Using a Focus Measure to Automate the Location of Biological Tissue Surfaces in Brightfield Microscopy

Daniel Toby Elozory

University of South Florida, delozory@verizon.net

Follow this and additional works at: http://scholarcommons.usf.edu/etd

Part of the American Studies Commons, Cell Anatomy Commons, and the Computer Sciences Commons

Scholar Commons Citation
Elozory, Daniel Toby, "Using a Focus Measure to Automate the Location of Biological Tissue Surfaces in Brightfield Microscopy" (2011). Graduate Theses and Dissertations. http://scholarcommons.usf.edu/etd/3089

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Using a Focus Measure to Automate the Location of Biological Tissue Surfaces in Brightfield Microscopy

by

Daniel Toby Elozory

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Co-Major Professor: Dmitry B. Goldgof, Ph.D.
Co-Major Professor: Lawrence O. Hall, Ph.D.
Peter R. Mouton, Ph.D.

Date of Approval:
June 29, 2011

Keywords: Stereology, Thresholded Absolute Gradient, Z-stack, Contrast, Depth From Focus

Copyright © 2011, Daniel Toby Elozory
DEDICATION

To Link Elozory, my father of blessed memory, from whom I learned to value education. It was an honor to walk in his shadow, for all who knew him enjoyed his goodness, his kindness, and his generosity. If I am half the man he was then I will live a happy and fulfilled life.
ACKNOWLEDGEMENTS

I am grateful to all the professors who I had the privilege of learning from these past few years. I was pleasantly surprised at the genuine effort all my professors at the University made at getting to know me as their student and their patience and good will in providing clarifications and better understanding of their lessons when I was in such need. I would like to particularly thank Dr. Dmitry Goldgof for his guidance and direction in my thesis and his willingness to accept me as his student. A thank you to Dr. Lawrence Hall as well, who, along with Dr. Goldgof, patiently listened to my thesis progress or lack of progress reports in weekly and frequently twice a week research meetings. I am also grateful for the impromptu lessons in microscopy, stereology, and biology that Dr. Peter Mouton happily gave me in our lab. A special thanks to Dr. Kurt Kramer the leader of our research group. He was always available and eager to help with programming questions or any problems that arose in my research.

I also would like to acknowledge the sacrifice and understanding of my wonderful family. if not for the love and support of my wife, Bonnie, and her unwavering faith in me, as well as the love and encouragement of my children: Cody, Brynne, Ruth, and Diana this thesis and indeed this graduate school journey would not have been possible.
TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. iii

LIST OF FIGURES ............................................................................................................... iv

ABSTRACT ............................................................................................................................. v

CHAPTER 1 INTRODUCTION .............................................................................................. 1
  1.1 Motivation ..................................................................................................................... 1
  1.2 Aims of Thesis ............................................................................................................. 2
  1.3 Novelty ......................................................................................................................... 4
  1.4 Thesis Organization ..................................................................................................... 5

CHAPTER 2 BACKGROUND ............................................................................................... 6
  2.1 Microscopy ................................................................................................................... 6
  2.2 Stereology .................................................................................................................... 9

CHAPTER 3 PREVIOUS WORK ......................................................................................... 14
  3.1 Focus Functions ......................................................................................................... 14
  3.2 Automated Stereology Programs ............................................................................... 16

CHAPTER 4 METHODOLOGY ............................................................................................ 17
  4.1 Methods Used ............................................................................................................ 17
  4.2 Common Focus Functions ......................................................................................... 20
  4.3 Common Non-Thresholded Focus Functions .............................................................. 22
  4.4 New Focus Functions ................................................................................................. 24
  4.5 Interpretation of Focus Functions .............................................................................. 26
  4.6 Ground Truth ............................................................................................................. 27
  4.7 Optimization of Threshold Parameters .................................................................... 28
    4.7.1 Nelder-Mead Simplex Search Method ................................................................. 28
    4.7.2 Golden Ratio Method .......................................................................................... 30

CHAPTER 5 DATASETS ....................................................................................................... 31
  5.1 Data Collection Equipment ......................................................................................... 33
  5.2 Initial Data Collection ................................................................................................. 34
  5.3 Training Set ................................................................................................................ 36
  5.4 Test Set #1 ................................................................................................................ 38
  5.5 Test Set #2 ................................................................................................................ 39
CHAPTER 6 EXPERIMENTAL RESULTS .......................................................... 41
  6.1 Results ............................................................................................... 41
  6.2 Statistical Significance ....................................................................... 57

CHAPTER 7 DISCUSSION .............................................................................. 61

CHAPTER 8 CONCLUSIONS ........................................................................ 64
  8.1 Conclusions ........................................................................................ 64
  8.2 Future Work ....................................................................................... 64

REFERENCES ............................................................................................... 66
LIST OF TABLES

Table 4.1 Citations for Common Focus Functions ................................................. 20
Table 4.2 Select Ground Truth of Color Image Z-stacks With 1.0 µm Step ........... 28
Table 5.1 Characteristics of Datasets Used for Evaluation ................................. 32
Table 6.1 Automated Focus Determination Training Set Optimization ............... 51
Table 6.2 Automated Focus Determination, Test Set #1 ..................................... 52
Table 6.3 Automated Focus Determination Test Set #2 ..................................... 52
Table 6.4 Automated Tissue Thickness Determination ...................................... 53
Table 6.5 Distribution of Deviation on Training Set .......................................... 55
Table 6.6 Distribution of Deviation on Test Set #1 ........................................... 56
Table 6.7 Distribution of Deviation on Test Set #2 .......................................... 56
Table 6.8 Distribution of Deviation on Combined Test Sets ............................. 57
Table 6.9 McNemar Chi Squared Test Measuring Statistical Significance ........... 59
Table 6.10 Paired t-Test Measuring Statistical Significance ............................. 60
LIST OF FIGURES

Figure 2.1 Comparison of Oil Immersion and Air ............................................. 8
Figure 2.2 The Corpuscle Problem ..................................................................... 11
Figure 2.3 The Physical Disector ...................................................................... 12
Figure 5.1 Unfocused Rat Hippocampus through Dirty Lenses ......................... 35
Figure 5.2 Rejected Z-stack from Test Set#2 ..................................................... 36
Figure 5.3 Rejected Z-stack from Training Set .................................................... 38
Figure 5.4 Test Set #2 Cresyl Violet Stained Pyramidal Neurons ....................... 39
Figure 6.1 Thresholded Gradient Focus Curves ................................................. 42
Figure 6.2 Focus Curves Sorted by Focus Results ............................................. 43
Figure 6.3 Thresholded “Intuitive” Focus Curves .............................................. 44
Figure 6.4 Zoomed in Thresholded “Intuitive” Focus Curves ......................... 45
Figure 6.5 Non-Thresholded Gradient Focus Curves ................................. 46
Figure 6.6 Non-Thresholded Statistical Focus Curves ..................................... 47
Figure 6.7 Thresholded Histogram Based Focus Curves ................................ 48
Figure 6.8 Modified Absolute Gradient Focus Curves ................................... 49
ABSTRACT

The study of microstructures in brightfield microscopy using unbiased stereology plays a large and growing role in bioscience research. Stereology enables objective quantitative analysis of biological structures within a tissue sample. A first step in the stereology process is to calculate the thickness of a tissue sample by locating the top and bottom surfaces of the sample. The aim of this project is to fully automate this location process by using the relative optical focus measure as an indicator of tissue surface boundary.

The current method for identification of focus bounding planes requires a trained user to manually select the top and bottom of the tissue at each sample position examined. To automate finding the correct focal planes, i.e. the “just out of focus” planes at the top and bottom surfaces of the tissue sections, a novel approach was developed. Several gray scale focusing functions were analyzed, but while the traditional emphasis of microscopy focus functions is to find global maximums on the focus curve, in this project the aim was to find the sharp “knees” on the focus curve. Starting with a low focus measure value when the focal plane of the objective lens is out of focus above the tissue sample, the objective focal plane is moved downward through the tissue. The ideal focus measure should increase sharply as the upper surface of the tissue passes into the depth of field of the objective lens. As the focal plane is moved through the
tissue, the focus measure value rises and falls as objects within the tissue come in and out of focus. As the bottom tissue surface passes into the depth of field the ideal focus measure should reflect some level of focus, dropping precipitously as the surface passes out of the depth of field into the unfocused region below the tissue.
CHAPTER 1

INTRODUCTION

1.1 Motivation

Current microscopy based computerized analysis of biological microstructures has been automated to a substantial degree [1]. The Stereologeter (Stereology Resource Center, Inc., Chester, MD), an advanced integrated hardware and software stereology system, can make accurate and precise estimates of first and second order stereological parameters, yet still requires considerable effort from trained users. Presently, automated processes include:

1. A step by step guide through the data collection process;
2. Systematic random location of probes as well as positioning of microscope stage on the probe (disector) locations;
3. Calculation of first and second order stereology parameters from manually collected data.

Manual processes requiring a trained user include:

1. Identification of the reference space to study;
2. Optimization of probe parameters to minimize error rates while also minimizing sampling size;
3. Determination of upper and lower tissue surfaces of each Z-stack examined;
4. Count of the objects within the Z-stacks.

Automating these processes would increase the throughput of computerized stereology analysis by reducing analysis from hours to minutes, thus reducing labor costs while increasing the ability of bioscientists to complete more scientific testing in a given time frame. An automated algorithm to locate the top and bottom tissue surface would be an important first step towards the goal of a completely automatic system for computerized analysis of microscopic biological structures using unbiased stereology [2].

1.2 Aims of Thesis

The aims of this thesis were threefold. The first goal was to develop an automated surface location algorithm to locate the top and bottom surface of a prepared stained biological tissue sample at particular X-Y locations within an anatomically defined reference space by employing a passive focus function. Using a brightfield microscope at high magnification (100x objective), the upper and lower surfaces of the tissue section were defined as the “just out of focus” optical planes, i.e., above and below the first in-focus optical planes. The desired level of accuracy for these automatic determinations was within 1.0 µm on average from the section thickness determined by manual measurement (ground-truth). With a typical step size of 1.0 µm between images collected along the Z-axis, this level of accuracy requires determining surface location within one image from a manually determined surface location. No attempt was made to
interpolate between Z-stack images to locate a more refined boundary. If better accuracy than 1.0 μm is required, Z-stacks with submicron step size, as small as 0.05 μm, can be captured and tested for this purpose.

