., [2, 3, 13]), the 3-D toroid model (e.g., [27]), and the *n*-D intensity ridge model (e.g., [59]). Weaker/more-relaxed models include deformable models [62].

Intensity models, on the other hand, correspond to intensity variation within and around the structure. The simplest form of intensity model is the intensity threshold [23, 60] used in intensity threshold-based segmentation algorithms. The intensity model used in region growing methods employs intensity similarity, rather than an explicit intensity threshold, accompanied by spatial proximity criteria [63]. More restrictive intensity models/profiles with structural/spatial variations include the Gaussian profile (e.g., [64, 65]) and the Gaussian-convolved-pulse profile [13]. Other than spatial assumptions, one may use temporal assumptions as appropriate, e.g., when vessel structures are spatially stationary over time and intensity model based on the temporal standard deviation is suffice to segment the structures [37]. Regardless, these intensity models may go further to incorporate imaging noise models [22].

2.1.2 Local Model Fitting

Given the local models, there are various strategies to fit them locally to the image data, called "*local fitting*" strategies. They include matched filtering (e.g., [64]), generalized likelihood ratio test (e.g., [22]), and formulations of measures that calculate the goodness-of-fit of the models to the image content (e.g., [2, 3, 13, 24]). Also, the way models are fitted differ when: (i) only structure boundary information is used (e.g., [3]), or when (ii) the entire local region is used (e.g., [21]). The latter approach is said to be less sensitive to image noise since the local model-fitting step is performed by integrating over a larger extent of the tube-like object rather than using small-scale measures as in the former [21].

There are also differences in terms of the spatial locality/orientation of the model being fitted. Active-contour/snakes/deformable-models are usually fitted on the plane tangent to the vessel direction [27]. Intensity models (for the vessel cross section) are also usually fitted to the aforementioned plane for methods that estimate the tube boundary using a different method than the one used to estimate the tube direction, such as in [13, 21].

The scales of the tube structures to be detected may be relatively constant as in neuronal axon images specifically those of hippocampal pyramidal neurons since they are known not to biologically vary in diameter throughout their lengths [66]. The tube structures may also exhibit known scale changes, e.g., tapering of dendrite diameters starting from the soma [66]. On the other extreme, the scales may vary significantly in images containing pathological vasculature, e.g., those in close proximity to a tumor [12], or in vascular stenosis images [27]. Therefore, in these cases, single-scale methods (e.g., [67-73]) must be extended to cope with width variability [24, 34, 74-76], and the multi-scale analyses need to be normalized with respect to scale [27, 61, 77-81]. Note that the size scale is just a parameter of the local structure model.

2.1.3 Global Model Fitting

To fit the local models to the entire image content, called "*global fitting,*" the segmentation algorithm fits the local models either: (i) to every image pixel, also known as pixel-wise/pixel-by-pixel/image-wide algorithms (e.g., [20, 22, 23]); (ii) iteratively by "exploring" the image regions containing the tube-like structure of interest, also known as tracing/tracking/traversing/vectorization/exploratory algorithms (e.g., [2, 3, 13, 14, 21, 24]), which includes region growing methods as well (e.g., [82]); or, (iii) by wave front propagation methods, also known as least-cost/energy-minimizing path methods (e.g., [25-27, 83]).

Another way to classify these global fitting strategies is whether the fitting is performed directly on the intensity image (e.g., [2, 3, 14, 37, 84]), or on a synthesized image where the pixel values correspond to likelihood of being part of a tube-like structure, sometimes called the "vessel-enhanced image" (e.g., [21, 34, 74]). The vessel-enhanced image may be obtained by replacing the pixel intensity values with the (normalized) response of goodness-of-fit measures of local models.

In terms of what is actually segmented, it can either be the medial axis/centerline, or the entire tube volume/network, or any one of these initially and then followed by an estimation of the other. They are also termed as indirect/skeleton-based and direct/nonskeleton-based methods[6], respectively [85]. For example, the centerlines are extracted/detected first in [21, 27] followed by boundary estimates and refinements to the estimates. Conversely, the boundaries are detected first in [2, 3] followed by the centerline/center-point estimations. In methods that extract the entire tube volume/network, skeletonization steps (e.g., [37, 86]) and (sometimes) branch point analyses (e.g., [20, 87]) are required to extract the medial axis [82].

---

[6] In indirect methods, the structures are reconstructed by computing the cross-section from the centerlines and width information.

On the question of structure scale, the same global fitting can be performed either on a "scale-projection" image containing the maximal local fitting response across all scales [88], or on each scale individually [89]. Using the latter approach, scale selection is postponed from the local fitting step to the global fitting phase [89].

2.1.4 Level of Automation

Given the modeling assumptions above, there exist several levels of automation in executing the algorithm implementation, ranging from user-interactive/semi-automated methods (e.g., [89]) to fully automated ones (e.g., [14, 34]). When viewing image analysis systems as an estimation system that estimates which pixel belongs to the biological structures, initial estimates can be provided by the user as in the interactive methods, or obtained automatically [3]. Also, automation can be performed either at the local or global model fitting stages, or both.

2.2 Segmentation Evaluation

The traditional approach to evaluating segmentation algorithms is by visual inspection [90]. Some progress has been made on more automated approaches [91]. Metrics for segmentation evaluation [39, 42, 46, 90, 92-98] can be either goal-oriented, i.e., evaluation based on the performance of post-segmentation analyses [95, 99, 100], based on other application-guided criteria, such as the probability of false detection [39, 93, 96], or based on mismatch with manually obtained ground truth segmentations (e.g., [35, 44, 45, 98]). A majority of them are specifically tuned for low-level vision tasks, such as edge detection [38-40, 90, 92, 94, 96] and region growing [29, 41-43] but are mentioned here nevertheless since they usually form the foundations of the more complex and specialized image segmentation algorithms (e.g., [14, 101]).

Specific to biomedical images, segmentation evaluation using ground truth segmentations has been identified as a trade-off between accuracy of the ground truth segmentations and how much they reflect the characteristics of segmentation problems in practice [46]. For example, synthetic images have high accuracy but typically lack characteristics encountered in practice [46]. On the other end, manual segmentation by experts suffers from inter-expert and intra-expert variability [44]. Somewhere in the middle is the use of physical and digital phantoms, but it is acknowledged that creating such realistic phantoms remains a difficult task [102, 103]. Using expert manual segmentations as the ground truth, the comparison between ground truth and automated segmentation ranges from measuring spatial overlap (e.g., [104]), inter-expert comparisons (e.g., [44, 46, 105]), and assessment of boundary differences (e.g., [106]).

Another method to evaluate segmentations is in terms of the description length of the image, given a particular segmentation or image partition [29]. This is introduced next.

2.3 MDL Principle

The MDL principle [32] offers a systematic way to obtain an objective balance between segmentation conciseness and coverage [29, 30, 41]. The search of optimal parameters is analogous to the variational/energy formulation for image segmentation [42, 43], where the energy term to be minimized is evaluated using the description lengths [29, 41]. Descriptive languages and representations to encode both the segmentation and the image given the segmentation vary between applications [41, 107-110]. Since this is the underlying principle used to measure the segmentation optimality in this work, the detailed description is deferred to the next chapter.

2.4 Global Optimization

To ensure modularity of the segmentation-optimization framework, the choice is focused on optimization methods that require little *a priori* domain-specific information. For this, genetic algorithms [111], simulated annealing methods [112], controlled random search [113, 114], and multi-start pattern search [115, 116] are widely used [114, 117, 118]. These are also known as stochastic optimization algorithms, mainly based on random sampling methods [119, 120], as opposed to deterministic algorithms (e.g., [121]). To improve efficiency, they are normally combined with local search techniques, such as steepest descent [122], and pattern search [115] despite that these local search methods being very susceptible to noise [123].

Limitations of global optimization algorithms are reviewed in [124, 125], known as the “no free lunch” theorem for optimization. It states that any elevated performance over one class of problems is offset by performance over another class [125].

2.5 Chapter Summary

Although this work draws upon four large bodies of literature, its primary focus is on automated segmentation of tube-like biological objects. Instead of identifying strengths/weaknesses of existing segmentation algorithms to create another one, the literature was surveyed to identify the key components of a biological image segmentation algorithm in general terms and how to obtain the optimal segmentation from such existing algorithms.

The survey of segmentation evaluation in the literature, both general to broad computer vision algorithms to the specific applications in biomedical images, eventually led to the MDL principle for automated (unsupervised) segmentation optimality measure. Then, to ensure modularity of the segmentation-optimization framework, the choice of optimization methods is focused to those that require little *a priori* domain-specific information.

3. Methods

We are in effect attempting to decide which segmentation to be optimal among the set of all *computable* segmentations—obtained by varying segmentation algorithm parameters. When concretely specified, the notion of optimality becomes a metric that compares between segmentations and decides which one is better than the other. This permits the execution of the segmentation algorithm within the presented segmentation-optimization framework (Figure 6).

Following the previous introduction in Section 1.4, the traditional *maximum-a-posteriori* (MAP) criterion that minimizes the probability of error[1] [28] can be opted if prior probabilities on the segmentations can be specified, but unfortunately these priors are not well defined [29]. If a uniform probability distribution is assumed for the priors, then the MAP criterion reduces to the *maximum-likelihood* (ML) criterion that measures only the goodness-of-fit and has a tendency to overfit models to the data.

Let us step back and consider an automated image analysis system to be a system that performs a non-trivial data reduction using a segmentation algorithm. These systems are designed to extract as much of the structural content in the image as possible and concisely expressing it in terms of instances and descriptive parameters of object models. In other words, the resulting segmentation is expected to model the biological structures present in an image. In the fitting of these models to the image data, a systematic trade-off must be made between the fitting error, and conciseness of the representation. This requires a quantitative metric to evaluate the segmentation quality/optimality along these two terms.

Returning to the problem at hand, it turns out that the MAP criterion can still be used if optimal descriptive languages for a segmentation and the image given the segmentation can be specified. Then the minimum description length (MDL) [30] criterion can be adopted to choose the optimal segmentation as these two criteria have

---

[1] In this work, error is defined as the disagreement between a segmentation and the ground truth at each image pixel. To avoid introducing further subjectivity in this work, ground truth is obtained from public-domain databases of manual segmentations and by consensus of manual observers when provided by these databases.

been shown to be equivalent by Leclerc [29] once the descriptive languages are specified.

Next, the vague notion of optimality is progressively made concrete from the general-purpose image partitioning problems to segmentation of tube-like structures in biological images.

3.1 Notion of Optimality

First, let a segmentation algorithm be a function $f:\{\mathcal{I},\xi_i\}\rightarrow\mathcal{M}_i$ that maps an image $\mathcal{I}$ to a segmentation $\mathcal{M}_i$ consisting of a set of objects in the image, using a parameter vector ("settings") $\xi_i\in\Omega$ (see Figure 6a). The goal is to automate the search for the optimal $\hat{\xi}$ that yields the optimal $\hat{\mathcal{M}}$. In practice, the goal is to be approached the other way around, where the optimal segmentation $\hat{\mathcal{M}}$ is searched, and the corresponding set of segmentation algorithm parameters $\hat{\xi}$ is obtained.

As introduced and explained earlier in Section 1.4, both MDL and MAP strategies are equivalent [29]. For the problem at hand, the more straightforward MDL strategy by specifying a descriptive language for the segmentations is chosen.

3.1.1 MDL Criterion for Optimality

An image segmentation algorithm *partitions* an image into foreground and background regions, where foreground pixels belongs to the structure of interest [44]. Therefore, finding the optimal segmentation is equivalent to finding the optimal partitioning of the image.

Let $\mathcal{L}_m$ be the optimal descriptive language to describe the partitioning $\mathcal{M}_i$ and $\mathcal{L}_d$ be the optimal descriptive language to describe the image $\mathcal{I}$ given the partitioning $\mathcal{M}_i$, the MDL strategy is to choose $\mathcal{M}_i$ that minimizes

$$\underbrace{\left|\mathcal{L}_d\left(\mathcal{I}\middle|\mathcal{M}_i\right)\right|}_{\text{coverage}}+\underbrace{\left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right|}_{\text{conciseness}}, \tag{7}$$

where $|\cdot|$ denotes the description length in bits. After specifying the optimal descriptive languages, we only need to calculate the number of bits had we used them to describe/represent/encode both the partitioning and the image given the partitioning [29].

To describe the image given a particular partitioning, a certain regularity characteristic within each partition is assumed. As an example, assume that each region has constant intensity. Then, after specifying the constant intensity, only the differences between the pixel intensity values $\mathcal{I}(\mathbf{x})$ at pixel $\mathbf{x}$ and the region's constant intensity value need to be encoded. Note that the constant intensity value is described later by the language $\mathcal{L}_m$. Without loss of generality, let each region $\mathcal{R}_j$ be a region of constant intensity value of zero, and the differences that need to be encoded becomes the observed image intensity value $\mathcal{I}(\mathbf{x})$ itself. To encode the intensity value $\mathcal{I}(\mathbf{x})$ for a pixel $\mathbf{x}$ in a region $\mathcal{R}_j$, the optimal descriptive language $\mathcal{L}_d\left(\mathcal{I}(\mathbf{x})\middle|\mathcal{M}_i\right)$ can be designed such that the resulting description length is given by $\left|\mathcal{L}_d\left(\mathcal{I}(\mathbf{x})\middle|\mathcal{M}_i\right)\right| = -\log_2\left[P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{R}_j\right)\right]$, where $P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{R}_j\right)$ is the probability of observing the intensity value $\mathcal{I}(\mathbf{x})$ at pixel $\mathbf{x}$ in the region $\mathcal{R}_j$. However, there is no need to do so. The number of bits had this optimal descriptive language been used is all that is needed [29]. Then, assuming pixel independence, with identical distribution assumption within a region, the first term in (7) becomes:

$$\left|\mathcal{L}_d\left(\mathcal{I}\middle|\mathcal{M}_i\right)\right| = \sum_{\mathcal{R}_j\in\mathcal{I}}\sum_{\mathbf{x}\in\mathcal{R}_j} -\log_2\left[P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{R}_j\right)\right]. \tag{8}$$

Now, the optimal descriptive language $\mathcal{L}_m$ to describe the partitioning $\mathcal{M}_i$ needs to be specified. Leclerc suggested the use of the pixel-chain code to optimally describe the region boundaries [29]. The description length to describe the partitioning $\left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right|$ is then the number of bits needed to represent the regions of the partitioning $\mathcal{M}_i$ using the pixel-chain code.

At this point, the description length in (7) fully defined, and can be used as an optimization metric for currently-available optimization algorithms.

3.1.2 Adapting Criterion to Biological Image Segmentation

A partitioning $\mathcal{M}_i$ of the image $\mathcal{I}$ is equivalent to the segmentation $\mathcal{M}_i$ of the image $\mathcal{I}$—it gets simpler since an image segmentation algorithm for biological images partition an image into only two regions, the foreground region $\mathcal{F}$ and the background region $\mathcal{B}$. Therefore, two probability distribution functions (pdf) must be specified: (i) $P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{F}\right)$; and (ii) $P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{B}\right)$. However, at this point, notice that the local structure and intensity models are still absent from the MDL-optimality criterion.

To account for a neighborhood of pixels $N(\mathbf{x})$ around $\mathbf{x}$—implicitly capturing inter-pixel correlations—the pdf at each pixel $\mathbf{x}$ can be estimated using Besag's pseudo-likelihoods [126]. Therefore the pdf's at each pixel becomes:

$$\begin{aligned} &P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{F},N(\mathbf{x})\right), \text{ and} \\ &P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{B},N(\mathbf{x})\right), \end{aligned} \tag{9}$$

for the foreground and background regions respectively. These pdf's as a function of a neighborhood of pixels preferably indicate the likelihood of the object of interest to be present at that pixel given that the pixel belongs to the foreground or the background region.

3.1.3 Local Structure-Indicator Function and Computable Segmentations

To incorporate the local structure and intensity models, a local structure-indicator function $\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})$, defined at each pixel $\mathbf{x}$, is proposed to indicate the likelihood of the pixel belonging to the biological object. The object is locally modeled by the structure model $\mathcal{G}$ and the intensity model $\mathcal{T}$ in a neighborhood $N(\mathbf{x})$. The likelihood value is obtained by measuring the goodness-of-fit of the models, requiring us to find the best fit models and their corresponding parameters. By doing this, we have the pixel pdf's previously defined in (9) to be:

$$\begin{aligned} &P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{F},N(\mathbf{x})\right)=P\left(\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})\middle|\mathbf{x}\in\mathcal{F}\right), \text{ and} \\ &P\left(\mathcal{I}(\mathbf{x})\middle|\mathbf{x}\in\mathcal{B},N(\mathbf{x})\right)=P\left(\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})\middle|\mathbf{x}\in\mathcal{B}\right), \end{aligned}$$

where the neighborhood $N(\mathbf{x})$ is now one of the parameters of the local structure-indicator function $\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})$ through the local structure model $\mathcal{G}$. Given this, the first term in (7) becomes:

$$\left|\mathcal{L}_d\left(\mathcal{I}\middle|\mathcal{M}_i\right)\right| = \sum_{\mathbf{x}\in\mathcal{F}} -\log_2\left[P\Big(\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})\Big|\mathbf{x}\in\mathcal{F}\Big)\right] + \sum_{\mathbf{x}\in\mathcal{B}} -\log_2\left[P\Big(\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})\Big|\mathbf{x}\in\mathcal{B}\Big)\right], \tag{10}$$

for the special case of biological image segmentation. Notice that the first term in (7) now incorporates three out of four components of an image analysis system, namely the local structure model, the local intensity model, and the local fitting.

To incorporate the last component of image analysis systems, i.e., the global fitting component, the search space for the optimal segmentation $\hat{\mathcal{M}_i}$ can simply be limited to the set of segmentations $\{\mathcal{M}_\mathrm{i}\}$ that are *computable*[2] by a segmentation algorithm that incorporates the other three components. After all, there are up to $2^{(\text{number of pixels})}$ possible partitions of an image.

3.1.4 MDL-Optimality Criterion Adapted to Tube-like Biological Objects

In essence, the local structure-indicator function $\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})$ specializes the abstract MDL-optimality criterion in (7) to the structures of interest in the image, where the structures are specified in terms of the local structure model $\mathcal{G}$ and the local intensity model $\mathcal{T}$—around a neighborhood $N(\mathbf{x})$. At each pixel $\mathbf{x}$, it tells how much structure content that can be inferred from the neighboring pixels. The local structure-indicator function $\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})$ can also be generalized to segmentation problems for other image types such as texture images by incorporating the texture model into the intensity model $\mathcal{T}$ of $\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T})$.

Two different approaches to compute the local structure-indicator function are described next. The first approach is adapted from the statistical robust generalized log-likelihood ratio test formulation described by Mahadevan *et al.* [22]—to improve

---

[2] A segmentation is computable if it satisfies the local structure model, local intensity model, and the local fitting strategy used by the segmentation algorithm.

computation speed for practical use, it is approximated down to a matched filter formulation. If the single-vessel-in-a-neighborhood assumption is limiting for the application at hand, multiple-vessel models can be specified and a model selection formulation using the Akaike Information Criterion can be used, as shown by Wang and Bhalerao [127]. The second approach is the multi-scale vesselness formulation by Frangi *et al.* [74]. Either way, the range of the local structure-indicator function values must be normalized to allow interchangeability of these approaches to the specific application at hand.

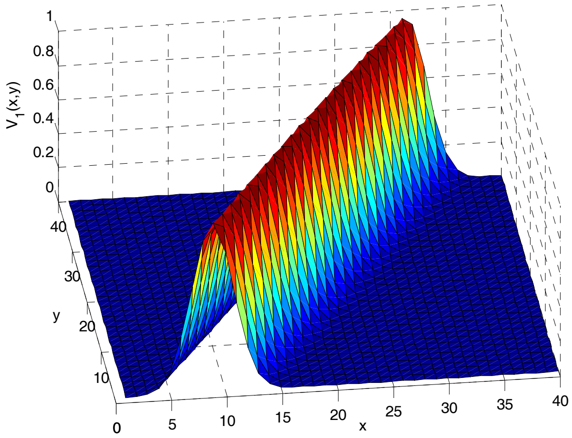


Figure 9. Displays the intensity profile in a 40×40 window using structure model of two parallel edges, and the Gaussian intensity model. One of the variances of the 2-D Gaussian is empirically set to $\sigma^2 = \left({}^{r}\!/_{3} \right)^2$ for the radius $r = 5$, and the other variance was set arbitrarily high. The direction of the Gaussian is defined in terms of the eigenvectors of the covariance matrix of the 2-D Gaussian function.

3.1.4.1 Generalized Log-Likelihood Ratio Test

Figure 9 displays an instance of the local model in a 40×40 neighborhood. Here, the generalized log-likelihood formulation is used to evaluate the local structure-indicator function $\mathcal{S}\left(\mathbf{x}, \mathcal{G}, \mathcal{T}\right)$.

To fit the parallel-edge structure model and Gaussian intensity model, a faster non-censored version of the robust vessel detection work in noisy retinal vasculature images by Mahadevan *et al.* [22] is proposed, motivated by practical reasons. Two hypotheses,

$\mathcal{H}_0$ (structure absent), and $\mathcal{H}_1$ (structure present), are tested at each pixel using the generalized log-likelihood ratio test. Using Poisson noise model for the observed image intensity at pixel $\mathbf{x}$, denoted $\mathcal{I}(\mathbf{x})$, the two hypotheses are:

$$\begin{aligned} &\mathcal{H}_0 : \mathcal{I}(\mathbf{x}) \sim Po(\mu_0), \mathbf{x} \in N(\mathbf{x}_0) \\ &\mathcal{H}_1 : \mathcal{I}(\mathbf{x}) \sim Po(\mu_1), \mathbf{x} \in N(\mathbf{x}_0) \end{aligned},$$

given a window (local neighborhood) $N(\mathbf{x}_0)$ such that $\mathbf{x}_0$ is the window's center pixel. For a particular pixel $\mathbf{x}_0 \in \mathcal{I}$ the likelihood ratio L at the pixel $\mathbf{x}_0$ assuming independent pixels is:

$$L(\mathbf{x}_0) = \frac{P_{\mathcal{H}_1}\left[\mathcal{I}(\mathbf{x}) \middle| N(\mathbf{x}_0)\right]}{P_{\mathcal{H}_0}\left[\mathcal{I}(\mathbf{x}) \middle| N(\mathbf{x}_0)\right]}.$$

The decision now becomes, reject $\mathcal{H}_0$ if $L(\mathbf{x}_0) > 1$, or reject $\mathcal{H}_0$ if $l(\mathbf{x}_0) \triangleq \ln\left(L(\mathbf{x}_0)\right) > 0$. Next, the test statistic (log-likelihood value $l(\mathbf{x}_0)$) is derived as follows:

$$\begin{aligned} l(\mathbf{x}_0) &= \ln\left[\frac{\prod\limits_{\mathbf{x}\in N(\mathbf{x}_0)} \dfrac{e^{-\hat{\mu}_1}\hat{\mu}_1^{\mathcal{I}(\mathbf{x})}}{\mathcal{I}(\mathbf{x})!}}{\prod\limits_{\mathbf{x}\in N(\mathbf{x}_0)} \dfrac{e^{-\hat{\mu}_0}\hat{\mu}_0^{\mathcal{I}(\mathbf{x})}}{\mathcal{I}(\mathbf{x})!}}\right]. \\ &= \underbrace{\sum_{\mathbf{x}\in N(\mathbf{x}_0)} \mathcal{I}(\mathbf{x}) \ln \hat{\mu}_1 - \hat{\mu}_1}_{\text{foreground fit}} - \underbrace{\sum_{\mathbf{x}\in N(\mathbf{x}_0)} \mathcal{I}(\mathbf{x}) \ln \hat{\mu}_0 - \hat{\mu}_0}_{\text{background fit}} \end{aligned}$$

Next, the foreground intensity estimate for pixel $\mathbf{x} \in N(\mathbf{x}_0)$ is locally modulated in the neighborhood $N(\mathbf{x}_0)$ using the vessel profile $V_1(\mathbf{x}) \in [0,1]$:

$$\hat{\mu}_1 = \hat{\mu}_1(\mathbf{x}) = \hat{\mu}_0(\mathbf{x}_0) + \left(\hat{\mu}_1(\mathbf{x}_0) - \hat{\mu}_0(\mathbf{x}_0)\right)V_1(\mathbf{x}), \mathbf{x} \in N(\mathbf{x}_0).$$

The vessel profile has two parameters, i.e., $V_1(\mathbf{x}) = V_1(\mathbf{x};\theta,r)$, where θ is the angle, and r is the radius. Figure 9 displays a vessel profile with $\theta = \pi/3$ and $r = 5$.

