Automatic Detection and Characterization of Parasite Eggs by Image Processing

Lindsey Eubank Ostergaard

Thesis submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science

in

Mechanical Engineering

Mary E. Kasarda, Chair

Shaadi F. Elswaifi

James R. Palmieri

Bahareh Behkam

August 2, 2013

Blacksburg, VA

Keywords: Parasite, Parasite Eggs, Image Processing, Parasite egg recognition, Characterization, Microscopic Image

Automatic Detection and Characterization of Parasite Eggs by Image Processing

Lindsey Eubank Ostergaard

Committee Chair: Dr. Mary Kasarda

ABSTRACT

The accurate identification of parasites allows for the quick diagnosis and treatment of infections. Current state-of-the-art identification techniques require a trained technician to examine prepared specimens by microscope or other molecular methods. In an effort to automate the process and better facilitate the field identification of parasites, approaches are developed to utilize LabVIEW™ and MATLAB™, which are commercially available image processing software packages, for parasite egg identification. The goal of this project is to investigate different image processing techniques and descriptors for the detection and characterization of the following parasite eggs: *Ascaris lumbricoides*, *Taenia* sp., and *Paragonimus westermani*. One manual approach and four automated approaches are used to locate the parasite eggs and gather parasite characterization data. The manual approach uses manual measurements of the parasite eggs within the digital images. The four automated approaches are LabVIEW™ Vision Assistant™ scripts, MATLAB™ separation code, MATLAB™ cross-section grayscale analysis, and MATLAB™ edge signature analysis. Forty-four separate measurements were analyzed through the four different approaches. Two types of statistical tests, single factor global Analysis of Variance (ANOVA) test and Multiple Comparison tests, are used to demonstrate that parasite eggs can be differentiated. Thirty-six of the measurements proved to be statistically significant in the differentiation of at least two of the parasite egg types. Of the thirty-six measurements, seven proved to be statistically significant in the differentiation of all three parasite egg types. These results have shown that it is feasible to develop an automated parasite egg detection and identification algorithm through image processing. The automated image processing techniques have proven successful at differentiating parasite eggs from background material. This initial research will be the foundation for future software structure, image processing techniques, and measurements that should be used for automated parasite egg detection.

Acknowledgements

First, I would like to thank my research advisor, Dr. Mary Kasarda. She has consistently provided motivation and direction to me throughout the course of my graduate studies. Even beyond research, she has given me life skills that I have applied to my current role in industry. I would also like to thank Dr. James Palmieri and Dr. Shaadi Elswaifi for providing the idea that my research was based on. They opened my eyes and were the experts to the world of microbiology and parasites. Thanks to Dr. Bahareh Behkam for answering any questions. During my first year of graduate studies, the Powell Fellowship provided such a blessing through their financial support. Thanks to the Department of Mechanical Engineering for providing the opportunity to be a Graduate Teaching Assistant and for their financial support. My parents, Ronnie and Teresa Eubank, were there along the way to offer their love and support. The biggest "thanks" goes to my husband, Erik. He was able to be my encouragement and support, even in the midst of completing his own graduate studies. Finally, to my Lord and Savior Jesus Christ, be all glory, honor, and praise forever and ever.

Table of Contents

List of Figures

List of Tables

Table **Description** **Page**

Abbreviations

Symbol	Description
μ_1	Mean of measurement for Parasite 1
μ_2	Mean of measurement for Parasite 2
μ_3	Mean of measurement for Parasite 3
α	Significance level
σ	Standard deviation
σ^2	Variance
$2D$	Two Dimensional
$3D$	Three Dimensional
A	Binary Image
a	Individual pixel
ANOVA	Analysis of Variance
A_{CH}	Area of a Convex Hull
A_T	Area of Particle and Holes
A	Area of the Particle
B	Structuring Element
df	Degree of freedom
df_b	Degree of freedom between groups
df_b	Degree of freedom between samples
df_w	Degree of freedom within groups
df_w	Degree of freedom within samples
df	Degrees of freedom
D	Distance between pixels
dx	Distance between x-coordinate pixels
dy	Distance between y-coordinate pixels
E	Elongation Factor
F_d	Max Feret Diameter
F	Test Statistic
$F\ crit$	Critical Test Statistic
H_a	Alternative Hypothesis

H_o	Null Hypothesis
k	Number of groups
MSB	Mean square between groups
MSB	Mean square between samples
MSW	Mean square within groups
MSW	Mean square within samples
n_T	Total number of samples
n	Number of samples
n_i	Number of samples from Parasite "i"
n_1	Number of samples from Parasite 1
n_2	Number of samples from Parasite 2
n_3	Number of samples from Parasite 3
P	Number of Peaks
$P(5)$	Number of Peaks (5)
P	Perimeter
p-value	Probability
Q_{crit}	Q-critical value
Q	Q-test statistic
Q	Q-test statistic
RF_b	Equivalent Rectangular Short Side
${s_W}^2$	Mean square variability within groups
${s_W}^2$	Mean square within
SSB	Sum of squares between-group variability
SSW	Sum of squares within-group variability
T	Number of groups
TSS	Total sum of squares
x	Pixel location of the x-coordinate
$\bar{y}$	Mean
$\bar{y}_{..}$	Overall mean
y	Pixel location of the y-coordinate
y_i	Value of the ith sample

Chapter 1 Project Introduction

The World Health Organization [1] states that the parasite *Ascaris lumbricoides* alone is the suspected cause of 60,000 deaths worldwide per year. The presence of parasites such as *Ascaris lumbricoides* is usually detected by trained medical personnel who examine fecal samples for any indication of parasite eggs. Specialized training and experience are required to identify the egg and diagnose a parasitic infection. This research investigates an alternative method for automatic detection of parasite eggs in fecal matter through the use of image processing techniques. The alternative method will involve image processing of digital microscope slide pictures in order to locate parasite eggs. The motivation for the project and background information is discussed. A literature review that provides insight to this project is presented. Previous works are explained with a short overview and divided into three main categories: Hardware & Software, Feature Detection, and Classification. With these previous works in mind, the goals and scope of the project are explained. A brief introduction to parasitology is presented with a focus on the following parasites of interest used in this study: *Paragonimus westermani, Taenia* spp., and *Ascaris lumbricoides*. It is important to understand fundamentals of the parasitology related to this project so that the research can better be understood. The parasite automatic detection method utilizing image processing techniques is outlined at the end of the chapter. Slides are prepared, and digital images are taken by a microscope-mounted camera. These images of the parasite eggs will go through calibration, pre-processing steps, and an application of morphological operators. All of the image processing steps take place in either LabVIEW™ or MATLAB™. Some manual, non-computer, measurements will also be recorded to gather data on the size of the parasite eggs.

1.1 Introduction and Motivation

With helminth parasite infections, symptoms may be present depending on the severity of the infection. The World Health Organization [1] states that *Ascaris lumbricoides* alone is the suspected cause of 60,000 deaths worldwide per year. These deaths are highest among children. The best way to prevent parasite infections is to implement proper hygiene and avoid consuming undercooked food. Patient education, rapid and accurate diagnosis, and proper medication are the best weapons against the spread of parasitic diseases. Areas of the world that are not educated in proper hygiene or food preparation will continue to have higher rates of parasite infections.

For the laboratory diagnosis of intestinal parasitic infections, there are three main steps. These steps are the following: gathering fecal samples, preparing a sample for slide analysis, and analyzing the

sample. The first two steps can be completed if sufficient instructions are provided. The last step, analyzing the sample, must be performed by a trained specialist.

In most cases it is not hard to find willing individuals who would support the cause of helping people with parasitic infections in areas where resources to diagnose and treat are not available. The difficulty is finding trained individuals who can effectively analyze the samples, identify the parasites or their eggs, and determine a diagnosis. If the untrained individual could be equipped to detect and distinguish parasite eggs, our ability to identify intestinal parasitic infections and the health of more individuals would be enhanced. The goal of this research is to forge a path that can eventually lead to an automatic detection method for the identification of intestinal parasite eggs. Automatic detection would enable non-specialized individuals to perform parasitic egg identification and assist in the diagnosis of parasitic diseases. Automatic detection would also enable the trained specialist to be quicker and more efficient during sample analysis. Bovik [2] states, "Computer-assisted microscopy provides the ability to enhance the speed of microscope data acquisition and data analysis, thus relieving humans of tedious tasks.

To some groups, such as children, intestinal parasitic infections may be considered life-threatening or lead to life-long complications, especially if left untreated. Parasites have a repeatable life-cycle in humans. It is important to discover the helminth parasite and treat the host organism with anthelminthic medications. Motivation for this work came from Dr. Shaadi Elswaifi and Dr. James R. Palmieri of the Virginia College of Osteopathic Medicine (VCOM). They have experienced first-hand the overwhelming sense of need in countries with high rates of human parasite infections and small numbers of trained health workers.

1.2 Parasitology Introduction

The general definition of a parasite is an organism that consumes nutrients from, lives on or lives in a host and also causes harm to that host. A specific group of parasites, which can be found in humans, are helminth parasites. There are three classifications within the helminth parasites category. They are the following: Trematodes, Cestodes, and Nematodes. From each classification, one type of parasite that causes intestinal infections was chosen for analysis. From the Trematodes group, *Paragonimus westermani* was chosen. *Taenia* spp. represents the Cestodes group. From the Nematodes, *Ascaris lumbricoides* was selected for analysis. Several concerns arise when a parasite infects a human. The main concern is for the health of the individual. Identifying the parasite, making a diagnosis, and administering medication are necessary for treating the host. Several processing methods exist to locate

parasite eggs. The processing methods typically involve creating a solution with a fecal sample. The process continues by making slides of the solution for microscopic analysis. Concerns of parasite infections as well as detection methods will be discussed. Each individual parasite type and their lifecycle will be explained.

1.2.1 Detection Methods

There is one main detection method for determining if a parasite infection is present. The detection can be completed by several other molecular methods, however. The main method is that a fecal sample must be taken and analyzed, but the fecal sample may be prepared for analysis through several different mediums. Fecal samples are always diluted to make a suspension. The fecal sample may be mixed with mediums such as saline, sugar solution, or a fixative like Formalin. The saline and Formalin suspensions allow the fecal matter to settle in order to get the best sample. In the sugar solution method, sugar is added to the suspension. The parasite eggs will then float and accumulate at the top quicker than the rest of the suspension material. A sample of the fecal matter suspension is then retrieved from the top layer. Once the suspension is created, a microscopic slide is prepared. The slide must then be analyzed. Due to the complexity of preparing a fecal sample suspension and the vast differences between parasite egg types, historically, detection has been most effective if the slide is analyzed by a trained individual examining the microscope slides of the fecal sample.

1.2.2 Paragonimus westermani

Paragonimus westermani is classified as Trematodes, which are also known as flukes or flatworms. They are common in Asia, South America, and India. *Paragonimus westermani* eggs are 80-120 micrometers in length and 45-70 micrometers in width [3]. It is an elliptical shape with a moderately thick shell. The parasite worms are carried in the intestines of carnivorous animals. *Paragonimus westermani* eggs are released into the environment through fecal excrement and make their way to water sources. Snails, which are intermediate hosts, eat the eggs and become infected. The cercarial stage of the parasite is released from the snail. A released cercarial parasite penetrates the the crab or crustacean, which is the second intermediate host. The now-infected crab or crustacean develops cysts; and if eaten raw or undercooked by humans, they will effectively transfer the parasite to the human host. In the host, the parasite egg progresses a worm. The worm travels from the intestines to the lungs. Parasite eggs, which are laid in the lungs, are coughed up and swallowed. This results in a parasite life-cycle within the human. Some symptoms may be the following: coughing, fever, diarrhea,

hives, abdominal pain, and an enlarged liver or spleen. Once diagnosed, treatment can involve a prescription dosage of either Praziquantel or Bithionel [3].