A second goal of this thesis was to develop a method for training the automated surface location algorithm across a range of threshold parameters to optimize the performance of a given focus function. Fourteen focus functions found in previous literature [3-16] as well as four additional focus functions developed in this thesis were employed in turn in the automated surface location algorithm. Eighteen Z-stacks of images were identified arbitrarily prior to analysis for use as the training set. Since each Z-stack has an upper and lower surface, there were 36 tissue surfaces to be measured in the training set of Eighteen Z-stacks. The automated surface location algorithm requires selection of a threshold so as to determine whether the focus measure output by the selected focus function is considered in or out of focus. This threshold is referred to as the focus threshold. In addition to the focus threshold parameter used by the surface location algorithm, some of the functions analyzed require another threshold within the function. This threshold within the function is referred to as the contrast threshold. The contrast threshold is used to evaluate each pixel at the local level. The eighteen focus functions were divided into two categories - thresholded functions and non-thresholded functions. Non-thresholded functions do not have a contrast threshold. For thresholded functions, the Nelder-Mead simplex search method [17] for two dimensional parameter spaces was used to find the optimal pairing of a focus threshold and contrast threshold to yield the
lowest average error rate on the training set. For non-thresholded functions the golden ratio optimization method [17] was used to find the focus threshold again yielding the lowest average error rate on the training set.

The third goal was to find the focus function that required the least amount of parameter tweaking when changing from study to study. Because of variation in the appearance of biological microstructures on tissue sections, focus function threshold parameters typically require retraining for different datasets. The most desirable focus function would have thresholds that perform well across differing datasets, without retraining, for different hardware, biological variation in microstructures, and different staining methods routinely used to process tissue sections for computerized stereological analysis. Thus, the third aim of this thesis was to evaluate the robustness of the threshold selection over different datasets.

1.3 Novelty

There has been much research in the analysis of focus functions in microscopy. Fourteen of the eighteen functions analyzed in this thesis have appeared in several well cited papers. Many comparative studies include from just a few to as many as eighteen of these common functions in their evaluations. The novelty in some works comes solely from the subject matter being studied or the type of microscope being used. Because there is great variety in the images being analyzed, a focus function that works well for one task may not work well for the next task. So selecting the best performing function has largely been task specific.
Yet one thread that every microscopy image focusing study has in
common is that the focused image desired is the image with the maximum clarity
of focus. The maximum focal plane is not, however, the aim of this thesis. The
aim of this thesis is to find that minimum measure of focus that indicates that the
objective lens optical plane no longer resides in the empty space above or below
the tissue but is coincident with the tissue surface.

1.4 Thesis Organization

Chapter 2 gives a brief overview of the elements of microscopy and
stereology that are helpful in understanding the methods and motivations for this
thesis. Chapter 3 reviews previous work that justifies our choice of focus
functions to analyze, as well as explains the novelty of this thesis. Chapter 4
describes the functions analyzed and how they were used within a learning
program and in an automated surface location algorithm. The generation of
ground truth is also discussed. The data sets used are described in Chapter 5. In
Chapter 6 the performance results as well as their statistical significance are
discussed. The achievement of our objectives is reviewed in Chapter 7.
Conclusions and future work are provided in Chapter 8.
CHAPTER 2

BACKGROUND

2.1 Microscopy

The purpose of this thesis is to determine the top and bottom of a specimen from images captured using a compound light microscope. The three dimensional Cartesian coordinates are set with the X and Y coordinate plane parallel to the stage of the microscope. The stage is independently movable in all three coordinate directions. This movement can be manually or computer controlled. The slide being examined is assumed to be parallel to this plane X-Y plane. The Z-axis is perpendicular to the microscope stage (and slide being examined) and parallel to the axis passing through the center of the objective. The Z coordinate increases in the downward direction as it moves away from the objective lens. In order to calculate the top and bottom of a specimen, a series of images must be captured along the Z-axis at a particular X-Y position. Starting with the microscope objective focal plane above the specimen, images are captured at a manually determined step size (typically 1.0 µm) as the stage is raised and the objective focal plane moves down along the Z-axis. The image should start to come into focus as the focal plane approaches and passes through the top of the specimen. As the focal plane is incrementally lowered through the specimen, portions of the image will come in and out of focus. When the focal plane is below the bottom of the image there will be no more regions on
the image that are in focus. This set of images, at a set X-Y position but varying Z position is called a Z-stack (see Figure 2.2).

In stereology, although we speak of a focal plane, the region of the object that is in focus at a given time is not two but three dimensional. The depth of field (DOF) is the measure along the Z-axis where the object appears in focus, this corresponds to the depth of focus as seen in the image. Finding focus in microscopy can be difficult because within the field of view there are generally objects at varying depth, so if one object is within the DOF another object may not be. For measuring depth from focus, DOF should be as small as possible.

\[
DOF = \frac{\lambda \eta}{(NA)^2} + \frac{e \eta}{M(NA)} \tag{1}
\]

\[
NA = \eta \sin \alpha \tag{2}
\]

\[
e = \text{Max}(2\rho, 240\text{nm}) \tag{3}
\]

Using Equation (1), \(\lambda\) is the wavelength of illuminating light (550 nm for average visible light); \(\eta\) is the refractive index (1.516 for immersion oil); NA is the objective numerical aperture. M is the lateral magnification; and, \(e\) is the lateral resolving power. A high numerical aperture is desirable to minimize DOF. The numerical aperture is the product of the refractive index and the sine of the angle, \(\alpha\), that is the half angle of the cone of light entering the objective lens (see Figure
2.1). The working distance is the distance from the front element of the objective lens to the top of the cover slip. The smaller the working distance is the larger the angle $\alpha$ will be and therefore the closer the $\sin(\alpha)$ is to its maximum value of 1. For high magnification oil immersion lenses, the working distance is minimal (170 $\mu$m in this thesis). The resolving power is constrained to a minimum of 240 nm by the Abbe diffraction limit and is also constrained to twice the image pixel length, $p$.

The second term of Equation (1) becomes insignificant at high magnification and high numerical aperture (100x oil immersion, NA 1.4). Oil immersion objectives also help reduce the depth of field. Without an immersion objective light travels from below the sample, through the slide, up through the cover slip, through the air, and into the objective lens. It is diffracted when it

![Image of optical setup]

Figure 2.1 Comparison of Oil Immersion and Air

The second term of Equation (1) becomes insignificant at high magnification and high numerical aperture (100x oil immersion, NA 1.4). Oil immersion objectives also help reduce the depth of field. Without an immersion objective light travels from below the sample, through the slide, up through the cover slip, through the air, and into the objective lens. It is diffracted when it
enters the air and again when it enters the glass objective lens. The oil used in an oil immersion objective is designed to have the same refractive index as glass ($\eta$ of air $= 1.000$, $\eta$ of glass $= 1.515$, $\eta$ of immersion oil $= 1.516$). So there is no refraction between the cover slip and the oil or between the oil and the objective lens. So with the proper choice of objective the practical light microscope minimum DOF can be achieved. A shallow depth of field results in a shallow depth of focus which is desirable as reducing depth of focus is equivalent to enhancing axial resolution [18]. A depth of field of less than 1.0 µm is desired to attain the aim of 1.0 µm precision.

2.2 Stereology

In brightfield microscopy, the study of microstructures using unbiased designed based stereology plays a large and growing role in bioscience research. Stereology enables objective quantitative analysis of biological structures within a tissue sample. This allows development of normative standards for easy and efficient comparison to normative healthy samples [19]. As more bioscientists discover the experimental advantages of stereology it is anticipated that it will become the state of the art in histology and cytometry. As T. M. Mayhem and G. J. Burton stated in a 1997 paper in *Microscopy Research and Technique*, “Our wish to emphasize stereology stems from the fact that, currently, no other form of section-based morphometry can match it in terms of the unbiasedness and efficiency of its estimators.” [20]

The origins of modern stereology can be traced to 1960 in the Black Forest of Germany where a multidisciplinary group of geologists, biologists, and
materials scientists met to discuss their common problem of quantification of 3D objects from their appearance in 2D. The biologist Hans Elias coined the term “Stereology” to describe this subject [19] and the science of stereology was born. The next year the first International Congress for Stereology was held. A central problem with which stereologists struggled is the corpuscle problem, illustrated in Figure 2.2, described in 1925 by S. D. Wicksell [21]. The corpuscle problem illuminates the difficulty of estimating the distribution of objects in 3D from the object profiles in 2D. The problem was not solved until 1984 by D. C. Sterio (a pseudonym and anagram for disector) and his idea of the physical disector [22], a 3D probe consisting of two planes with a known separation between them (see Figure 2.3) and a set of counting rules.

The disector principle enabled counting of objects without modeling the object shape and with no assumptions as to their shape, size, orientation or distribution. Biological objects are generally not well represented by classical Euclidian shapes (such as spheres, ellipsoids, cubes, lines). Modeling biological objects as classical shapes requires making assumptions that are simply not true. Correction factors were added to models with less than satisfactory results and systemic errors. With the physical disector and the many stereology probes that were designed following the advent of the disector principle, stereology shifted from model based to design based. Probes are systematically yet randomly placed within the reference space. The number of samples to measure can be efficiently chosen to estimate the stereology parameter within a desired minimum coefficient of error. By implementing stereological principles a
Figure 2.2 The Corpuscle Problem. Images 1 through 8 represent a Z-stack of the tissue sample shown. There are four 3D objects in the tissue sample, yet fourteen 2D profiles of the objects in the eight images.

A researcher can develop a clear picture of region volume, object count, orientation, and distribution, and variation between specimens by utilizing a relatively small number of manual samples in a region. Typically studies require counting less than a few hundred objects for each subject and generally only ten to twenty subjects are required from control and test groups combined. This
compares to non-stereology studies where tens or even hundreds of thousands of objects are manually counted in a single specimen, with no easy or meaningful way of comparing data parameters. Free from assumptions, poor models, and correction factors, and with workload reduced by a factor of a hundred or more, the use of stereology as the preferred method of histological analysis was inevitable.