Next, let $\hat{c}(\mathbf{x}_0) = \hat{\mu}_1(\mathbf{x}_0) - \hat{\mu}_0(\mathbf{x}_0)$ be the contrast estimate at pixel $\mathbf{x}_0$. Then, the log-likelihood term becomes,

$$l\left(\mathbf{x}_0\right)=\left(\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\mathcal{I}\left(\mathbf{x}\right)\ln\left[\hat{\mu}_0\left(\mathbf{x}_0\right)+\hat{c}\left(\mathbf{x}_0\right)V_1\left(\mathbf{x}\right)\right]-\left[\hat{\mu}_0\left(\mathbf{x}_0\right)+\hat{c}\left(\mathbf{x}_0\right)V_1\left(\mathbf{x}\right)\right]\right) \\ -\left(\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\mathcal{I}\left(\mathbf{x}\right)\ln\hat{\mu}_0\left(\mathbf{x}_0\right)-\hat{\mu}_0\left(\mathbf{x}_0\right)\right).$$

To simplify, let $l\left(\mathbf{x}_0\right)=l_f\left(\mathbf{x}_0\right)+l_b\left(\mathbf{x}_0\right)$, where

$$l_f\left(\mathbf{x}_0\right)=\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\left(\mathcal{I}\left(\mathbf{x}\right)\ln\left\{\hat{\mu}_0\left(\mathbf{x}_0\right)\left(1+\frac{\hat{c}\left(\mathbf{x}_0\right)}{\hat{\mu}_0\left(\mathbf{x}_0\right)}V_1\left(\mathbf{x}\right)\right)\right\}-\left[\hat{\mu}_0\left(\mathbf{x}_0\right)+\hat{c}\left(\mathbf{x}_0\right)V_1\left(\mathbf{x}\right)\right]\right),$$

and

$$l_b\left(\mathbf{x}_0\right)=\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\mathcal{I}\left(\mathbf{x}\right)\ln\hat{\mu}_0\left(\mathbf{x}_0\right)-\hat{\mu}_0\left(\mathbf{x}_0\right).$$

Without loss of generality, let $\hat{c}\left(\mathbf{x}_0\right)=\hat{\mu}_0\left(\mathbf{x}_0\right)$ to permit the use Taylor series approximation (1$^{\text{st}}$ order) for the term $\ln\left(1+\frac{\hat{c}\left(\mathbf{x}_0\right)}{\hat{\mu}_0\left(\mathbf{x}_0\right)}V_1\left(\mathbf{x}\right)\right)=\ln\left(1+V_1\left(\mathbf{x}\right)\right)\approx V_1\left(\mathbf{x}\right)$.

Using these, the test statistic becomes:

$$l_f\left(\mathbf{x}_0\right)\approx\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\mathcal{I}\left(\mathbf{x}\right)\left(\ln\hat{\mu}_0\left(\mathbf{x}_0\right)+V_1\left(\mathbf{x}\right)\right)-\left[\hat{\mu}_0\left(\mathbf{x}_0\right)+\hat{\mu}_0\left(\mathbf{x}_0\right)V_1\left(\mathbf{x}\right)\right].$$

After term cancellations, the test statistic is:

$$l\left(\mathbf{x}_0\right)\approx\sum_{\mathbf{x}\in N\left(\mathbf{x}_0\right)}\left(\mathcal{I}\left(\mathbf{x}\right)-\hat{\mu}_0\left(\mathbf{x}_0\right)\right)V_1\left(\mathbf{x}\right).$$

This effectively reduces the log-likelihood ratio to a single convolution operation. Interestingly, it reduces the log-likelihood ratio to a matched-filter with $\mathcal{I}\left(\mathbf{x}\right)-\hat{\mu}_0\left(\mathbf{x}_0\right)$ as the input signal and $V_1\left(\mathbf{x}\right)$ as the known/deterministic signal.

Therefore, using the log-likelihood ratio formulation, the local structure-indicator function becomes:

$$\mathcal{S}\left(\mathbf{x},\mathcal{G},\mathcal{T}\right)\triangleq l\left(\mathbf{x}\right) \qquad (11)$$

3.1.4.2 The Vesselness Formulation

Here the intensity ridge local structure model is adopted for tube-like biological structures, treating the pixel intensity at each pixel as the pixel height, and viewing the image as a “landscape.” In the $n-Dimensional$ Euclidean space, for values $1 \le d \le n$, the pixel $\mathbf{x} \in \mathbb{R}^n$ is a ridge point of type $n-d$ if and only if $\left[\mathbf{v}_1 \cdots \mathbf{v}_d\right]^t \nabla \mathcal{I}\left(\mathbf{x}\right) = 0$ and $\lambda_d < 0$, where $\left\{\mathbf{v}_i\right\}_{i \in [1,n]}$ are the eigenvectors of the Hessian matrix evaluated at $\mathbf{x}$ with the corresponding ordered eigenvalues $\lambda_1 \le \cdots \le \lambda_n$ and $\nabla \mathcal{I}\left(\mathbf{x}\right)$ is the image gradient evaluated at $\mathbf{x}$ (for review, see [59]). The eigenvector of the Hessian with the largest eigenvalue tells the direction of highest gradient, i.e., the direction perpendicular to the tube’s medial axis, where the other eigenvector with the smaller eigenvalues is oriented in the direction of the tube’s medial axis—the latter eigenvector must be perpendicular to the image gradient at that pixel to qualify the pixel as a ridge point.

No intensity model is explicitly assumed, but can be incorporated into the vesselness formulation by a multiplicative factor that relates the scale (read: the standard deviation of the differential Gaussian kernel) and the radius of the object.

Using the intensity ridge local structure model, the multi-scale vesselness formulation [27, 74, 128] is adopted in this work. It has formed the basis for numerous tube-segmentation algorithms [6, 21, 129]. Denoted $\mathcal{V}_\sigma\left(\mathbf{x}\right) \in \left[0,1\right)$, it measures the likelihood that a local group of pixels centered at $\mathbf{x}$ belong to a tube of size scale σ [27, 74, 128]. The range of size scales specified in the vesselness formulation is a parameter of the intensity ridge structure model—the search for ridges are limited to these scales. No intensity model is explicitly assumed, but can be incorporated into the vesselness formulation by a multiplicative factor that relates the scale [13]. For the Gaussian intensity model the relation between the scale σ and the actual vessel radius r has been derived analytically by Krissian *et al.* [13] as $\sigma = 0.5r^2$, and will be used throughout to translate between scale and radius values.

The vesselness formulation inspects second-order features [59] of the image $\mathcal{I}$, at scale σ, at each pixel $\mathbf{x}$, which is obtained from the Hessian matrix $H_\sigma\left(\mathbf{x}\right) = \sigma^\gamma \left[\nabla^2 \left(\mathcal{I}\left(\mathbf{x}\right) * G_\sigma\left(\mathbf{x}\right)\right)\right]$. In the $n-Dimensional$ Euclidean space, eigenvalues of the corresponding $n \times n$ $H_\sigma(\mathbf{x})$ matrix are denoted $\lambda_d, d = 1,2,...n$ and arranged in increasing magnitude, i.e., $|\lambda_1| \le |\lambda_2| \le \cdots \le |\lambda_n|$. The Gaussian function with standard deviation σ is denoted as G_σ, the Lindeberg constant [81] for a family of scale-normalized derivatives is denoted γ, and "*" is the convolution operator.

For 3-D images, the vesselness measure at scale σ, denoted $\mathcal{V}_\sigma(\mathbf{x})$, is given by:

$$\mathcal{V}_\sigma(\mathbf{x}) = \begin{cases} 0, \text{ if } \lambda_2 > 0 \\ \left[1 - \exp\left(-\frac{R_A^2}{2a^2}\right)\right]\left[\exp\left(-\frac{R_B^2}{2b^2}\right)\right]\left[1 - \exp\left(-\frac{S^2}{2c^2}\right)\right] \end{cases}, \tag{12}$$

where $R_B = \lambda_1 \big/ \sqrt{|\lambda_2 \lambda_3|}$ distinguishes between blob-like and other structures, $R_A = |\lambda_2| / |\lambda_3|$ distinguishes between plate-like and tube-like structures, and $S = \sqrt{\sum_{d \le D} \lambda_d^2}$ is the image intensity contrast factor ($D = 3$ for 3-D images). The terms a, b, and c in (12) are weighting parameters for R_A, R_B, and S.

For 2-D images, $\mathcal{V}_\sigma(\mathbf{x})$ is given by:

$$\mathcal{V}_\sigma(\mathbf{x}) = \begin{cases} 0, \text{ if } \lambda_2 > 0 \\ \exp\left(-\frac{R_C^2}{2b^2}\right)\left[1 - \exp\left(-\frac{S^2}{2c^2}\right)\right] \end{cases}, \tag{13}$$

$R_C = \lambda_1 / \lambda_2$ is the likelihood of being on a tube in 2-D images. The terms b and c in (13) are weighting parameters for R_C and S.

The multi-scale vesselness measure, denoted $\mathcal{V}_{[\sigma_{\min}, \sigma_{\max}]}(\mathbf{x})$, is then defined as

$$\mathcal{V}_{[\sigma_{\min}, \sigma_{\max}]}(\mathbf{x}) \triangleq \max_{\sigma_{\min} \le \sigma \le \sigma_{\max}} \mathcal{V}_\sigma(\mathbf{x}), \tag{14}$$

for a given range of size scales $[\sigma_{\min}, \sigma_{\max}]$ which is essentially one of the parameters of the structure model $\mathcal{G}$.

As a side note, Frangi *et al.* [74] have described approaches to adapt this measure to other key structure models in biological images. Table I is a simplified summary linking the models and the corresponding eigenvalues. In 3-D, the 3×3 $H_\sigma(\mathbf{x})$ matrix has three eigenvalues, similarly ordered as $|\lambda_1| \le |\lambda_2| \le |\lambda_3|$. Notice that the vesselness value in Table I increase with decreasing magnitude of R_B, effectively capturing the eigenvalues-shape associations for tube-like objects. This suggests that (13) can be modified accordingly for other geometrical models.

Table I. Indicates the relationship between the eigenvalues of the Hessian matrix with common structure models in biological cell and tissue-level imagery [74]. For dark objects on bright backgrounds, the signs are reversed. In 2-D plate-like structures can not be detected. The first two eigenvalues λ_1 and λ_2 are used in 2-D. All three eigenvalues are used in 3-D.

Structure Model $\mathcal{G}$	λ_1	λ_2	λ_3
Tube-like	≈ 0	$\ll 0$	$\ll 0$
Blob-like	$\ll 0$	$\ll 0$	$\ll 0$
Plate-like	≈ 0	≈ 0	$\ll 0$

Returning to the specific case for tube-like biological objects, the local structure-indicator function that indicates the likelihood that a pixel $\mathbf{x}$ belongs to a tube-like structure is defined as:

$$\mathcal{S}(\mathbf{x},\mathcal{G},\mathcal{T}) \triangleq \mathcal{V}_{[\sigma_{\min},\sigma_{\max}]}(\mathbf{x}). \tag{15}$$

3.1.5 Universal Parameter and Segmentation Quality Metric

At this point, the problem of image segmentation is ready to be wrapped into an optimization framework using the optimization metric based on (7). Before proceeding, the users' desire to override/bias the trade-off between conciseness and segmentation coverage is acknowledged. For this, a "universal parameter," denoted α, is incorporated in the optimization metric based on the MDL principle:

$$q(\mathcal{M}_i|\mathcal{G},\mathcal{I},\alpha,\xi) = \alpha\underbrace{\left|\mathcal{L}_d(\mathcal{I}|\mathcal{M}_i)\right|}_{\text{coverage}} + (1-\alpha)\underbrace{\left|\mathcal{L}_m(\mathcal{M}_i)\right|}_{\text{conciseness}}, \tag{16}$$

for a given structure model $\mathcal{G}$ and intensity model $\mathcal{T}$. From this point onwards, we will refer instances of the evaluated optimization metric simply as the "q-values." For segmentation of tube-like biological structures, the segmentation coverage term is as defined in (10), with the local structure-indicator function $\mathcal{S}\left(\mathbf{x},\mathcal{G},\mathcal{T}\right)$ defined in (11) or (15). When $\alpha = 0.5$, the conciseness-coverage trade-off is a balanced one.

With this, the segmentation quality metric Q is defined as:

$$Q \triangleq -q\,, \tag{17}$$

to equate objectives of minimizing the α-weighted description length q and maximizing the segmentation quality Q.

3.2 Recursive Random Search Strategy

Several considerations motivate the selection of the recursive random search (RRS) as a preferred strategy compared to other available alternatives. First, exhaustive search is time-prohibitive. Second, the objective function is not differentiable with respect to the parameter vector, mainly because the mapping between parameter settings and the resulting segmentation itself is not differentiable. Third, for the image segmentation problems of interest, it is rational to seek approximate solutions within a defined time frame, rather than truly global optima. Fourth, it is common for segmentation software to have several irrelevant/ineffective settings for a given application, and search algorithms that are able to minimize the computational effort in such dimensions are preferable. Finally, we expect the search algorithm to be robust to minor noise-like fluctuations in the objective function. Interestingly, Ye *et al.* [119] developed this algorithm motivated by similar parameter spaces encountered in computer networking – specifically, automatic and dynamic configuration of network components to maximize network throughput.

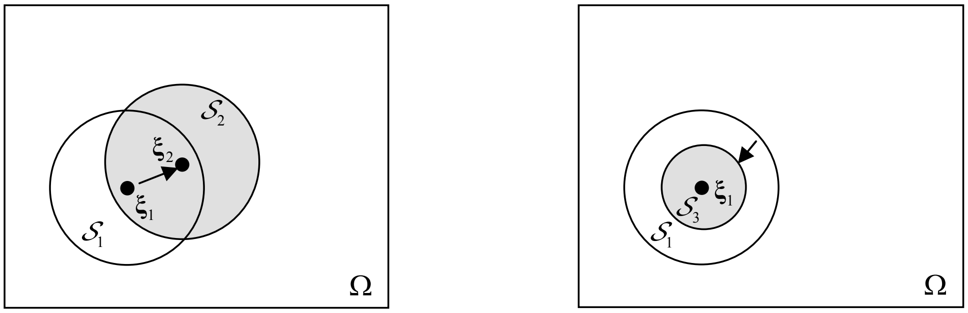
As indicated by the name, the RRS algorithm is based on random sampling. This algorithm searches the parameter space in two recursive steps: exploration and exploitation, respectively. The exploration step examines the macroscopic features of the objective function (e.g., globally convex or "big valleys" structure [64]) and attempts to identify promising areas in the parameter space Ω that are subsequently "exploited" intensively by the second step, called the exploitation step. For the random sampling, a uniform distribution over Ω is used. This has been shown to be the simplest search technique for similar non-linear problems, and is widely used [113, 119, 120, 130]. It has been shown to be more efficient for exploring high-dimensional parameter spaces compared to deterministic exploration methods [119, 131], and can be shown to converge to the global optima [132].

We show below that random sampling is in fact very efficient in its initial steps and only starts to become inefficient in the later sampling steps. Given a measurable objective function $q(\xi)$ over the parameter space Ω, we can define the distribution of objective function values for some $q_0 \in [q_{\min}, q_{\max}]$ as $\phi_\Omega(q_0) = m(\{\xi \in \Omega \mid q(\xi) \le q_0\}) / m(\Omega)$, where $m(.)$ is the Lebesgue measure. Hence, the distribution function $\phi_\Omega(q_0)$ represents the portion of the points in the parameter space whose objective function values are smaller than a certain level q_0. Furthermore, it has a maximum value of 1 when $q_0 = q_{\max}$ and a minimum value of 0 when $q_0 = q_{\min}$ corresponding to the (set of) global optimum (optima). Without loss of generality, assume that $q(\xi)$ is a continuous function and $m(\{\xi \in \Omega \mid q(\xi) = q_0\}) = 0,\ \forall q_0 \in [q_{\min}, q_{\max}]$.

Assuming a $q_r \in [q_{\min}, q_{\max}]$ value such that $\phi_\Omega(q_r) = r$, $r \in [0,1]$, an r-percentile subspace $\mathcal{S}_\Omega(r)$ in the parameter space can be defined as $\mathcal{S}_\Omega(r) = \{\xi \in \Omega \mid q(\xi) \le q_r\}$. Note that $\mathcal{S}_\Omega(1)$ is just the whole parameter space and $\lim_{\delta \to 0} \mathcal{S}_\Omega(\delta)$ converges to the global optima. Suppose the sample sequence generated by n steps of random sampling is $\{\xi\}_{i=1}^{n}$ and $\xi^{(i)}$ is the one with the minimum objective function value, then the probability of $\xi^{(i)}$ in $\mathcal{S}_\Omega(r)$ is given by $P(\{\xi^{(i)} \in \mathcal{S}_\Omega(r)\}) = 1 - (1-r)^n = p$. Stated in another way, the value of r for which $\xi^{(i)}$ will be reached with probability p is given by $r = 1 - (1-p)^{1/n}$.

For any $0 < p < 1$, r will tend to 0 with increasing n, implying that random sampling will converge to the global optima with increasing numbers of samples. Since r decreases exponentially with increasing n, the efficiency of random sampling is high at initial samples but falls sharply at later samples. This observation led to the idea of restarting the sampling before its efficiency drops off [119], either by moving or resizing the sample space according to sample history.

Specifically, RRS performs exploitation in two iterative steps: (i) random sampling within the current space, and (ii) realign or shrink. As illustrated in Figure 10a, after drawing a certain number of random samples (we used 7 in this paper) within the current space $\mathcal{S}_1$, if a superior sample ξ_2 is found, then $\mathcal{S}_1$ centered at the current sample ξ_1 is realigned (moved) to the sample space $\mathcal{S}_2$. If no better sample is found during the random sampling, the parameter space $\mathcal{S}_1$ is shrunk to $\mathcal{S}_3$ instead of realigning to $\mathcal{S}_2$ as shown in Figure 10b.



(a) Re-align (b) Shrink

Figure 10. Venn diagrams illustrating the re-align and shrink operations in the exploitation step of the recursive random search (RRS) algorithm. The current sample is denoted ξ_1, and the local exploitation subspace is depicted as an unshaded circle around it. After drawing a certain number of random samples within the current space $\mathcal{S}_1$, if a better sample ξ_2 is found, then the search is re-aligned to the sample space $\mathcal{S}_2$. If no better sample is found during the random sampling, the parameter space $\mathcal{S}_1$ is shrunk to $\mathcal{S}_3$ instead of realigning to $\mathcal{S}_2$.

The strategy for limiting the exploitation step to promising subspaces is based on identifying an r-percentile subspace for exploitation, as described above. In this way, most trivial subspaces will be excluded from exploitation, improving the overall efficiency of the search. In contrast, algorithms such as multi-start [130], do not distinguish between subspaces and hence may waste time in trivial areas.

The RRS is efficient at handling an objective function with a subset of ineffective parameters [119] because random samples maintain a uniform distribution within the subspace composed of only the effective parameters, minimizing the computational effort invested on negligible parameters. In contrast, local search methods are affected by unimportant parameters because of their high dependency on dimensionality of the search space. For more details, the interested reader is referred to [119].

Table II. The components of the parameter vector $\xi \in \mathbb{Z}^8$, their respective ranges, default values, and constraints on their values for the tracing algorithm.

Parameter	Range	Default Value	Constraint on Values
Grid spacing g	10 – 30	15	divisible by 5
Minimum template length $L_{\min}$	8 – 20	10	$L_{\min} \le L_{\max}$
Maximum template length $L_{\max}$	8 –30	18	$L_{\min} \le L_{\max}$
Relative shift distance n_{shift}	2 –10	2	none
Directional degree of freedom n_{rotate}	3 –7	7	odd-numbered
Maximum step size $s_{\max}$	3 –10	8	$s_{\max} \le L_{\min}$
Contrast threshold multiplier τ_c	1 –10	3	none
Maximum allowed stopping violations ν	1 –10	1	none

3.3 Automated Tracing Algorithm

The tracing algorithm [3, 14], used in generating all experimental results, is briefly summarized here with graphical illustrations of several key intermediate steps taken by the algorithm. In principle, the tracing algorithm models tube-like biological structures locally as piecewise-linear generalized cylinder segments (Figure 11) [3, 14]. In 2-D, the generalized cylinder model reduces to the parallel-edge model (Figure 12). Parameter settings for this algorithm are denoted:

$$\xi = \begin{pmatrix} g & L_{\min} & L_{\max} & n_{shift} & n_{rotate} & s_{\max} & \rho & \nu \end{pmatrix}.$$

They are summarized in Table II along with their default values. The traces are not merged for illustrative purposes since the traced segments are merged after tracing, i.e., as a post-processing step in the implementation of the tracing algorithm [87].