1.2.3 Taenia spp.

Taenia spp. are classified as Cestodes, which are also known as tapeworms. *Taenia* spp. are between 30-35 micrometers in both length and width [3]. They are circular in shape with a thick shell. If the parasite eggs with a hexacanth larva are eaten, the hexacanth larva can then form into a cystericercous larva. These cystericercous larva cannot produce parasite eggs within a human host. The parasite is also transferred as larvae cysts found in animal tissues. Humans contract the adult parasite worms when cystericercous larva is eaten from raw or undercooked meats; and these larvae will penetrate the gastroinstinal area. The larvae have the potential to develop into adult tapeworms and produce parasite eggs. Larval cycles found in beef are from *Taenia saginata.* Larval cycles found within pork are from *Taenia solium. Taenia* spp. can only be visually identified to the family (*Taeniidae*) not the species (*Taenia solium, Taenia saginata).* If the parasite egg infected fecal matter is ingested by swine or cattle, the life-cycle of the parasite will continue. The symptoms experienced vary slightly dependent on the *Taenia* spp. type of infection. A symptom of *Taenia solium* may be cysticercosis, the development of cysts in the body. A symptom of *Taenia saginata* may be mild abdominal problems. In acute instances, appendicitis or an infection of the common bile duct may occur. Once discovered, treatment can involve a prescription dosage of the drug Praziquantel [3].

1.2.4 Ascaris lumbricoides

Ascaris lumbricoides is classified as a Nematode, which is also known as roundworm. They are prevalent and found worldwide. *Ascaris lumbricoides* eggs are between 45-75 micrometers in length and 35-50 micrometers in width [4], depending on the stage in the lifecycle and whether the egg is fertilized. They are round in shape with a thick rough shell. *Ascaris lumbricoides* eggs are released into the environment through fecal matter. Only the fertilized eggs that are released are harmful. If these fertilized eggs are ingested by humans, they will hatch into larvae and dwell in the intestines. The larvae are able to transfer from the intestines to the lung. From the lungs the larvae can be coughed up and granted access into the throat, where they are swallowed. Upon reception into the intestines, the larvae grow into worms. A mature *Ascaris lumbricoides* worm can produce up to 200,000 eggs per day. This results in a vigorous parasite life-cycle within the human. Some symptoms may be abdominal pain or, in acute situations, intestinal obstruction. If the parasite reaches the lungs, the person subject to the infection may experience coughing, trouble breathing, and may produce pneumonia. Once identified,

treatment involves one of the following drugs: Albendazole (with mebendazole), Ivermectin, or Nitazoxanide [3].

1.3 Scope of Work

This research will focus on feature detection of parasite eggs within a digital image. The digital images will be taken with a camera mounted to a microscope. Pictures will be taken of the parasite eggs on prepared microscopic slides. The following three parasite eggs are chosen for analysis: *Paragonimus westermani, Taenia* spp., and *Ascaris lumbricoides*. Microscopic slides will have background matter in the parasite egg suspension. Therefore, the new processes will distinguish between parasite eggs and debris. LabVIEW™ Vision Assistant™ will perform image processing steps. The image processing will help distinguish between background matter and objects of interest based on several quantified values such as shape, size, and grayscale pixel value. Each species of parasite egg has a specific image processing flow, also called a script. A multi-species parasite egg locator script is beyond the scope of this work. MATLAB™ will also be used to gather additional data on other features of the parasite eggs. With all methods of data collection combined, there will be 44 measurement values between all three types of parasite eggs. Statistical analysis of the results proved that parasite eggs can be identified based on the measurements. Of the 44 measurement values, 36 proved to be statistically significant enough to determine parasite type. Therefore, approximately 82% of the egg measurements recorded in this research are effective at distinguishing between parasite species. These statistical results will be the extent of the parasite egg classification and are used to demonstrate the viability of the four techniques.

1.4 Literature Review

A search of previous research was conducted on automatic parasite egg detection and identification. Insight was provided into techniques that were tested and their associated results. There were three reoccurring themes discovered in the research articles. The themes are the following: Hardware and Software, Feature Detection, and Classification. Each of these themes will be discussed; and the similarities and differences, with respect to the goals of this project, will be addressed. The scope of the research in this paper is based on the detection and characterization of the parasite eggs within digital images through image processing.

1.4.1 Hardware and Software

Hardware refers to the physical elements that are utilized in the research. This could be, for example, the type of microscope, computer, or camera used. Software refers to the computer system or programs that are used to process the digital images. The first research that claims the use of MATLAB™ for automatic feature extract was by Dogantekin et. al [5]. This research did not mention any hardware used. Pre-captured images in JPEG files from Kansas State University were utilized in research done by Dogantekin et. al [5]. Feature extraction quantifies the features of the parasite eggs and provides a way to differentiate between parasite types. A second research project by Avci et. al [6] also used MATLAB™ as the software platform for their image processing and locating of parasite eggs. The same pre-captured images in JPEG format from Kansas State University were utilized in research; therefore, no hardware was discussed. In these two projects, MATLAB™ was used for feature extraction, classification, and active learning of the system for parasite egg identification. Yang, et. al. [7] conducted research prior to this that used MATLAB™ only as a manual tool to gather feature data. A microscope and color video camera captured images at 200x magnification. The images were captured at 560x420 pixels on an image processing board attached to a computer. Sommer [8] had similar goals of identifying and classifying parasite eggs but did not use the MATLAB™ platform. Instead un-disclosed algorithms were used to collect measurement data from the parasite eggs. The images used in his research were captured at 512x512 pixels by a charge-coupled device (CCD) camera, microscope, and PC. No research projects were discovered that used National Instrument LabVIEW™ Vision Assistant™ software or other commercially available coding software to process digital images. National Instrument's LabVIEW™ Vision Assistant™ was a desirable program due to its user-friendly interface and image processing abilities. The National Instrument platform is also easily translated to many hardware applications. In this paper, a combination of LABVIEW™ Vision Assistant™ and MATLAB™ are used. The combination allowed the research to use Vision Assistant™'s user friendly image processing scripts and MATLAB™'s detailed coding applications.

1.4.2 Feature Detection

All research papers which were reviewed used some form of parasite egg feature detection in order to locate the parasite egg. One of the earlier journal articles by Nugent et. al [9] used an Eigenimage approach to parasite detection. The Eigenimage approach was originally created for facial recognition. There were two main problems with this method. The first is that parasites can vary in their position on the microscope slides, whereas facial recognition is completed every time at the same angle. The second issue was that faces are not typically separated from the background. Separating the

parasites from the background material in the slide is a crucial step. A process is necessary that can separate objects of interest from background material. Orientation of the objects must be a non-issue for detection.

Two published reports [5, 6] used invariant moment measurements as the catalysts for the feature extraction. In both projects, image processing is done by median filters, contrast enhancement, segmentation, conversion to binary, and use of morphological operators. Median filtering slightly blurs the images, but the integrity of the image is maintained while background noise is blurred away. This is one step in the process of distinguishing the parasite egg from the background material. Contrast enhancement takes an image and improves the edges of the objects. Segmentation extracts features from the image, such as edges. A binary image converts the image to foreground and background material, resulting in a binary colored image. Morphological operators manipulate the shapes of the objects of interest in the binary image. This research [5, 6] is based on Seven Invariant Moments that can be found through MATLAB™ and used to describe shapes. While MATLAB™ will be used for some feature detection in this paper, the invariant moment technique will not be used. MATLAB™ and LabVIEW™ Vision Assistant™ will rely purely upon image processing steps to detect the parasite eggs and extract real-life measurements.

Most previous research used a form of morphological operators, to perform the feature detection. In a work by Yang, et. al. [7], morphological features were extracted through the use of median filtering, binary thresholding, and segmentation. Yang's feature extraction method plots angular trajectory of the parasite egg. In the research of this paper, the same technique will be implemented in MATLAB™ as the Edge Signature. The edge signature is the length from the centroid of the object to the edge for the entire perimeter of the object. Sommer's method [8] of feature extraction began by measuring the size from the parasite egg's shell outline. In later work [10], Sommer was able to detect texture features of the parasite eggs. Texture features, along with shape and size, are able to quantify the parasite eggs' measurements. This work was for quantification purposes only; no attempt was made to identify the parasite egg types based on the results.

There was also applicable feature-detection research which was found that did not include parasite eggs. Broyles [11] used a dual-camera system to automatically detect the sizes and shapes of aggregate (crushed stones). This research was done so that industry could guarantee a high quality of crushed stones to customers. Although this project has a very different application, the main goals are the same. Image processing, similar to the parasite eggs, is able to detect shapes and sizes of the stones. Elongation, surface irregularity, jaggedness, and flatness are the main measurements of Broyles' thesis. For the parasite eggs, elongation will be measured and compared. The surface irregularity will also be

measured through the method of the cross-section gray-scale analysis on the parasite eggs. Jaggedness is a form of the parasite's edge signature measurement that will be found through MATLAB™. Flatness is not a measurement of interest due to the small size of the parasite eggs and the 2D digital images from a one-camera system.

Although all of the previous works had some form of feature extraction, there were varying levels of difficulty in locating the parasite eggs. Yang, et. al. [7] discovered methods to discriminate between the parasite eggs and debris in the slide. Whereas, in Sommer's research [8, 12], there was no identification that took place between the parasite and background material. Sommer used a 400x magnification setting on these digital images even though normal analysis magnification is 100x. The parasite egg was also the only object in the image. No background material, to differentiate against, was present. In images where the parasite egg is the main focus (400x) and no background material is present, it is simple to locate, focus upon, and extract details from the parasite egg. In images of lesser magnification (100x), background material will be present. Not only must details be extracted from the parasite eggs, but the image must first be filtered of all background material. Sommer [8] refers to the idea of lowering the magnification levels so that more than one parasite egg can be inluded in the image. This would require the difficult task of sorting through the debris and fecal matter to find the parasite eggs. In this work, methods to determine the parasite egg from the fecal matter and suspension debris are done using LabVIEW™ Vision Assistant™ image processing techniques. To help with this, a method was adopted from Liu, et. al. [13] to separate overlapping or touching objects. Liu, et. al. [13] specifically used this approach to separate liver cells that were touching and overlapping. Since the parasite eggs and fecal matter in the digital images also may be touching or overlapping, a similar technique was applied through a development of MATLAB™ code. This technique works through a series of morphological operators and a scanning algorithm that locates the core of objects and will be described in Section 3.2.2.

1.4.3 Classification

Previous works [7] were not just concerned with detection of parasites egg, but also the identification of parasite eggs. The identification process took place through classification of the parasite egg features. Yang et. al [7] used the results from the feature extraction measurements to classify the egg type. Classification uses stored results of previous parasite egg measurements to identify future parasite eggs. Artificial neural networks were then used to train the computer to identify future images. Nugent et. al [9] used active learning of the feature extraction results so that future images are processed more successfully. The initial steps for creating a classification method will be

applied in this work. Feature extraction results will be captured and analyzed statistically. The statistical analysis will show that a classification method, if implemented for future parasite egg results, would be successful in the determination of parasite egg types.

1.5 Project Design

This project was designed to investigate the feasibility of an automated parasite detection and identification method based on image processing techniques. There are four main components of this research. First, the parasite eggs must be separated from the background matter. Secondly, the image processing approach to identify parasite eggs will be developed. Third, identification metrics will be measured. Finally, the experimental results and statistics of the identification metrics will evaluate the projects viability.

Based on the functional desires, several specifications for the research of this paper were determined. Three parasite eggs are used in the initial program. They are the following: *Paragonimus westermani, Taenia* spp., and *Ascaris lumbricoides.* A light-field microscope, which is a common and inexpensive instrument, with digital image capturing capabilities is used to capture digital images of the parasites. Image processing techniques are applied to these images so that the parasite eggs can be detected. Commercially available software such as MATLAB™ and National Instrument's LabVIEW™ Vision Assistant™ are used to process the digital images. Image processing steps are the basis for feature extraction of the parasite eggs within this project. Filtering, thresholding and the implementation of morphological operators take place within LabVIEW™ Vision Assistant™. MATLAB™ is used to extract the gray-scale cross-section, which is considered a form of texture detection. MATLAB™ also incorporates morphological operators in the image pre-processing.

After the detection of the parasite eggs, the data from the measurements will differentiate between the parasite types. If a measurement is proven to be statistically different between parasite egg types then that measurement metric can be used to differentiate between parasite egg types. Statistical analysis has been done on each measurement metric to determine if differentiation is possible. A flow diagram based on the desired functions and specifications of this project can be seen in Figure 1.1.