Figure 2.3 The Physical Disector. Objects are only counted if they are within the disector lower frame and do not touch the look-up plane or the (red) exclusion line.

Each stereology parameter to be measured or estimated required a probe designed specifically for the task. First order stereology parameters include count, length, surface area, and volume. All the design base probes used to estimate first order stereology parameters require knowledge of the thickness of the sample tissue as well as the location of the top and bottom surface of the
tissue corresponding to the probe placement. The thickness is required to accurately calculate the fraction of the tissue that the probe covers. The surface location is required to set guard or buffer planes to avoid bias caused by physical cutting artifacts.
CHAPTER 3
PREVIOUS WORK

3.1 Focus Functions

In typical microscopy autofocus studies, the goal is to find the depth at which the biological microstructures appear at maximal focus [2, 4, 7, 8, 10-14, 23, 24]. This seems reasonable for a traditional histology or cytometry study and even in the rapidly advancing science of stereology [25] the counting of objects of interest occurs when such objects are in optimal focus. Yet for analysis of such counts in stereology, knowledge of the tissue thickness as well as tissue surface locations is essential.

A shared aim of each of the above referenced studies is determining the optimal focus location by finding the global maximum of a focus curve. However, each study is unique in some other aspect, such as:

1. Equipment used: brightfield, darkfield, fluorescent, confocal, differential interference contrast (DIC), or phase contrast microscopes;
2. Method of preparation: various stains, time-lapse, live;
3. Specimen subjects: Pap smears, blood smears, sputum, diatoms, etc.

The selection and application of an appropriate focus function for a specific sample and application is task dependent [10, 11].
Utilizing a focus measure to determine the optical plane locations corresponding to the top and bottom tissue surfaces is a different application of the focus function. The tissue surface is delineated as the optical plane that separates the unfocused region from the focused region of the focus curve. Accurately locating the boundaries between the focused and unfocused depth of field for microstructures in the reference space is a requirement for estimation of both first-order (e.g., number, volume, length, surface area) and second-order (e.g., variation, spatial orientation) stereological parameters [19].

Therefore, what differentiates this thesis from previous studies of autofocusing is not the focus functions analyzed, but the portion of the focus curve produced by the focus functions that is considered significant. For the purposes of stereology, the aim is to determine the thickness of the biological tissue and the location of the tissue surfaces. In typical autofocus studies, emphasis is on finding a focus function that yields a global maximum on the focus curve where optimal focus is attained. Whereas in this thesis, emphasis is on finding the focus function that yields the most drastic change when transitioning from unfocused to focused regions of the focus curve. Two sharp bends or “knees” in the focus curve near the local minima are desired (see Figure 6.1). When traversing the focus curve from the right (i.e. moving down through the Z-stack from above the upper tissue surface) the curve should bend sharply as it changes from a low (ideally the focus value should be zero) flat section to a steep rise. Traversing the focus curve from the left (beginning below the bottom surface) a similar “knee” should be present as the tissue surface is
crossed. The low focus measure images on the right and left should correspond to the images where the focal plane is either above the top surface of the tissue sample or below the bottom surface of the sample. Approaching from either right or left, the last image where no significant focus measure is detected is the “just out of focus” bounding image. The very next image examined should exhibit a sharp rise in the focus measure indicating that there is now something in the focal plane. The ideal focus curve for this purpose should consistently and unambiguously demarcate this boundary. Such a boundary should be the transition on the focus curve from an unfocused region to the focused region and back to the other unfocused region.

3.2 Automated Stereology Programs

For this thesis, the work of collecting Z-stacks of images was accomplished using a Stereologer system. The Stereologer can automatically capture and store Z-stacks of images for a complete case study. The user must select the reference space to be analyzed on each tissue section. The Stereologer can then generate the systematically random probe locations and move the microscope stage precisely to each X-Y location and then precisely move through the Z-axis capturing the Z-stack of images at each location. This is a tremendous time saver compared to collecting Z-stack images manually.

Typically in practice, the Stereologer is used with live video images with no need to store the still images. In this live mode, the trained user must select the top and bottom surface manually at each X-Y location. This is the process that this thesis looks to automate.
CHAPTER 4
METHODOLOGY

4.1 Methods Used

Fourteen common grayscale focus functions were analyzed. Well-studied focus functions exist for finding the maximal point of focus as required for stereology studies (see Table 4.1). Each function interprets focus as some measure of intensity variance between the pixels (either locally or globally) of the image. In this thesis, the qualities of these common functions that are of interest differ from the qualities of interest identified in past works. On the focus curve, as opposed to a smooth yet narrow change to an apex for finding the location of maximum focus, the ideal focus-curve function for use in stereology probe placement and parameter measures should have sharp or abrupt changes (knees) near the minima. This is where the images change from unfocused to focused regions and from focused to unfocused regions. It is this transition that marks the surface of the tissue sample.

All the focus functions studied were designed for grayscale images. The images captured were originally color images with the color of each pixel represented by three color channels, RGB (8 bits per channel). There are several acceptable methods to convert to grayscale. Equation (4) [26, 27] is the conversion from color to grayscale used in this thesis. Equation (4) was developed from the YUV color space approximating human perception. YUV
color space separates luminance (our perception of intensity or brightness), $Y$, from the color components, $U$ and $V$. This formula is common to Adobe’s Photoshop and the popular freeware image processing program Irfanview.

$$I(i,j) = .30R + .59G + .11B,$$
where $R, G, B$ are the color channel intensity values \( (4) \)

An automated surfaced location algorithm was implemented in a C++ program. Any of the fourteen focus functions can be chosen as the driver for the surface location algorithm. The focus functions were divided into thresholded functions and non-thresholded functions. Thresholded functions require a contrast threshold that determines on a pixel to pixel level which data to include in the focus measure. The non-thresholded functions do not require a contrast threshold, yet the automated surface location algorithm requires a global focus threshold parameter for every function. This focus threshold is used to determine if the image is in focus.

The focus threshold is a constant derived for each focus function by a parameter learning program also implemented in C++. The parameter learning program takes as input a set of Z-stacks of images as well as the classification of the images in the Z-stacks. The images are classified as either surface images or non surface images. The parameter learning program seeks to minimize the error in determining the surface location for the set of Z-stacks. In the case of thresholded functions, the parameter learning program must selection of the contrast threshold as well as the focus threshold that together yield the lowest error rate. This involves searching the 2D parameter space for a minimum. For
this problem, the Nelder-Mead simplex search method was utilized to converge on a minimum. In the case of the non thresholded functions, the parameter learning program must search the 1D parameter space of the focus threshold. For this problem, the golden ratio search method was used for its simplicity and rapid convergence to an optimal value.

The focus threshold was originally proposed to be a function of the focus function. The focus threshold as a function of the maximum focus measure or the range of focus measure did not seem reasonable because a Z-stack could have several finely detailed objects in a single focal plane that would return a high focus value or the stack could have little to focus on within a single focal plane. In either case this would have little bearing on the just out of focus images. The focus threshold as a function of the minimum focus measures seemed more reasonable. Yet, ideally, the unfocused images should have a zero measure of focus, so taking a multiple of the minimum could (or ideally should) return zero. The focus threshold as a function of the rate of change of the focus function also seemed like a good idea. However, although ideally there is a large change at the tissue surface, the maximum rate of change often occurs within the tissue and not at the surface. Nonetheless, the first and last slope over an optimized threshold would seem to work as indications of tissue surface. Yet, in practice the idealized instantaneous increase in slope often did not occur, instead a more gradual increase in slope over a few images was observed. In the case of the gradient functions, which appeared the most promising, the focus function was already a derivative, and taking a second derivative was perhaps subjecting
the algorithm to too much noise. Thus a constant threshold was used and this worked well for several functions.

In addition to the fourteen common focus functions four additional focus functions were developed. Based on the observation of the common focus functions performance with the training set and the success of the parameter learning program, modifications were made to the thresholded absolute gradient function to yield four novel functions.

4.2 Common Focus Functions

Table 4.1 Citations for Common Focus Functions

<table>
<thead>
<tr>
<th>Equation</th>
<th>Name</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thresholded Functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>[2, 4-6, 8, 10, 12-15]</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>[2, 4, 5, 8-10, 12-14]</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>[2, 3, 7, 8, 10-15]</td>
</tr>
<tr>
<td>(9)</td>
<td>Content</td>
<td>[4, 7-10, 12-14]</td>
</tr>
<tr>
<td>(10)</td>
<td>Pixel Count</td>
<td>[4, 8-10, 12-14]</td>
</tr>
<tr>
<td>(11)</td>
<td>Image Power</td>
<td>[4, 7-10, 12-14]</td>
</tr>
<tr>
<td>Non-Thresholded Functions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12)</td>
<td>Tenenbaum</td>
<td>[6, 8, 10-16, 23, 24, 28]</td>
</tr>
<tr>
<td>(13)</td>
<td>Energy Laplace</td>
<td>[4, 6, 10-14, 16]</td>
</tr>
<tr>
<td>(14)</td>
<td>Variance</td>
<td>[2, 4, 6, 8, 10, 12-14, 16]</td>
</tr>
<tr>
<td>(15)</td>
<td>Normalized Variance</td>
<td>[2, 4, 7, 8, 10-15]</td>
</tr>
<tr>
<td>(16)</td>
<td>Autocorrelation</td>
<td>[2, 5, 8, 10-15, 24]</td>
</tr>
<tr>
<td>(17)</td>
<td>Standard Deviation Based Autocorrelation</td>
<td>[2, 8, 10, 12-15]</td>
</tr>
<tr>
<td>(18)</td>
<td>Range</td>
<td>[7, 8, 10, 12-14]</td>
</tr>
<tr>
<td>(19)</td>
<td>Entropy</td>
<td>[6-8, 10, 12-14]</td>
</tr>
</tbody>
</table>
The equations below briefly describe the fourteen common grayscale focus functions analyzed and listed in Table 4.1. First the common thresholded functions are discussed, followed by the common non thresholded functions. Finally the functions developed for this thesis – all of which are thresholded – are described.