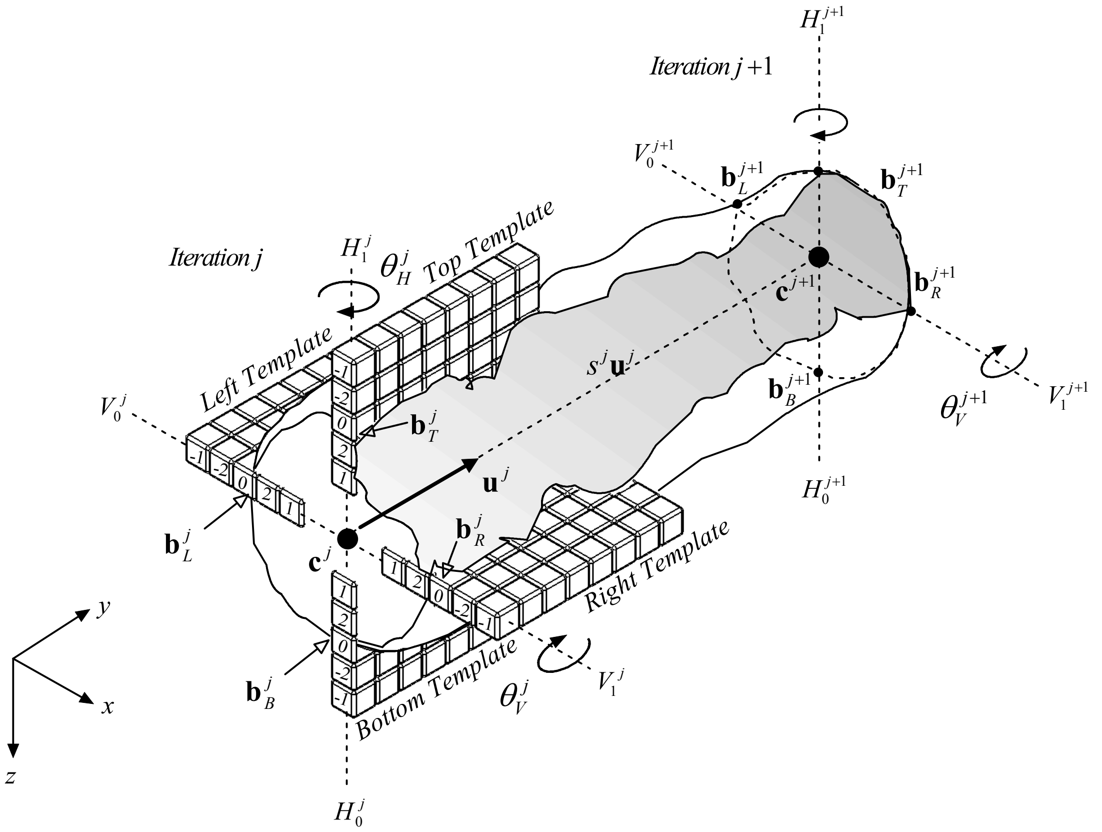


Figure 11. Illustrates, in 3-D, the generalized cylinder model and the iterative procedure of the tracing algorithm. Starting with initial seeds, the algorithm estimates the next location based on a robust estimates of the local boundaries of the cylinder. In 2-D, illustrated separately in Figure 12, only the left and the right templates are needed, collapsing the generalized cylinder model to the parallel-edge model.

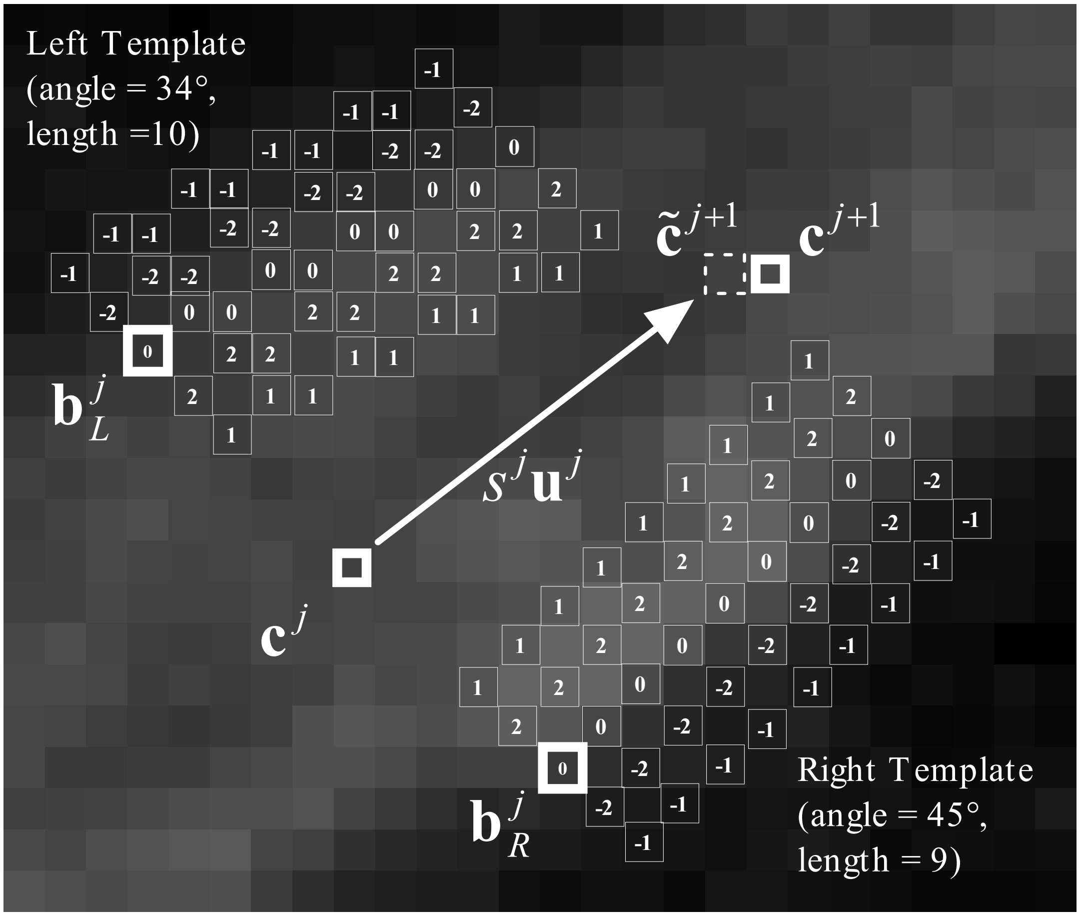
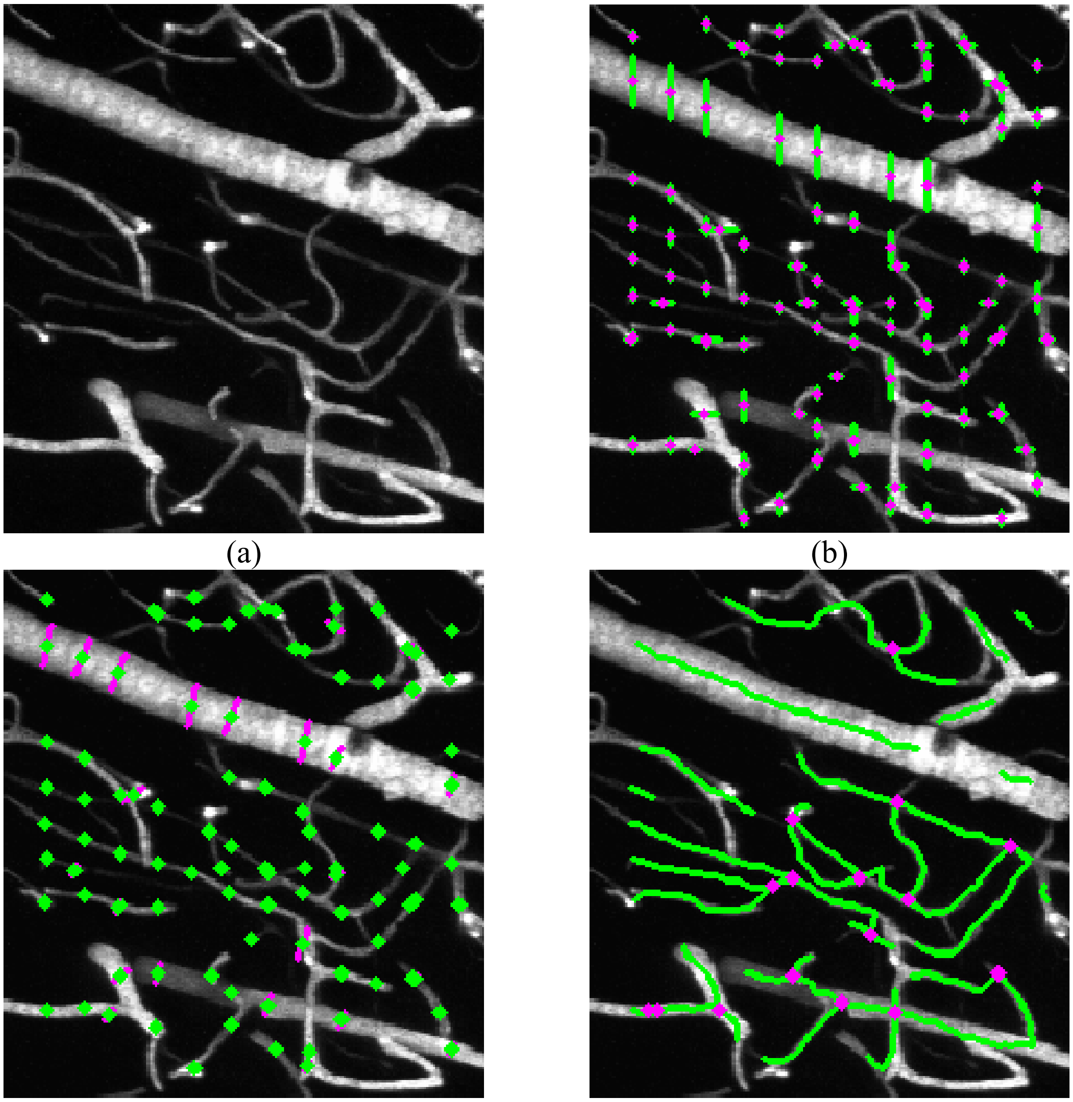


Figure 12. Illustrates, in 2-D, the generalized cylinder model (effectively reduced to the parallel-edge model) and the iterative procedure of the tracing algorithm. The 3-D version is illustrated separately in Figure 11.



(a) (b)

(c) (d)

Figure 13. Illustrates the intermediate steps/stages of the automated tracing algorithm. Results displayed are of default algorithm parameters. (a) Input image (displayed is a partial view of a 512×480×151 image) of a neurovascular cast imaged using confocal microscopy (Image source: C. Bjornsson, K. Smith, W. Shain, The Wadsworth Center). (b) The detected seed candidates in magenta, along with the detected width in green relative to the grid used to detect the seeds. (c) The verified seed candidate, after being fitted to the generalized cylinder model. The verified seeds are displayed in green with the width estimates in magenta. The direction of the verified seeds—to initiate tracing from—is perpendicular to the magenta lines. (d) The resulting traces in green, with the detected branch points in magenta.

Figure 13 illustrates the intermediate steps/stages of the automated tracing algorithm. Results displayed are of default algorithm parameters. Panel (a) displays the input image containing a neurovascular cast imaged using confocal microscopy. First, seed points are detected along rectangular grids g pixels apart overlaid on the image. A pair of 1-D edge-detector kernels of the form $\left(-1,-2,0,+2,+1\right)^T$ are correlated to the pixel intensity values along each 1-D grid profile, and correlation values less than an adaptive threshold times the contrast threshold multiplier ρ are discarded. Then the remaining seeds are called as "seed candidates." Panel (b) displays the detected seed candidates, along with the detected widths in green relative to the grid used to estimate the seed locations. Second, the seed candidates are verified using the generalized cylinder model (3-D: Figure 11, 2-D: Figure 12), turning them into "verified seeds." Panel (c) displays the verified seeds in green, with the width estimates in magenta. The direction of the verified seeds—to initiate tracing from—is perpendicular to the magenta lines. Panel (d) displays the resulting traces in green, with the detected branch points in magenta.

In the 3-D implementation of the algorithm, seed detection is performed on the axial maximum-intensity-projection image. The axial coordinate is found later by using the same previously-described correlation method axially at each seed point's coordinate in the x-y plane of the 3-D image.

Next, at each tracing iteration j, four boundary points $\left\{\mathbf{b}_L^j,\mathbf{b}_R^j,\mathbf{b}_T^j,\mathbf{b}_B^j\right\}$ (only $\left\{\mathbf{b}_L^j,\mathbf{b}_R^j\right\}$ in 2-D) corresponding to the left, right, top, and bottom tube boundaries are each found using directional correlation kernels called templates (3-D: Figure 11, 2-D: Figure 12). The center point $\mathbf{c}^j$ is simply defined as the center of these boundary points. Each template consists of linearly stacked 1-D edge-detector kernels of the form $\left(-1,-2,0,+2,+1\right)^T$ anchored at the image point $\mathbf{b}^j$ along a particular direction $\mathbf{u}^j$. The edge-strengths from each 1-D edge detector kernel r are averaged through the entire template length $l \in L$, using median statistics [133] for robustness. The set $L=\left[L_{\min},L_{\max}\right]$ contains all template lengths. The correlation between the templates and the image is called the "template response," denoted R, as below:

$$R\left(\mathbf{b}^j, \mathbf{u}^j, L\right) = \arg\max_{l \in L} \left\{ \underset{t=1,2,\ldots,l}{median} \left\{ r\left(\mathbf{b}^j + t\mathbf{u}^j\right) \right\} \right\}. \tag{18}$$

Each boundary point corresponds to the maximal response template parameters $\left\{\mathbf{b}^j, \mathbf{u}^j, l^j\right\}$:

$$(\mathbf{b}^j, \mathbf{u}^j, l^j) = \underset{\{(\mathbf{b},\mathbf{u},l) \mid \mathbf{b}=\mathbf{c}^j + m\mathbf{u}_\perp, m=1,..,\frac{M}{2}, \mathbf{u}\in\mathbf{U}, l\in L\}}{\arg\max} R(\mathbf{b}, \mathbf{u}, L), \tag{19}$$

where $\mathbf{U}$ is the set of unit vectors along directions in the neighborhood of $\mathbf{u}^j$. The radius of the widest expected vasculature is denoted as M—specified earlier as one of the parameters of the local structure model. Each template is elongated, shifted from $\mathbf{c}^j$, and rotated about two axes $H_0^j H_1^j$ and $V_0^j V_1^j$ to find the corresponding boundary points. To save computation, the template is only shifted in a neighborhood n_{shift} of previously calculated width at iteration $j-1$ and rotated in a neighborhood n_{rotate} around previously calculated tracing direction $\mathbf{u}^{j-1}$. The next center point $\mathbf{c}^{j+1}$ is estimated by scaling $\mathbf{u}^j$ with the adaptive step-size s^j, limited by the parameter $s_{\max}$. The same boundary-finding process is repeated until a stopping criterion containing a contrast-based threshold multiplier ρ is met v consecutive times. The interested reader is referred to [3, 14] for details of the algorithm.

3.4 Improving Execution Speed for 3-D Images

Practical difficulties are expected for 3-D images when incremental changes in the segmentation algorithm parameters do not translate to incremental changes in the segmentation, i.e., each RRS function evaluation corresponds to complete re-segmentation of an image. Given the closed-loop design in Figure 6, speeding up the segmentation algorithm directly reduces the overall execution time. In practice, computing segmentations for smaller images are faster than for larger images due to reduction of access time (disk and memory). This section presents a systematic way to create image subvolumes and using them to automatically find the segmentation algorithm parameters for the entire image.

Two approaches are proposed to achieve this goal. The first approach is to obtain the segmentation algorithm parameters from a single subvolume. A similarity measure using image and structural characteristics is presented to systematically choose among possible subvolumes. The second approach generalizes the first by obtaining segmentation algorithm parameters from multiple subvolumes. This enables adaptation of segmentation algorithm parameters originally designed for the entire image (called as the non-adaptive parameters), effectively making the segmentation algorithm spatially adaptive.

3.4.1 Creating the Subvolumes

Tube-like structures may not be present at all parts of the image. Therefore, simple uniform image partitioning into non-overlapping subvolumes of similar size may produce empty subvolumes. To avoid this, the local model-fitting stage of the segmentation algorithm is used to obtain non-empty subvolumes.

The local model-fitting stage of the 3-D segmentation algorithm [3] (Section 3.3) used in generating the experimental results, is the seed verification, where each seed candidate obtained from the grid-search is verified using the generalized cylinder model (Figure 11, p. 38). These seed points are clustered into k clusters using the k-means clustering algorithm [134]. The bounding-boxes of the clusters define the subvolumes $\left\{v_j\right\}_{j=1}^{k}$ that are structure-wise non-empty. A subvolume, instead of the entire image, is used as the input image for the loop in Figure 6 (p. 9).

3.4.2 Representative Subvolume

A distance measure between two image volumes, e.g., a subvolume and the entire image is needed in order to measure the similarity between the two. More importantly, it will measure how representative the subvolume is to the whole image. The distance measure should ideally incorporate both image intensity information and structural information. Given the modular approach of this work, this can be obtained from the local model fitting stage of the segmentation algorithm. For this particular algorithm [3], it is the seed verification step.

Each verified seed point contains information that characterizes the locally best-fitted model. Referring to the generalized cylinder model shown in Figure 11, let $\|\cdot\|$ denote the Euclidean distance and the following best-fitted model information are used:

- Horizontal width w_H, computed as $\|\mathbf{b}_L - \mathbf{b}_R\|$.
- Vertical width w_V, computed as $\|\mathbf{b}_T - \mathbf{b}_B\|$.
- Cylinder orientation characterized by the unit vector $\mathbf{u}$ defined by two angles θ_H, and θ_V [3]. The horizontal angle θ_H is the rotation around the z axis. The vertical angle θ_V is the rotation angle around the y axis that has been rotated by θ_H around the z axis.
- Estimated local contrast $c = \frac{1}{3}\left(\frac{1}{4}\sum_{p\in\{L,R,T,B\}} R\left(\mathbf{b}_p, \mathbf{u}\right)\right)$ from the four best-fit template at boundary points $\{\mathbf{b}_L, \mathbf{b}_R, \mathbf{b}_T, \mathbf{b}_B\}$ using the template-response to contrast conversion ratio of $\frac{1}{3}$ as presented in [3].

The next step is to construct a feature (column) vector $\mathbf{m}$ that summarizes the information obtained from each seed point:

$$\mathbf{m} = \begin{bmatrix} c & w_H & w_V & \cos 2\theta_H & \sin 2\theta_H & \cos 2\theta_V & \sin 2\theta_V \end{bmatrix}^T. \tag{20}$$

The angles are doubled since anti-parallel angles separated by π are considered equal for the purpose of characterizing the tube/cylinder orientation [47] as adapted from angular statistics [135].

The mean feature vector, denoted $\bar{\mathbf{m}}$, is used to characterize a set of N seeds. For the set of all seeds in the image $\mathcal{I}$, the (sample) image mean feature vector $\bar{\mathbf{m}}_{\mathcal{I}}$ is:

$$\bar{\mathbf{m}}_{\mathcal{I}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{m}_i, \tag{21}$$

and the (sample) covariance matrix $\mathbf{C}_{\mathcal{I}}$ is computed as:

$$\mathbf{C}_{\mathcal{I}} = \frac{1}{N-1}\sum_{i=1}^{N}\left[\mathbf{m}_i - \bar{\mathbf{m}}_{\mathcal{I}}\right]\left[\mathbf{m}_i - \bar{\mathbf{m}}_{\mathcal{I}}\right]^T. \tag{22}$$

For a subvolume v containing n seeds, the mean feature vector for the subvolume, denoted $\bar{\mathbf{m}}_v$, is defined as:

$$\bar{\mathbf{m}}_v = \frac{1}{n}\sum_{i=1}^{n}\mathbf{m}_i \,. \tag{23}$$

The squared Mahalanobis distance between a subvolume v and the image $\mathcal{I}$ is computed as:

$$d\left(v,\mathcal{I}\right) = \left[\bar{\mathbf{m}}_v - \bar{\mathbf{m}}_{\mathcal{I}}\right]^T \mathbf{C}_{\mathcal{I}}^{-1}\left[\bar{\mathbf{m}}_v - \bar{\mathbf{m}}_{\mathcal{I}}\right]. \tag{24}$$

Using the distance measure in (24), the representative volume v^* among k subvolumes is defined as the one that minimizes the squared Mahalanobis distance:

$$v^* = \arg\min_{v}\left[d\left(v,\mathcal{I}\right)\right], \tag{25}$$

and the corresponding optimal number of subvolumes, denoted k^*, is defined as:

$$k^* = \arg_{k}\left[v^* \in \left\{v_j\right\}_{j=1}^{k}\right] \tag{26}$$

Although intuitively the larger subvolumes should be more representative compared to the smaller ones, experimental observations (Chapter 4, Figure 22, p. 63) have shown that this is not always the case. Recall that the subvolumes are defined as the bounding boxes of the seed clusters. Increasing the number of seed clusters, k, results in smaller seed clusters, and hence, smaller subvolumes. Therefore, to find the representative subvolume v^*, the number of subvolumes k is varied until the number of seeds in the smallest seed cluster is too low, e.g., less than 12 seeds. While varying k, the subvolume v^* is updated as more representative subvolumes are found.

3.4.3 Single Representative Subvolume Parameter Selection

A representative subvolume v^* as defined in (25) is chosen between k^* subvolumes $\left\{v_j\right\}_{j=1}^{k^*}$. The subvolume v^* is then used as the input image for the segmentation-optimization framework (Figure 6, p. 9). Automatically-selected segmentation algorithm parameters for this subvolume are applied for the entire image.

Next, rather than using just one subvolume, a systematic method to use all k^* subvolumes is described.

```
Detect seed points;
Cluster seed points into $k*$ clusters;
Obtain bounding-boxes of the clusters;
Create circular array of $k*$ subvolumes $\{v_j\}_{j=1}^{k*}$;
$j \leftarrow 1, i \leftarrow 1, S \leftarrow 0$; $\xi_{next} \leftarrow \xi_{default}$; // $S$ is the counter for the stopping criteria
while $S < k*$ do
    for all subvolumes $j$ do
        if $\xi_{next} = \hat{\xi}_j$ then   // the best parameter in circulation is not new
            if $i > 1$ then   // to allow the first iteration to complete
                $S \leftarrow S + 1$;
            endif
        else
            Explicitly include $\xi_{next}$ in RRS exploration stage;
            // Execute the loop in Figure 6 with subvolume $v_j$ as input to obtain $\xi_{j,i}$;
            $\xi_{j,i} \leftarrow \arg\max_{\xi \in \Omega} Q(\xi, v_j)$;
            if $Q(\xi_{j,i}, v_j) > Q(\hat{\xi}_j, v_j)$ then // found improved parameters
                $\hat{\xi}_j \leftarrow \xi_{j,i}$;
                $\xi_{next} \leftarrow \xi_{j,i}$;
                $S \leftarrow 0$;
            else
                $S \leftarrow S + 1$;
            endif
        endif
        if $S \geq k*$ then
            exit while loop;
        endif
    endf
    $i \leftarrow i + 1$;   // next iteration of the circle
endw
Choose $\hat{\xi}_j$ such that $Q(\hat{\xi}_j, \mathcal{I}) < Q(\hat{\xi}_a, \mathcal{I})$ for all $a \neq j, a \in 1,...,k*$ to be applied to the entire image $\mathcal{I}$;
```

Figure 14. The pseudo-code for systematically obtaining the segmentation algorithm parameters using a set of subvolumes. Subscript *i* denotes the ith iteration, i.e., for each complete circle. Optimization is performed on each subvolume in a circular ring where each subvolume contributes its best parameters to the optimization step in the next subvolume. The search for segmentation parameters terminates when the subvolumes are not benefiting from each other. The last step can be skipped to produce adaptive settings.

3.4.4 Multiple-Subvolumes Coordinated Parameter Selection

The idea is to automatically obtain segmentation algorithm parameters for all $k*$ subvolumes $\left\{v_j\right\}_{j=1}^{k*}$, denoted ξ_j, in a coordinated manner and then choose between their best parameters to be applied for the entire image. Ideally the set of subvolumes $\left\{v_j\right\}_{j=1}^{k*}$ should contain the representative subvolume $v*$ found by varying k.

The pseudo-code for the coordination procedure is displayed in Figure 14. Before proceeding, note that given the same number of function evaluations per subvolume, i.e., number of iterations of the loop in Figure 6, using more subvolumes will result in an increase in computation time. Therefore, for this section, the limit on the number of function evaluations per subvolume must be reduced.