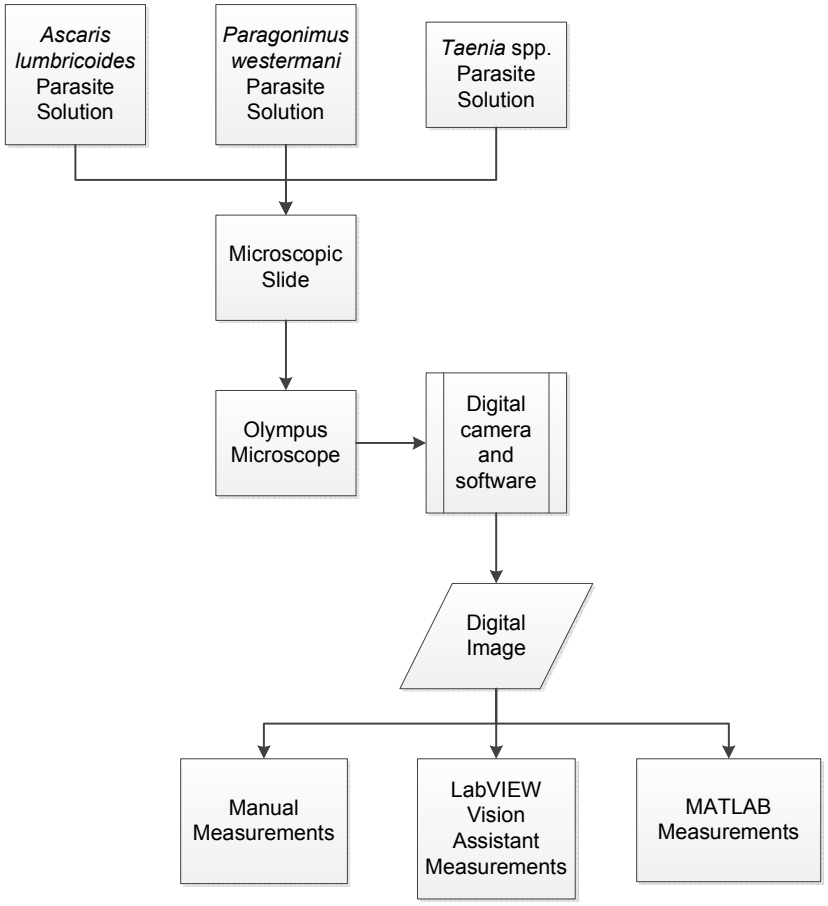

Figure 1.1 A system block diagram displays the process flow for gathering data on the parasite eggs.

1.6 Initial Results

Several approaches were used to analyze the microscopic slide images and characterize the parasite eggs. The digital images were processed four different ways. The first approach was manual. Manual measurements were recorded by two techniques, either through the microscope eyepiece (pre-image capture) or directly after it was captured (post-image capture). The manual approach took into consideration the length and width of the parasite egg. The remaining approaches to characterize the parasite eggs involve several different types of image processing techniques. The second approach was through LabVIEW™ Vision Assistant™ scripts. The scripts were used to record measurements once the parasite egg had been located within the image. LabVIEW™ has built-in measurement functions that were utilized to perform these measurements. A third approach recorded results through an analysis of the gray-scale intensity of the parasite egg cross-section that was found through a MATLAB™ code. The fourth, and final method, analyzed the parasite's edge signature which was found through another

MATLAB™ code. All four methods were applied to the three parasite egg types. Thirty images per parasite type were analyzed using each of the four methods.

Each of the approaches collected data, in the form of quantifiers, which can be utilized in locating and distinguishing between parasite egg types. Out of the 44 quantifiers in the manual and automated approaches, 36 quantifiers proved to be statistically significant in characterizing the parasite eggs. These 36 measurements, along with the summary of data collected in this research, are used as the basis for parasite egg differentiation.

Chapter 2 Experimental Setup to Obtain Microscopic Images of Parasites

In order to reach the goal of this research, microscopic digital images of the parasite eggs were obtained for the image processing applications. A basic understanding of microscopy and laboratory procedures was needed to set up the experiment. This section discusses the basics of microscopy to better understand the operation of the microscope. The microscope used for the research was an Olympus CX21 Microscope, a compound light microscope, with digital image capturing capability. The basic terminology of microscopy and the features of this specific microscope is discussed. Microscopic slides were also prepared, and the applied technique will be explained. The parasite eggs were mixed in a suspension of fecal matter in saline and formalin, which is a preservative. The microscope employed a digital color camera from Southern Microscope to capture the slide images. Slide images are taken of the three parasite eggs discussed in Section 1.2. Details of the camera and the image capturing process are explained here in Chapter 2.

2.1 Introduction to Microscopy

Although many types of microscopy methods exist, compound lens light microscopy is the type of microscopy utilized in this research. Light microscopes simply use visible light, and compound microscopes use several lenses to refract the light within the microscope. The parts of the microscope can easily be compared to the human eye. The diaphragm on the microscope acts as the iris of the eye, by controlling the amount of light that enters. A microscopic condenser concentrates the light just as the lens of the eye would focus. The human eye would take this light and focus it onto the retina. The microscope focuses this light onto the microscopic slide. The field of vision is the area visible through the microscope eyepiece.

The microscope used for this research project is an Olympus CX21 [14] light microscope. There are four separate magnification levels of the objective lenses. They are the following: 4x, 10x, 40x, and 100x. The microscope is also fitted with a 10x magnification/18 mm reticule eyepiece. Therefore, while viewing the slide through the eyepiece, the total magnification values are as follows: 40x, 100x, 400X, and 1000x. The 40x magnification is referred to as the scanning power. The 100x magnification is referred to as the low power. The 400x magnification is referred to as the high power. In this project, images were captured at the total magnification of 100x, which is the objective lens of 10x magnification.

There are several important effects that the microscope lighting and focal depth have on the view of the slide. The amount of light will affect the brightness of the slide. A longer exposure to light

creates a brighter background, which results in more differentiation of objects on the slide. It is important to optimize the brightness of the image to create the contrast desired. Optical shutters help to control the exposure time of the image.[14] Evenness of the illumination is also an effect of the lighting.[2] The more light that is let in through the diaphragm, the less the focal depth can be. Vignetting is a condition of non-uniform illumination of the slide. It results in dark corners and a bright central area. For the parasite egg images, it is important to have enough brightness so that a contrast can be created between the background matter and the foreground objects. The evenness of the illumination is also important so that image processing applications can be applied evenly across the image. For these reasons, the diaphragm on the microscope was completely open, and the condenser was in the highest position while the images were taken.

Before use, it is important to calibrate the microscope. The calibration will allow for the manual measurement of lengths and widths of the parasite eggs through the eyepiece. To calibrate the microscope eyepiece, a stage micrometer is placed in the field of vision of the microscope. The stage micrometer picture, shown in Figure 2.1, is captured by the digital camera on the microscope. The reticule eyepiece of the microscope, which contains measurement divisions, is used to measure the stage micrometer. The stage micrometer is 1mm per every 100 divisions long; therefore, every division is 10 micrometers. For the Olympus CX21 microscope at 40x magnification, one ocular division corresponds to 25 microns. At 100x magnification, there were 10 microns for every ocular division. At 400x magnification, there were 2.5 microns per ocular division. There is also error in this calibration, and it will vary based on the level of magnification. These results are summarized in Table 2.1. This calibration will be applied in Section 4.1 to calculate the manual measurements of the parasite eggs.

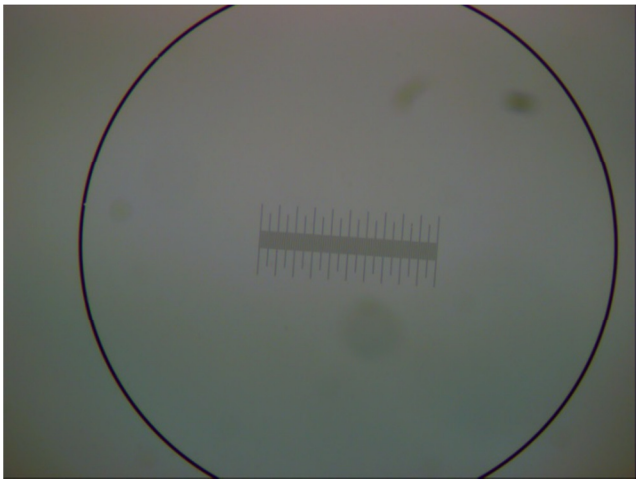

Figure 2.1 The micrometer was used to calibrate this view at 40x magnification.

Table 2.1 The micrometer calibrating results will be used in future calculations.

Magnification	Microns/ocular division	Error (+/- microns)
40x	25	12.5
100x	10	5
400x	2.5	1.25

The error included in the manual measurement can also be compared to the parasite eggs. The parasite eggs used in the experiment, *Paragonimus westermani, Taenia* spp., and *Ascaris lumbricoides,* have expected lengths and widths as specified in Table 2.2. By using the 100x magnification level, there will be less error than at 40x. The details of the parasite egg will also be seen, and the image will be able to include background material. A summary of the percent error on the manual measurement for each parasite egg can be seen in

Table **2.**.

Table 2.2 Expected lengths and widths in micrometers [3]

	Paragonimus westermani	*Taenia* spp.	*Ascaris lumbricoides*
Length	80-120	30-35	45-75
Width	45-70	30-35	35-50

Table 2.3 Expected percent error on lengths and widths from manual measurements

	Paragonimus westermani	*Taenia* spp.	*Ascaris lumbricoides*
Length Error (%)	4.2-6.25	14.3-16.7	6.7-11.1
Width Error (%)	7.1-11.1	14.3-16.7	10-14.3

2.2 Microscopic Slide Preparation

The parasite eggs, *Paragonimus westermani, Taenia* spp., and *Ascaris lumbricoides,* were purchased from Ward's Natural Scientific [15] in the preserved Platyhelminthes eggs and Nematode eggs collections. The parasite eggs are advertised as preserved and came in a solution that is special to Ward's Natural Scientific (Ward's). A diluted fecal sample taken from the Virginia College of Osteopathic Medicine microbiology laboratory was used in the experiment to provide some background material. It was diluted in a solution of formalin. The *Paragonimus westermani* suspension from Ward's

was 3.5 mL. To this, 2.5 mL of saline was added along with 1 mL of the fecal matter sample. The end product was a 7mL suspension. The *Taenia* spp. suspension from Ward's was 4.25 mL. To this, 2.75 mL of formalin solution (30% ethanol, 2% formalin, and 68% water) was added along with 1 mL fecal matter. The final Taenia spp. solution was 8 mL. The *Ascaris lumbricoides* suspension from Ward's was 4.75 mL. To this, 3 mL of formalin solution (30% ethanol, 2% formalin, and 68% water) was added along with 1.25 mL fecal matter. The final Ascaris lumbricoides suspension was 9mL. The *Paragonimus westermani* suspension was the first to be created and the only one to use the saline solution. The formalin solution was used in the other two parasite egg suspensions because it was believed that the formalin solution would provide a better preservative for the solution. Different values for the saline/formalin solution and the fecal matter were used so that the percent concentration values could be the same.

In the laboratory identification of intestinal parasite eggs, one of the fecal sample analysis methods necessitates a suspension to be made so that the feces can be placed on a slide. The suspension is created with the feces and saline to dilute the fecal matter. There are several current techniques to retrieve a good sample for slide analysis. The first is sedimentation. Simply put, the parasite eggs will eventually settle in the bottom of the solution. A sample is taken from the bottom of the saline and fecal matter solution for slide analysis by a pipette. A second technique is flotation. The suspected parasite-infested feces also has a sugar solution added with the saline. Given a period of time, the parasites will float to the top of the solution while the feces settles in the bottom. The sample is then taken from the surface of the solution with a pipette. For this research, the sedimentation technique was used.

Once the sample is retrieved, it is placed on a slide. The cover slip is placed on top. It is also good procedure, during the cover slip placement to make sure that no air bubbles are left in the sample. This ensures that the slide can be read completely and accurately. The slides must also be immediately observed. If not, the slide will dry out. Some re-wetting is possible, however this is not ideal. In the case of the parasite eggs, the dryer the slide becomes, the more likely it is that the parasite eggs will fracture. Efficient slide analysis begins in one corner of the slide. A sweep is made across the length of the slide until the opposite corner is reached. A new row is begun with very little overlap from the previous row as the user continues analysis across the full length of the slide. This row pattern sweep continues until the full height of the slide has been analyzed.