By requiring a threshold, the ten thresholded functions analyzed, six in Table 4.1 and four modified functions based on the thresholded gradient functions (discussed in Section 4.4 New Focus Functions), have an additional level of complexity compared with the eight non thresholded functions. Yet the threshold enables “fine tuning” of the function. Equation (5) is a simple thresholding function that returns a given value if that value is greater than or equal to the specified threshold. Equation (5) is used to add clarity to the thresholding equations that follow.

\[
Threshold(x,T) = \begin{cases} 
  x, & x \geq T \\
  0, & x < 0 
\end{cases} \text{, where } x, T \in \text{integers} 
\] (5)

Equations (6) through (11) show the thresholded focus functions implemented. The first three are gradient based functions. As gradients these functions attempt to measure the pixel to pixel intensity differences similarly to edge detectors. Such functions are often effective in determining when there is a boundary change (e.g. image moves from unfocussed to focused). The next three are so called “intuitive” [10, 14] equations: Note that the pixel count function uses the threshold as an upper limit, as opposed to all the other thresholded functions that use the threshold as a lower limit. All three intuitive
functions attempt to derive focus information from the overall brightness of the image. The intuition here is unclear.

**Absolute Gradient:** \[ F_{ag} = \sum_i \sum_j |\text{Threshold}(I(i, j + 1) - I(i, j), T)| \] (6)

**Squared Gradient:** \[ F_{sg} = \sum_i \sum_j \text{Threshold}((I(i, j + 1) - I(i, j))^2, T) \] (7)

**Brenner Gradient:** \[ F_{bg} = \sum_i \sum_j \text{Threshold}((I(i, j + 2) - I(i, j))^2, T) \] (8)

**Content:** \[ F_{ct} = \sum_i \sum_j \text{Threshold}(I(i, j), T) \] (9)

**Pixel Count:** \[ F_{pc} = \sum_i \sum_j \text{PCount}(I(i, j), T), \text{ where } \text{PCount}(x, T) = \begin{cases} 1, & x \leq T \\ 0, & x > T \end{cases} \] (10)

**Image Power:** \[ F_{ip} = \sum_i \sum_j \text{Threshold}(I(i, j)^2, T) \] (11)

where \( I(i, j) \) is the grayscale intensity value at pixel \((i, j)\).

### 4.3 Common Non-Thresholded Focus Functions

Equations (12) through (19) make up the conventional non-thresholded focus functions that were evaluated. Non-thresholded focus functions remove a level of complexity by removing the need to select a contrast threshold. Tenenbaum gradient and energy Laplace are gradient based functions. Equations (14) through (17) are statistically based functions, all four have been designated as top performing functions for optimal focus [5, 8, 10, 11, 14]. Equations (17) and (18) are histogram based functions.
Tenenbaum Gradient: 
\[ F_{tg} = \sum_{i} \sum_{j} (S_x(i, j)^2 + S_y(i, j)^2) \]  
where \( S_x \) and \( S_y \) are the convoluted images with Sobel operators. (12)

Energy Laplace: 
\[ F_{el} = \sum_{i} \sum_{j} L(i, j)^2 \]  
where \( L \) is the convoluted image with Laplace mask: 
\[
\begin{bmatrix}
-1 & -4 & -1 \\
-4 & 20 & -4 \\
-1 & -4 & -1 \\
\end{bmatrix}
\] (13)

Variance: 
\[ F_{var} = \frac{1}{H \cdot W} \sum_{i} \sum_{j} (I(i, j) - \bar{I})^2 \]  
where \( \bar{I} \) is the mean intensity, \( H \) and \( W \) are the number of pixel rows and columns (14)

Normalized Variance: 
\[ F_{nvar} = \frac{1}{H \cdot W \cdot \bar{I}} \sum_{i} \sum_{j} (I(i, j) - \bar{I})^2 \] (15)

Autocorrelation: 
\[ F_{acor} = \sum_{i} \sum_{j} I(i, j) \cdot I(i, j + 1) - \sum_{i} \sum_{j} I(i, j) \cdot I(i, j + 2) \] (16)

Standard Deviation Based Correlation: 
\[ F_{sdcor} = \sum_{i} \sum_{j} I(i, j) \cdot I(i, j + 1) - H \cdot W \cdot \bar{I}^2 \] (17)

Range: 
\[ F_{range} = \max \{ \text{hist}(i) > 0 \} - \min \{ \text{hist}(i) > 0 \} \]  
where \( \text{hist}(i) \) is the number of pixels of intensity \( i \) (18)

Entropy: 
\[ F_{entropy} = -\sum_{i} p_i \cdot \log_2 (p_i), \text{ where } p_i = \frac{\text{hist}(i)}{(H \cdot W)} \] (19)

The format of variance and normalized variance [Equations (14) and (15)] matches that of variance and normalized variance functions cited in the literature (see Table 4.1). In this format both functions require an initial pass to calculate the mean intensity. A more efficient yet equivalent calculation of variance and
normalized variance [Equations (22) and (23)], requiring only a single pass through each image was used in the automated surface location algorithm [29]:

\[
F_{\text{invar}} = \frac{1}{H \cdot W} \sum_i \sum_j I(i, j) - (\bar{I})^2
\]

where \( \bar{I} \) is the mean intensity, \( H \) and \( W \) are the number of pixel rows and columns

\[
F_{\text{invar}} = \frac{1}{H \cdot W \cdot \bar{I}} \sum_i \sum_j I(i, j) - \bar{I}
\]

4.4 New Focus Functions

Equations (22) through (24) represent intermediate functions that simplify the descriptions of four novel focus functions [Equations (25) through (28)] developed for this thesis. These four functions were created by modifying the thresholded absolute gradient function [Equation (6)]. Equation (22) is an indicator function that signifies whether the current pixel location \((i, j)\) is high contrast, and returns 1 if this is true or 0 otherwise. High contrast for this function occurs when the absolute value of the difference between the current pixel and either its neighboring pixel to the right or directly below is greater than or equal to the designated threshold value. Equation (23) describes the binary 3 x 3 median filter used in functions represented by Equations (27) and (28). Equation (23) determines the number of the eight immediate neighboring pixel locations with high contrast. Equation (24) returns the thresholded absolute gradient contrast value in both the horizontal and vertical direction for a given pixel location \((i, j)\).
\[
\text{Count}(i, j, T) = \begin{cases} 
1, & |I(i, j + 1) - I(i, j)| \geq T \text{ OR } |I(i + 1, j) - I(i, j)| \geq T \\
0, & \text{otherwise}
\end{cases}
\]  
(22)

\[
\text{Median}(i, j, T) = \begin{cases} 
\text{Count}(i - 1, j - 1, T) + \text{Count}(i - 1, j, T) + \text{Count}(i - 1, j + 1, T) \\
+ \text{Count}(i, j - 1, T) + \text{Count}(i, j + 1, T) + \\
\text{Count}(i + 1, j - 1, T) + \text{Count}(i + 1, j, T) + \text{Count}(i + 1, j + 1, T)
\end{cases}
\]  
(23)

\[
\text{ModAbsThr}(i, j, T) = \begin{cases} 
\text{Threshold}\{I(i, j + 1) - I(i, j)\}, T\} + \\
\text{Threshold}\{I(i + 1, j) - I(i, j)\}, T\}
\end{cases}
\]  
(24)

Equations (25) through (28) represent the four new focus functions introduced in this thesis. The modified absolute gradient [Equation (25)], is a rotationally invariant form of the absolute gradient [Equation (6)]. The modified absolute gradient count [Equation (26)] is also rotationally invariant and is a simple count of high contrast pixels as determined by Equation (22). Application of a 3 x 3 median filter to the modified absolute gradient and the modified absolute gradient count functions reduce spurious noise. Equations (27) and (28) are filtered versions of Equations (25) and (26).

Modified Absolute Gradient:

\[
F_{mag} = \sum_i \sum_j \text{ModAbsThr}(i, j, T)
\]  
(25)

Modified Absolute Gradient Count:

\[
F_{magc} = \sum_i \sum_j \text{Count}(i, j, T)
\]  
(26)
Filtered Modified Absolute Gradient:

\[
F_{fmag} = \sum_i \sum_j \begin{cases} 
ModAbsThr(i, j, T), & \text{if } \text{Count}(i, j, T) = 1 \text{ AND Median}(i, j) \geq 4 \\
0, & \text{otherwise}
\end{cases}
\]  

(27)

Filtered Modified Absolute Gradient Count:

\[
F_{fmagc} = \sum_i \sum_j \begin{cases} 
1, & \text{if } \text{Count}(i, j, T) = 1 \text{ AND Median}(i, j) \geq 4 \\
0, & \text{otherwise}
\end{cases}
\]  

(28)

4.5 Interpretation of Focus Functions

For each Z-stack of images, a focus curve was generated for each of the focus function (Equations 5-18, 24-27). The acquisition of Z-stacks of images began with the optical plane above the top surface of the tissue and ended at a focal plane below the bottom surface. The goal for accuracy for automated identification of tissue location was an average error rate to 1 µm. The incremental step-size through each Z-stack was 1 µm. Thus the image identified by the automated surface location algorithm had to be no more than a single image away from the manually determined surface image on average. At the micron level of precision, the tissue surface is not flat. For stereology, the surface planes are the closest X-Y planes above and below the tissue that completely contain the tissue. Practically, these are the "just out of focus" images with no interpolation between images. This approach differs markedly from the purpose of traditional approaches. For example, the work of Osibote et al. [11] seeks to locate the optical plane of peak focus. Osibote reasoned that focus functions behave like Gaussian curves near their maximum [30] and exploited the knowledge that the logarithm of a Gaussian is quadratic, enabling
fitting of a parabolic curve between adjacent images to determine optimal peak focus along the focus curve. Since the "just out of focus" area of the focus curve, the primary emphasis of this thesis, does not approach a Gaussian or any other well-defined function, no interpolation or curve fitting was carried out.

To locate the just out of focus images, the ideal focus curve would be capable of differentiating between three regions in each Z-stack:

1. The unfocused images above the tissue sample
2. The images within the tissue sample – assumed to be bounded by in focus images
3. The unfocused images below the tissue sample.

It is assumed that the two unfocused regions on either side of the in focus region behave similarly. Specifically, these regions of the focus curve response should ideally be zero, with no points of focus in the image. In practice, these regions would be at or close to a minimum and relatively flat compared to the in focus region, with no assumption that these regions would be monotonic. In the "in focus" region, the only requirement is that the bounding images of the region have a higher focus measure than every image from the unfocused regions. There was no assumption of unimodality in the region and no requirement for all these images to yield higher focus measures than those of the unfocused images.