First, recall that the global optimization algorithm in Figure 6 (p. 9) performs *exploration* of the entire space of possible segmentation algorithm parameter combinations Ω and *exploitation* of the local characteristics. The RRS algorithm used in this work performs random sampling for exploration and realign-shrink random sampling for local exploitations, mainly to maintain the high search efficiency of the exploration step [119]. Search efficiency is important given the limited computing resources. During the realign-shrink procedure, after drawing a certain number of locally-confined random samples, the current local sample space is either: (i) realigned/moved if a better sample is found; or (ii) shrunk if no better solution is found.

From experimental results (Section 4.3.3, p. 68), it was the exploration step that yielded the largest amount of improvement in segmentation quality, while the contribution of the exploitation step was less significant. With this, the exploration step performed on subvolume j is followed by minimal exploitation, and the best estimate $\hat{\xi}_j$ is passed on to the next subvolume $j+1$ to be included as one of points in its exploration step. This is performed in a circular ring, i.e., $j+1=1$ if $j=k$. This terminates when no better parameters are found for all subvolumes, i.e., the subvolumes are not benefiting from each other. Then each subvolume parameters are applied to the image $\mathcal{I}$ and the one that gives the best segmentation is chosen. The last step can be skipped if the parameters are to be used adaptively within the image $\mathcal{I}$. Then, the entire image can be segmented using nearest-neighbor lookup, or if possible, interpolation, for the segmentation algorithm parameters.

3.5 Chapter Summary

We are attempting to decide and choose the optimal segmentation among the set of all *computable* segmentations—obtained by varying segmentation algorithm parameters. The optimal segmentation can be chosen using the traditional MAP criterion that minimizes the probability of error if prior probabilities on the segmentations can be specified. Since it is more practical to specify the optimal descriptive language for the segmentation and the image given the segmentation, the equivalent MDL criterion is chosen.

Since the MDL criterion was first applied to the problem of image partitioning, the vague notion of optimality is progressively made concrete from that general-purpose computer vision application to the field this work—segmentation of tube-like structures in biological images. In adapting the MDL criterion to tube segmentation, a local structure-indicator function is proposed and presented to replace the intensity value at each pixel with a neighborhood-based formulation that indicates how likely the pixel to belong to the biological structure of interest. It also implicitly captures the correlation between pixels. Two ways to evaluate the local structure-indicator function are described—the generalized log-likelihood ratio test and the vesselness formulation.

In essence, the MDL-optimality criterion in (7) incorporates the local structure and intensity models along with the local fitting strategy in the first term of (7) through the local structure-indicator function. In addition, the set of segmentations is limited to those computable by a segmentation algorithm that globally fits predefined local models.

To choose the optimal segmentation—to explore the parameter space of the segmentation algorithm—the RRS algorithm was chosen since it was developed along requirements similar to those of the segmentation-optimization framework. Given the context and wide applicability of this framework, suboptimal solutions within a reasonable time frame are acceptable. This emphasizes the need for the optimization algorithm to be efficient, i.e., capable to get closer to the global optimum quickly.

Since a single function evaluation for RRS corresponds to executing the segmentation algorithm from start to finish, execution time for 3-D images may take much longer than 2-D images. A systematic method to speed up computations in obtaining the segmentation algorithm parameters for 3-D images is described using non-empty image subvolumes. The notion of representative subvolume is defined using image intensity information and structural information obtained from the local model fitting stage of a segmentation algorithm. Two approaches on using the subvolumes to obtain the parameters for the entire image are presented, either using one representative subvolume or using a set of subvolumes containing the representative subvolume. When a set of subvolumes is used, a procedure to coordinate the parameter selection process is presented where the subvolumes collaborate in a circular array until they no longer benefit from each other. Rather than choosing between the subvolume segmentation algorithm parameters to be applied to the entire image, these parameters can be used adaptively by the segmentation algorithm. This enables adaptation of segmentation algorithm parameters originally designed for the entire image (called as the non-adaptive parameters), effectively making the segmentation algorithm spatially adaptive.

4. Experimental Results

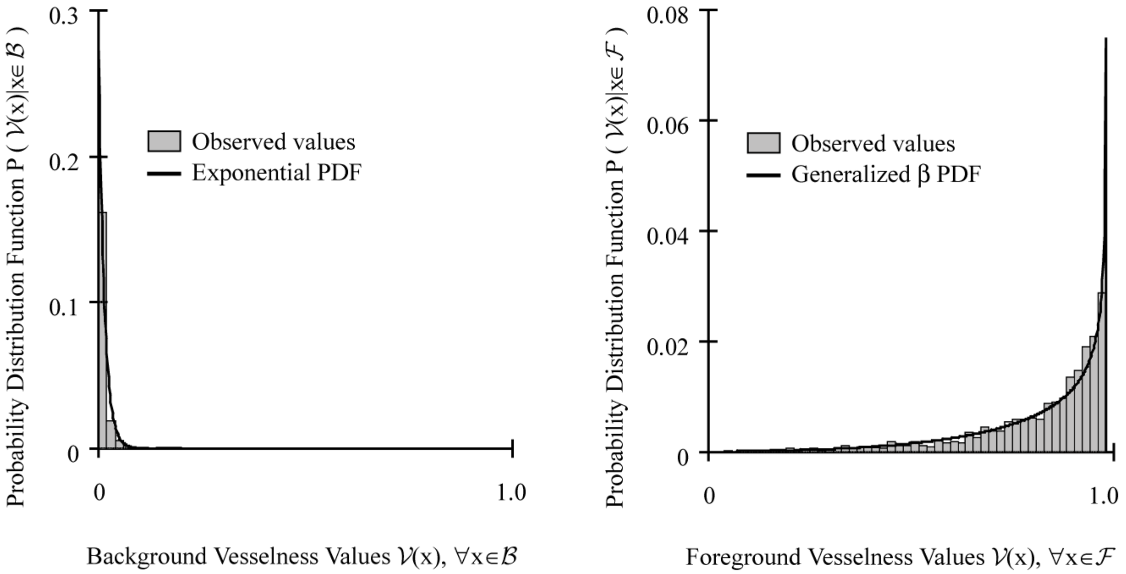
This chapter provides a series of examples of progressively increasing complexity, starting with the simplest case of segmentation by global image thresholding [60]. The thresholding example is followed by a neuron tracing algorithm [2, 3] in which just two parameters out of eight are optimized, and concluded with a full eight-dimensional parameter search. In all examples, $\alpha = 0.5$ is chosen for pure MDL-balanced optimization. Also in all examples: (i) the size scales are manually obtained from the image content; and, (ii) the parameters for the vesselness measure are set to the values published in [74].

The vesselness measure described in Section 3.1.4.2 (p. 31) is chosen as the local structure-indicator function (see Sections 3.1.3-3.1.4, pp. 26-27) for all presented results since it is readily generalized to both 2-D and 3-D images [74], unlike the currently 2-D [22] generalized log-likelihood ratio method in Section 3.1.4.1 (p. 28). The vesselness measure also exhibits faster computation speed compared to the generalized log-likelihood ratio method since the latter requires more computations per pixel [22, 74].

To evaluate the segmentation quality metric Q (Section 3.1.5, p. 33) using the description length in (16), the probability distribution function (pdf) of the vesselness values are estimated using 20 ground truth segmentations from the Digital Retinal Images for Vessel Extraction (DRIVE) database [136] and used in generating all presented results. Figure 15 displays empirical and best-fit pdf of the vesselness values at the background regions $\mathcal{B}$ in Panel (a) and at the foreground regions $\mathcal{F}$ in Panel (b). Ranked by the Kolmogorov-Smirnov (KS) test statistic (e.g., see [137]), the exponential distribution for $\mathcal{B}$ (KS value 0.27) and the generalized-beta distribution for $\mathcal{F}$ (KS value 0.05) were determined to be the best fit out of 15 distributions considered. Table III and Table IV list the detailed fitting results. Values of KS statistic close to 0 suggest accepting the hypothesized distribution and values close to 1 suggest rejecting it. The parameters of the best-fitted distributions were obtained using maximum-likelihood estimation.

Due to the limited scope to tube-like structures, the tracing algorithm's soma detection module was disabled. Therefore, somas in neuron images were manually segmented and supplied to the tracing algorithm. Furthermore, the trace merging step of the algorithm was also disabled since it is a post-processing step [87].

The recursive random search (RRS) algorithm was limited to 1000 function evaluations. This computational budget/limit was chosen empirically. An exhaustive search—without constraints on the parameter values—would require 6,804,000 function evaluations using the parameter ranges in Table II (p. 37).



(a) (b)

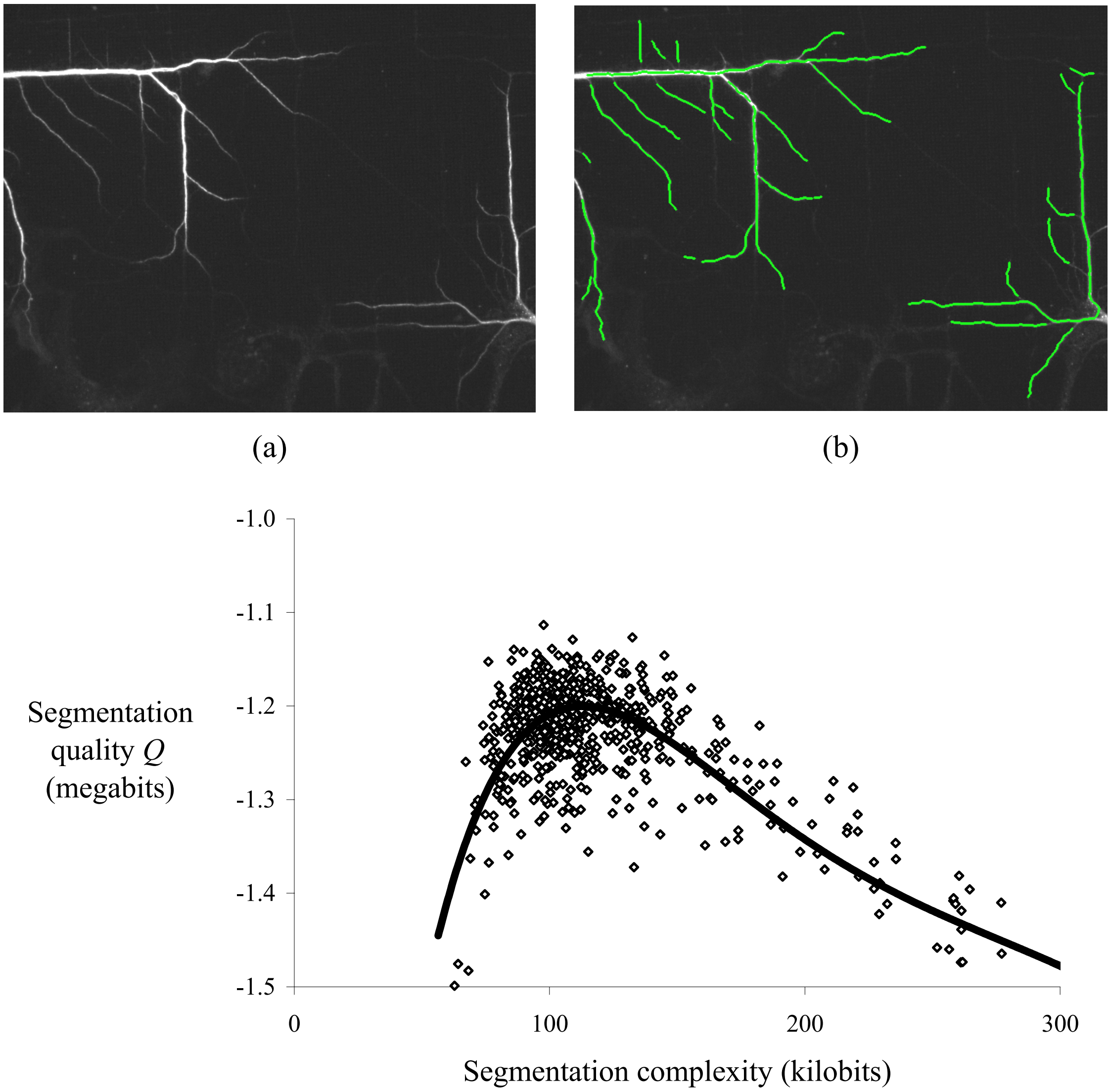
Figure 15. Displays empirical and best-fit probability distribution function (pdf) of the vesselness values at the background regions $\mathcal{B}$ in Panel (a) and at the foreground regions $\mathcal{F}$ in Panel (b). Ranked by the Kolmogorov-Smirnov (KS) test statistic, the exponential distribution for $\mathcal{B}$ (KS value 0.27) and the generalized-beta distribution for $\mathcal{F}$ (KS value 0.05) were determined to be the best fit out of 15 distributions considered. The parameters of the best-fitted distributions were obtained using maximum-likelihood estimation.

Table III. Lists the results for fitting distributions to the observed histogram of the local structure-indicator function (vesselness) values at the background region, $P\left(\mathcal{V}(\mathbf{x}) \middle| \mathbf{x} \in \mathcal{B}\right)$, where the best fit is displayed in Figure 15a. Only valid fits are listed. There were 8 invalid fits. The Kolmogorov-Smirnov goodness-of-fit measure was used and fits were ranked according to the KS statistic.

Distribution	Fitting Rank	KS Statistic	Fitted Distribution Parameters $\boldsymbol{\theta}$
Exponential	1	0.2698	$\boldsymbol{\theta} = (0.0142)$
Extreme value	2	0.3224	$\boldsymbol{\theta} = (0.0062 \quad 0.0106)$
Logistic	3	0.3759	$\boldsymbol{\theta} = (0.0090 \quad 0.0094)$
Inverse Gaussian	4	0.3980	$\boldsymbol{\theta} = (0.0148 \quad 0.0019)$
Normal	5	0.4246	$\boldsymbol{\theta} = (0.0142 \quad 0.0313)$
Triangular	6	0.7847	$\boldsymbol{\theta} = (0.0000 \quad 0.0000 \quad 0.5061)$
Uniform	7	0.8578	$\boldsymbol{\theta} = (0.0000 \quad 0.5061)$

Table IV. Lists the results for fitting distributions to the observed histogram of the local structure-indicator function (vesselness) values at the foreground region, $P\left(\mathcal{V}(\mathbf{x}) \middle| \mathbf{x} \in \mathcal{F}\right)$, where the best fit is displayed in Figure 15b. Only valid fits are listed. There were 7 invalid fits. The Kolmogorov-Smirnov goodness-of-fit measure was used and fits were ranked according to the KS statistic.

Distribution	Fitting Rank	KS Statistic	Fitted Distribution Parameters $\boldsymbol{\theta}$
Generalized beta	1	0.0490	$\boldsymbol{\theta} = (4.7771 \quad 0.6462 \quad -0.4354 \quad 1.0000)$
Normal	2	0.1791	$\boldsymbol{\theta} = (0.8302 \quad 0.1848)$
Logistic	3	0.1857	$\boldsymbol{\theta} = (0.8615 \quad 0.0937)$
Extreme value	4	0.2811	$\boldsymbol{\theta} = (0.7254 \quad 0.2477)$
Triangular	5	0.3366	$\boldsymbol{\theta} = (0.0188 \quad 1.0000 \quad 1.0000)$
Exponential	6	0.4033	$\boldsymbol{\theta} = (0.8106)$
Uniform	7	0.5153	$\boldsymbol{\theta} = (0.0191 \quad 1.0000)$
Pareto	8	0.5156	$\boldsymbol{\theta} = (0.2698 \quad 0.0196)$



(a) (b)

(c)

Figure 16. Displays a plot of segmentation quality versus segmentation complexity. Increased segmentation complexity means less concise segmentation, and vice-versa. (a) The input image, containing segments of neurons grown on topographically-modified semiconductor surface (Image source: N. Dowell, Wadsworth Center). (b) Traces obtained with 1000 RRS function evaluations (18.5% improvement in segmentation quality metric Q versus using default segmentation algorithm settings). (c) The plot of the segmentation quality Q in megabits versus the segmentation complexity $\left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right|$ in kilobits, obtained from 1000 RRS function evaluations. A polynomial trend line is also plotted. As expected, the segmentation quality increases as segmentation complexity increases up until a point where segmentation quality decreases since the segmentation becomes overly complex.

Segmentation quality is expected to increase as the segmentation gets more complex up until a point of diminishing return where more complex segmentations exhibit degrading quality. Figure 16 displays a plot of segmentation quality versus segmentation complexity. Increased segmentation complexity means less concise segmentation, and vice-versa. Panel (a) displays the input image, containing segments of neurons grown on topographically-modified semiconductor surface. (b) The optimal trace obtained with 1000 RRS function evaluations. (c) The plot of the segmentation quality Q in megabits versus the segmentation complexity $\left|\mathcal{L}_m\left(\mathcal{M}_i\right)\right|$ in kilobits, obtained from 1000 RRS function evaluations. A polynomial trend line is also plotted. As expected, the segmentation quality increases as segmentation complexity increases, up until a point where segmentation quality decreases since the segmentation becomes overly complex.

4.1 Global Thresholding Example

This is intended to illustrate the automated parameter selection method in a trivial and readily-understood context – global intensity thresholding [60]. A global intensity threshold τ is applied to images containing tube-like structures, and the goal is to find the optimal threshold value $\hat{\tau}$ that yields the optimal segmentation $\hat{\mathcal{M}}$. For objects brighter than the background, the segmentation function is given by:

$$f\left(\mathcal{I}\left(\mathbf{x}\right),\xi\right)=\begin{cases}\mathbf{x}\in\mathcal{F} & if\,\mathcal{I}\left(\mathbf{x}\right)\geq\tau\\ \mathbf{x}\in\mathcal{B} & otherwise\end{cases}. \tag{27}$$

For 8-bit grayscale images, $\mathcal{I}\left(\mathbf{x}\right)\in\left[0,255\right]$, the global intensity threshold $\xi=\tau\in\left[0,255\right]$ is a one-dimensional parameter vector in the parameter space $\Omega\in\mathbb{Z}^1$. Figure 17 shows the results of applying the proposed methodology to this case. Panel (a) shows an image of neuronal dendrites captured by fluorescence microscopy. The vesselness measure $\mathcal{V}(\mathbf{x})$ is displayed in Panel (b). Panel (c) is a plot of the segmentation quality metric Q versus the threshold value τ. The optimal segmentation $\hat{\mathcal{M}}$ (using the optimal threshold value $\hat{\tau}=55$) is shown in Panel (d).

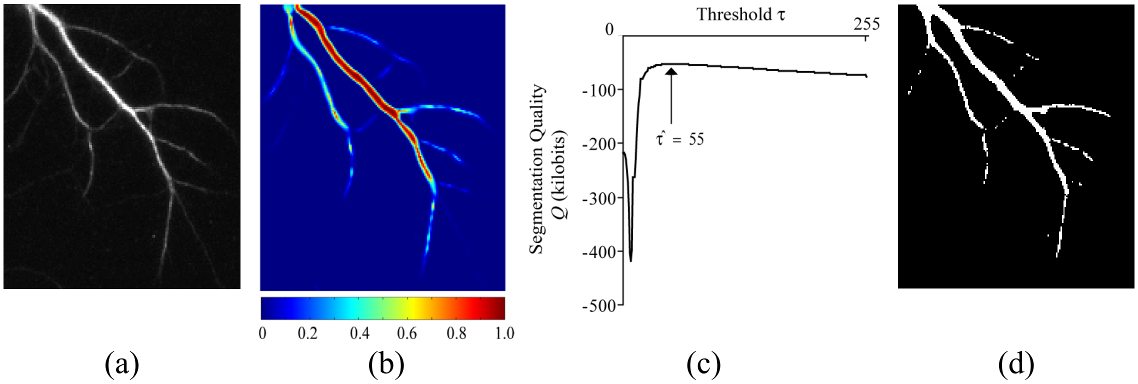


Figure 17. Trivial automatic thresholding example illustrating the behavior of the segmentation quality metric Q. (a) Image of a fluorescently labeled neurite captured by a widefield microscope. (b) The multi-scale vesselness measure. (c) Plot of the segmentation quality metric Q against the threshold τ value. (c) The optimal segmentation $\hat{\mathcal{M}}$ using $\hat{\tau} = 55$.

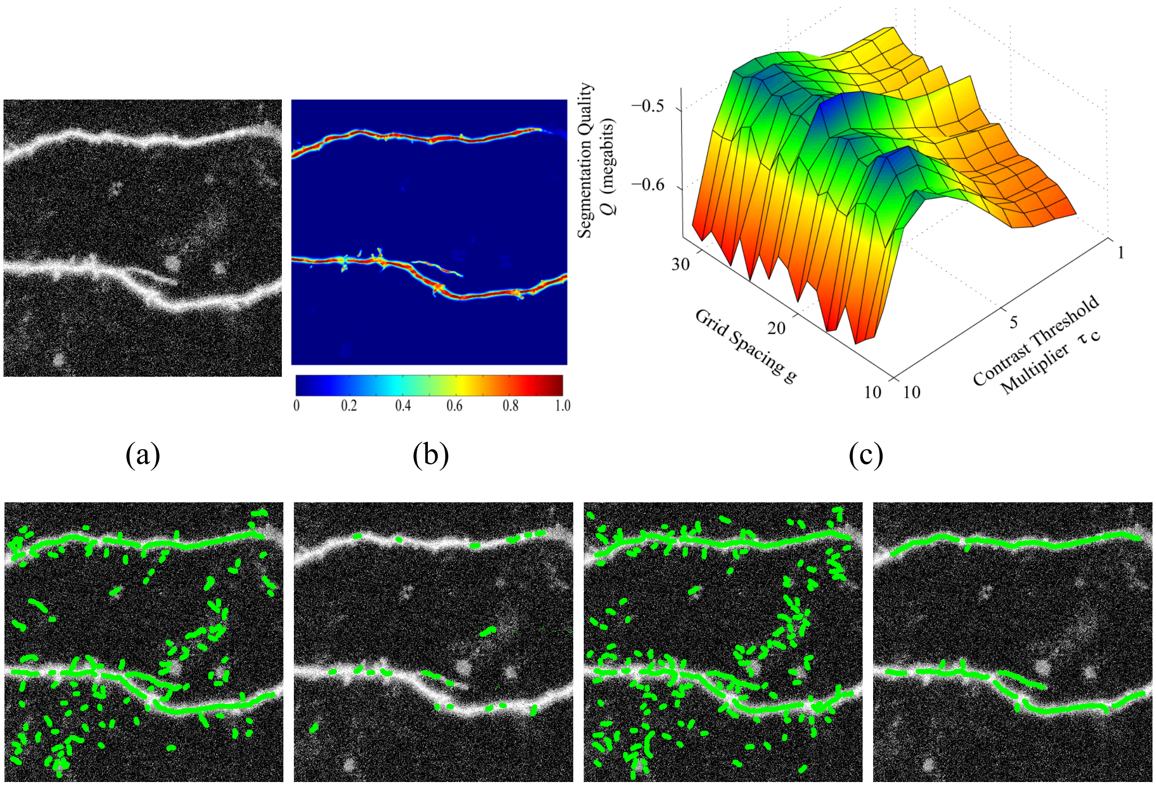
4.2 Tracing Algorithm: Two Parameters

In this example, the use of the proposed methodology is demonstrated in choosing the parameters $\hat{\boldsymbol{\xi}}$ for an automated neuron tracing algorithm [2, 3] (summarized in Section 3.3, p. 37). The parameter settings for this algorithm can be combined into a vector $\boldsymbol{\xi}$ as follows:

$$\boldsymbol{\xi} = \begin{pmatrix} g & L_{\min} & L_{\max} & n_{shift} & n_{rotate} & s_{\max} & \tau_c & \nu \end{pmatrix}.$$

They are summarized in Table II (p. 37), along with their default values. Note that the traces are not merged for these examples since the segments are merged after tracing by the tracing software, i.e., as a post-processing step.