2.3 Digital Image Capture and Analysis

After the microscope slide has been analyzed and the parasite located in the field of view, the digital image is ready to capture. The digital image is necessary for the image processing. On the Olympus CX21 Microscope, there is a color digital camera by Southern Microscope, Inc. The camera is mounted by a standard C mount lens on the head of the microscope. A sensor within the digital camera is used to process and transmit the images. The digital images are captured via a computer USB port that interfaces with the microscope camera. Image capturing software that was provided by Southern Microscope, Inc. was installed into the computer. Preview images are available for viewing on the computer. This camera can also record videos, but for the application of this research only the still color image feature was utilized. The images taken are 2048 x 1536 pixel images that are saved as 9MB bitmap files on the computer. The image shown in Figure 2.2 is the result of one of the *Paragonimus westermani* egg images. Thirty images for each of the three parasite eggs species discussed in Section 1.2 were captured.

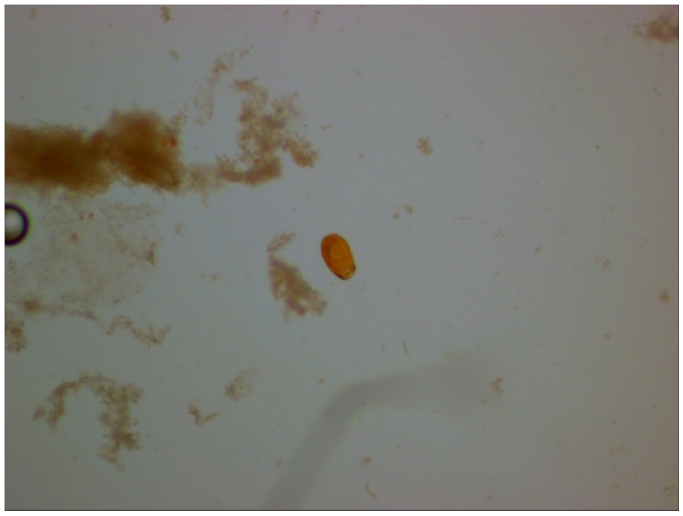

Figure 2.2 Original digital image was captured through the microscope-mounted camera

There are several features available in the image capturing software while a still image is being taken on the microscope. An autofocus feature is available on the digital camera. There are two main types of autofocus: active autofocus and passive autofocus. Active autofocus is typically a motor-driven application that emits either light or sound. The return time of the emission is used to judge the distance and adjust the focus accordingly. Passive autofocus is a computer-driven analysis. The camera computer actuates the camera lens up and down looking for the setting that creates the maximum intensity

difference. At this maximum intensity difference, the details are the clearest and the image, as viewed in the image capture software, should be in focus. The microscope camera utilizes passive autofocus through the help of a Scientific CMOS (sCMOS) sensor. Several features, such as, gain and RBG (Red, Blue, and Green Values) have the option of being adjusted manually. For all of the parasite images taken, the autofocus feature was used. No manual adjustments were made. Autofocus merely finds the image that displays the most detail. No matter what autofocus determines the setting to be, the total pixel size of the image is maintained at 2048 x 1536 pixels. Once the digital images are taken and saved, they are ready to be processed.

Chapter 3 Image Processing Approach

Image processing is a broad subject matter with a wide variety of applications. [2] For the extent of this paper, it can also be referred to specifically as computer vision. Computer vision is the application of image processing tools through a computer interface or other digital media. In this chapter, discussion will be centered on the image processing steps used to retrieve characterization information and differentiate between parasite eggs. Within image processing, there exist several methodologies to obtain information from the images, and this chapter elaborates on several that were chosen to study. The implementation of each image processing method is crucial for producing accurate data that can be used in the characterization of the parasite eggs. Common methods, such as thresholding, filtering, and the application of morphological operators, will be explained both at the general level and at a level specific to this project.

The explained image processing steps will then be applied to the digital images of parasites through LabVIEW™ [16] and MATLAB™ [17] with the main goal of being able to characterize parasite egg types. LabVIEW™ and MATLAB™ are commercially available, programming software packages. Several applications were created in the LabVIEW™ and MATLAB™ coding languages to gather characterization data on the size of the eggs, separate the eggs from background noise and other objects, collect cross-sectional gray-scale information, and identify the edge signature of the parasite eggs. These applications will be discussed in detail with image results shown of each. The following four applications are the following: LabVIEW™ scripts, MATLAB™ separation code, MATLAB™ cross-section grayscale analysis, and MATLAB™ edge signature analysis.

Three scripts were written within LabVIEW™ Vision Assistant™, one for each parasite egg type. The scripts utilize the explained image processing techniques from Section 3.1. Each parasite egg type has its unique script, which will be explained, compared, and contrasted. All three LabVIEW™ scripts follow the same process flow. There are only a few steps that had to be customized for the particular parasite that was being identified. Creating these three separate scripts is the first step in reaching the end goal of a single LabVIEW™ script that could distinguish between all three parasites with no previous knowledge of what parasite type is present in the digital image.

A MATLAB™ separation code was also utilized to separate the parasite eggs from the background matter and any touching objects. Overlapping of both parasite and background particles may be present. Minor overlapping can be reduced with separation and filtering. The success of separating overlapping images relies on the following two steps: finding the interior cores of the foreground objects and dilating the objects to find their linking points.

Through MATLAB™, the parasite cross-section grayscale values were also analyzed. The gray-scale intensities across the parasite eggs will differ between parasite egg types. One distinguishing factor is the outer edge shell thickness of the parasite egg. Since the parasite shell data will be included in the cross-section grayscale, this data will be a determining factor in distinguishing the parasite egg types. The interior of the parasite egg is also a distinguishing factor. In the interior, eggs will have embryonic cells or larvae. This will surely contribute to a difference in the cross-section grayscale intensities. MATLAB™ was utilized to collect data about the cross-section grayscale intensities of the parasite eggs. Because the interior and shell of parasite eggs can be distinctively different, this program was implemented to gain more data to better characterize the different parasite egg types.

The final MATLAB™ approach, edge signature, provides a way to capture the shape of an object quantitatively. The distance between the boundary and the centroid creates the edge signature plot. Several measurement are recorded from the edge signature plot in order to characterize the parasite eggs.

3.1 Image Processing

Within the image processing field, there are many ways to extract data from an image. The steps taken to apply image processing to our digital image will be explained in general throughout this section. First, digital images must be calibrated. This calibration allows for the proper measurement information to be collected on objects within the image. The second step is multi-staged pre-processing. Before the characterization data from the parasite eggs in the digital image can be retrieved, the image needs to be pre-processed in several steps to enhance the image. The image enhancements occur through color plane extraction, thresholding, and filtering. The preprocessing allows for enhanced data retrieval. The third step applies morphological operators to the enhanced image. The morphological operators process the image in order to complete the fourth and final step of gathering shape, size, and measurement data for parasite characterization. All of the image processing steps in this research take place in either LabVIEW™™ or MATLAB™. Each of the steps will be discussed both generally and specifically to the context of this project. The similarities of the image processing steps between applications will be addressed, as well as any differences.

3.1.1 Image Calibration for Size Measurements

Calibration of the digital images allows measurements to be calculated from pixel measurements from within a digital image. All digital images are measured in terms of pixels. Pixels are small square elements of color data that make up the picture. The images within this paper were taken

with a pixel measurement of 2048 pixels across and 1536 pixels tall, for a total pixel count of approximately 3.1 million pixels per image. In order to have any size-measured results, these pixels values must be calibrated at the beginning of the image processing steps into a unit of length. The method chosen to apply a calibration to the images was through LabVIEW™ Vision Assistant's™ calibration. LabVIEW™ has several calibration tools that are readily available. Most of the calibrations available assume a flat object is the focus of the image. The calibration tool chosen was a simple calibration. When the simple calibration is used in microscopic inspections, the results are referred to as coarse measurements. Simple calibrations measure the pixel distance between two points of a known length, also called the point coordinates method. The example of a simple calibration application is found in Appendix A. Since all of the images were the same pixel size; the same calibration can be used. Shapiro and Stockman [18] state that "camera calibration is conceived to be parameter recovery." Two types of camera parameters can be recovered: Extrinsic and Intrinsic. Extrinsic factors are found in 3D images in order to measure translation and rotation. For the scope of this project, the only concern is to recapture the intrinsic parameters, such as scale factors. Other intrinsic parameters are the focal length and the lens distortion factor. Since the focal length of the camera, which represents the length from the optical center of the lens to the focal point, is so small on the microscope; this value is assumed not to affect the calibration process. The lens distortion factor is also neglected. In order to take the lens distortion factor into account, a micro calibration grid would be required. If the calibration was completed with a micro calibration grid, the resultant would be considered a precise calibration. Instead, a stage micrometer was used with a point coordinate calibration method. The micrometer is built onto a slide, and it is viewed while resting on microscope stage. The total length of the stage micrometer is 100µm. By using the micrometer, the resultant is referred to as a coarse calibration. Depending on the type of image that the calibration is being applied to, other forms of calibration measurement may need to be used. More complicated types of image processing may require calibration transformation from 2D images to 3D coordinates. One instrument of calibration is a calibration jig. Calibration jigs record the 3D objects and creates a 2D template. This 2D template can then be used to take images and convert them back into 3D coordinates.

3.1.2 Color Plane Extraction

Most digital images are captured as color images. They ultimately, though, need to be processed as gray-scale images. Gray-scale images are simplified versions that allow for easy pixel manipulation and processing. The specific images that were taken in this experiment are all color images. One method to convert color images to grayscale images is color plane extraction. In order to transform a color

image to a gray-scale image, one color plane must be removed. There are three color planes that make up a color image: red, blue, and green. These three color planes are often referred to as RBG (red, blue, green). Through these three color planes, humans are able to distinguish shades of various colors. The original color image is simply a representation of all three color planes stacked together. The three color planes can be expressed in decimal format. Each color plane is represented by a value between 0 and 255. For example, the true red plane is (255, 0, 0). The true blue plane is (0, 0, 255); and the true green plane is (0, 255, 0). Overlaying two or more planes results in various colors. For example, a true red plane and a true green plane overlaid become yellow (255, 255, 0) [18]. The options of color plane combinations results in over 16 million colors. The analysis of one color plane by itself will result in a gray scale image. The gray-scale image is a simplified representation of the color image. Every pixel within a gray-scale image is represented by a single numerical value between 0 and 256. The numerical value represents the spectrum between white (255) and black (0) in varying scales of gray. An image with the simplicity of a single value per pixel is desirable within image processing for its ease of processing. For the scope of this project, a green color plane was extracted from the original image as seen previously in Figure 2.2. The resultant gray-scale image is shown in Figure 3.1. For the application of this research, the color plane that will be extracted depends upon the type of parasite that was analyzed in order to maximize the contrast in the image for better results. The gray-scale images retrieved in the color plane extraction step will continue on to receive more image processing.

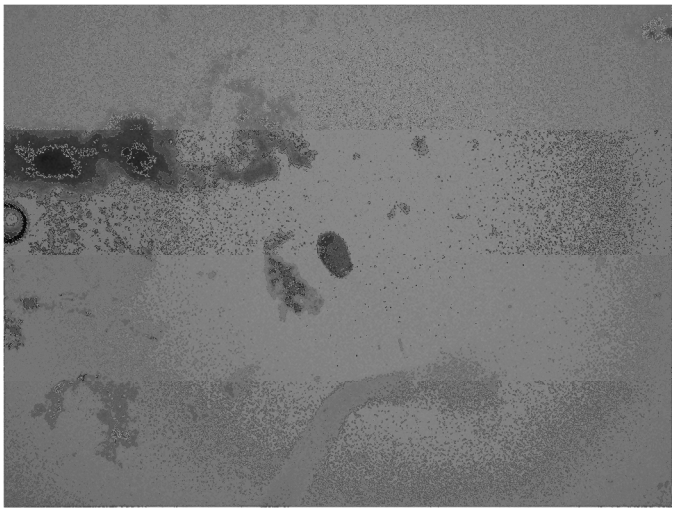

Figure 3.1 The green color plane extracted from a color digital image of *Paragonimus westermani* results in a gray-scale image.

3.1.3 Thresholding

Thresholding is an operation on gray-scale images that separates background and foreground pixels. The threshold step is crucial because it has the potential to eliminate most background material in the image. The level at which the foreground and background is separated is called the threshold. First, the image must be gray-scale before the threshold operation can be performed. The gray-scale is based on which color plane extraction is selected as discussed in Section 3.1.2. The pixels are all values between 0 and 255 in the grayscale image. The threshold operation is performed on the image determining at what grayscale value to set the threshold. All pixels that are located below the threshold are now considered background pixels. All pixels that are located at or above the threshold are foreground pixels, or pixels of interest.