4.6 Ground Truth

There is a subjective nature to manually determining focus level. Some variation between trained stereologists is expected [25]. Table 4.2 illustrates the
variance of manual surface determination. Three students were trained in locating just out of focus images in Z-stacks and then asked to independently ground truth our training set as well as test set #1. The dataset designation of training set and test set pertain to the parameter learning program and are not relevant to ground truthing. The students were unaware of any distinction between the two sets of data.

Table 4.2 Select Ground Truth of Color Image Z-stacks With 1.0 µm Step.

<table>
<thead>
<tr>
<th>Manual Focus Determination (Ground Truth)</th>
<th>Average Variance over the Dataset with Three Independent Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trials</td>
<td>Variance (µm²)</td>
</tr>
<tr>
<td>Training (36 surfaces measured)</td>
<td>0.69</td>
</tr>
<tr>
<td>Test set #1 (32 surfaces measured)</td>
<td>0.82</td>
</tr>
</tbody>
</table>

4.7 Optimization of Threshold Parameters

4.7.1 Nelder-Mead Simplex Search Method

The Nelder-Mead simplex search method works well in multi dimensional parameter spaces [17, 31]. Training the thresholded focus function in the parameter learning program requires two parameters: the contrast threshold and the focus threshold. A simplex is a shape that has one more vertex than the dimensions of the space in which it lies. So a triangle is a simplex in 2D. After initialization of the simplex vertices, the simplex either expands or contracts as needed to search the parameter space for a minimum. When a minimum is found the simplex tries to expand in the same direction to find a better minimum.
When no better minimum can be found by expanding, the simplex begins to contract around its current minimum.

Nelder-Mead works well even when parameters are of different scale as is the case with the thresholded functions in this thesis. Contrast threshold values for the focus functions are integer values within a small range, e.g. for absolute gradient the possible range is from 0 to 255, while the practical range is from 0 to about 160. The focus threshold, representing an accumulation of contrast values is much larger. For example, in an 800 x 600 pixel image, the absolute gradient focus threshold has a possible integer range from 0 to 122,400,000, while the practical range is from 0 to about 1,000,000. With more training and experience in the lab, perhaps the contrast threshold search range can be further limited to just a few choices in which case using a 1D parameter space search algorithm, such as the golden ratio method might be an option.

For Nelder-Mead, the initial vertices must be chosen and a reasonable choice is to use orthogonal vectors that cover about 20% of the space. Because the possible range of threshold parameters can be calculated for each function a reasonable initial simplex can be easily chosen. The other parameters that must be set for the Nelder-Mead are stopping criteria. For the parameter learning program the stopping conditions could be set for each focus function individually. For the absolute gradient function the stopping conditions were a contrast threshold change of less than two levels of intensity as well as a change in the focus threshold that was less than 100. As an additional fail safe, if the Nelder-Mead has not converged after one hundred points are analyzed, the learning
program automatically stops. After the algorithm stops, it is reinitialized with one point at the minimum and the process is repeated a second time.

4.7.2 Golden Ratio Method

Automation of the non-thresholded functions requires optimization in only 1D parameter space. The golden ratio optimization method [17] was for its simplicity and efficiency. The golden ratio method works by choosing evaluating the function at the end point at each extreme of the 1D parameter space and then a third point between the endpoints. The ratio of the distance between each end points and the point between is fixed as the golden ratio φ (1.618…). A fourth point is then evaluated between the larger segment also maintaining the golden ratio. The points around the lower inner point are kept and the distant end point is dropped. And the process is repeated, until the distance between end points falls below the stopping criterion. The end points were selected by taking the maximum and minimum focus curve values on the training set data. The stopping criterion was when the distance between converging end points was less than 1% of the distance between the originally selected end points.
A Z-stack was defined as a series of images taken along the optical axis at incremental steps in the X-Y plane. The Z-axis is coincident with the optical axis and perpendicular to the X-Y plane. Three separate datasets (Table 5.1) consisting of Z-stacks of images were acquired at high magnification from sections stained with tyrosine hydroxylase to reveal dopaminergic neurons in the rat substantia nigra or cresyl violet stained pyramidal neurons in the CA region of the rat hippocampus [sections supplied by the Stereology Resource Center (Chester, MD)]. Using the Stereologer, Z-stacks were captured as follows:

1. Anatomically defined reference spaces (the substantia nigra and the hippocampus) were manually selected at low magnification (2.5x objective).

2. After switching to high magnification, a series of X-Y locations within the reference space were selected in a systematic-random manner.

3. Top and bottom optical planes at each X-Y location were found by manually locating boundaries between unfocused and focused images at the top and bottom of each tissue section.

4. The step increment and buffer in Z-axis were set to ensure acquisition of unfocused images above and below tissue.
5. The Stereologer system automatically acquired Z-stacks.

6. The system moved through the X-Y plane of the reference space repeating steps 3, 4 and 5 above.

The acquired Z-stacks were converted to grayscale and divided into three datasets – a training set and two test sets. The training set and test set #1 were comprised of Z-stacks from tissue sections through a single brain region (substantia nigra) from a single rat subject, taken in the same sections. A second test set (test set #2) was comprised of Z-stacks from a different rat subject, from a different case study, a different brain region (hippocampus), and using a different stain. Certain Z-stacks were rejected from the thesis during manual determination of surfaces. Reasons for these rejections will be discussed later in this chapter.

Table 5.1 Characteristics of Datasets Used for Evaluation

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Description</th>
<th>Z-stacks</th>
<th>Images Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Set</td>
<td>Cryostat 010610 Substantia Nigra tyrosine hydroxylase to reveal dopaminergic neurons</td>
<td>18</td>
<td>455</td>
</tr>
<tr>
<td>Test Set #1</td>
<td>Cryostat 010610 Substantia Nigra tyrosine hydroxylase to reveal dopaminergic neurons</td>
<td>16</td>
<td>373</td>
</tr>
<tr>
<td>Test Set #2</td>
<td>Nissl 041696 Hippocampus cresyl violet stained pyramidal neurons</td>
<td>18</td>
<td>490</td>
</tr>
</tbody>
</table>

Even with calibration of the microscope in the X, Y and Z axes it would be extremely difficult to acquire the exact same stacks of images – mapping pixel to pixel – in two separate trials. Slight unrecordable movement in the microscope stage can be caused during placement and removal of slides. The slide holder
as a movable part has a slight degree of positional variance. The coordinate system can be calibrated for extremely accurate measurement, but the zero position is not referenced to any fixed point and is easily set and reset during microscope use. These are significant reasons why it is common practice in microbiology stereology not to save location data within a case study. In this thesis, although spatial coordinates were saved, it is not possible to use this information to relocate or exactly reconstruct any Z-stack. However, the specific slide, the section on that slide, and the biological region (hippocampus or substantia nigra) are locatable.

5.1 Data Collection Equipment

The images analyzed were captured using a Stereologer system equipped with a motorized X-Y-Z stage (ASI, Eugene, Oregon) with capability of both manual and automated control via the system software. The specific configuration of the system included a Zeiss Axioskop 20 brightfield microscope with objectives for low magnification (Zeiss Plan Neofluar 2.5x, numerical aperture 0.075) and high magnification (Zeiss Plan-Apochromat 100x oil immersion, numerical aperture 1.4); an Optronics Microfire camera that captured 800 x 600 pixel images in 8 bit by three channel (RGB) color; and, the Stereologer system software was operated on both an iMac G4 platform (32 bit 1.25GHz PowerPC) as well as a Dell Optiplex GX280 (32 bit 3.20 GHz Intel Pentium4) running Microsoft Windows 7 Enterprise.

The charged-coupled device (CCD) for the Optronics Microfire camera is a single ¾” 1600 X 1200 pixels array. It uses a Bayer filter color filter array to
capture color information. This common Bayer filter places a mosaic filter over the CCD so that each 2 X 2 pixel square has an RGBG (25% Red, 50% Green, 25% Blue) color filter over it.

**5.2 Initial Data Collection**

The initial data collected over the first few months of the thesis was error prone and problematic in many ways. With a policy of not destroying any data collected, this set includes all data that was collected early in the research incorrectly. Out of the first 13 stacks that were attempted, only two stacks were deemed usable. Reasons for unusable stacks were:

1. The microscope and camera lenses were not clean. A significantly large layer of dirt and debris was visible on many images (see Figure 5.1). The dirt residing on the lens surfaces occludes the light and creates sharp edges that appear in focus regardless of the location of the focal plane of the microscope. Although this noise is relatively constant across a Z-stack, the level of noise in the measure of contrast focus can be so large that it obscures the intended measure of focus level of the tissue sample.

2. The microscope stage motor was not engaged with the software control. The z-coordinate did not change during image acquisition.

3. Images in the Z-stack became misaligned. The image processing lab where the microscope resides is less than ideal. The table on which the microscope is positioned is neither anchored to the wall nor to the floor and is therefore not completely stable. The floor in
the lab is raised, and this adds another point of instability. Any bump to the table or heavy movement on the floor will cause misalignment in the Z-stack being acquired.

4. Images were stored as jpeg (Joint Photographic Expert Group) files. Jpeg is a lossy compression method with artifacts that can affect the focus algorithms. All usable images are stored as bitmaps (BMP) files and processed as portable gray map (PGM) files.

![Figure 5.1 Unfocused Rat Hippocampus through Dirty Lenses. The black “strings and blobs” in this image are all caused by dirt on the surfaces of the lenses, blocking the light from reaching the camera sensor. There are also tens of small gray circles of dust over the entire image]
5.3 Training Set

This dataset is comprised of 22 stacks of images captured on a study of rat brain tissue prepared on January 6, 2010. The rat brain tissue was cut along the coronal axis in 40µm sections using a cryostat microtome. The tissue was stained. The substantia nigra was located and outlined as the region of interest using the 25X objective. The Stereologer created random systematic probes within the region of interest. Images were captured using the 100X oil immersion objective. Images were taken from a single rat case, on sections close in proximity. Images from sections 05, 07, 08, and 09 were included in this dataset. While ground truthing these Z-stacks it was realized that three of the Z-stacks
were not complete Z-stacks. Either there were no unfocused images at the top of the Z-stack or no unfocused images at the bottom. This is in breach of the assumptions made in designing the autofocus algorithm. Incomplete Z-stacks are problematic because without images in the unfocused region there is no unfocused region and without an unfocused region there is no boundary to find. Therefore, these three Z-stacks were removed from this training set.