Figure 18 shows the result of an exhaustive search for just two parameters, grid spacing g in the range $[10, 30]$, and contrast threshold multiplier τ_c in the range $[1, 10]$ with other parameters set at default values (Table II, p. 37). Panel (a) shows the input image containing fluorescently-labeled neurites imaged using a multi-photon microscope. Panel (b) shows the computed vesselness values. Panel (c) displays the entire segmentation quality metric Q versus g and τ_c. Panel (d) displays the traces obtained using the default parameter values $(g = 15, \tau_c = 3)$. The worst under-segmentation $(g = 26, \tau_c = 10)$, is displayed in panel (e). Panel (f) displays the worst over-segmentation $(g = 10, \tau_c = 1)$. The optimal segmentation $(g = 21, \tau_c = 8)$ is displayed in panel (g). Since only two parameters are being searched here, it is feasible to search exhaustively. These two parameters are related directly to the initial sampling of the image and the stopping criteria of the tracing algorithm. They are chosen to illustrate the effect of these aspects of the tracing algorithm on trace conciseness and coverage.



(a) (b) (c)

(d) (e) (f) (g)

Figure 18. An example varying just two parameters, $g \in [10, 30]$ and $\tau_c \in [1, 10]$, with others fixed at default values (Table II). Panel (a) shows a maximum intensity projection of a 512×512×55 multi-photon microscope image of fluorescently-labeled neurites. Panel (b) displays the vesselness measure. Panel (c) displays the entire segmentation quality metric Q, demonstrating the nontrivial optimization landscape, versus the parameters g and τ_c. (d) Using the default parameter values $(g = 15, \tau_c = 3)$. (e) The worst under-segmentation $(g = 26, \tau_c = 10)$. (f) The worst over-segmentation $(g = 10, \tau_c = 1)$. (g) The optimal segmentation $(g = 21, \tau_c = 8)$.

4.3 Tracing Algorithm: Eight Parameters

This example shows the search in the full eight-dimensional parameter space of the tracing algorithm, using at most 1000 RRS function evaluations. To begin, 40 2-D test images of human retinas (with ground truth), and 183 2-D images of neurites were gathered from four sources (Table V). Results for 3-D images are reported separately in Section 4.3.2. For the retinal images, the pixel-wise agreement to the ground truth was also computed (See Section 4.4, p. 70). Improvements in description lengths and agreement with the ground truth were found to be strongly correlated ($r = 0.78$), and statistically significant ($p < 0.0005$ to reject the hypothesis $r = 0$). All reported improvements in this section are in terms of the segmentation quality metric Q compared to using the algorithm's default parameter settings (Table II, p. 37).

Figure 19 displays the applications of the proposed method to a human retinal vasculature image shown in panel (a) and to images of cultured neurons in panels (b)-(d). Traces using default settings are shown in panels (e)-(h). Traces using automatically-selected settings using 1000 RRS function evaluations are shown in panels (i) through (l). The quality improvement is 4% for the retina image in panel (a), 6% for the neuron image in panel (b), 7% for the neuron image in panel (c), and 38% for the neuron image on the micro-fabricated surface in panel (d).

4.3.1 Segmentation of Image Batches

Table V summarizes the results for all 223 test images. The first column lists the image source followed by the number of images in the second column. The third column shows the improvements in segmentation quality metric Q when the parameter settings are automatically-selected for each image using the parameter selection method presented in this work. The fourth column shows the improvements when automatically-selected settings for the first image from the same source are applied to the rest of the images. Reported improvements are statistically significant ($p < 0.0005$) as concluded from paired *t*-tests on all 223 test images (see Section 4.4.2, p.72).

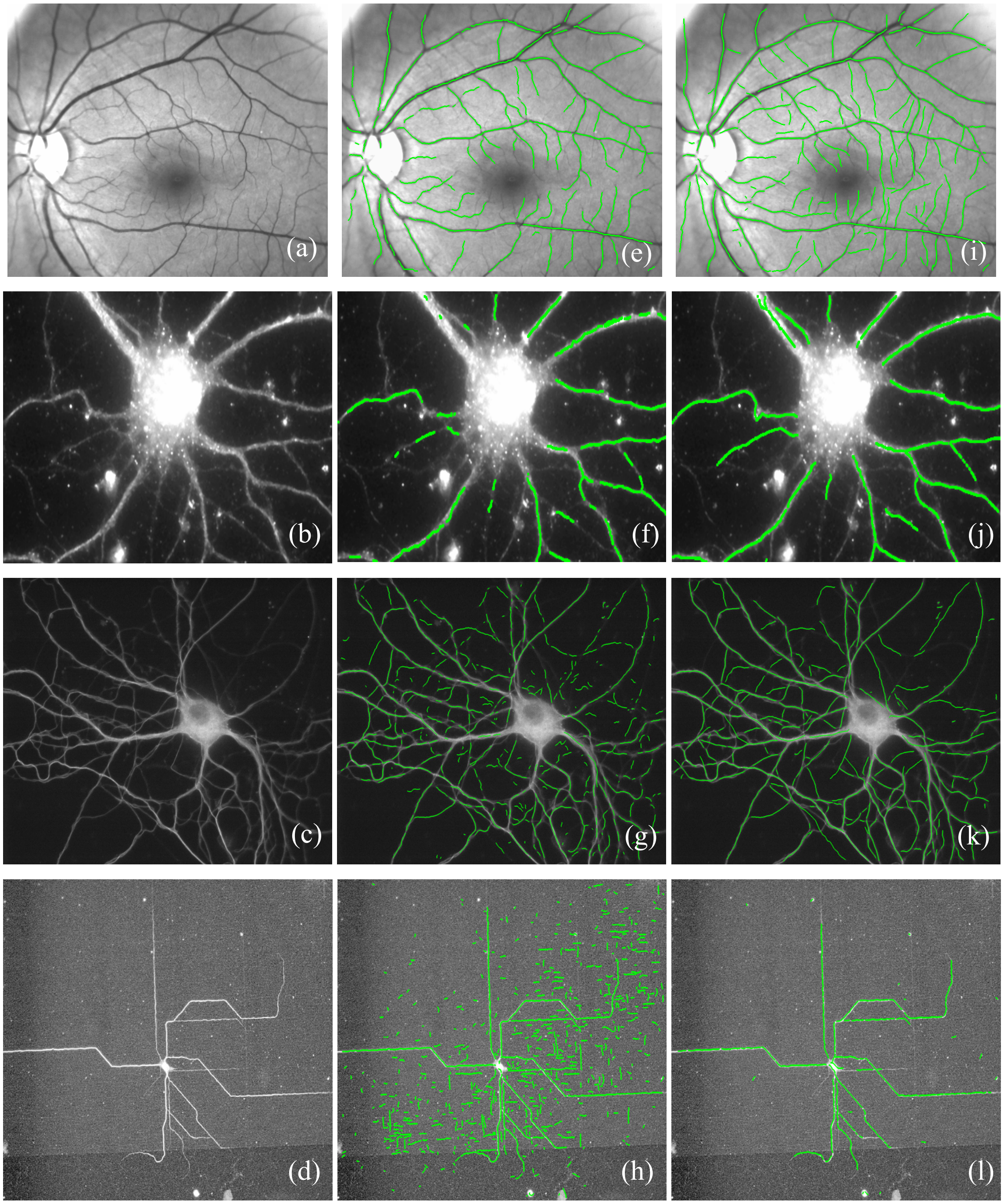


Figure 19. Illustrating applications of the proposed method to: (a) a human retinal vasculature image; (b)-(d) images of cultured neurons. Default traces are shown in panels (e) through (h). Traces using automatically-selected settings using 1000 RRS function evaluations are shown in panels (i) through (l). The quality improvement is 4% for the retina image in panel (a), 6% for the neuron image in panel (b), 7% for the neuron image in panel (c), and 38% for the neuron image on the topographically-modified surface in panel (d).

Table V. Summary of experimental results with 223 images from four sources. The first two columns list the image sources, and number of images. For all experiments, 1000 RRS function evaluations were used, and an 8-dimensional parameter space was searched. The third column shows the improvements in the segmentation quality metric Q when parameter settings are automatically-selected for each image. The fourth column shows the improvements when settings are optimized for just one randomly selected image, and then applied to the rest of the images in the batch.

Image Source and Reference	Number of Images	Improvement in segmentation quality metric Q with 1000 RRS function evaluations, compared to default settings	
		Improvement when each image is optimized (%)	Improvement when just one image is optimized (%)
Directionality of Neurite Outgrowth Study (Neuron)	154	21.3	18.5
The STARE Project (Retina)	20	4.7	1.3
The DRIVE Database (Retina)	20	9.0	4.1
Synaptic Distribution Study (Neuron)	29	10.9	10.0

Figure 20 graphically displays the segmentation quality improvements on the batch of images from the Synaptic Distribution Study [138] as summarized in Table V. The fourth image from this batch is shown in Figure 19b. The shaded bars show the percentage improvements in the segmentation quality metric when parameter settings automatically selected for each image. Twelve images were found to share the same settings after 1000 function evaluations. The blank bars show the percentage improvements when settings are optimized for just one randomly selected image, and then applied to the rest of the images in the batch.

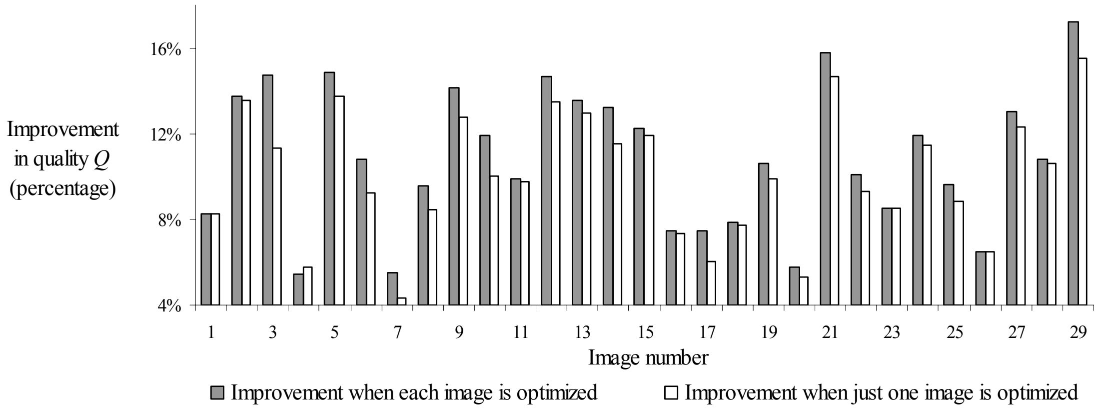


Figure 20. Displays the segmentation quality improvements on the batch of images from the Synaptic Distribution Study [138] as summarized in Table V. The fourth image from this batch is shown in Figure 19b. The shaded bars show the percentage improvements in the segmentation quality metric when optimal parameter settings are computed for each image. The blank bars show the percentage improvements when settings are optimized for just one randomly selected image, and then applied to the rest of the images in the batch.

Table VI. The automatically-selected parameter vector $\hat{\xi} = \left(\hat{g} \quad \hat{L}_{\min} \quad \hat{L}_{\max} \quad \hat{n}_{shift} \quad \hat{n}_{rotate} \quad \hat{s}_{\max} \quad \hat{\rho} \quad \hat{v} \right)$ values specific to each image. Notice that only 11 distinct parameter vectors were obtained.

Image number	$\hat{g}$	$\hat{L}_{\min}$	$\hat{L}_{\max}$	$\hat{n}_{shift}$	$\hat{n}_{rotate}$	$\hat{s}_{\max}$	$\hat{\rho}$	$\hat{v}$
1,4,5,6,7,8,9,16,19,23,25,28	2	10	14	3	3	5	2	10
2,13,15,21,24	2	16	28	3	3	5	3	10
3	2	10	14	3	5	6	1	10
10	2	12	14	3	5	4	2	15
11,12,26	2	10	18	3	4	4	1	15
14	2	20	24	3	5	4	3	15
17,20	2	16	26	3	4	3	3	15
18	2	10	24	3	6	5	2	10
22	2	8	14	3	9	4	1	10
27	2	10	28	3	7	7	2	10
29	2	10	22	3	6	7	2	10

Table VI displays the values of the automatically selected parameter vector $\hat{\xi} = \left(\hat{g} \quad \hat{L}_{\min} \quad \hat{L}_{\max} \quad \hat{n}_{shift} \quad \hat{n}_{rotate} \quad \hat{s}_{\max} \quad \hat{\rho} \quad \hat{v} \right)$ specific to each image from the Synaptic Distribution Study [138] as summarized in Table V. Only 11 distinct parameter vectors were obtained. In one case, 12 out of 29 images in the set resulted in the same optimal parameters. These images are displayed in Figure 4b. This may suggest that these 12 images exhibit similar underlying characteristics from which the parameters are derived, i.e., they "look similar" to the tracing algorithm, given that their automatically selected parameters are identical.

The optimized parameter settings vary from one application to another. For example, when the parameter settings from the DRIVE database containing retinal images are applied to the neuron images from the Synaptic Distribution Study, the observed segmentation quality metric Q worsens 30% on average.

4.3.2 Using 3-D Images

A total of 22 3-D images are used. To create image subvolumes, seed points are generated using the default segmentation algorithm settings (Table II, p. 37) and clustered. The bounding boxes of the seed clusters define the subvolumes. Computation times to execute 1000 RRS function evaluations for different non-empty subvolume sizes obtained with seed clusters are plotted in Figure 21. The resulting plot seems clustered too, where the right-most cluster corresponds to using 1 seed cluster, the middle cluster corresponds to using 2 seed clusters, and the left-most cluster corresponds to using between 3 to 10 seed clusters. Execution times and subvolume sizes are found to be almost-perfectly linearly related (correlation coefficient value $r = 0.98$). This motivated the subvolume-based speedup methods as described in Section 3.4 (p. 42).

The two subvolume-based methods in Section 3.4 are compared by observing both speedup and improvements in segmentation quality metric Q. The optimal number of subvolumes k^* is found to vary between 2 to 5 for the test images. Parameter selection performed on the entire image averages at 61 minutes and improvements in segmentation quality metric Q averages at 1.30%. The single representative subvolume method (Section 3.4.3, p 45) produces on average 3.1X speedup with 1.07% average improvement in segmentation quality metric Q. The multiple-subvolumes coordinated parameter selection method (Section 3.4.4 p. 47) produces comparable average segmentation quality improvement of 1.05%, but with much higher speedup of 11.2X on average.

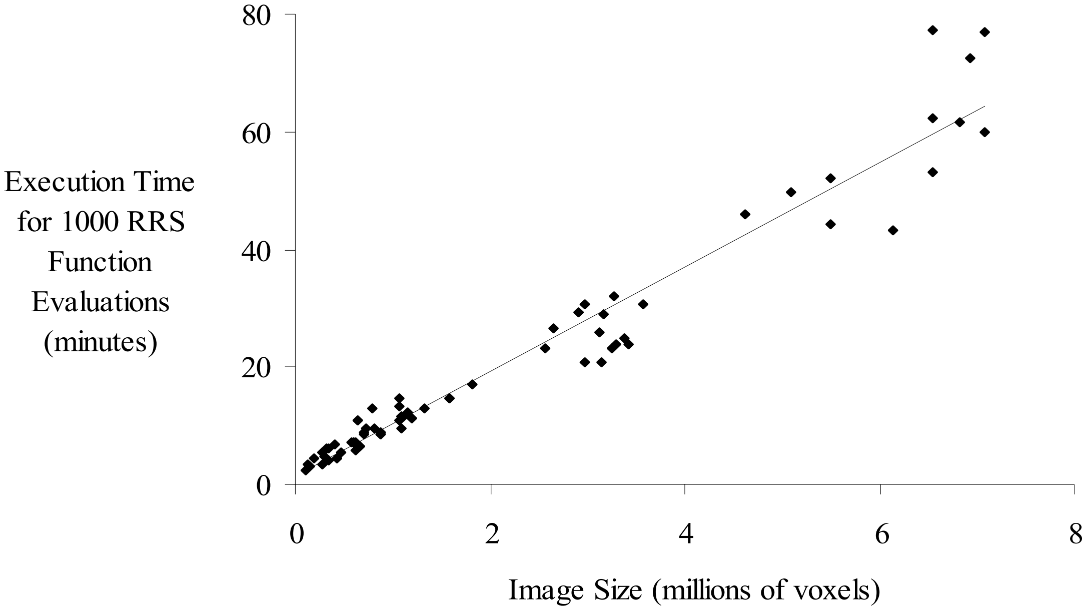


Figure 21. Displays execution times for the automated parameter selection method using 1000 RRS function evaluations and non-empty 3-D subvolumes of different sizes. A linear trend line is also plotted. Execution time and image size is almost-perfectly linearly related (correlation coefficient $r = 0.98$).

Based on this observation, the multiple-subvolumes coordinated parameter selection is chosen to speed up the parameter selection process for 3-D images and also to investigate the results of using spatially adaptive segmentation algorithm parameters. Using this speedup method, 65 RRS function evaluations are performed on each subvolume before the best parameters are passed on to the next subvolume that will conduct another 65 RRS function evaluations. The 99% confidence-level used by RRS limits the number of function evaluations in the exploration step to 44. The number of function evaluations for RRS exploitation stage is limited to 21, corresponding to 3 realign-shrink steps with 7 function evaluations each [119].

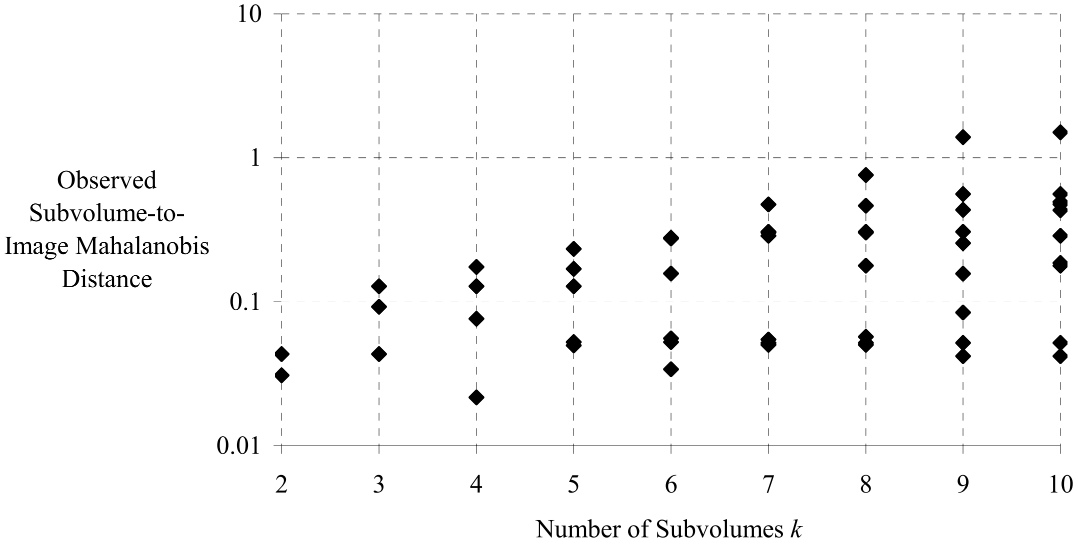


Figure 22. Displays the observed Mahalanobis distance for k subvolumes corresponding to k seed clusters, where the optimal number of subvolumes is $k^* = 4$ since it contains the subvolume v^* with the smallest Mahalanobis distance to the entire image. The input image is shown in Figure 24a. Beyond 10 clusters, the smallest seed cluster contains less than 12 seeds. Notice that larger subvolume sizes (smaller k) do not necessarily correspond to smaller distance to the entire image.

The optimal number of subvolumes k^* is determined by varying the number of seed clusters k until the smallest seed cluster contains less than 12 seeds. While varying k, the representative subvolume v^* is updated. Once the process completes, the optimal number of subvolumes k^* corresponds to the cardinality of the set of subvolumes $\left\{v_j\right\}_{j=1}^{k^*}$ that contains v^*. Recall that v^* is defined as the closest subvolume to the entire image based on the Mahalanobis distance of the seed point characteristics and image contrast. Figure 22 displays the observed Mahalanobis distance for k subvolumes as k is varied using the 3-D image shown later in Figure 24a, where $k^* = 4$. Using the multiple-subvolume coordinated parameter selection method for the same input image, Figure 23 displays the improvements in segmentation quality metric Q using different number of subvolumes. Most improvement is observed when the set of subvolumes contain the most representative subvolume v^*. In this case, it is when $k^* = 4$.

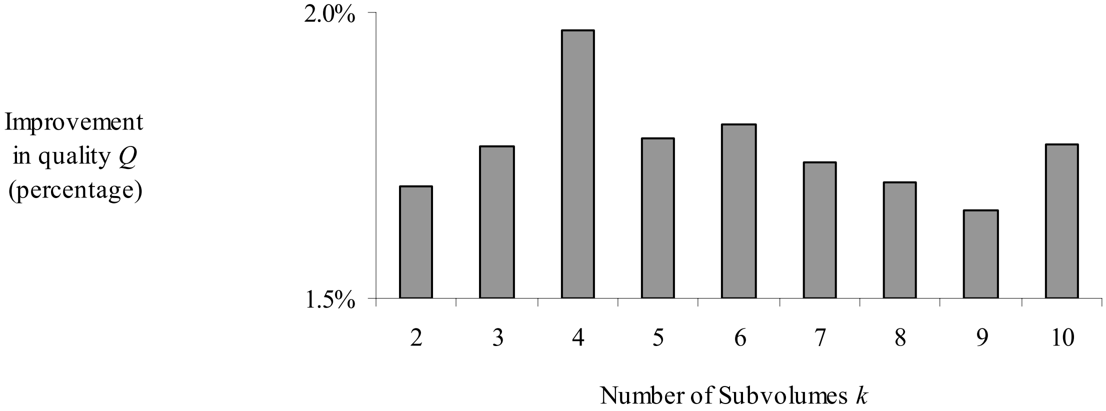
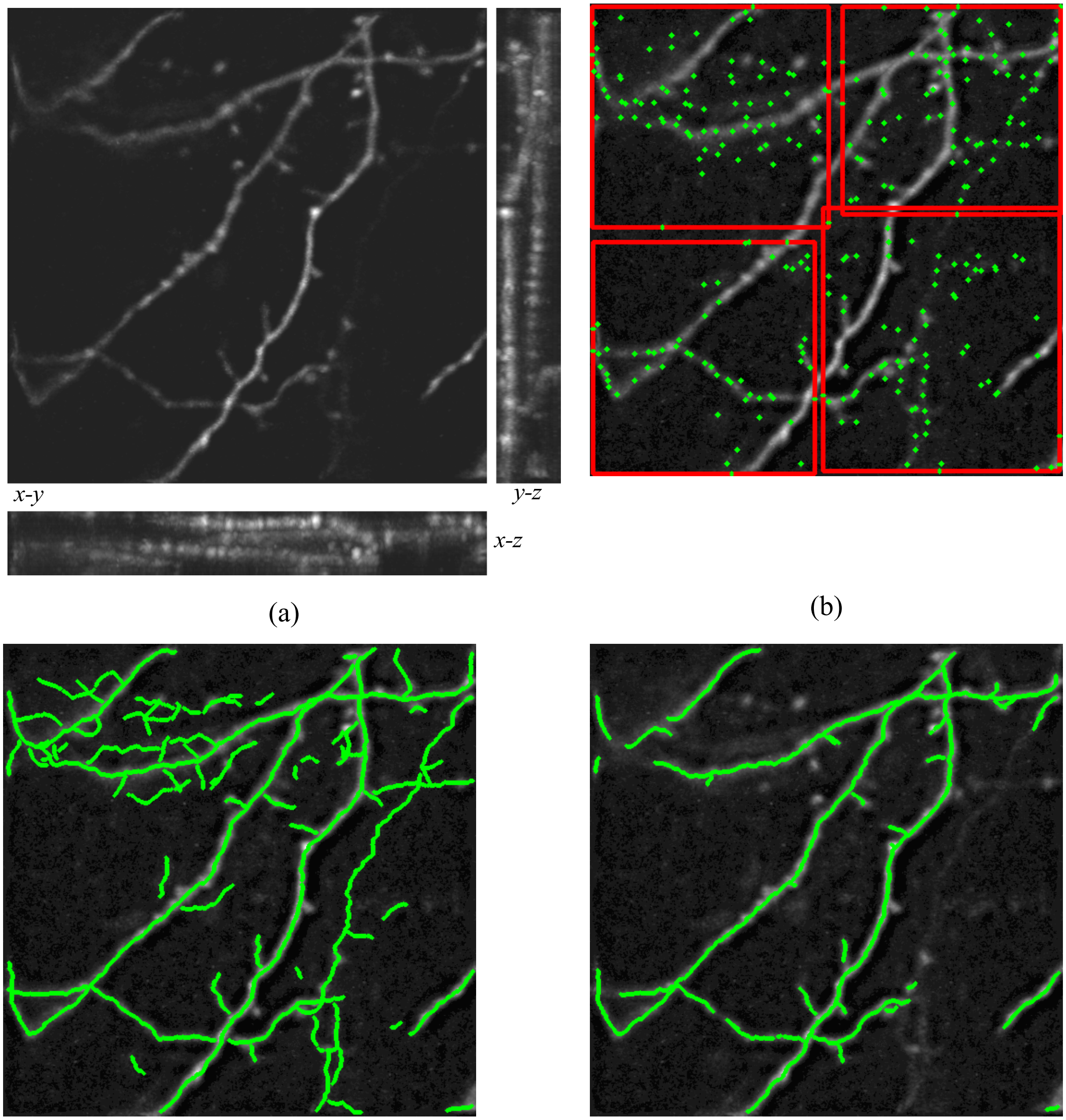


Figure 23. Displays the amount of improvement in segmentation quality Q using different number of subvolumes using the multiple-subvolumes coordinated parameter selection method. Notice that most improvement occurs at $k^*=4$ when the set of subvolumes contain the most representative subvolume v^*.