The following are three methods to perform thresholding: manually, Otsu's global approach, and adaptive thresholding. A manual selection refers to the user defining the thresholding value. The manual thresholding approach is done by trial-and-error. The user wants the significant objects in the image to remain as foreground material, so the threshold will be set in a manner to retain the objects while turning the unwanted objects into the background. In this particular research, the manual method was used, and the threshold was selected so that the parasite egg's exterior shape would maintain its form. A second common method to apply thresholding was developed by Otsu [19] and uses a global thresholding approach. Otsu's global approach assumes that a bimodal distribution is present within the image. A bimodal gray-scale distribution is present when the background is uniform, and there is a significant difference in the grayscale values of the background and foreground objects. Otsu's global thresholding approach will automatically choose a threshold value based on the analysis of the whole image. The third method of thresholding is called adaptive thresholding [20], which analyzes sections of the image to determine the best thresholding value by section. Adaptive thresholding operates under the pretense that the background is non-uniform. In this case, there is typically a unimodal distribution of the gray-scale values. Adaptive thresholding will automatically choose a threshold value for the whole image based on the results from the sections' analysis. Both of these methods, global and adaptive, will be applied to the digital images within this project. The results of each will be analyzed and compared. The new image is referred to as a binary image.

The basis for understanding thresholding lies in the histogram of an image. A histogram is the graphical representation of the results of the threshold operation. A histogram displays the number of pixels that represent each grayscale value from 0 to 255 for an entire image. A larger peak indicates that there are many pixels in the image that correspond to that grayscale. The threshold is placed at the

average grayscale value. A manual selection of the threshold can be implemented as in Figure 3.2 where the threshold is represented by the vertical line at the gray-scale value of 90. An example of a histogram is seen in Figure 3.2 with a threshold value of 90.

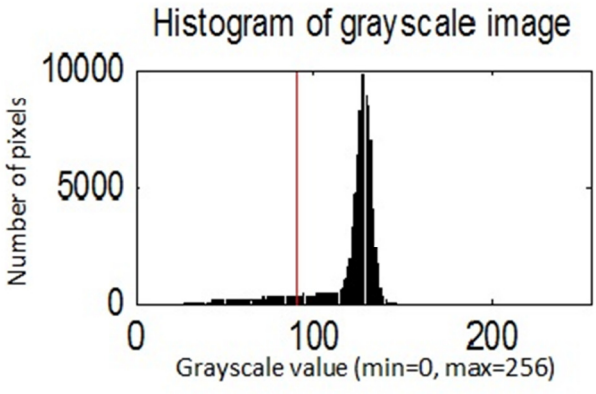

Figure 3.2 A basic unimodal histogram with a manually applied threshold of 90.

In Figure 3.2, the pixels with a gray-scale value greater than 90 will be considered foreground material. Depending on the computer programming being used, the foreground will automatically be colored red or white. Gray-scale values less than 90, in Figure 3.2, will be classified as background material and will be colored black. Once the value for the threshold has been chosen, the pixel values are changed to reflect the binary result. A pixel, if determined to be background, will be set to '0'. All foreground pixels will be set to '1'. The binary image from Figure 3.2 is shown in Figure 3.3.

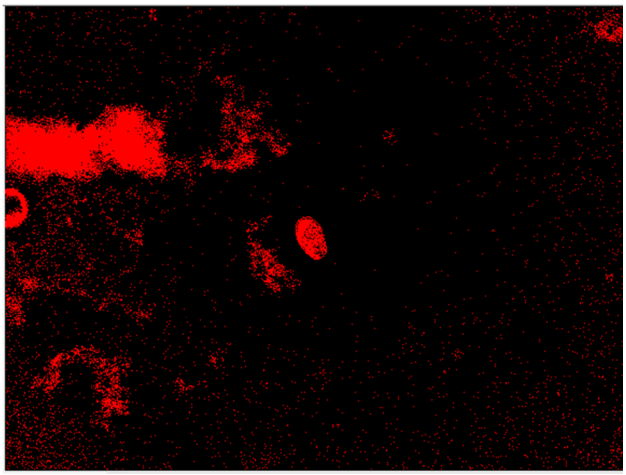

Figure 3.3 The binary image of a digital image of *Paragonimus westermani*.

Converting the original image to a binary image is an important step in pre-processing the image, and the step can be considered the weak leak of the process since so much importance is placed upon it. A binary image in Figure 3.2 is made up of the foreground pixels, which are values of '1', and the background pixels, which are values of '0'. In the LabVIEW™ results of Figure 3.2, the foreground is colored red; the background is colored black. Other programs, such as MATLAB™, show the foreground pixels as white and the background pixels as black. The foreground is composed of the particles of interest to the viewer. The background can be noise or other particles not of interest. Sometimes the foreground pixels will include objects that are actually noise. Thresholding can eliminate most noise, but it does rely on other image processing steps to completely rid the image of noise.

3.1.4 Filtering

The filtering step manipulates a gray-scale image in order to separate and define objects. There are many ways to filter an image. One way is through a segmentation operation. Segmentation extracts features from the image such as edges, blobs, or even partitioning of the image based on certain values. Segmentation is based on pixel connectivity [21], which is the touching of pixels. For example, a 4-connected pixel is evaluated based on the four pixels above, below, and on each side. An 8-connected pixel is evaluated by all eight pixels that touch. An example of both the 4- and 8-connected pixel can be seen in Figure 3.4 and Figure 3.5, respectively.

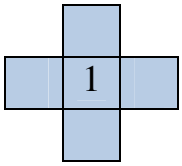

Figure 3.4 A 4-connected pixel analysis overlays the pixels at the sides, top, and bottom**.**

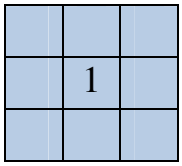

Figure 3.5 An 8-connected pixel analysis overlays the pixels at all sides.

Bovik [2] references an edge-based segmentation which locates the edges of the foreground objects and separates it from background objects. The edged-based segmentation can be performed by edge detection filters like the Laplacian of Gaussian (LoG) [22]. The Laplacian of Gaussian can be broken down into two parts. The first step is to apply the Gaussian step of blurring. The second step is to

apply the Laplacian operator for edge detection. The MATLAB™ applications used Laplacian of Gaussian to define edges and find the parasite eggs.

Convolution is a type of filtering that sharpens edges and is based on the kernel size. Kernels are various size squares of pixels, which can be 3x3, 5x5, or 7x7. The size of the kernel is user-defined. The convolution feature establishes sharp transitions from black to white pixels in areas that were previously gray. The kernel size is chosen depending on the amount of detail that the user would like to filter. Choosing a 3x3 kernel size would result in more edges being defined because the area being analyzed by the kernel is small. If the larger 7x7 kernel size is chosen, the overall results of the filtering will have only the major edges defined. This is because the individual kernels were larger and cannot capture all of the edge-details. In LabVIEW™ Vision Assistant™, convolution is used to find the edges of particles. Figure 3.6 shows the results from Vision Assistant™'s application of convolution to a microscope digital image. Comparing this image to the original image in Figure 3.1, the edges around the parasite egg have become more defined.

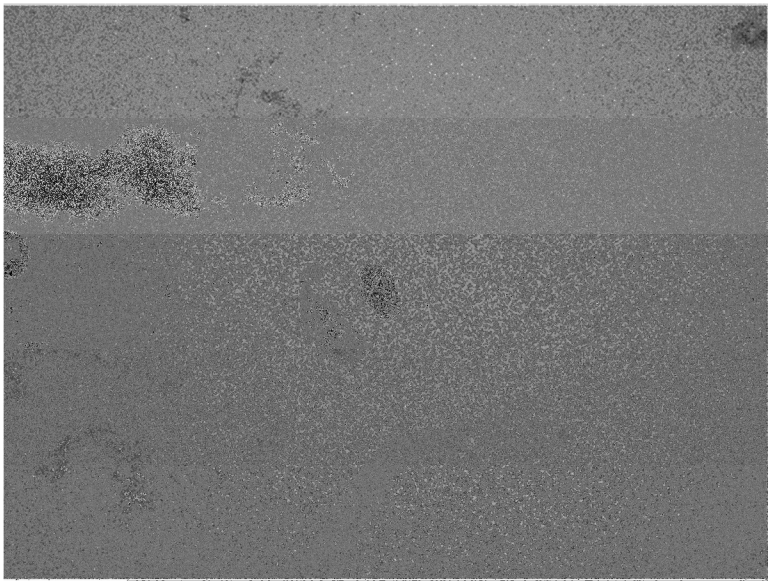

Figure 3.6 A filtered image of *Paragonimus westermani* that has been involved in a convolution step with a kernel size of 5x5.

A third filter type is smoothing, which is also based on kernel size. Smoothing takes place through a high pass frequency domain filter. A high-pass filter rids the image of small, ill-defined background particles, specifically objects that are not parasite eggs, by only allowing the well-defined objects of interest to translate. The size of the chosen kernel will affect the results. If the largest kernel size is selected, the most smoothing will take place. This leaves only the main objects of interest defined

in the image. The smaller kernel size will result in minimal smoothing and most objects of interest will remain defined in the image.

A fourth filter is blob detection, a region-based filtering method. Blobs are connected similar pixels. There are several forms of blob detection such as regions of interest and histogram analysis. Blob detection is easily achievable if there are higher intensity changes, discontinuity among pixels or great similarity in regions according to a set criterion. A simple filtering step can be background subtraction. An image of a blank field of vision is taken and subtracted from the analyzed data image. This type of filtering may reduce the gray-scale values and destroy any background noise that may be present due to lighting conditions or dirt particles.

Filtering can also take place by image enhancement through increasing the contrast and resolution levels. Low, median, and high-pass filters are another option to filtering. Each has varying effects on the image. Low-pass filtering tends to blur the image. Median filter slightly blurs the images, but the integrity of the image is maintained while background noise is blurred away. High-pass images intensify the edges of objects. It also strengthens the gray-scale intensity of the background.

3.1.5 Morphological Operators

Morphological operators manipulate the shapes of the objects of interest in a binary image. There are two main reasons that morphological operators would need to be applied to digital images. The first is that the foreground pixels determined during thresholding are typically not representative of the objects' true shapes. Secondly, the foreground pixels of two or more objects may also be touching or overlapping. This creates the false impression of one large object. Morphological operators come in the following four main forms for binary images: dilation, erosion, closing, and opening. [18, 23] Additionally shape matching can be performed based on the desired objects shapes and the shapes of the objects in the image.

Dilation enlarges the foreground pixels. Dilation is represented by the following mathematical equation,

$$\mathbf{A} \oplus \mathbf{B} = \cup_{a \in A} \boldsymbol{B}_{a}, \tag{3.1}$$

where A represents the binary image, B is the structuring element [18] of a fixed kernel size, a is the individual pixel of analysis in the binary image. Dilation processes the kernel of a set size through the image by scanning every line beginning with the top right corner. The structuring element, if it were a 3x3 kernel size, would look like the object in Figure 3.7. Through the pass, the union of binary pixels is looked for between the structuring element and the image's pixels. If the structuring element matrix was

placed on the binary image and it touched any pixel with a value of 1, the central pixel would turn into a value of 1, which classifies it as part of the object.

1	1	1
1	**1**	1
1	1	1

Figure 3.7 A structuring element, such as this 3x3 example, is used in the morphological operations.

Erosion is another morphological operator that ultimately reduces the foreground objects at their edges. It uses the following equation,

$$\mathbf{A} \ominus \mathbf{B} = \bigcap_{a \in A} \boldsymbol{B}_{-a}, \quad 3.2$$

where A represents the binary image, B is the structuring element of a fixed kernel size, a is the individual pixel of analysis in the binary image. Erosion completes the same pass of the structuring element, seen in Figure 3.7, across the binary image as the Dilation operator. Instead of looking for unions of pixels, the erosion operation watches for intersections. Through this operation, the foreground object's outlier pixels are effectively deleted as part of the object. To complete erosion, all pixels in the structuring element must intersect with pixels that belong to a foreground object in the binary image. If this condition is met, the center pixel remains a value of 1. If even one overlapped pixel does not equal the value of 1, the center pixel becomes a 0 value, which acknowledges this pixel as background material. Thus, the foreground objects are effectively eroded.