The intensity range of a gray scale image is the difference of the maximum pixel intensity present in the image and the minimum intensity present in the image plus one (to include both the maximum and minimum). The maximum intensity possible in the bitmap format used for this thesis is 255 representing pure white. The minimum intensity is 0 representing pure black. So the maximum intensity range possible is 256. In the training set derived from this dataset the average intensity range for all 18 Z-stacks was 174. One Z-stack from the test set was rejected during ground truthing because it was too dark to derive meaningful data. The maximum intensity present in this Z-stack was only 68. The average intensity range per image in this Z-stack was 57 and the maximum range per image was only 64. Although it may be difficult to utilize the entire intensity range when imaging, using only 25% of the possible intensity range was deemed cause for rejecting this Z-stack (see Figure 5.3).

As a training set, this data was used to validate and develop the parameter learning program and the automated surface location algorithm. This data set was the only data used to affect the choice of parameters.
Figure 5.3 Rejected Z-stack from Training Set. This Z-stack of images was captured at a low light intensity. Consequently, only 25% of the available intensity range was used. Making accurate detection of focus level difficult both by manually observation as well as algorithmically.

5.4 Test Set #1

This dataset is comprised of 18 stacks of images acquired from the same study of rat brain tissue dated January 6, 2010. The rat brain tissue was cut along the coronal axis in 40µm sections using a cryostat microtome. The tissue was stained. The region of interest captured in these images is the substantia nigra. The images were captured using the 100X oil immersion objective. Z-stacks for this dataset were taken from the same study, the same rat and the same sections 05, 07, 08, and 09 as the training set. Two Z-stacks from this data
set were rejected during ground truthing for incompleteness, just like those rejected in the training set.

Figure 5.4 Test Set #2 Cresyl Violet Stained Pyramidal Neurons. In the Hippocampus.

5.5 Test Set #2

This dataset is comprised of 18 stacks of images captured from a study of rat brain tissue dated April 16, 1996. The rat brain tissue was cut along the coronal axis in 40µm sections. The tissue was stained using cresyl violet stain to reveal pyramidal neurons. The hippocampus was located and outlined as the region of interest using the 25X objective. The Stereologer was used to create random systematic probes within the region of interest. Images were captured
using the 100X oil immersion objective (see Figure 5.4). Images were taken from a single rat case over two adjacent sections. One Z-stack was rejected from this dataset because during slide preparation a flap of tissue was folded over the section creating a significant disparity in surface depth making ground truth difficult (see Figure 5.2).
CHAPTER 6
EXPERIMENTAL RESULTS

6.1 Results

Figure 6.1 through Figure 6.7 show the focus curves for a typical Z-stack in the training set for each of the fourteen commonly used functions analyzed. Figure 6.8 shows the focus curves for the four modified functions for the same Z-stack. All the functions were optimized over the training set to minimize surface location deviation from ground truth. The two red vertical bars in every graph in these figures indicate the manually determined top and bottom surfaces (i.e., ground truth). These ground truth bars separate each focus curve into three regions: the middle region between the ground truth bars, that is, the in-focus region; and the two out-of-focus regions on either side of the in-focus region that include the ground truth bars. The range of acceptable thresholds is shaded in green horizontally across the graph (see Figure 6.1, Figure 6.4b, and Figure 6.8) when it is possible to select correct focus thresholds that, correctly separate the focus curve at the ground truth bars.

The three thresholded gradient functions appear to be near ideal in Figure 6.1. Close inspection of the focus measure at the ground truth bar at 26 µm reveals that both squared gradient and Brenner gradient are beginning to rise more so than absolute gradient. This narrows the acceptable threshold band raising it above zero. Also of note, the "in focus" region between the ground truth
bounds is determined by the bounds and the in focus points adjacent to them, therefore the behavior of the curve between these points (in this case 27 µm and 41 µm) is not of interest. Figure 6.2 sorts the functions by their focus measure output and only the first and last in focus images are included.

Figure 6.1 Thresholded Gradient Focus Curves from a typical Z stack in the Cryostat Training Set (Rat C1 Sec09 XY02). The two red vertical bars depict the manually determined surface depth. The green region depicts the range of threshold values that correctly identify the surface depth by partitioning the Z-stack into a focused region bounded by an unfocused region on either side.
The thresholded intuitive focus functions did not fare as well. Figure 6.3 a) and b) are graphed with the same full scale (0 to 1 normalized) for the focus measure as that used in Figure 6.1, Figure 6.5, and Figure 6.6. Because there is

![Figure 6.2 Focus Curves Sorted by Focus Results for a typical Z stack. The distance between the two green bars is the range of acceptable thresholds. By maximizing this distance for each Z-stack the odds of better performance increase.](image)
not as much variation in these focus curves, which is problematic in its own right, a second graph of each was produced (Figure 6.4) with significantly smaller scale in order to reveal characteristics of the curves. These focus curves do

Figure 6.3 Thresholded "Intuitive" Focus Curves from a typical Z-stack in the Cryostat Training Set (Rat C1 Sec09 XY02). The two red vertical bars depict the manually determined surface depth. For comparison, these graphs are at the same scale as those of Figure 6.1.
suggest three regions, yet they do not match the manually determined ground truth. Also of note as the scale is changed to augment the variation in the curve so too is the noise level augmented.

Figure 6.4 Zoomed in Thresholded “Intuitive” Focus Curves. These are the same focus functions that were plotted in Figure 6.3 with the vertical scale magnified to better show curve characteristics.
The next group of focus curves includes the two gradient based functions that do not require a contrast threshold (see Figure 6.5). Both the Tenenbaum gradient and the energy Laplace produce focus curves that resemble the three thresholded gradient focus curves of Figure 6.1. However, on close comparison, it can be seen that the unfocused regions are not as close to zero, and not as flat as the thresholded gradient functions. This is significant because the functions with unfocused regions substantially at zero will have less variance from Z-stack to Z-stack as the unfocused regions of each Z-stack should also show no measure of focus. Whereas functions that are not zero in unfocused regions are detecting some measure, perhaps noise, and identify it as focused content. From Z-stack to Z-stack this measure will vary making placing a threshold that works

![Graphs showing Tenenbaum Gradient and Energy Laplace focus curves.]

Figure 6.5 Non-Thresholded Gradient Focus Curves. As in Figure 6.1, the green region shown in the energy Laplace function depicts the range of threshold values that correctly identify the surface depth by partitioning the Z stack into a focused region bounded by an unfocused region on either side. There is no threshold that can correctly identify both surface depths for the Tenenbaum gradient function.
for all Z-stacks more difficult. The “flatness” of the unfocused regions should be a minimal as well. This makes finding the “knee” bend easier as it is an easily detected change. For functions with a gradually increasing slope, there is no clearly definitive point that determines a boundary, but more of an arbitrary limit that is reached. The statistical functions too share these same problems (Figure 6.6).

Figure 6.6 Non-Thresholded Statistical Focus Curves
The range and entropy histogram based curves (Figure 6.7), have the same problem as the intuitive functions. As we zoom in on the scale, we increase the effect of the noise. Even when these focus curves are optimized using the

![Figure 6.7 Thresholded Histogram Based Focus Curves. a) and b) are at full scale while c) and d) show the same data but the vertical scale has been magnified to better show curve characteristics.](image)
training set, the average boundary detection error rate (deviation from ground truth) is 2.6 µm for thresholded pixel count, 4.2 µm for thresholded content, and 4.3 µm for image power (see Table 5.2). These error rates are unacceptable; therefore no further testing of these functions was necessary.

Figure 6.8 Modified Absolute Gradient Focus Curves. These functions were modified in an attempt to better locate the surface depth. As in Figure 6.1, the green region depicts the range of threshold values that correctly identify the surface depth by partitioning the Z-stack into a focused region bounded by an unfocused region on either side.
The Modified focus functions of Figure 6.8 look nearly ideal and quite similar to the thresholded gradient functions. The modified absolute gradient count (MAGC) was the highest ranked function and in Figure 6.2 shows its slight edge over the absolute gradient.

Of the eighteen focus functions analyzed, the thresholded gradient based functions (Figure 6.1 and Figure 6.8) were the only functions to achieve the performance requirements for this thesis. These seven focus functions were each independently incorporated into the automated surface location algorithm. Each was trained to find the top and bottom tissue surfaces within an average tolerance of 1.0 µm using the arbitrarily selected training set (Table 6.1). With trained threshold parameters, all seven functions identified the top and bottom tissue surfaces within 1.0 µm on a test set of similar Z-stacks, test set #1 (Table 6.2), and on a test set #2, a set of Z-stacks from a different rat brain case study (Table 6.3).
### Automated Focus Determination Training Set Optimization