Figure 24 displays a 512×512×50 image of a neuron captured using multiphoton microscopy. For this image, the number of seed clusters is computed to be $k^*=4$. Panel (b) displays the detected seed points on the maximum intensity projection of the 3-D image along with the subvolumes obtained from seed point clusters. Panel (c) shows the traces obtained using default segmentation algorithm settings. Panel (d) shows the traces obtained using the automatically-selected (non-adaptive) settings, with 1.92% improvement in segmentation quality Q. Recall that by default, the segmentation algorithm settings are non-adaptive since they are used for the entire image. When using adaptive settings, improvement in segmentation quality metric Q is almost similar at 1.90%. Computational speedup using multiple subvolumes versus using the entire image averages at 12.8X (average reduction from 77 minutes to 6 minutes) on a 2 GHz AMD Opteron processor. More detailed progress of the multiple-subvolumes coordinated parameter selection method is described next.



(a) (b)

(c) (d)

Figure 24. Results using a 512×512×50 3-D image of a neuron captured using multiphoton microscopy. (Image source: J. Trachtenberg, UCLA). (b) Detected seed points on the maximum intensity projection of the 3-D image along with the four regions obtained by clustering the seed points. (c) Traces obtained using default segmentation algorithm settings. (d) Traces obtained using the automated parameter selection method (coordinated multiple-subvolumes). The improvement in segmentation quality metric Q is 1.92%.

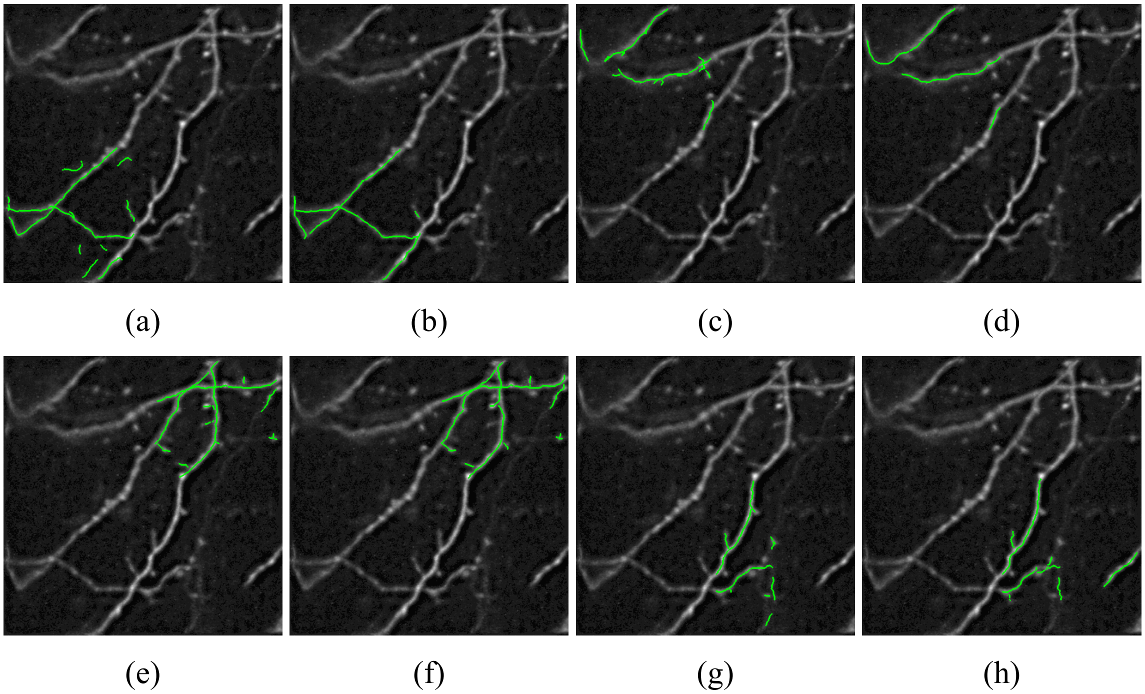


Figure 25. Multiple-subvolumes coordinated parameter selection, during the first iteration. The input 512×512×50 3-D image is displayed in Figure 24a. Results are displayed on the *x-y* maximum intensity projection image. Panel (a) displays the first subvolume. It begins with the default parameter settings, and searches for its best parameters $\xi_{j=1}$, where the result is shown in panel (b). Panel (c) displays the traces for the second subvolume, using $\xi_{j=1}$. It includes $\xi_{j=1}$ as part of its exploration stage, and obtains the traces with its updated parameters $\xi_{j=2}$ in panel (d). Panel (e) displays the traces in the third subvolume with $\xi_{j=2}$. Its updated traces using $\xi_{j=3}$ is shown in panel (f). Similarly, panel (g) displays the traces in the fourth subvolume with $\xi_{j=3}$. Its updated traces using $\xi_{j=4}$ is shown in panel (f). Next, the first subvolume will include $\xi_{j=4}$ in its exploration stage. Note that each subvolume may not always obtain updated/better parameters. The process terminates when all subvolumes no longer produce better parameters.

For the input 512×512×50 3-D image displayed in Figure 24a, observed progress of the multiple-subvolumes coordinated parameter selection during the first iteration is displayed in Figure 25. Recall that subscript j denotes the subvolume index and subscript i denotes the iteration number. Panel (a) displays the first subvolume. It begins with the default parameter settings, and searches for its best parameters $\xi_{j=1,i=1}$, where the result is shown in panel (b). Panel (c) displays the traces for the second subvolume, using $\xi_{j=1,i=1}$. It includes $\xi_{j=1,i=1}$ as part of its exploration stage, and obtains the traces with its updated parameters $\xi_{j=2,i=1}$ in panel (d). Panel (e) displays the traces in the third subvolume with $\xi_{j=2,i=1}$. Its updated traces using $\xi_{j=3,i=1}$ is shown in panel (f). Similarly, panel (g) displays the traces in the fourth subvolume with $\xi_{j=3,i=1}$. Its updated traces using $\xi_{j=4,i=1}$ is shown in panel (f). Next, the first subvolume will include $\xi_{j=4,i=1}$ in its exploration stage. Note that each subvolume may not always obtain updated/better parameters. The process terminates when all subvolumes no longer produce better parameters.

Segmentation quality for 3-D images is also expected to increase as the segmentation gets more complex up until a point of diminishing return where more complex segmentations exhibit degrading quality, similar to the case for 2-D images (illustrated in Figure 16, p. 53). Figure 26 displays a plot of segmentation quality versus segmentation complexity for the image shown in Figure 24a, obtained from 1000 RRS function evaluations. The behavior of segmentation quality Q is found to be similar between 2-D and 3-D images.

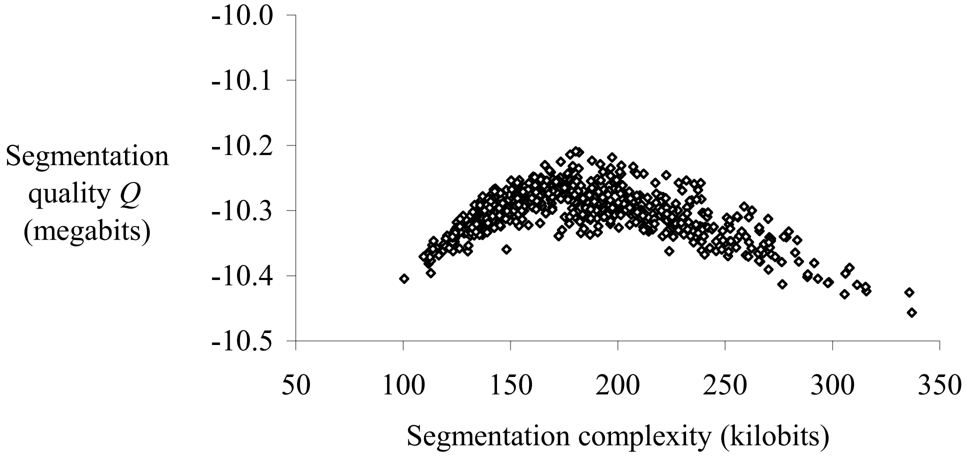


Figure 26. Displays a plot of the segmentation quality versus segmentation complexity for the image shown in Figure 24a obtained from 1000 RRS function evaluations. As seen for 2-D images (Figure 16, p. 53), the segmentation quality increases as segmentation complexity increases, up to a point until the segmentation quality decreases since the segmentation becomes overly complex.

Figure 27 displays the improvements in segmentation quality for a time-series of eight 3-D images (first image shown in Figure 24a). The first bar in Figure 27 displays the improvement when each image in the time-series is optimized separately, using $k^* = 4$ subvolumes and non-adaptive segmentation settings. The second bar displays the improvement obtained when the non-adaptive settings obtained for the first image are applied to the rest of the series. The third bar displays the improvement using the adaptive settings obtained from the first image applied adaptively to the rest of the series.

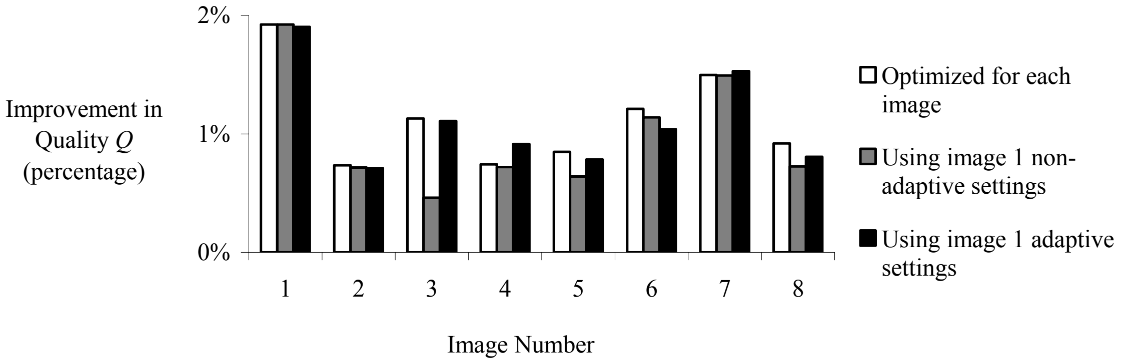


Figure 27. Displays the improvements in segmentation quality for a time-series of eight 3-D images (first image shown in Figure 24a). The first bar displays the improvement when each image is optimized separately, using $k^* = 4$ subvolumes and non-adaptive segmentation settings. The second bar displays the improvement obtained when the non-adaptive settings obtained for the first image are applied to the rest of the series. The third bar displays the improvement using the adaptive settings obtained from the first image applied adaptively to the rest of the series.

4.3.3 RRS Performance

A full eight-dimensional exhaustive search, comprising of 6,804,000 function evaluations using the parameter values in Table II (p. 37), is conducted using the image shown in Figure 18a (p. 56). The exhaustive search took 9 days to complete on three servers each with two 700 MHz Pentium III processors, a total of six processors running in parallel. On average of 50 RRS runs with 1000 function evaluations each, RRS is within 3.56% from global optimum.

The execution time is dependent on the time for one RRS function evaluations, i.e., one run of the segmentation algorithm. Execution times for 1000 RRS function evaluations to obtain the optimal parameter settings vary from 12 minutes per image on average from the DRIVE database [136] to 53 minutes on average for images from the Synaptic Distribution Study [138] on a 2GHz AMD Opteron processor.

The performance of the RRS algorithm is plotted in Figure 28, displaying the segmentation quality improvement in percentage relative to using default segmentation algorithm parameter settings, on average, for all 223 test images (Table V, p. 59). The average improvement relative to default settings increases only by 0.22% beyond 500 RRS function evaluations up to 1000 RRS function evaluations. This plot also illustrates the high efficiency of RRS during its exploration of the global parameter space. Furthermore, it shows that the default parameter settings are very unsuitable for the majority of the test images, especially from the neurite outgrowth directionality study [11] (see Table V). The minimum number of RRS function evaluations is determined by the confidence-level of finding the optimal value within the global sample space, i.e., during its exploration step. For the 99% confidence-level used for all presented examples, RRS needs 44 function evaluations [119] before it can begin identifying the promising subspaces for the exploitation step.

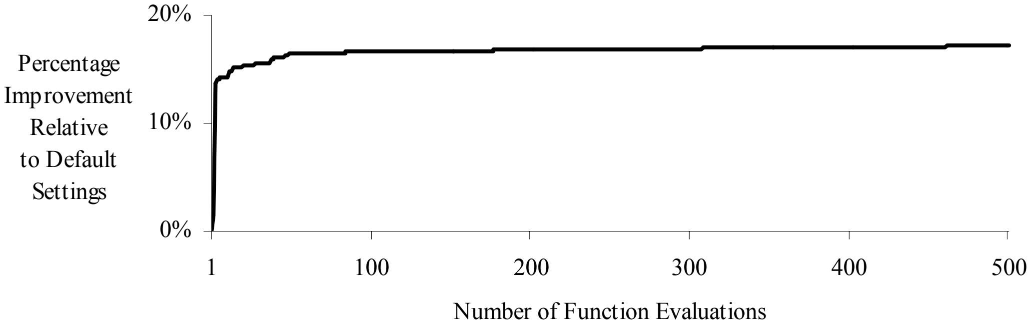


Figure 28. Average percentage improvement in segmentation quality for all 223 test images (Table V). This plot illustrates the high efficiency of RRS during its early exploration of the global parameter space. The improvement between 500 and 1000 RRS function evaluations is only 0.22%.

4.4 Validation

In the context of validating the automatically-selected segmentation algorithm parameters, the reported improvements in segmentation quality Q need to be shown to relate to improvements in terms of agreement between automated segmentation and manual segmentation. Once the relation is validated and improvement in segmentation quality Q is concluded to be meaningful, then the statistical significance of observed improvements in Q for the test images needs to be validated.

4.4.1 Relationship between Improvement in Segmentation Quality and Agreement with Ground Truth

Improvements in segmentation quality metric Q must be validated in terms of its relationship to improvements in agreement between automated segmentation $\mathcal{M}_{auto}$ and ground truth segmentation $\mathcal{M}_{GT}$. Recall that both improvements are relative to using the segmentation algorithm default settings. Currently, the widely used method to obtain the ground truth segmentation is by using human observers, i.e., manual segmentation [39, 44]. Manual segmentation, other than being tedious and very time consuming, suffers from intra-observer and inter-observer subjectivity and disagreements [3]. Since the segmentation quality metric Q is based on partition of the image, the ground truth must label all pixels in the image, not just the centerlines (as in [3]). This makes manual segmentation more tedious. Rather than introducing more subjectivity, the validation is performed on currently-available, published sets of 40 manual segmentations, from the STARE Project [65] and the DRIVE Database [136].

Let $\mathcal{M}_{auto}\left(\xi_{default}\right)$ be the automated segmentation using default segmentation algorithm settings $\xi_{default}$, and $\mathcal{M}_{auto}\left(\hat{\xi}\right)$ be the automated segmentation using automatically-selected algorithm settings $\hat{\xi}$. The number of pixels where the automated segmentation $\mathcal{M}_{auto}$ agrees with the ground truth segmentation $\mathcal{M}_{GT}$ is denoted $\left|\mathcal{M}_{auto} \cap \mathcal{M}_{GT}\right|$. The improvement in agreement between automated segmentation and ground truth segmentation becomes:

$$\Delta GT = \frac{\left|\mathcal{M}_{auto}\left(\hat{\xi}\right)\cap\mathcal{M}_{GT}\right|-\left|\mathcal{M}_{auto}\left(\xi_{default}\right)\cap\mathcal{M}_{GT}\right|}{\left|\mathcal{M}_{auto}\left(\xi_{default}\right)\cap\mathcal{M}_{GT}\right|}. \tag{28}$$

The segmentation quality metric Q of $\mathcal{M}_{auto}\left(\xi_{default}\right)$ is denoted $Q\left(\xi_{default}\right)$, and the segmentation quality metric of $\mathcal{M}_{auto}\left(\hat{\xi}\right)$ is denoted $Q\left(\hat{\xi}\right)$. Since Q is negative valued, the improvement in segmentation quality is computed as:

$$\Delta Q = \frac{Q\left(\xi_{default}\right)-Q\left(\hat{\xi}\right)}{Q\left(\xi_{default}\right)} \tag{29}$$

The (Pearson) correlation between two variables X and Y, each with N observations is:

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\left(\sum X^2\right)-\frac{\left(\sum X\right)^2}{N}\right)\left(\left(\sum Y^2\right)-\frac{\left(\sum Y\right)^2}{N}\right)}} \tag{30}$$

Substituting in (30) the variables ΔQ and ΔGT, and the value $N = 40$ for the 40 images with ground truth used in this validation section, the correlation is computed to be $r = 0.78$. A scatter plot showing observed ΔQ and ΔGT pairs is shown in Figure 29.

Next, a two-sided hypothesis test is conducted with the following hypotheses:

$$H_0 : r = 0$$
$$H_1 : r \neq 0$$

where the significance level of the test is set at 0.05. The significance level is the probability of rejecting a true H_0, i.e. concluding that there is a linear relationship between ΔQ and ΔGT where there is actually no linear relationship between the two. The p-value for this test is computed to be less than 0.0005. The p-value tells the smallest significance level of the test at which H_0 can be rejected. Since it is lower than the significance level of 0.05, hypothesis H_0 saying that there is no linear relationship between observed improvements in segmentation quality and improvements in agreement with ground truth can be rejected.

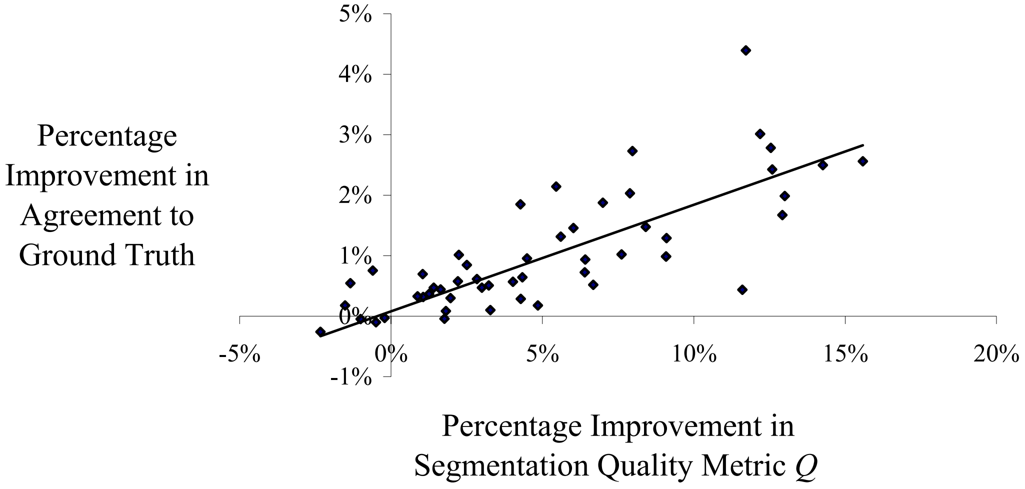


Figure 29. Displays a scatter plot of observed improvements in segmentation quality metric Q and improvements in agreement to ground truth. A linear trend line is also plotted. The improvements are strongly correlated with correlation coefficient $r = 0.78$.

4.4.2 Significance of Improvement in Segmentation Quality

Significance of the improvements observed for all 223 test images (Table V, p. 59), can be determined by conducting a paired *t*-test using the observed values for segmentation quality using default segmentation settings $Q\left(\xi_{default}\right)$ and segmentation quality using automatically-selected settings $Q\left(\hat{\xi}\right)$. Paired *t*-test is used because it is reasonable to expect dependence between observations within a pair because they are measurements from the same image.

First the difference between $Q\left(\xi_{default}\right)$ and $Q\left(\hat{\xi}\right)$ needs to be tested to see whether the distribution of the differences is normally distributed. Using the Anderson-Darling test for normality [139], we can reject the hypothesis that the differences are not normally distributed with $p < 0.005$.

The paired *t*-test is conducted using the pair of $Q\left(\xi_{default}\right)$ and $Q\left(\hat{\xi}\right)$ values for each of the 223 test images. The null hypothesis says that the difference between $Q\left(\xi_{default}\right)$

and $Q\left(\hat{\xi}\right)$ is zero, and the alternative hypothesis says that the difference is greater than zero. In other words, the null hypothesis says that the automatically-selected parameters result in no improvement in segmentation quality metric Q. From the observed pairs of $Q\left(\xi_{default}\right)$ and $Q\left(\hat{\xi}\right)$ values, the null hypothesis is rejected at the 0.05 significance level with $p < 0.0005$.

4.5 Comparison of Optimization Algorithms

Comparative performance evaluation for the optimization algorithm is conducted using four other global optimization algorithms. Other than the RRS algorithm used to generate all results reported in this work, the four evaluated optimization algorithms are: (i) controlled random search; (ii) multi-start pattern search; (iii) simulated annealing; and (iv) genetic algorithm. These algorithms are chosen since they are recommended for black-box optimization problems [114, 117, 118, 140]. They are briefly described next.