Closing is a combination of a dilation operation followed by an erosion operation. It uses the following equation,

$$\mathbf{A} \bullet \mathbf{B} = (\mathbf{A} \oplus \mathbf{B}) \ominus \mathbf{B}, \quad 3.3$$

which is previously explained in Equations 3.1 and 3.2. By dilating the image first and then performing an erosion operation, the original binary object will be able to close some missing particles that were left out during thresholding. An object that is a good candidate for the close operation would have a recognizable shape. It would be missing some interior and edge pixels. The dilation operation would fill in the missing pixels of the interior and edges. At the same time, some additional pixels may be added to the edges. The erosion operation would then follow to trim back these extra pixels that were added to

the edges of the object. The closing operation maintains the integrity of the object's shape without enlarging it.

Opening is a combination of an erosion operation followed by a dilation operation. It is performed through the following equation,

$$\mathbf{A} \circ \mathbf{B} = (\mathbf{A} \ominus \mathbf{B}) \oplus \mathbf{B}, \tag{3.4}$$

which is previously explained in Equations 3.1 and 3.2. The erosion operation would then follow to trim back these extra pixels that were added to the edges of the object. An object that is a good candidate for the opening operation would have a recognizable shape, but there would be outlier particles present that may not be part of the object. In other words, extra pixels are present that do not represent the particles of interest. If outlying pixels are present, they will be turned to background pixels during the erosion step along with the whole edge of the object. The dilation operation would fill in the missing pixels of the interior and edges, while restoring any edge pixel that were altered in the erosion step. This step retrieves the edges of the object to maintain the proper edge without invoking new outlier pixels. Through the opening operation, a binary image is able to rid itself of outlying pixels while maintaining the true shape of the objects.

The morphological operations results depend on the selected kernel size. The four operations, dilation, erosion, closing, and opening, allow image processing to successfully remove objects or fill holes in objects. The dilation, erosion, and closing operations will be applied in Vision Assistant™. An example of a closing operation can be seen in Figure 3.8. Image (a) of Figure 3.8 represents a binary image of a *Paragonimus westermani* egg. In addition to the parasite, there are many particles that are not of interest. Image (b) of Figure 3.8 shows the same image after a closing operation has been performed. The first step, dilation, fills pixel gaps of the remaining objects. The second step, erosion, rids the image of unwanted particles.

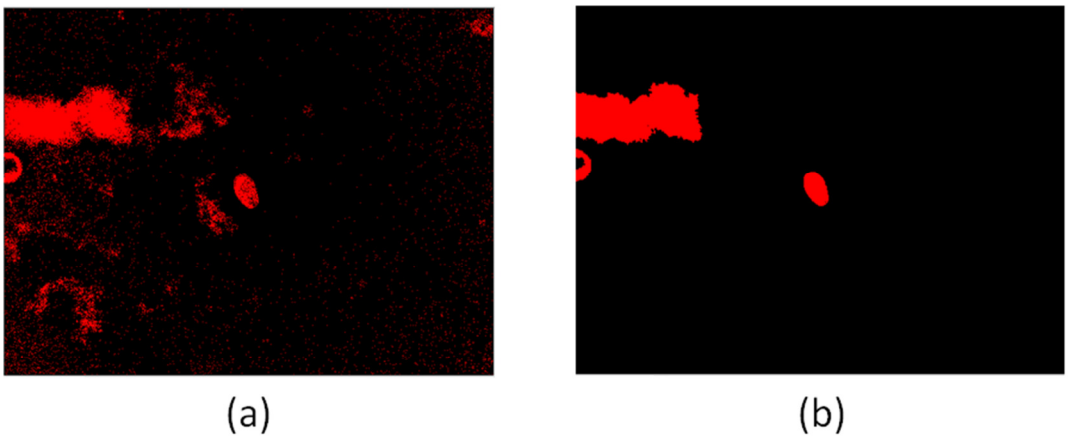

Figure 3.8 A filtered image of *Paragonimus westermani* that has been involved in a closing operation with a kernel size of 5x5 pixels.

Shape matching is the final morphological operation that will be discussed. An object can be saved within a computer program, such as LabVIEW™, where measurements such as width, length, and area are stored. This same data is gathered from objects found in a binary image, and the data is compared to the stored image. This is called shape matching. Based on the results of the comparison, a value will be produced that shows the likelihood that the two objects are related. The user can set the lower limit for the value of object relatedness. If the value is above this lower limit, the shapes are assumed to be matching. Shape matching proves to be an acceptable image processing step, if the user knows the shape of the object that is desired from the binary image.

3.2 Methods of Specific Processing Applications

One of the many applications of image processing is directed at the field of computer-assisted microscopy. Bovik [2] states, "Computer-assisted microscopy provides the ability to enhance the speed of microscope data acquisition and data analysis, thus relieving humans of tedious tasks." In Bovik's book, an entire chapter is solely dedicated to the application of image processing to computer-aided microscopy. For the scope of this project, image processing techniques from Section 3.1 are applied using two commercially available computer programs, LabVIEW™ Vision Assistant™ and MATLAB™. The following four applications will be discussed in detail: LabVIEW™ scripts, MATLAB™ separation code, MATLAB™ cross-section grayscale analysis, and MATLAB™ edge signature analysis. These four applications were chosen because the results from each image processing application can be used to characterize and differentiate between the parasite egg types. The images that were captured in Section 2.3 will be processed through the scripts, separation code, cross-section

grayscale analysis, and edge signature analysis applications that were developed. The general results will be analyzed and discussed. The specific results of the LabVIEW™ scripts, MATLAB™ Cross-section Grayscale, and MATLAB™ Edge Signature are saved for discussion within Chapter 4.

3.2.1 LabVIEW™ scripts

LabVIEW™ Vision Assistant™ is the program used to process individual images for locating parasite eggs and characterizing them. Vision Assistant™ is a National Instruments™ application program that performs image processing on the platform of the LabVIEW™ language. Three scripts were created within LabVIEW™ Vision Assistant™, one for each parasite egg type. A script is a flow diagram for the image processing steps that take place. The author designed the image processing steps order along with the specifications for each step. All image processing techniques from Section 3.1 were utilized to determine a method for distinguishing the parasite egg in each image. Each parasite egg type has its unique script. The initial goal was to determine a way to locate a single specific parasite egg by using a script that is ideal for that parasite type. This step was followed up by applying the same image processing flow to all three types of parasite eggs. The three scripts were streamlined and made similar as much as possible within the scope of this project. In the following sections, the three parasite egg scripts will be explained, and their similarities and differences will be discussed. The original digital images, which were collected in Section 2.3, were opened within LabVIEW™ for processing. Figure 3.9 shows us the layout of an example microscope image in LabVIEW™ Vision Assistant™.

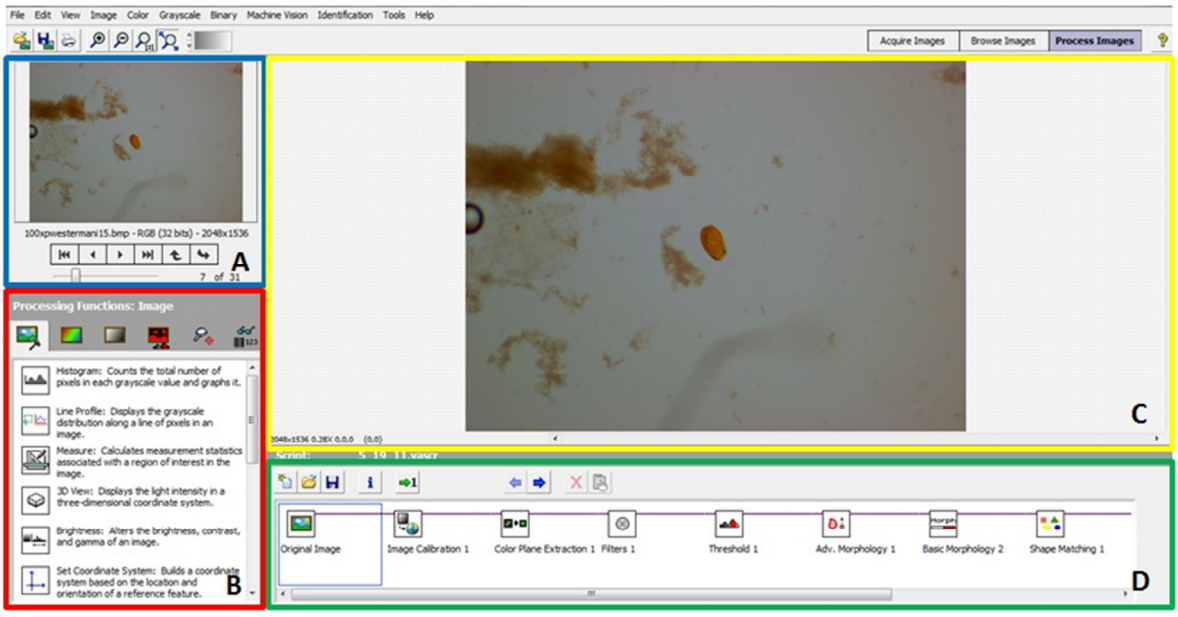

Figure 3.9 LabVIEW™ Vision Assistant™ executes image processing steps through user-built scripts.

Area A in Figure 3.9 displays the digital microscopic images that have been opened within LabVIEW™ and are available for processing. Multiple images can be opened within LabVIEW™ at the same time. They are processed individually, however. Area B in Figure 3.9 gives the user control to implement the image processing steps. Color, Grayscale and Binary image processing steps are available as well as some measurement processes. The detailed image processing steps described in Section 3.1 are available within Vision Assistant™. Area C of Figure 3.9 displays the selected image that was chosen to be processed from Area A. The number of pixels is shown in the bottom left corner of Area C. Area D of the image displays the script for the image processing steps. When a step from Area B is chosen, it appears in Area D. The user also has the option of highlighting existing script steps from Area D. This allows for the user to edit the step within Area B. When a step is highlighted, the resultant of the script step is instantaneously shown as the image in Area C. Figure 3.9 shows the "Original Image" step highlighted in Area D; thus, the original image is displayed in Area C.

3.2.1.1 *Paragonimus westermani* LabVIEW™ script

For *Paragonimus westermani,* the first step in Vision Assistant™ is to load the original parasite egg images. These images are stored so that all images are easily accessible one at a time by the script. The second step of the script is to calibrate the image. The calibration is shown in more detail in Appendix A. The calibration is the same for all parasite images no matter which parasite type is being analyzed. It is based on the point coordinate method calibration discussed in Section 3.1.1. The third step is the color plane extraction. In the case of the *Paragonimus westermani* eggs, the color plane that is extracted is the green plane. The green color plane was chosen so that there would be a maximum contrast in the image. The parasite would appear much darker than the background material. This makes it easier to differentiate in further image processing. This step is also necessary to achieve a gray-scale image for further processing. The fourth step is the first filtering step, which is convolution with a 5x5 pixel kernel. This type of filtering allows Vision Assistant™ to begin the process of locating edges. The fifth step implements thresholding, which is manually set at a value of 88. The threshold value was selected so that significant objects in the image would remain as foreground material, and the threshold would retain the parasite eggs while turning the unwanted objects into the background. Advanced Morphology, a type of filtering discussed in Section 3.1.4, was then implemented. The advanced morphology option of removing smaller particles was completed in four iterations with a pixel connectivity of eight, as discussed in Section 3.1.4. A basic morphology step is then executed to close all remaining objects. The closing step is accomplished with a 5x5 pixel kernel. A 5x5 kernel size results in a medium amount of edges being defined since the area being analyzed by the kernel is a

medium size. This kernel size works for this parasite since the shell of the parasite is already well-defined in the image. Shape matching then takes place with a *Paragonimus westermani* egg template. The minimum score for the shape matching is a value of 750. The ninth step is a particle filter. The particle filter looks for a Heywood Circularity Factor between 0.5 and 2 pixels with a connectivity of eight pixels. Heywood Circularity Factor, a measure of roundness, is a ratio of the perimeter of the object divided by the perimeter of a circle with the same area as the object. The particles that remain then go through a particle analysis step. All of the data that will be analyzed in Chapter 4 was gathered in this script step. The eleventh step of the script involves an automatic particle analysis on the remaining objects. This step analyzes and gathers information such as area, length, and width of the parasite eggs. The twelfth and final step is another particle filter. This particle filters double-checks the size of the remaining particles. If the particle's area is between 5000 and 12000 pixels, it is kept. At the conclusion of this step the script is finished. The user will be left with a binary image as the result. A flow chart comparing this script to the other parasite egg scripts will be shown in Figure 3.10 of Section 3.2.1.4.