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>Contrast Threshold</th>
<th>Focus Threshold</th>
<th>Average Error from G.T. (µm)</th>
<th>Rank</th>
<th>Standard Deviation from G.T. (µm)</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>16</td>
<td>700</td>
<td>0.72</td>
<td>1</td>
<td>0.80</td>
<td>6</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>119</td>
<td>197,878</td>
<td>0.89</td>
<td>7</td>
<td>1.56</td>
<td>7</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>486</td>
<td>130,444</td>
<td>0.75</td>
<td>4</td>
<td>0.79</td>
<td>5</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>18</td>
<td>239</td>
<td>0.72</td>
<td>1</td>
<td>0.69</td>
<td>3</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>18</td>
<td>18</td>
<td>0.75</td>
<td>4</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>17</td>
<td>118</td>
<td>0.72</td>
<td>1</td>
<td>0.77</td>
<td>4</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>16</td>
<td>10</td>
<td>0.75</td>
<td>4</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>(9)</td>
<td>Content</td>
<td>96</td>
<td>159,139</td>
<td>4.22</td>
<td>16</td>
<td>2.76</td>
<td>17</td>
</tr>
<tr>
<td>(10)</td>
<td>Pixel Count</td>
<td>42</td>
<td>13,429</td>
<td>2.56</td>
<td>11</td>
<td>2.65</td>
<td>14</td>
</tr>
<tr>
<td>(11)</td>
<td>Image Power</td>
<td>80</td>
<td>695,994,640</td>
<td>4.33</td>
<td>17</td>
<td>2.69</td>
<td>15</td>
</tr>
<tr>
<td>(12)</td>
<td>Tenenbaum Gradient</td>
<td>N/A</td>
<td>517,555,008</td>
<td>1.50</td>
<td>10</td>
<td>2.13</td>
<td>8</td>
</tr>
<tr>
<td>(13)</td>
<td>Energy Laplace</td>
<td>N/A</td>
<td>393,364,455</td>
<td>1.17</td>
<td>8</td>
<td>2.23</td>
<td>10</td>
</tr>
<tr>
<td>(14)</td>
<td>Variance</td>
<td>N/A</td>
<td>948</td>
<td>2.86</td>
<td>12</td>
<td>2.63</td>
<td>13</td>
</tr>
<tr>
<td>(15)</td>
<td>Normal Variance</td>
<td>N/A</td>
<td>89,505</td>
<td>2.97</td>
<td>13</td>
<td>2.51</td>
<td>12</td>
</tr>
<tr>
<td>(16)</td>
<td>Autocorrelation</td>
<td>N/A</td>
<td>6,027,949</td>
<td>1.44</td>
<td>9</td>
<td>2.15</td>
<td>9</td>
</tr>
<tr>
<td>(17)</td>
<td>Standard Deviation Autocorrelation</td>
<td>N/A</td>
<td>758,447,304</td>
<td>4.50</td>
<td>18</td>
<td>2.69</td>
<td>16</td>
</tr>
<tr>
<td>(14)</td>
<td>Range</td>
<td>N/A</td>
<td>180</td>
<td>3.31</td>
<td>15</td>
<td>2.95</td>
<td>18</td>
</tr>
<tr>
<td>(19)</td>
<td>Entropy</td>
<td>N/A</td>
<td>6.8484</td>
<td>3.00</td>
<td>14</td>
<td>2.44</td>
<td>11</td>
</tr>
</tbody>
</table>

Table 6.1 Automated Focus Determination Training Set Optimization
### Table 6.2 Automated Focus Determination, Test Set #1

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>Contrast Threshold</th>
<th>Focus Threshold</th>
<th>Average Error from G.T. (µm)</th>
<th>Standard Deviation from G.T. (µm)</th>
<th>Rank</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>16</td>
<td>700</td>
<td>0.63</td>
<td>0.42</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>119</td>
<td>197,239</td>
<td>0.69</td>
<td>0.53</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>486</td>
<td>130,341</td>
<td>0.78</td>
<td>0.48</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>18</td>
<td>230</td>
<td>0.66</td>
<td>0.41</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>18</td>
<td>18</td>
<td>0.56</td>
<td>0.43</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>17</td>
<td>118</td>
<td>0.59</td>
<td>0.43</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>16</td>
<td>20</td>
<td>0.56</td>
<td>0.43</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

### Table 6.3 Automated Focus Determination Test Set #2

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>Contrast Threshold</th>
<th>Focus Threshold</th>
<th>Average Error from G.T. (µm)</th>
<th>Standard Deviation from G.T. (µm)</th>
<th>Rank</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>16</td>
<td>700</td>
<td>0.42</td>
<td>0.71</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>119</td>
<td>197,239</td>
<td>1.00</td>
<td>0.71</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>486</td>
<td>130,341</td>
<td>0.53</td>
<td>0.68</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>18</td>
<td>230</td>
<td>0.47</td>
<td>0.65</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>18</td>
<td>18</td>
<td>0.39</td>
<td>0.65</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>17</td>
<td>118</td>
<td>0.47</td>
<td>0.65</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>16</td>
<td>20</td>
<td>0.44</td>
<td>0.67</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Ranking the algorithms by average error rate, as well as standard deviation from ground truth for each test sets, showed that the modified absolute gradient count (MAGC) outperformed the others by finding tissue section surfaces within 0.56 µm on average for test set #1 and within 0.39 µm on average for test set #2. Surpassing the tolerance goal of ±1.0 µm for each surface, tissue thickness was also determined within ±1.0 µm on average. Because two surface locations for each Z-stack are required to determine one thickness measure, the two test sets were combined to determine tissue thickness. On the resulting set of 34 tissues samples, six of the seven thresholded gradient functions yielded an error rate of less than 1.0 µm (Table 6.4). Once again, MAGC had the lowest error rate (0.76 µm).

Table 6.4 Automated Tissue Thickness Determination

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>Contrast Threshold</th>
<th>Focus Threshold</th>
<th>Average Error from G.T.</th>
<th>Standard Deviation from G.T.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>16</td>
<td>700</td>
<td>0.85</td>
<td>1.13</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>119</td>
<td>197,239</td>
<td>1.41</td>
<td>1.13</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>486</td>
<td>130,341</td>
<td>0.94</td>
<td>1.04</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>18</td>
<td>230</td>
<td>0.94</td>
<td>0.98</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>18</td>
<td>18</td>
<td>0.76</td>
<td>1.00</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>17</td>
<td>118</td>
<td>0.88</td>
<td>0.96</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>16</td>
<td>20</td>
<td>0.82</td>
<td>0.99</td>
</tr>
</tbody>
</table>
Supervised sets of Z-stack images stored as portable gray scale maps (pgm files) on hard disk were used to develop a training method to optimize the automated surface location algorithm. The automation of the thresholded gradient functions requires selection of two thresholds: The pixel-to-pixel contrast threshold to decide whether to include the contrast between two pixels in the function summation; and, the total image focus threshold to decide whether an image is in focus or not. Since the Nelder-Mead optimization outcome is dependent on the initial selection of simplex coordinates [17], the method was run three times for each focus function, with a different initialization each time. The automated surface location algorithm used each of the hundreds of threshold pairs selected by Nelder-Mead with each focus function to locate 36 tissue surfaces within 18 Z-stacks consisting of a total of 480 images. This analysis led to seven focus functions that located correct tissue surfaces within the desired error rate tolerance after application of the optimized thresholds.

Test set #2, a set of rat brain coronal sections taken from a different study than the rat brain coronal sections from the training set, was used to analyze the robustness of the optimized thresholds. Importantly, the histochemical stain (cresyl violet) and region of interest (hippocampus) for sections in test set #2 were different from the tyrosine hydroxylase immunostain of substantia nigra in the training set. Nonetheless, with threshold parameters optimized on the training set, the same seven focus functions that performed well on test set #1, performed equally well on test set #2 (see Table 6.3), with MAGC outperforming others with an average error rate of 0.76 µm.
The distribution of the deviation of the automated location of tissue surfaces from the manually determined tissue surfaces is shown in Table 6.5 through Table 6.8. The deviation is the positive difference between the automated surface location algorithms determination of surface location and the manually determined value. This difference is measured in whole micrometers and as the step size through the Z-stacks was 1 µm, this difference is essentially the number of images away from the manually determined “just out of focus” image that the algorithm returned. On both test set #1 and test set #2 MAGC was never more than a single image away from the manually determined focus. For thickness determination this was not the case, as measuring thickness requires measuring both the upper and lower surface it is more difficult to maintain the tolerance level. The two Z-stacks that deviated more than 2 µm were both off by 3 µm.

### Table 6.5 Distribution of Deviation on Training Set

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>0 µm Count</th>
<th>0 µm %</th>
<th>1 µm Count</th>
<th>1 µm %</th>
<th>2 µm Count</th>
<th>2 µm %</th>
<th>&gt;2 µm Count</th>
<th>&gt;2 µm %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>17</td>
<td>47</td>
<td>13</td>
<td>36</td>
<td>5</td>
<td>14</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>17</td>
<td>47</td>
<td>14</td>
<td>39</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>16</td>
<td>44</td>
<td>14</td>
<td>39</td>
<td>5</td>
<td>14</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>15</td>
<td>42</td>
<td>16</td>
<td>44</td>
<td>5</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>14</td>
<td>39</td>
<td>17</td>
<td>47</td>
<td>5</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>16</td>
<td>44</td>
<td>15</td>
<td>42</td>
<td>4</td>
<td>11</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>14</td>
<td>39</td>
<td>17</td>
<td>47</td>
<td>5</td>
<td>14</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### Table 6.6 Distribution of Deviation on Test Set #1

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>0 µm Count</th>
<th>0 µm %</th>
<th>1 µm Count</th>
<th>1 µm %</th>
<th>2 µm Count</th>
<th>2 µm %</th>
<th>&gt;2 µm Count</th>
<th>&gt;2 µm %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>15</td>
<td>47</td>
<td>14</td>
<td>44</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>15</td>
<td>47</td>
<td>12</td>
<td>38</td>
<td>5</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>12</td>
<td>38</td>
<td>15</td>
<td>47</td>
<td>5</td>
<td>16</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>14</td>
<td>44</td>
<td>15</td>
<td>47</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>17</td>
<td>53</td>
<td>12</td>
<td>38</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>16</td>
<td>50</td>
<td>13</td>
<td>41</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>17</td>
<td>53</td>
<td>12</td>
<td>38</td>
<td>3</td>
<td>9</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Table 6.7 Distribution of Deviation on Test Set #2

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>0 µm Count</th>
<th>0 µm %</th>
<th>1 µm Count</th>
<th>1 µm %</th>
<th>2 µm Count</th>
<th>2 µm %</th>
<th>&gt;2 µm Count</th>
<th>&gt;2 µm %</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>23</td>
<td>64</td>
<td>12</td>
<td>33</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>9</td>
<td>25</td>
<td>19</td>
<td>53</td>
<td>7</td>
<td>19</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>19</td>
<td>53</td>
<td>16</td>
<td>44</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>21</td>
<td>58</td>
<td>13</td>
<td>36</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(26)</td>
<td>Modified Absolute Gradient Count</td>
<td>23</td>
<td>64</td>
<td>12</td>
<td>33</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>21</td>
<td>58</td>
<td>13</td>
<td>36</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>22</td>
<td>61</td>
<td>12</td>
<td>33</td>
<td>2</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
6.2 Statistical Significance