4.5.1 Optimization Algorithm Descriptions

Controlled random search, also called as Price's algorithm [113, 114, 141], randomly generates a population of n points, randomly chooses a subset of the n points and performs a downhill simplex move with that subset. The population of n points is updated by removing the worst element as better points are obtained in the downhill simplex moves. A simplex move is used to find the local optima. The simplex, a generalization of a triangle in high dimensional space, is moved using systematic reflections, expansions, and contractions [142].

Multi-start pattern search performs random sampling to generate new starting points whenever it reaches local optima [115]. Pattern search is an exploitation method that maintains a pattern (information) that guides the speed of convergence to the local optima [116]. A pattern is updated by performing a series of exploratory steps, one in each dimension of the search space [115].

Simulated annealing [112], also known as Monte Carlo annealing, statistical cooling, probabilistic hill-climbing, stochastic relaxation, and probabilistic exchange algorithm [140], is based on the concept of annealing in metallurgy (study of metals and their mixture, alloys, in materials science and materials engineering). Annealing in metallurgy is a method involving heating and controlled cooling of a material for allowing a system such as atoms in an alloy to find a low-energy configuration to increase the size of its crystals and reduce defects. To avoid local optima, simulated annealing begins with a pure random search and gradually resembles hill-climbing methods [140], where the transition from random search to hill-climbing is achieved by varying its "temperature" (e.g., see [143]).

The genetic algorithm is based on the theory of evolution, driven by the survival of the fittest [144-146]. An initial random population of individuals is created, and each individual's fitness value is a scaled version of the optimization objective function [147]. To create new generations of the population, multiple individuals are stochastically selected from the current population based on their fitness to create new members of the population, either by breeding (crossover of two parent individuals) or by mutation.

The interested reader is referred to the cited references for further details on these algorithms.

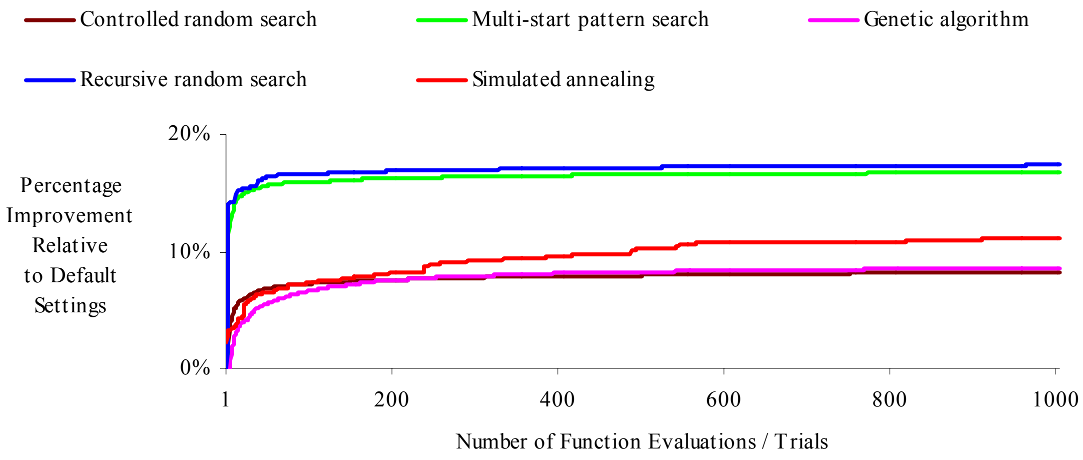


Figure 30. Average percentage improvement in segmentation quality for all 223 test images (Table V) using controlled random search, multi-start pattern search, genetic algorithm, simulated annealing, and RRS. RRS outperforms all other optimizations tested at all 1000 function evaluations.

4.5.2 Experimental Results for Comparative Study

This is similar to the result shown earlier in Figure 28 (p. 69) for RRS, but here the performances of other optimization algorithms are included. Figure 30 displays the average percentage improvement in segmentation quality for all 223 test images (Table V) using controlled random search, multi-start pattern search, simulated annealing, genetic algorithm, and RRS. This plot illustrates the relatively higher efficiency of RRS during its early exploration of the global parameter space compared to the other optimization algorithms. It also demonstrates the relatively higher efficiency of RRS during later function evaluations compared to the others. RRS outperforms the other 3 optimization algorithms at all function evaluations for all 223 test images.

4.6 Chapter Summary

Since the goal is to obtain the optimal result from a segmentation algorithm, and therefore automatically obtain the segmentation algorithm parameter settings, segmentation algorithms ranging from as simple as the global thresholding [60] to the eight-parameter exploratory tracing algorithm [3, 19] are adopted to demonstrate the applicability of the automatic parameter selection method. A gallery of results that extends across applications is displayed. These results also demonstrate that *a single* implementation/version of the tracing algorithm can be used across applications just by using different parameter settings. Capability for 3-D images is also demonstrated using a systematic method to speed up computation since incremental changes in the tracing algorithm parameters do not translate to incremental changes in the traces—each RRS function evaluation corresponds to complete retrace of an image.

The method is validated by investigating the relationship between improvements in the proposed segmentation quality metric and improvements in agreement to ground truth segmentation. The reported improvements in segmentation quality are found to be statistically significant and strongly correlated ($r = 0.78$) with improvements in agreement between automated and ground truth segmentation.

Performance of the RRS algorithm used in finding the segmentation algorithm parameters is compared with four other search algorithms generally recommended for black-box optimization: (i) controlled random search; (ii) multi-start pattern search; (iii) simulated annealing; and (iv) genetic algorithm. On average for 223 test images, RRS outperforms the other four optimization algorithms at all 1000 function evaluations.

5. Discussion and Conclusions

An automated method to select segmentation algorithm parameter settings is presented and demonstrated across applications based on the core idea of trading off segmentation conciseness and segmentation coverage. Using automatically selected parameters resulted in improved segmentations on all test images. This translates to more accurate analyses performed on the improved segmentations. Current parameter selection methods using empirically-chosen parameters (e.g., [2, 3, 19, 20]) or heuristic formulations (e.g., [21]) are not only time-consuming but are also application-specific with no assurance towards optimality.

5.1 Automated Parameter Selection

Using an optimization algorithm, possible combinations of segmentation algorithm parameter settings are efficiently searched using an objective function based on segmentation quality. This is performed in an unsupervised manner, unlike in other methods such as in reinforcement-learning segmentation systems [36]. This is possible because the segmentation quality assessment module is an online component of the automated parameter selection. Current segmentation quality assessment methods performs offline evaluations of the segmentations (e.g. [44, 46]).

5.1.1 Segmentation Quality Metric

The segmentation quality metric is based on the minimum description length (MDL) principle [32]. It objectively trades off segmentation coverage of image content and its conciseness. The optional universal parameter α allows a user to bias the trade-off. This effectively reduces the many number of parameters to one that intuitively adjusts the aggressiveness of the segmentation algorithm. Improvements in this metric are demonstrated to strongly correlate with improvements in agreement with ground truth segmentations.

5.1.2 Recursive Random Search Algorithm

Use of direct search methods is necessary to avoid excessive function evaluations needed for estimating objective function derivatives through finite differences. Direct search methods involve only examining trial solutions coupled with a strategy for the next trial [115]. This is necessary for the presented automated parameter selection method since one function evaluation, or one trial, corresponds to one complete execution of a segmentation algorithm. This becomes the computational bottleneck especially for slower segmentation algorithms involving 3-D images.

The recursive random search (RRS) optimization algorithm, used in generating all presented experimental results is shown to outperform other optimization algorithms recommended for black-box applications [114, 117, 118], such as the genetic algorithm [144-146], multi-start pattern search [115, 116], and controlled random search [113, 114, 141]. As demonstrated from experimental results, it has the highest efficiency at producing better solutions sooner than other algorithms compared, which is important for this work.

5.2 Batch Segmentation

On batches of similar images, the experimental results demonstrated that significant improvements can still be gained when the automated parameter selection method is performed on a single representative image (or perhaps, a modest subset) and the obtained parameter settings are applied to the rest of the images. This yields a better operating point for the segmentation algorithm compared to application-independent default values.

5.3 Speedup Methods

Even though the execution speed of the tracing algorithm used in producing experimental results scales with image content, not with image size [3], disk and memory access times for different image sizes are shown to make the execution time for 1000 RRS function evaluations (1000 runs of the segmentation algorithm each time possibly with different parameters) to almost linearly scale with image size (correlation coefficient $r = 0.98$). This can also be explained by the observations that different parameter settings produce different segmentations, ranging from overly-simple to overly-complex, effectively canceling the image-content execution-time scaling factor.

This motivated the idea of performing parameter selection on representative subvolume of the image when the execution speed of the segmentation algorithm is the bottleneck, as for 3-D images for example. The first method is to use a single most representative subvolume. The second one generalizes this idea to use a set of subvolumes containing the most representative subvolume in a coordinated manner.

Speedup of 11.2X on average is observed for 22 3-D images using the second method, versus 3.0X using the first. Compared to performing parameter selection using the entire image, both speedup methods on average sacrifice only 0.23%-0.25% of segmentation quality improvements. The observed difference in segmentation quality improvements between the two methods is only 0.02%. Although this speedup method has been presented only for 3-D images, it can be adapted to 2-D images as well.

5.3.1 Representative Subvolume

The representative measure is inspired by related work of Shen *et al.* [47] and Lin *et al.* [48] that prioritizes segmentation to achieve real-time performance for subsequent analyses. The proposed representative measure is based on both intensity and structural information. Use of the Mahalanobis distance measure effectively captures the dynamic ranges of each component of the representative measure. Representative measures presented in the literature for content-based image retrieval systems (e.g., [49-51]) are mainly based only on image intensity with future research heading towards using structural information [51].

5.3.2 Multiple-Subvolumes Coordinated Parameter Selection

This method recognizes the observation that image subvolumes are part of larger data that are highly correlated, i.e., the subvolumes are not much different from each other. Under this observation, it is expected that the subvolumes can “collaborate” between each other towards achieving optimal entire-image segmentation. Simply by sharing their automatically selected segmentation algorithm parameters, each subvolume then includes received parameters as part of their search and notifies other subvolumes when they have found better parameters for themselves. The entire process terminates when the subvolumes are not benefiting from each other.

5.3.3 Making a Global Segmentation Algorithm Spatially Adaptive

By using the multiple-subvolumes coordinated parameter selection method, subvolume segmentation parameter settings can be used by themselves, i.e., by dynamically changing and adapting the segmentation algorithm parameters as the image is being segmented. If possible, the parameters can also be interpolated over the entire image. This adapts the otherwise global segmentation algorithm parameters.

5.4 Future Research

5.4.1 Generalization to Other Geometrical Models

The core idea of trading off segmentation conciseness and segmentation coverage is expected to be extensible to other applications and other geometrical models, which remains as future work. Segmentation algorithms for other key biological objects types such as blobs [148, 149] would require matching quality metrics. Nevertheless, the modularity of the presented framework allows insertion of alternative segmentation algorithms and/or global optimization algorithms, as long as the interfaces between modules are maintained [150].

5.4.2 Selecting Parameters for Other Segmentation Algorithms

Beyond the global thresholding and tracing algorithms used in generating all presented experimental results, other algorithms for automatically segmenting tube-like structures need to have their interface modified to be able to use the presented parameter selection method. Other than the interface, they may also use different local model fitting methods than the tracing algorithm used in this work. This means that the seed-clustering method used for constructing non-empty subvolumes for parameter selection speedup may need to be adapted to their local model fitting stages.

5.4.3 Which Image to Use for Batch Segmentation[9]?

The batch segmentation method described in this work uses the automatically-selected segmentation parameter settings for one image selected at random and applied

---

[9] D. Thompson, BMED Dept., Rensselaer Polytechnic Institute, Troy, NY, USA, personal communication, 2005.

the parameter settings to the rest of the images in the batch. More systematic method of using the most representative image in the batch, that can be obtained using methods described in Section 3.4.2, can be used. The subimage-to-image distance used in that method needs to be modified to correspond to image-to-image distance.

As a side note, if the segmentation algorithm is linear (incremental changes in parameter settings correspond to incremental changes in resulting segmentation), then a group of representative images can be used. When this is not the case, a joint estimation of the segmentation algorithm parameter settings needs to be performed on the group of representative images. This can be done by performing multiple-objective optimization instead of single-objective optimization used in this work, where each objective corresponds to each image in that group.

In addition, the distribution of the observed distances to the representative image indicates whether batch segmentation can be performed with reasonable loss of potential increase in segmentation quality, i.e., by selecting parameter for every image in the batch. Recall that the representative measure is based on the local structure and intensity models used by the segmentation algorithm. Intuitively the higher the variance of that distribution, the higher the loss would be since the images are not very similar to each other, i.e., “in the eye” of the segmentation algorithm based on the fitted local models. On the other hand, lower variance indicates that the images in the batch are very similar in terms of the fitted local models.

5.4.4 Improved Segmentation Quality Metric

The presented segmentation quality metric uses a local structure-indicator function that measures the likelihood of a pixel to belong to structures of interest and implicitly captures inter-pixel correlation. Currently, there exist a number of possibilities for the local structure-indicator function, such as the vesselness measure [74], the generalized log-likelihood ratio method [22], matched filters [64, 65], and the recent likelihood ratio vesselness measure [151]. In this work, the use of the vesselness measure may be improved in future research where more accurate local structure-indicator functions are substituted in [151]. Once they are substituted, a comparative study between using different local structure-indicator functions needs to be performed.

5.4.5 Alternative Speedup Methods

Other methods for more efficient multiple-subvolume coordinated parameter selection other than the presented circular array method remains to be explored in future research. They may borrow ideas from computer networking (e.g. [152]) where the parameters of each network component such as a router are optimized to the benefit of the entire network.

As an alternative to using fixed subvolumes, the idea of systematic subvolume growing/expansion as applied in view-based image registration problems [153] remains to be investigated in future research. Applying that idea to this work essentially says that increasing subvolume sizes will increase improvements in segmentation quality, using the underlying assumption that larger subvolumes better represent the entire image. However, since counter-examples are found in the presented experimental results, this idea needs to be carefully refined and must be demonstrated to outperform the presented method both in terms of segmentation quality improvement and in observed speedup.

5.4.6 Temporally Adaptive Segmentation Algorithm

Beyond spatial adaptation of segmentation algorithm parameters as demonstrated in this work, temporal adaptation can be achieved in time-series images when the image-to-image transformation parameters are available. Adaptive segmentation algorithm parameters can then be transformed using the transformation parameters across the time series. Multiple-passes may be needed however when feature-based image registration algorithms are used (e.g., [153]) since the extracted features from which the image transformation parameters are computed depend on segmentation accuracy (which depends on segmentation algorithm parameters).

5.4.7 Comparative Studies

Comparative studies between segmentation algorithms on specific applications using the proposed segmentation quality metric, without relying on time-consuming and subjective manual segmentations as the ground truths [44-46], also remain as future research. These studies will enable users to make informed decisions on comparing between segmentation algorithms for their respective applications.

5.4.8 Automated Feature Detection: Pathological Conditions and Landmarks[10]

Disagreements between automated segmentations produced while performing parameter selection, i.e., while segmenting 1000 times to produce the segmentation algorithm parameters in this work, may be used to point out areas of interest to the user. For example, the tracing algorithm used in this work [3, 19] will encounter difficulties at image regions where the generalized cylinder model does not fit well. From observations, these regions may contain pathological conditions such as aneurysms (abnormal blood-filled dilatation of a blood vessel that may cause rupture) and also branch/crossover points. Therefore, out of the 1000 segmentation produced while searching for the segmentation algorithm parameters, finding locations where they disagree may elucidate these regions of interest.

5.4.9 Estimation of the Distribution of Local Structure-Indicator Function Values in Absence of Ground Truth Segmentations

Methods that estimate the ground truth segmentation (e.g., [46, 96]) can be used to estimate the ground truth from a collection of segmentations, possibly obtained by randomly varying the segmentation algorithm parameter settings. This will increase the objectivity of the parameter selection method. Recall that in this work, the distribution of the local structure-indicator function values are obtained from manual segmentations. One may proceed one step further where the ground truths are estimated for each application (image source, or image batches). With this, the distribution of the local structure-indicator function values is specialized to each application. This leads to the next item in potential future research.

5.4.10 Sensitivity Analysis for the Distributions of the Local Structure-Indicator Function Values[11]

In this work, a single distribution of the local structure-indicator function is used when automatically-selecting the segmentation algorithm parameters for all test images.

---

[10] A. Mercado, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, USA, personal communication, 2005.

[11] Q. Ji, ECSE Dept., Rensselaer Polytechnic Institute, Troy, NY, USA, personal communication, 2005.

Ideally the structure-indicator function should be estimated for each image, or at the very least, for a representative image or group of images from a particular application. It remains to be investigated in future research on how sensitive the resulting (near-) optimal segmentations are to changes in the distribution of the local structure-indicator function values.

5.4.11 Post Hoc Validation[12]

For routine use of this application, post hoc validations need to be performed on the resulting (near-) optimal segmentations. In other words, improvements in segmentation quality must be validated in terms of its relationship to improvements in agreement with ground truth for the specific application. Other criteria of importance to the application beyond pixel-by-pixel agreement can be included in the validation process, such as accuracy of branch point detection, or even the accuracy of extracted structural measurements. This essentially makes sure that observed high correlation between segmentation-quality-improvements and ground-truth-agreement-improvements in this work holds for new applications in order for this method to be adapted for routine use in those applications.

5.5 Conclusions

The presented experimental results have demonstrated the practicality of automatically tuning complex segmentation algorithms, using automatic segmentation quality assessment and global optimization, guided by the MDL principle. The proposed approach can greatly simplify the external interface of segmentation software packages, enable adaptation across large image batches, and reduce the need for expensive technical support.

For segmentation algorithm users, the proposed parameter selection method enables them use these algorithms without knowledge of the underlying algorithms. It also reveals the limitations of the segmentation algorithms. Therefore, it allows segmentation algorithm users to objectively compare between segmentation algorithms for their

---

[12] W. Shain, Wadsworth Center, NYS Dept. of Health, Albany, USA, personal communication, 2005.

specific applications. This is a step forward than comparing segmentation results of segmentation algorithms obtained using default settings. Instead, the comparison will be performed on the best-possible result obtainable by each segmentation algorithm given the limitation in computational time. Once the limit on accuracy is reached, segmentation errors beyond acceptable rates then justify investing tedious efforts for conducting the subjective manual segmentation for certain applications (e.g., [154]).

For image analysis systems that segments the structures of interest and extract morphometrics based on the segmented structures, improved segmentation accuracy directly translates into improved accuracy of the morphometrics. Obtaining morphometrics quickly and objectively are the reasons why segmentation algorithms are used in image-based biomedical studies in the first place. The proposed parameter selection method frees the users from manual hand-tuning of segmentation algorithm settings that does not guarantee increased accuracy of extracted-morphometrics.

Literature Cited

[1] R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. Upper Saddle River, N.J.: Prentice Hall, 2002.

[2] M.-A. Abdul-Karim, K. Al-Kofahi, E. B. Brown, R. K. Jain, and B. Roysam, "Automated tracing and change analysis of angiogenic vasculature from in vivo multiphoton confocal image time series," *Microvasc. Res.*, vol. 66, no. 2, pp. 113-125, 2003.

[3] K. A. Al-Kofahi, S. Lasek, D. H. Szarowski, C. J. Pace, G. Nagy, J. N. Turner, and B. Roysam, "Rapid automated three-dimensional tracing of neurons from confocal image stacks," *IEEE Trans. Inform. Technol. Biomed.*, vol. 6, no. 2, pp. 171-187, 2002.

[4] I. Y. Koh, W. B. Lindquist, K. Zito, E. A. Nimchinsky, and K. Svoboda, "An image analysis algorithm for dendritic spines," *Neural Comput.*, vol. 14, no. 6, pp. 1283-310, 2002.

[5] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, and R. L. Kennedy, "Measurement of retinal vessel widths from fundus images based on 2-D modeling," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1196-1204, 2004.

[6] E. Meijering, M. Jacob, J. C. Sarria, P. Steiner, H. Hirling, and M. Unser, "Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images," *Cytometry A*, vol. 58A, no. 2, pp. 167-76, 2004.

[7] M. J. Cross, M.-A. Abdul-Karim, G. Lin, B. Roysam, and D. M. Thompson, "Guidance of neurite outgrowth by non-neural cells," presented at Annual Fall Meeting of the Biomedical Engineering Soc., Baltimore, MD, 2005.

[8] J. V. Watson, *Introduction to flow cytometry*. Cambridge, England: Cambridge University Press, 2004.

[9] M. Rieseberg, C. Kasper, K. F. Reardon, and T. Scheper, "Flow cytometry in biotechnology," *Appl. Microbiol. Biotechnol.*, vol. 56, no. 3-4, pp. 350-360, 2001.

[10] G. Lin, C. S. Bjornsson, K. L. Smith, M. A. Abdul-Karim, J. N. Turner, W. Shain, and B. Roysam, "Automated image analysis methods for 3-D

quantification of the neurovascular unit from multichannel confocal microscope images," *Cytometry A*, vol. 66A, no. 1, pp. 9-23, 2005.

[11] N. M. Dowell-Mesfin, M.-A. Abdul-Karim, A. M. P. Turner, S. Schanz, H. G. Craighead, B. Roysam, J. N. Turner, and W. Shain, "Topographically modified surfaces affect orientation and growth of hippocampal neurons," *J. Neural Eng.*, no. 2, pp. 78-90, 2004.

[12] R. K. Jain, L. L. Munn, and D. Fukumura, "Dissecting tumour pathophysiology using intravital microscopy," *Nat. Rev. Cancer*, vol. 2, no. 4, pp. 266-276, 2002.

[13] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset, "Model-based detection of tubular structures in 3D images," *Comput. Vis. Image Underst.*, vol. 80, no. 2, pp. 130-71, 2000.

[14] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam, "Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms," *IEEE Trans. Inform. Technol. Biomed.*, vol. 3, no. 2, pp. 125-138, 1999.

[15] A. Dellas, H. Moch, E. Schultheiss, G. Feichter, A. C. Almendral, F. Gudat, and J. Torhorst, "Angiogenesis in cervical neoplasia: Microvessel quantitation in precancerous lesions and invasive carcinomas with clinicopathological correlations," *Gynecol. Oncol.*, vol. 67, no. 1, pp. 27-33, 1997.

[16] M. Dellian, B. P. Witwer, H. A. Salehi, F. Yuan, and R. K. Jain, "Quantitation and physiological characterization of angiogenic vessels in mice - Effect of basic fibroblast growth factor vascular endothelial growth factor vascular permeability factor, and host microenvironment," *Am. J. Pathol.*, vol. 149, no. 1, pp. 59-71, 1996.

[17] L. M. Kirchner, S. P. Schmidt, and B. S. Gruber, "Quantitation of angiogenesis in the chick chorioallantoic membrane model using fractal analysis," *Microvasc. Res.*, vol. 51, no. 1, pp. 2-14, 1996.

[18] J. J. Capowski, *Computer techniques in neuroanatomy*. New York: Plenum Press, 1989.

[19] K. A. Al-Kofahi, A. Can, S. Lasek, D. H. Szarowski, N. Dowell-Mesfin, W. Shain, J. N. Turner, and B. Roysam, "Median-based robust algorithms for tracing

neurons from noisy confocal microscope images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 7, no. 4, pp. 302-317, 2003.