3.2.1.2 *Taenia* spp. LabVIEW™ script

For *Taenia* spp. parasite type, the first step in Vision Assistant™ is to load the original parasite egg images. These images are stored so that all images are easily accessible one at a time by the script. The second step of the script is to calibrate the image. The calibration is shown in more detail in Appendix A. The third step is the color plane extraction. In the case of the *Taenia* spp. egg, the color plane that is extracted is the red plane. The red color plane was chosen so that there would be a maximum contrast in the image. The parasite would appear much darker than the background material. This makes it easier to differentiate in further image processing. This step is also necessary to achieve a gray-scale image for further processing. The fourth step implements thresholding. The manual threshold was set between grayscale values of 72 and 99. The threshold value was selected so that significant objects in the image would remain as foreground material, and the threshold would retain the parasite eggs while turning the unwanted objects into the background. Advanced Morphology, a type of filtering discussed in Section 3.1.4, was then implemented. In this filtering, a "fill holes" step was completed with a connectivity of eight pixels, as discussed in Section 3.1.4. A basic morphology step is then executed to erode all objects. The erosion step is accomplished with a 3x3 pixel kernel. A 3x3 kernel size results in more edges being defined since the area being analyzed by the kernel is small. Another Advanced Morphology script step takes place to fill holes again with a connectivity of eight pixels. The eighth step is a particle filter. The particle filter looks for a Heywood Circularity Factor between 0.5 and

2.5 pixels with a connectivity of eight pixels. The particles that remain then go through a particle analysis step. This step analyzes and gathers information such as area, length, and width of the parasite eggs. All of the data that will be analyzed in Chapter 4 was gathered in this script step. The tenth step is a particle filter. This particle filters double-checks the size of the remaining particles. If the perimeter of the particle is between 150 to 210 pixels, it is kept. Another particle filter is also implemented in the eleventh and last script step to check the area of the remaining objects. If the objects area is between 1800 and 4200 pixels, the object is kept. At the conclusion of this step the script is finished. The user will be left with a binary image as the result. A flow chart comparing this script to the other parasite egg scripts will be shown in Figure 3.10 of Section 3.2.1.4.

3.2.1.3 *Ascaris lumbricoides* LabVIEW™ script

For *Ascaris lumbricoides,* the first step in Vision Assistant™ is to load the original parasite egg images. These images are stored so that all images are easily accessible one at a time by the script. The second step of the script is to calibrate the image. The calibration is shown in more detail in Appendix A. The third step is the color plane extraction. In the case of *Ascaris lumbricoides*, the color plane that is extracted is the green plane. The green color plane was chosen so that there would be a maximum contrast in the image. The parasite would appear much darker than the background material. This makes it easier to differentiate in further image processing. This step is also necessary to achieve a gray-scale image for further processing. The fourth step implements a "square root" filter over the entire image. The original pixel values are thus reduced to their square root value. Thresholding was completed in the fifth script step. The threshold was set manually to a value of 152. The threshold value was selected so that significant objects in the image would remain as foreground material, and the threshold would retain the parasite eggs while turning the unwanted objects into the background. Advanced Morphology, a type of filtering discussed in Section 3.1.4, was then implemented. In this step, smaller particles were removed in five iterations with a pixel connectivity of eight, as discussed in Section 3.1.4. A basic morphology step is then executed to dilate all objects. The dilation step is accomplished with a 3x3 kernel pixel. A 3x3 kernel size results in more edges being defined since the area being analyzed by the kernel is small. The eighth step is a particle filter. The particle filter looks for a Heywood Circularity Factor between 1 and 1.5 pixels with a connectivity of eight pixels. The particles that are left then go through a particle analysis step. All of the data that is analyzed in Chapters 4 and 5 was gathered in this script step. The tenth and last step is a particle filter. This particle filters double-checks the size of the remaining particles. If the Maximum Feret Diameter of the particle is between 40-70 micrometers, the particle is kept as a foreground object. The Maximum Feret Diameter (F_d) is the distance between the

two perimeter points that are located the farthest away from each other. It will be discussed in more detail in Section 4.2. At the conclusion of this step the script is finished. The user will be left with a binary image as the result. A flow chart comparing this script to the other parasite egg scripts will be shown in Figure 3.10 of Section 3.2.1.4.

3.2.1.4 LabVIEW™ script summary

All three LabVIEW™ scripts follow the same process flow. There are only a few steps that had to be customized for the particular parasite that was being identified. The parasite eggs have different scripts because originally the main focus was on creating the best script based solely on a single parasite type. Over the course of the research, the focus changed to the possibility of having one script for all parasite types. Changes were made to create similar scripts, but more research is needed to consolidate the scripts further. A summary of the LabVIEW™ Vision Assistant™ scripts is shown in Figure 3.10.

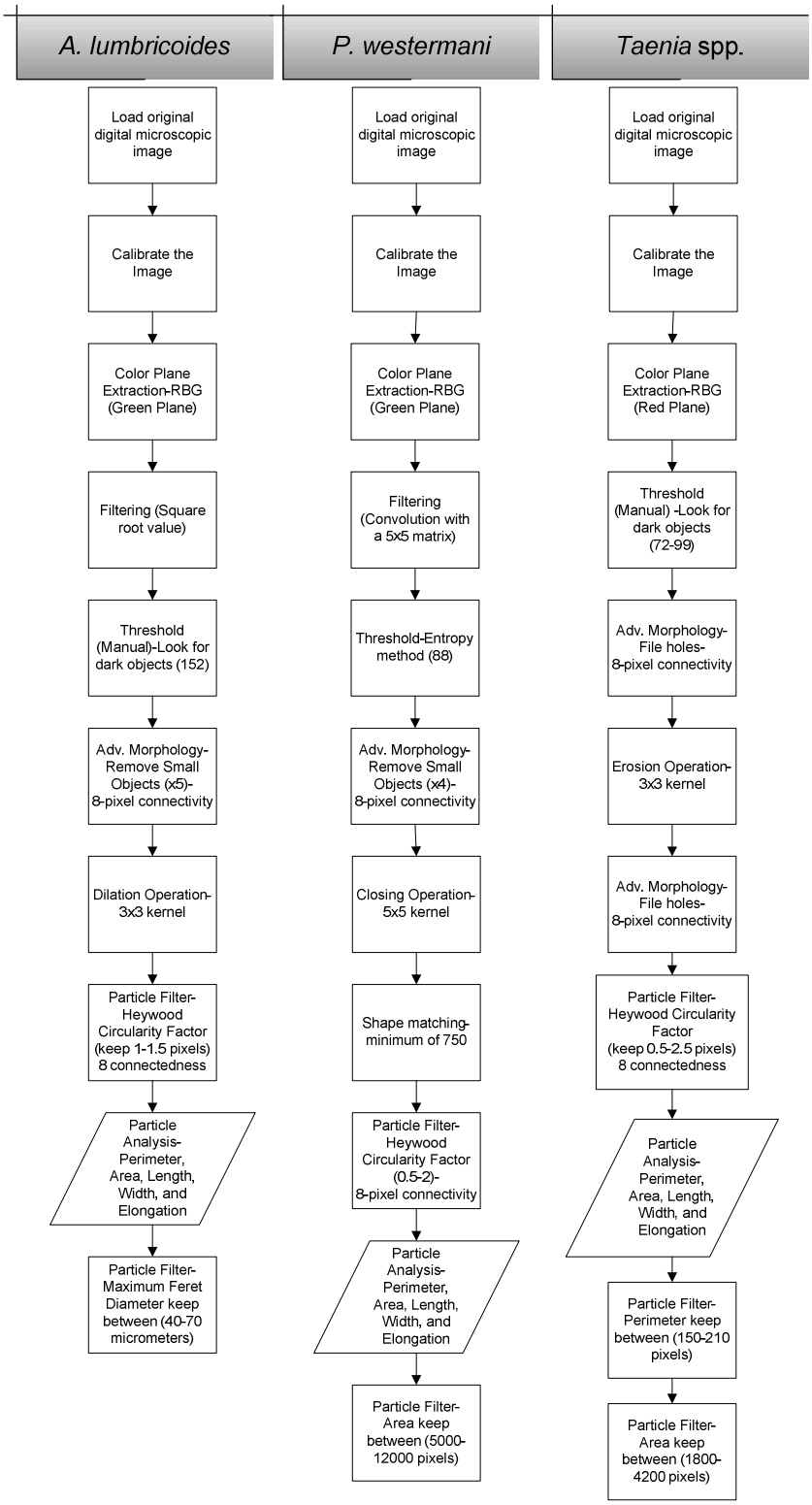

Figure 3.10 Flow diagram of the LabVIEW™ scripts for the three parasite types

There are several similarities in the scripts. Each script requires a color plan extract to obtain the gray-scale image. Each script also uses a threshold operation to separate the background and foreground. Smaller particles are also removed in all three scripts through a filtering step. Near the end of each script, the image of analysis goes through a particle filter. Within these scripts, there were also differences. Both red and green color planes were extracted from the images depending on which parasite was being analyzed. The best color plane was chosen based on the contrast in the extracted plane. A desirable color plane would have high contrast; for example, dark objects on a light background. The threshold values also vary by parasite egg type. The particle filters are the last step, but they require some knowledge of the parasite morphology. Two scripts use the area of the particles to filter, but the third script uses the diameter of the parasites to filter objects. Creating these three separate scripts is recognized as the first step in reaching the end goal of a single LabVIEW™ script that, with no previous knowledge of what parasite type is present in the digital image, could distinguish between all three parasites.

3.2.2 MATLAB™ Separation Code

In some of the digital images that were captured, the parasite cannot be visibly separated or defined. Overlapping of both parasite and background particles is present. Minor overlapping can be fixed by some separation and filtering steps presented in Section 3.1.4. Sometimes though, these steps are not successful in separating objects. Bingham et al. [24] developed a method to separate these overlapping objects. Based on Bingham et al.'s [24] research, the author developed a MATLAB™ code to separate any overlapping in the microscopic parasite images. The success of separating overlapping images relies on the following two steps: finding the interior cores of the foreground objects and dilating the objects to find their linking points. These steps were implemented in a MATLAB™ code developed specifically for this project by the author. Prior to putting the digital image through the separation code, the images needed to be pre-processed within MATLAB™ using standard thresholding and morphological operations. This MATLAB™ separation code is not connected to the LabVIEW™ scripts; therefore, all of the image processing steps were completed in MATLAB™. Figure 3.11 shows the pre-processing steps that took place in MATLAB™ to prepare the image for the separation code.

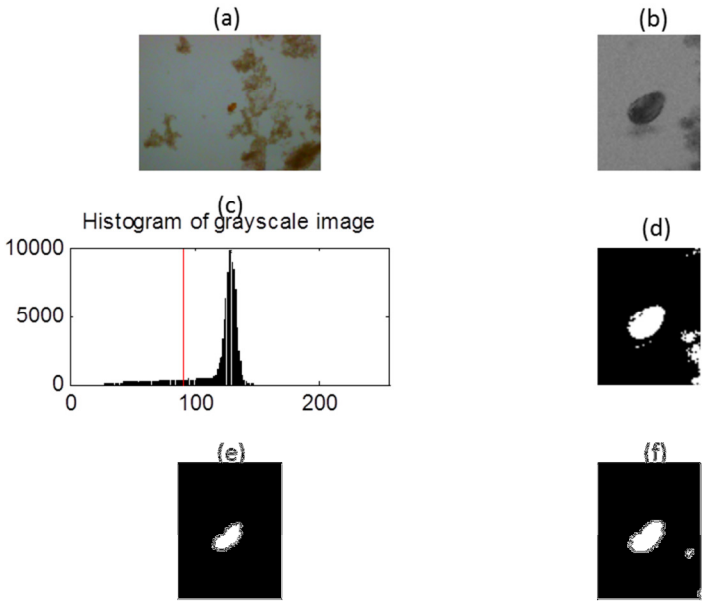

Figure 3.11 Pre-processing for MATLAB™ Separation Code *Paragonimus westermani* Image #1

In Figure 3.11, Image A represents the original digital image that is entered into MATLAB™. Image B is a cropped, close-up gray-scale image of the parasite egg of interest. A MATLAB™ histogram of Image B is shown in Image C with the threshold value set at the solid vertical line. The threshold was determined by applying the global threshold method to Image B. Image D is the binary image result of the threshold being applied. An erosion operator was applied in Image E to rid the image of some background noise. Image F shows the results of a dilation operation. Each of the steps was performed within MATLAB™ prior to the separation code. The separation code was then applied to the image. First, a scan is performed from the top left corner to the bottom right. This first pass takes into account the pixels of the objects and calculates the distance to the edge of the object. This value is assigned to the pixel. The same pass is completed from the bottom right corner of the image to the top left. Once again, the distance from the pixel of interest to the edge of the object is assigned. This calculation is called the Distance Transform. Through both passes, the cores of the objects are able to be located as the pixels of highest value. Only the core pixels of the objects are assigned a gray-scale value of one. An example of the final result can be seen in Figure 3.12.