Although the modified absolute gradient count function (MAGC) appeared to perform best over the test sets, it was not clear if this better performance was statistically significant. To check statistical significance the success rate of all seventeen other focus functions were compared to the modified absolute gradient function. Success was defined as location of a tissue surface within 1 µm of the manually determined surface location. The seven thresholded gradient functions were successful at attaining this goal, while the remaining functions were not. It was ascertained using the McNemar chi-squared test [32] that there was no statistical difference between MAGC and the other six thresholded gradient functions in finding tissue surfaces within 1 µm of the correct location (see Table 6.9). Further, using the paired t-Test (Tamhame,
2000), we hypothesize that MAGC is statistically the same as each of the other thresholded gradient functions at determining tissue surface location as well as tissue thickness. The results of the paired t-Test show MAGC is indeed statistically different (better) than the other functions with a confidence interval of better than 99% (see Table 6.10). Note though that the confidence intervals given in both Table 6.9 and Table 6.10 are for the null hypothesis which is that MAGC is statistically the same as each of the other functions. So a confidence level of 1% means we are 99% confident the functions are not the same.
Table 6.9 McNemar Chi Squared Test Measuring Statistical Significance

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Function</th>
<th>Training</th>
<th>Test Set #1</th>
<th>Test Set #2</th>
<th>Thickness Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\chi^2$</td>
<td>Confidence</td>
<td>$\chi^2$</td>
<td>Confidence</td>
</tr>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>0.25</td>
<td>38% NSD*</td>
<td>0% NSD*</td>
<td>0% NSD*</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>0.13</td>
<td>28%</td>
<td>0.56</td>
<td>55%</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>0.25</td>
<td>38%</td>
<td>0.56</td>
<td>55%</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>NSD*</td>
<td>0%</td>
<td>0.28</td>
<td>40%</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>NSD*</td>
<td>0%</td>
<td>0.28</td>
<td>40%</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>NSD*</td>
<td>0%</td>
<td>NSD*</td>
<td>0%</td>
</tr>
<tr>
<td>(9)</td>
<td>Content</td>
<td>22.23</td>
<td>100%</td>
<td>24.01</td>
<td>100%</td>
</tr>
<tr>
<td>(10)</td>
<td>Pixel Count</td>
<td>25.01</td>
<td>100%</td>
<td>19.01</td>
<td>100%</td>
</tr>
<tr>
<td>(11)</td>
<td>Image Power</td>
<td>25.01</td>
<td>100%</td>
<td>24.01</td>
<td>100%</td>
</tr>
<tr>
<td>(12)</td>
<td>Tenenbaum Gradient</td>
<td>1.23</td>
<td>73%</td>
<td>8.03</td>
<td>100%</td>
</tr>
<tr>
<td>(13)</td>
<td>Energy Laplace</td>
<td>0.56</td>
<td>55%</td>
<td>1.13</td>
<td>71%</td>
</tr>
<tr>
<td>(14)</td>
<td>Variance</td>
<td>13.35</td>
<td>100%</td>
<td>22.01</td>
<td>100%</td>
</tr>
<tr>
<td>(15)</td>
<td>Normal Variance</td>
<td>17.28</td>
<td>100%</td>
<td>22.01</td>
<td>100%</td>
</tr>
<tr>
<td>(16)</td>
<td>Autocorrelation</td>
<td>0.28</td>
<td>40%</td>
<td>4.69</td>
<td>97%</td>
</tr>
<tr>
<td>(17)</td>
<td>St Dev Based Autocorrelation</td>
<td>23.22</td>
<td>100%</td>
<td>24.01</td>
<td>100%</td>
</tr>
<tr>
<td>(18)</td>
<td>Range</td>
<td>14.33</td>
<td>100%</td>
<td>22.01</td>
<td>100%</td>
</tr>
<tr>
<td>(19)</td>
<td>Entropy</td>
<td>15.31</td>
<td>100%</td>
<td>20.01</td>
<td>100%</td>
</tr>
</tbody>
</table>

*NSD = No Statistical Difference
Table 6.10 Paired t-Test Measuring Statistical Significance

<table>
<thead>
<tr>
<th>Eq.</th>
<th>Functions</th>
<th>Training</th>
<th>Test Set #1</th>
<th>Test Set #2</th>
<th>Thickness Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>35 df</td>
<td>31 df</td>
<td>35 df</td>
<td>33 df</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Confidence</td>
<td>Confidence</td>
<td>Confidence</td>
<td>Confidence</td>
</tr>
<tr>
<td>(6)</td>
<td>Absolute Gradient</td>
<td>2.02</td>
<td>5%</td>
<td>2.10</td>
<td>4%</td>
</tr>
<tr>
<td>(7)</td>
<td>Squared Gradient</td>
<td>3.02</td>
<td>0%</td>
<td>4.00</td>
<td>0%</td>
</tr>
<tr>
<td>(8)</td>
<td>Brenner Gradient</td>
<td>2.71</td>
<td>1%</td>
<td>4.61</td>
<td>0%</td>
</tr>
<tr>
<td>(25)</td>
<td>Modified Absolute Gradient</td>
<td>1.43</td>
<td>16%</td>
<td>1.44</td>
<td>16%</td>
</tr>
<tr>
<td>(27)</td>
<td>Filtered Modified Absolute Gradient</td>
<td>1.00</td>
<td>32%</td>
<td>1.79</td>
<td>8%</td>
</tr>
<tr>
<td>(28)</td>
<td>Filtered Modified Absolute Gradient Count</td>
<td>1.78</td>
<td>8%</td>
<td>1.79</td>
<td>8%</td>
</tr>
</tbody>
</table>

df = degrees of freedom
Of the fourteen common functions initially tested, the three thresholded gradient functions – absolute gradient, squared gradient, and Brenner gradient – successfully achieved the goal of locating tissue surfaces with an average tolerance of 1.0 µm. Furthermore, absolute gradient and Brenner gradient were the only two conventional functions to determine tissue thickness with an average tolerance of 1.0 µm. Out of this set of commonly used functions, absolute gradient achieved the highest level of performance.

Three improvements made to absolute gradient, including rotational invariance, improved weighting, and filtering, led to further improvement in performance. Rotational invariance was accomplished by adding a pixel-to-pixel comparison in the vertical and horizontal directions. For weighting improvement, the squared gradient gave higher weight to relatively high contrast pixels than absolute gradient; nevertheless, absolute gradient performed better than squared gradient. We tested whether the performance could be further improved by eliminated weighting of higher contrast pixels. A simple count function was introduced to count pixels over the absolute contrast threshold. Moreover, under the reasoning that isolated high contrast pixels were more likely salt and pepper noise rather than points of focus, a median filter was applied to the intermediary binarized image of the modified absolute gradient.
These modifications to the absolute gradient function generated four new functions for this thesis (see Figure 6.8). All four modified functions achieved the average tolerance goal of ±1.0 µm for surface location with similar test and training sets, and all four met the robustness goal by achieving the same tolerance goal for a dissimilar test set. Finally, all four modified functions determined tissue thickness within ±1.0 µm average. Thus, the modified absolute gradient count, developed here, achieved the best performance as assessed by lowest error rate and highest rank for robustness across different training sets.

The computational complexity for analysis for determining focus of an image using the modified functions is $O(n^2)$ where $n$ is the number of rows or columns of the image. For rotational invariance, the vertical comparison is done in parallel with the horizontal comparison with no increase in complexity. The median filter, however, requires a second pass through the image raster. During the first pass the high contrast pixels are identified, and in the second pass they are filtered. However, the second pass can be pipelined one row and column behind the first pass, with only a slight constant time increase. With current processing speeds of standard computers, a Z-stack is captured and analyzed in real time, with the step motor speed of the automated stage as the time limiting factor.

As discussed in Chapter 5, a few Z-stacks were rejected from the training and test sets. Because they were rejected during ground truthing, they should have no affect on the validity of the thesis results. Nonetheless with experience and careful attention, few Z-stacks should ever need to be rejected from a study.
Care must be taken in the collection of Z-stacks. To assure that only complete Z-stacks are collected with unfocused images on both the top and bottom of the stack. Even the highest quality prepared slides can have an occasional problem as we had with a portion of one section folded over on itself. A possible solution to this problem is to set a threshold of surface variance so that when the change in surface location is over this threshold, the Z-stack is flagged for manual inspection. The image brightness level should be such that the range on intensity values in each image is large. To this end, a live dynamic range display could be added to the Stereologer to aid in manually setting the manual lamp intensity to an optimal value. Dirt in the optical path can also cause problems with focusing. It is assumed that the microscopy equipment as well as the specimen slides will be in a proper state of cleanliness.
8.1 Conclusions

In the large and growing field of automated stereology, consistent location of tissue surfaces is a prerequisite for automatic computerized stereoanalysis. This work shows that automation of section thickness analysis in support of this automation is possible using the modified absolute gradient count function. The new MAGC can locate tissues surfaces within 1.0 µm of ground truth on average. The MAGC can also determine tissue thickness within 1 µm of ground truth on average.

As part of the automation process a training algorithm was developed to optimize the function parameters. This optimization requires a sufficient set of manually located surfaces to train. With optimized parameters, the automated surface location algorithm using the modified absolute gradient count should be sufficiently robust for applications across a range of different studies with different tissues stains using a variety of techniques. However, the training algorithm can optimize the function parameters again when needed.

8.2 Future Work

A second, finer, step size parameter would refine the determination of the specimen top and bottom. By using two step sizes, the first, larger step would
allow a quick search over a large volume for an approximate specimen top and bottom. Followed by the smaller second step parameter to hone in on and acquire a Z-stack around the volume surrounding the approximate top and another Z-stack around the approximate bottom.

In live mode, the ability to adjust the light level to a consistent exposure is contemplated for future work. The light level is not controlled by the Stereologer and at present is left to the subjective judgment of the operator. Nonetheless, being able to adjust the light to assure intensities with the widest possible range within an image would also help to assure that the program returns accurate results. This would have to be a manual setting, but it could be done in real time in conjunction with the calculation of the intensity histogram algorithm.
REFERENCES


[31] F. Hanson, "MCS 471 Numerical Analysis --- Special Class Notes: Notes from Class on Special Topics," Fall 2005 ed. Chicago, 2005.