[20] W. He, T. A. Hamilton, A. R. Cohen, T. J. Holmes, C. Pace, D. H. Szarowski, J. N. Turner, and B. Roysam, "Automated three-dimensional tracing of neurons in confocal and brightfield images," *Microsc. Microanal.*, vol. 9, no. 4, pp. 296-310, 2003.

[21] S. R. Aylward and E. Bullitt, "Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 61-75, 2002.

[22] V. Mahadevan, H. Narasimha-Iyer, B. Roysam, and H. L. Tanenbaum, "Robust model-based vasculature detection in noisy biomedical images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, no. 3, pp. 360-376, 2004.

[23] D. L. Wilson and J. A. Noble, "An adaptive segmentation algorithm for time-of-flight MRA data," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 938-945, 1999.

[24] T. M. G. Koller, G.; Szekely, G.; Dettwiler, D., "Multiscale detection of curvilinear structures in 2-D and 3-D image data," presented at Fifth International Conference on Computer Vision, 1995, pp. 864-869.

[25] O. Wink, W. J. Niessen, B. Verdonck, and M. A. Viergever, "Vessel axis determination using wave front propagation analysis," *Lecture Notes in Computer Science*, no. 2208, pp. 845-853, 2001.

[26] F. K. H. Quek and C. Kirbas, "Vessel extraction in medical images by wave-propagation and traceback," *IEEE Trans. Med. Imag.*, vol. 20, no. 2, pp. 117-131, 2001.

[27] A. F. Frangi, W. J. Niessen, R. M. Hoogeveen, T. van Walsum, and M. A. Viergever, "Model-based quantitation of 3-D magnetic resonance angiographic images," *IEEE Trans. Med. Imag.*, vol. 18, no. 10, pp. 946-956, 1999.

[28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed. New York: Wiley, 2001.

[29] Y. G. Leclerc, "Constructing simple stable descriptions for image partitioning," *Int. J. Comput. Vision*, vol. 3, no. 1, pp. 73-102, 1989.

[30] J. Rissanen, *Stochastic complexity in statistical inquiry*. Singapore: World Scientific, 1989.

[31] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465-471, 1978.

[32] A. Barron, J. Rissanen, and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2743-2760, 1998.

[33] P. D. Grünwald, I. J. Myung, and M. A. Pitt, *Advances in minimum description length: theory and applications*. Cambridge, Mass.: MIT Press, 2005.

[34] S. Aylward, E. Bullitt, S. Pizer, and D. Eberly, "Intensity ridge and widths for tubular object segmentation and description," presented at IEEE/SIAM Workshop Mathematical Methods Biomedical Image Analysis, 1996, pp. 131-138.

[35] J. Min, M. Powell, and K. W. Bowyer, "Automated performance evaluation of range image segmentation algorithms," *IEEE Trans. Syst., Man, Cybern. B*, vol. 34, no. 1, pp. 263-271, 2004.

[36] J. Peng and B. Bhanu, "Closed-loop object recognition using reinforcement learning," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 2, pp. 139-154, 1998.

[37] Y. Sato, J. Chen, R. A. Zoroofi, N. Harada, S. Tamura, and T. Shiga, "Automatic extraction and measurement of leukocyte motion in microvessels using spatiotemporal image analysis," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 4, pp. 225-236, 1997.

[38] P. Meer and B. Georgescu, "Edge detection with embedded confidence," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 12, pp. 1351-1365, 2001.

[39] K. Bowyer, C. Kranenburg, and S. Dougherty, "Edge detector evaluation using empirical ROC curves," *Comput. Vis. Image Underst.*, vol. 84, no. 1, pp. 77-103, 2001.

[40] W. H. H. J. Lunscher and M. P. Beddoes, "Optimal edge detector design.1. Parameter selection and noise effects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 2, pp. 164-177, 1986.

[41] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 9, pp. 884-900, 1996.

[42] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational-problems," *Commun. Pur. Appl. Math.*, vol. 42, no. 5, pp. 577-685, 1989.

[43] J.-M. Morel and S. Solimini, *Variational methods in image segmentation: with seven image processing experiments*. Boston: Birkhèauser, 1995.

[44] V. Chalana and Y. Kim, "A methodology for evaluation of boundary detection algorithms on medical images," *IEEE Trans. Med. Imag.*, vol. 16, no. 5, pp. 642-652, 1997.

[45] C. Alberola-Lopez, M. Martin-Fernandez, and J. Ruiz-Alzola, "Comments on: A methodology for evaluation of boundary detection algorithms on medical images," *IEEE Trans. Med. Imag.*, vol. 23, no. 5, pp. 658-660, 2004.

[46] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 7, pp. 903-921, 2004.

[47] H. Shen, B. Roysam, C. V. Stewart, J. N. Turner, and H. L. Tanenbaum, "Optimal scheduling of tracing computations for real-time vascular landmark extraction from retinal fundus images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 5, no. 1, pp. 77, 2001.

[48] G. Lin, C. V. Stewart, B. Roysam, K. Fritzsche, Y. Gehua, and H. L. Tanenbaum, "Predictive scheduling algorithms for real-time feature extraction and spatial referencing: application to retinal image sequences," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 1, pp. 115, 2004.

[49] A. Natsev, R. Rajeev, and K. Shim, "WALRUS: a similarity retrieval algorithm for image databases," *IEEE Trans. Knowledge Data Eng.*, vol. 16, no. 3, pp. 301, 2004.

[50] W. Xiong, B. Qiu, Q. Tian, H. Mueller, and C. Xu, "A novel content-based medical image retrieval method based on query topic dependent image features

(QTDIF)," presented at Medical Imaging 2005: PACS and Imaging Informatics, San Diego, CA, USA, 2005, pp. 123.

[51] J. R. Parker and B. Behm, "Use of multiple algorithms in image content searches," presented at International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004., Las Vegas, 2004, pp. 246.

[52] M. K. Markey, M. V. Boland, and R. F. Murphy, "Toward objective selection of representative microscope images," *Biophys. J.*, vol. 76, no. 4, pp. 2230-2237, 1999.

[53] D. Aykac, E. A. Hoffman, G. McLennan, and J. M. Reinhardt, "Segmentation and analysis of the human airway tree from three-dimensional X-ray CT images," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 940-950, 2003.

[54] M. de Bruijne, W. J. Niessen, J. B. A. Maintz, and M. A. Viergever, "Localization and segmentation of aortic endografts using marker detection," *IEEE Trans. Med. Imag.*, vol. 22, no. 4, pp. 473-482, 2003.

[55] J. Canny, "A computational approach to edge-detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679-698, 1986.

[56] R. B. Adams, L., "Seeded region growing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, no. 6, pp. 641-647, 1994.

[57] S. L. Horowitz and T. Pavlidis, "Picture segmentation by a directed split-and-merge procedure," presented at 2nd Int. Joint Conf. Pattern Recognit., 1974, pp. 424-433.

[58] R. M. Haralick, "Ridges and valleys on digital images," *CVGIP*, vol. 22, no. 1, pp. 28-38, 1983.

[59] D. Eberly, *Ridges in image and data analysis*. Boston: Kluwer Academic Publishers, 1996.

[60] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen, "A survey of thresholding techniques," *CVGIP*, no. 2, pp. 233-60, 1988.

[61] Y. Sato, S. Nakajima, N. Shiraga, H. Atsumi, S. Yoshida, T. Koller, G. Gerig, and R. Kikinis, "Three-dimensional multi-scale line filter for segmentation and

visualization of curvilinear structures in medical images," *Med. Image Anal.*, vol. 2, no. 2, pp. 143-68, 1998.

[62] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: a survey," *Med. Image Anal.*, vol. 1, no. 2, pp. 91-108, 1996.

[63] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*. New York: McGraw-Hill, 1995.

[64] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," *IEEE Trans. Med. Imag.*, vol. 8, no. 0278-0062, pp. 263-269, 1989.

[65] A. D. K. Hoover, V.; Goldbaum, M., "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203-210, 2000.

[66] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience: exploring the brain*, 2nd ed. Baltimore, Md.: Lippincott Williams & Wilkins, 2001.

[67] Y. P. Du, D. L. Parker, and W. L. Davis, "Vessel enhancement filtering in 3-dimensional MR-Angiography," *J. Magn. Reson. Im.*, vol. 5, no. 2, pp. 151-157, 1995.

[68] Y. P. Du and D. L. Parker, "Vessel enhancement filtering in three-dimensional MR angiograms using long-range signal correlation," *J. Magn. Reson. Im.*, vol. 7, no. 2, pp. 447-450, 1997.

[69] M. M. Orkisz, C. Bresson, I. E. Magnin, O. Champin, and P. C. Douek, "Improved vessel visualization in MR angiography by nonlinear anisotropic filtering," *Magnet. Reson. Med.*, vol. 37, no. 6, pp. 914-919, 1997.

[70] H. G. Chen and J. Hale, "An algorithm for MR-Angiography image-enhancement," *Magnet. Reson. Med.*, vol. 33, no. 4, pp. 534-540, 1995.

[71] D. L. Wilson and J. A. Noble, "Segmentation of cerebral vessels and aneurysms from MR angiography data," *Lecture Notes in Computer Science*, no. 1230, pp. 423-428, 1997.

[72] R. Kutka and S. Stier, "Extraction of line properties based on direction fields," *IEEE Trans. Med. Imag.*, vol. 15, no. 1, pp. 51-58, 1996.

[73] C. Steger, "An unbiased detector of curvilinear structures," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 2, pp. 113-125, 1998.

[74] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," presented at Medical Image Computing and Computer-Assisted Intervention, 1998, pp. 130-137.

[75] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese, "Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images," presented at CVRMed-MRCAS'97 (Lecture Notes in Computer Science), 1997, pp. 233-242.

[76] P. E. Summers, A. H. Bhalerao, and D. J. Hawkes, "Multiresolution, model-based segmentation of MR angiograms," *J. Magn. Reson. Im.*, vol. 7, no. 6, pp. 950-957, 1997.

[77] T. Lindeberg, *Scale-space theory in computer vision*. Boston: Kluwer Academic, 1994.

[78] J. J. Koenderink, "The structure of images," *Biol. Cybern.*, vol. 50, no. 5, pp. 363-370, 1984.

[79] L. M. J. Florack, B. M. T. Romeny, J. J. Koenderink, and M. A. Viergever, "Scale and the differential structure of images," *Image Vision Comput.*, vol. 10, no. 6, pp. 376-388, 1992.

[80] R. Poli and G. Valli, "An algorithm for real-time vessel enhancement and detection," *Comput. Meth. Prog. Bio.*, vol. 52, no. 1, pp. 1-22, 1997.

[81] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *Int. J. Comput. Vision*, vol. 30, no. 2, pp. 117, 1998.

[82] P. J. C. Yim, P.L.; Summers, R.M., "Gray-scale skeletonization of small vessels in magnetic resonance angiography," *IEEE Trans. Med. Imag.*, vol. 19, no. 6, pp. 568-576, 2000.

[83] M. A. T. Figueiredo and J. M. N. Leitao, "A nonsmoothing approach to the estimation of vessel contours in angiograms," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 162-172, 1995.

[84] C. M. Weaver, P. R. Hof, S. L. Wearne, and W. B. Lindquist, "Automated algorithms for multiscale morphometry of neuronal dendrites," *Neural Comput.*, vol. 16, no. 7, pp. 1353-83, 2004.

[85] J. S. Suri and S. Laxminarayan, *Angiography and plaque imaging: advanced segmentation techniques*. Boca Raton: CRC Press, 2003.

[86] L. Lam, S. W. Lee, and C. Y. Suen, "Thinning methodologies - a comprehensive survey," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 9, pp. 869-885, 1992.

[87] C. L. Tsai, C. V. Stewart, H. L. Tanenbaum, and B. Roysam, "Model-based method for improving the accuracy and repeatability of estimating vascular bifurcations and crossovers from retinal fundus images," *IEEE Trans. Inform. Technol. Biomed.*, vol. 8, no. 2, pp. 122-130, 2004.

[88] O. Wink, A. F. Frangi, B. Verdonck, M. A. Viergever, and W. J. Niessen, "3D MRA coronary axis determination using a minimum cost path approach," *Magnet. Reson. Med.*, vol. 47, no. 6, pp. 1169-1175, 2002.

[89] O. Wink, W. J. Niessen, and M. A. Viergever, "Multiscale vessel tracking," *IEEE Trans. Med. Imag.*, vol. 23, no. 1, pp. 130-133, 2004.

[90] M. D. Heath, S. Sarkar, T. Sanocki, and K. W. Bowyer, "A robust visual method for assessing the relative performance of edge-detection algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, no. 12, pp. 1338-1359, 1997.

[91] M. Niemeijer, J. Staal, B. van Ginneken, M. Loog, and M. D. Abramoff, "Comparative study of retinal vessel segmentation methods on a new publicly available database [5370-69]," presented at Image Processing Conference, San Diego, Calif., 2004, pp. 648-656.

[92] M. Heath, S. Sarkar, T. Sanocki, and K. Bowyer, "Comparison of edge detectors: A methodology and initial study," *Comput. Vis. Image Underst.*, vol. 69, no. 1, pp. 38-54, 1998.

[93] T. Kanungo, M. Y. Jaisimha, J. Palmer, and R. M. Haralick, "A methodology for quantitative performance evaluation of detection algorithms," *IEEE Trans. Image Processing*, vol. 4, no. 12, pp. 1667-1674, 1995.

[94] M. C. Shin, D. B. Goldgof, and K. W. Bowyer, "Comparison of edge detector performance through use in an object recognition task," *Comput. Vis. Image Underst.*, vol. 84, no. 1, pp. 160-178, 2001.

[95] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 12, pp. 1191-1201, 1995.

[96] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, no. 8, pp. 1027-1033, 2003.

[97] Y. J. Zhang, "A survey on evaluation methods for image segmentation," *Pattern Recog.*, vol. 29, no. 8, pp. 1335-1346, 1996.

[98] A. Hammoude, "An empirical parameter selection method for endocardial border identification algorithms," *Comput. Med. Imag. Grap.*, vol. 25, no. 1, pp. 33-45, 2001.

[99] M. C. Shin, D. Goldgof, and K. W. Bowyer, "An objective comparison methodology of edge detection algorithms using a structure from motion task," presented at IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, 1998, pp. 190.

[100] M. C. Shin, D. Goldgof, and K. W. Bowyer, "Comparison of edge detectors using an object recognition task," presented at IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Fort Collins, CO, 1999, pp. 365.

[101] Y. Sun, R. J. Lucariello, and S. A. Chiaramida, "Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary arteriograms," *IEEE Trans. Med. Imag.*, vol. 14, no. 2, pp. 242-248, 1995.

[102] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463-468, 1998.

[103] M. Styner, C. Brechbuhler, G. Szekely, and G. Gerig, "Parametric estimate of intensity inhomogeneities applied to MRI," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 153-165, 2000.

[104] R. E. Hogan, K. E. Mark, L. Wang, S. Joshi, M. I. Miller, and R. D. Bucholz, "Mesial temporal sclerosis and temporal lobe epilepsy: MR imaging deformation-based segmentation of the hippocampus in five patients," *Radiology*, vol. 216, no. 1, pp. 291-297, 2000.

[105] A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: method and validation," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 716, 1994.

[106] G. Gerig, M. Jomier, and M. Chakos, "VALMET: A new validation tool for assessing and improving 3D object segmentation," presented at Medical Image Computing and Computer-Assisted Intervention, 2001, pp. 516-523.

[107] Y. Lin and B. Bhanu, "Object detection via feature synthesis using MDL-based genetic programming," *IEEE Trans. Syst., Man, Cybern. B*, vol. 35, no. 3, pp. 538, 2005.

[108] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *IEEE Trans. Med. Imag.*, vol. 21, no. 5, pp. 525, 2002.

[109] M. A. T. Figueiredo, J. M. N. Leitao, and A. K. Jain, "Unsupervised contour representation and estimation using B-splines and a minimum description length criterion," *IEEE Trans. Image Processing*, vol. 9, no. 6, pp. 1075, 2000.

[110] F. Galland, N. Bertaux, and P. Refregier, "Minimum description length synthetic aperture radar image segmentation," *IEEE Trans. Image Processing*, vol. 12, no. 9, pp. 995, 2003.

[111] M. Mitchell, *An introduction to genetic algorithms*. Cambridge, Mass.: MIT Press, 1996.

[112] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671-680, 1983.

[113] W. L. Price, "Global optimization by controlled random search," *J. Optimiz. Theory App.*, vol. 40, no. 3, pp. 333-348, 1983.

[114] P. Brachetti, M. D. Ciccoli, G. DiPillo, and S. Lucidi, "A new version of the Price's algorithm for global optimization," *J Global Optim*, vol. 10, no. 2, pp. 165-184, 1997.

[115] R. Hooke and T. A. Jeeves, "Direct search solution of numerical and statistical problems," *J. ACM*, vol. 8, no. 2, pp. 212-218, 1961.

[116] V. Torczon, "On the convergence of pattern search algorithms," *SIAM J. Optimiz.*, vol. 7, no. 1, pp. 1-25, 1997.

[117] M. M. Ali, C. Storey, and A. Torn, "Application of stochastic global optimization algorithms to practical problems," *J. Optimiz. Theory App.*, vol. 95, no. 3, pp. 545-563, 1997.

[118] Z. B. Zabinsky, "Stochastic methods for practical global optimization," *J Global Optim*, vol. 13, no. 4, pp. 433-444, 1998.

[119] T. Ye and S. Kalyanaraman, "Internet traffic engineering - A recursive random search algorithm for large-scale network parameter configuration," *ACM SIGMETRICS Perf. Eval. Rev.*, vol. 31, no. 1, pp. 10, 2003.

[120] A. H. Kan and G. T. Timmer, "Stochastic global optimization methods. part II: multi level methods," *Math. Program.*, vol. 39, no. 1, pp. 57-78, 1987.

[121] J. E. Falk and R. M. Soland, "Algorithm for separable nonconvex programming problems," *Manage. Sci.*, vol. 15, no. 9, pp. 550-569, 1969.

[122] L. Armijo, "Minimization of functions having lipschitz continuous first partial derivatives," *Pac. J. Math.*, vol. 16, no. 1, pp. 1-8, 1966.

[123] S. Rana, L. D. Whitley, and R. Cogswell, "Searching in the presence of noise," *Lecture Notes in Computer Science*, no. 1141, pp. 198, 1996.

[124] N. J. Radcliffe and P. D. Surry, "Fundamental limitations on search algorithms: Evolutionary computing in perspective," *Lecture Notes in Computer Science*, no. 1000, pp. 275-291, 1995.

[125] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67-82, 1997.

[126] J. Besag, "On the statistical-analysis of dirty pictures," *J. Roy. Stat. Soc. B. Met.*, vol. 48, no. 3, pp. 259-302, 1986.

[127] L. Wang and A. Bhalerao, "Model based segmentation for retinal fundus images," presented at Image analysis; SCIA 2003, Halmstad, Sweden, 2003, pp. 422-429.

[128] C. Canero and P. Radeva, "Vesselness enhancement diffusion," *Pattern Recog. Lett.*, vol. 24, no. 16, pp. 3141-3151, 2003.

[129] G. J. Streekstra and J. van Pelt, "Analysis of tubular structures in three-dimensional confocal images," *Network*, vol. 13, no. 3, pp. 381-95, 2002.

[130] A. Tèorn and A. Zhilinskas, *Global optimization*. Berlin; New York: Springer-Verlag, 1989.

[131] B. Xi, Z. Liu, M. Raghavachari, C. H. Xia, and L. Zhang, "A smart hill-climbing algorithm for application server configuration," presented at The 13th International Conference on World Wide Web, New York, NY, 2004, pp. 287 - 296.

[132] R. Y. Rubinstein, *Simulation and the Monte Carlo method*. New York: Wiley, 1981.

[133] P. J. Huber, *Robust statistics*. New York: Wiley, 1981.

[134] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "An efficient k-means clustering algorithm: analysis and implementation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, no. 7, pp. 881, 2002.

[135] G. S. Watson, *Statistics on spheres*. New York: Wiley, 1983.

[136] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501, 2004.

[137] L. J. Bain and M. Engelhardt, *Introduction to probability and mathematical statistics*, 2nd ed. Boston: PWS-KENT Pub., 1992.

[138] H. Narasimha-Iyer, N. M. Dowell-Mesfin, M.-A. Abdul-Karim, B. Roysam, J. N. Turner, and W. Shain, "Associative multiple-label image analysis method for synapse identification in neuronal cultures: application to comparative analysis of synapse formation efficiency & distribution on smooth and topographically modified surfaces," presented at Microsc. Microanal., Honolulu, 2005.

[139] M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *J. Amer. Statistical Assoc.*, vol. 69, no. 347, pp. 730-737, 1974.

[140] Z. Michalewicz and D. B. Fogel, *How to solve it: modern heuristics*, 2nd ed. Berlin; New York: Springer, 2004.

[141] W. L. Price, "Controlled random search procedure for global optimization," *Computer Journal*, vol. 20, no. 4, pp. 367-370, 1977.

[142] J. A. Nelder and R. Mead, "A simplex-method for function minimization," *Computer Journal*, vol. 7, no. 4, pp. 308-313, 1965.

[143] Y. Nourani and B. Andresen, "A comparison of simulated annealing cooling strategies," *J. Phys. A: Math. Gen.*, vol. 31, no. 41, pp. 8373-8385, 1998.

[144] K. A. Dejong and W. M. Spears, "An analysis of the interacting roles of population-size and crossover in genetic algorithms," *Lecture Notes in Computer Science*, no. 496, pp. 38-47, 1991.

[145] D. B. Fogel, *Evolutionary computation: the fossil record*. New York: IEEE Press, 1998.

[146] K. Dejong, "Special issue on genetic algorithms - Introduction," *Machine Learning*, vol. 5, no. 4, pp. 351-353, 1990.

[147] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*. Reading, Mass.: Addison-Wesley Pub. Co., 1989.

[148] G. Lin, U. Adiga, K. Olson, J. F. Guzowski, C. A. Barnes, and B. Roysam, "A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks," *Cytometry A*, vol. 56A, no. 1, pp. 23-36, 2003.

[149] C. Ortiz De Solorzano, E. Garcia Rodriguez, A. Jones, D. Pinkel, J. W. Gray, D. Sudar, and S. J. Lockett, "Segmentation of confocal microscope images of cell nuclei in thick tissue sections," *J. Microsc.*, vol. 193, no. 3, pp. 212-226, 1999.

[150] E. Gamma, *Design patterns: elements of reusable object-oriented software*. Reading: Addison-Wesley, 1995.

[151] M. Sofka and C. V. Stewart, "Retinal Vessel Extraction Using Multiscale Matched Filters, Confidence and Edge Measures," Department of Computer Science, Rensselaer Polytechnic Institute Technical Report # 05-20, 2005.

[152] J. F. Kurose and K. W. Ross, *Computer networking: a top-down approach featuring the Internet*, 3rd ed. Boston: Pearson/Addison Wesley, 2005.

[153] C. V. Stewart, T. Chia-Ling, and B. Roysam, "The dual-bootstrap iterative closest point algorithm with application to retinal image registration," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1379, 2003.

[154] G. M. G. Shepherd, A. Stepanyants, I. Bureau, D. Chklovskii, and K. Svoboda, "Geometric and functional organization of cortical circuits," *Nature Neuroscience*, vol. 8, no. 6, pp. 782-790, 2005.