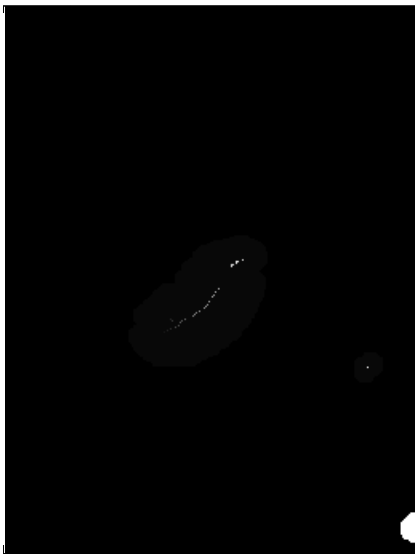

Figure 3.12 *Paragonimus westermani* Image #1 result from MATLAB™ Separation Code

Through the process of applying this separation code, it was realized that the code would work best on circular objects. However, the core of non-circular objects can still be found. For more elliptical parasites, like those in Figure 3.12, a line of pixels is recovered as the core of the object. This data can be a defining factor in helping determine not only the location of the parasite but also its type. Due to the fact that quantitative measurement data was not pulled from this researched method, there are no numerical results to include in Chapter 4. However, the image results from the MATLAB™ separation code, like those seen in Figure 3.12, will be shown and discussed in Appendix C. The MATLAB™ separation code can be seen in Appendix D. This MATLAB™ separation code is not connected to the LabVIEW™ scripts; therefore, all of the image processing steps were completed in MATLAB™. In the future, it would be desirable to use the output of LabVIEW™ Vision Assistant's™ scripts to be the input into the MATLAB™ separation code. The main goal was to test the validity of the separation method, so the extensive task of linking the script outputs to the separation code was not acted upon. From the results of the separation code, the particles of interest may be located based on the core pixels. This will require more research into the subject to determine if it is possible.

3.2.3 MATLAB™ Cross-section Grayscale

Through MATLAB™, the parasite cross-section grayscale values were also analyzed. The gray-scale intensities across the parasite eggs will differ between parasite egg types. One distinguishing factor is the shell thickness. Some parasites tend to have a thicker shell while others are thin and hardly noticeable. The shell is the protective outer edge of the parasite egg. Since the parasite shell data will be

included in the cross-section grayscale, this data will be a determining factor in distinguishing the parasite egg types. The interior of the parasite egg is also a distinguishing factor. In the interior, parasites will have differing larvae forming within their interior. This will surely contribute to a difference in the cross-section grayscale intensities. MATLAB™ was utilized to collect data about the cross-section grayscale intensities of the parasite eggs. Because the interior and shell of parasite eggs can be distinctively different, this program was implemented to gain more data to better characterize the different parasite egg types.

The author developed a MATLAB™ code to obtain the gray-scale intensities across the parasite eggs. First, the image of interest is called into the MATLAB™ program. A color plane extraction is completed by removing the green plane. This obtains a gray-scale image. After manually zooming in on the parasite egg, the user is able to select the cross-section of the parasite. This is a manual step. The user of the program draws a line through the mid-section of the parasite eggs. The midsection is characterized by the width of the parasite egg. Therefore, lines were drawn in such a way as to divide the parasites in half through the width. This action is based solely on program user's judgment of the cross-section location. In the image of Figure 3.13, one can see the diamond and circle as well as a dashed line through the parasite egg. The diamond represents the start position, initial click, of the user's manually-drawn line. On their second click, the user clicked the area of the circle. This represents the stop for MATLAB™ analysis. MATLAB™ automatically draws a straight line between the two clicks. The gray-scale intensity at every pixel across the line is captured in the graph of Figure 3.13.

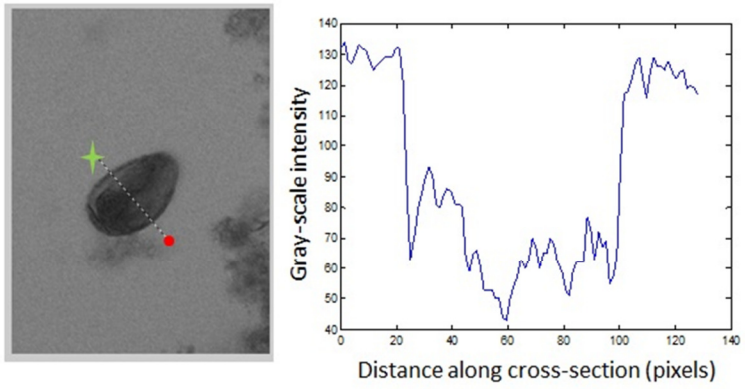

Figure 3.13 *Paragonimus westermani* results of the cross-section grayscale intensities.

From the data retrieved in Figure 3.13, MATLAB™ is able to make some calculations from the intensity values. Data such as the maximum peak value and number of peaks is captured from within MATLAB™. This data proved to be very insightful into the classification of the three parasite eggs.

More discussion and results will be presented in Section 4.3. The additional image results can be viewed in Appendix E. The MATLAB™ code for the gray-scale cross-section can be viewed in Appendix F.

3.2.4 MATLAB™ Edge Signature

One characteristic that differs between parasite eggs is the shape. The author developed a MATLAB™ code to obtain the edge signature of the parasite eggs located in the image. Edge signatures provide a way to capture the shape of an object in quantitative data form. The first step in characterizing the edge signature of an object is to locate the centroid. The centroid is the center of the object. It gives a reference point to base the data. The data is then retrieved by calculating the distance from the centroid to all outer edges of the object. A good mental image of this concept is measuring a single line from the centroid of the object to an arbitrary point of the object's edge. Imagine how the length of this line would change if it were rotated 360° around the object. Through narrow sections, the length would be shorter. In wider portions, the length would need to be longer. This length data is gathered and plotted on a graph versus the angle of rotation around the object. An example of this graph is shown in Figure 3.14.

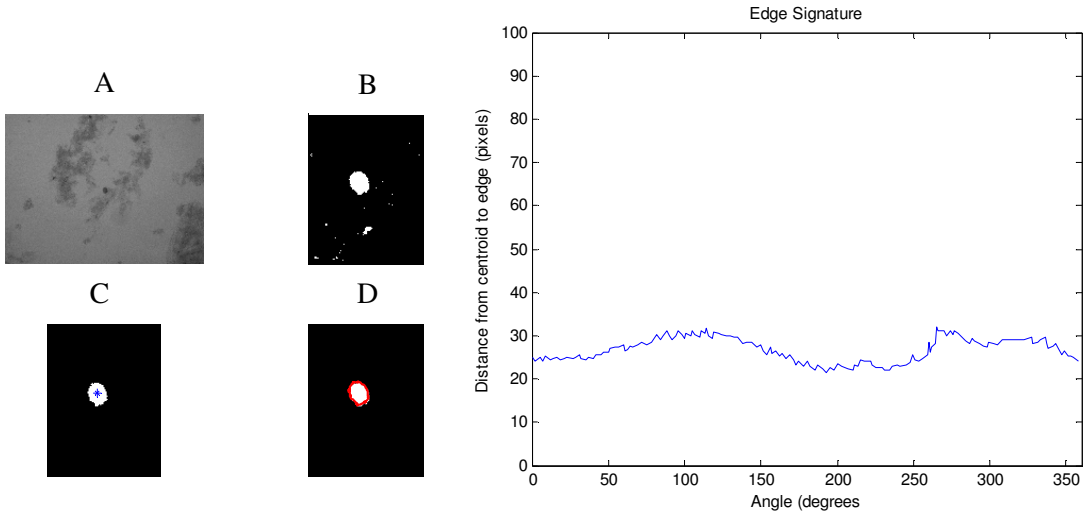

Figure 3.14 Edge Signature results of *Taenia* spp. Image #1 at a threshold of 90.

The left portion of Figure 3.14 represents the MATLAB™ steps to process the image in order to get the edge signature results. Image A is calling the image into MATLAB™ and converting the image to grayscale. MATLAB™ extracted the green plane to gain a grayscale image. The threshold was set to a value of 100 in Image B. A subsection of the image which contains the parasite was also user-found and enlarged in this step. This subsection image was then analyzed to find the centroid of the largest

object. Image C in Figure 3.14 represents the binary subsection image of the parasite where the centroid has been marked. Boundaries of the object are then marked in the Image D. The distance between the boundary and the centroid creates the edge signature graph. As the boundary is examined in rotated single degree increments from the centroid, the distance is plotted. In order to distinguish between types of eggs, the following data is also collected from the edge signature graph: Number of Peaks, Peaks per Degree of rotation, Mean Distance from the centroid to the boundary edge, maximum distance between the boundary and centroid, minimum distance between the centroid and the boundary, and the difference between the maximum and minimum distances. These values are calculated within MATLAB™. They will be discussed further in Section 4.4.

The thirty images of each of the three parasite types were put in the Edge Signature code. From these pre-numbered images, three images were selected to show in this paper (Images #1, 15, and 30). Image #1 results for *Paragonimus westermani* and *Ascaris lumbricoides* are seen in Figure 3.15 and Figure 3.16, respectively. Image #15 and #30 results can be seen in Appendix G. The edge signature code developed in MATLAB™ can be seen in Appendix H.

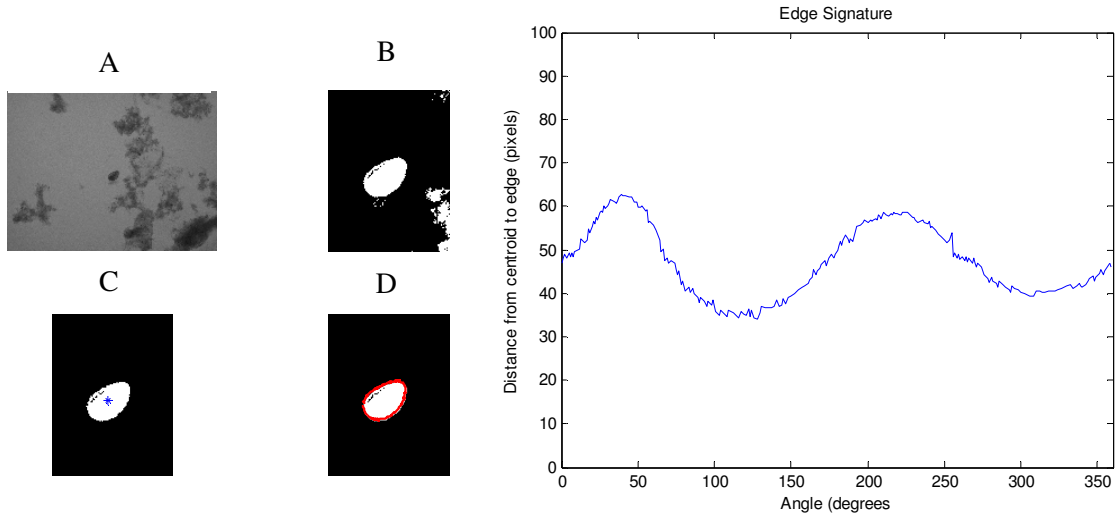

Figure 3.15 Edge Signature result of *Paragonimus westermani* Image #1 at a threshold of 90.

The *Paragonimus westermani* results are greatly different than those presented in Figure 3.14 for *Taenia* spp. The dissimilarity is due to the difference in shape. Since *Paragonimus westermani* is more elliptical in overall shape, the edge signature produces greater maximum distance values. Also, the difference between the maximum and minimum values will be larger than the smaller circular *Taenia* spp.

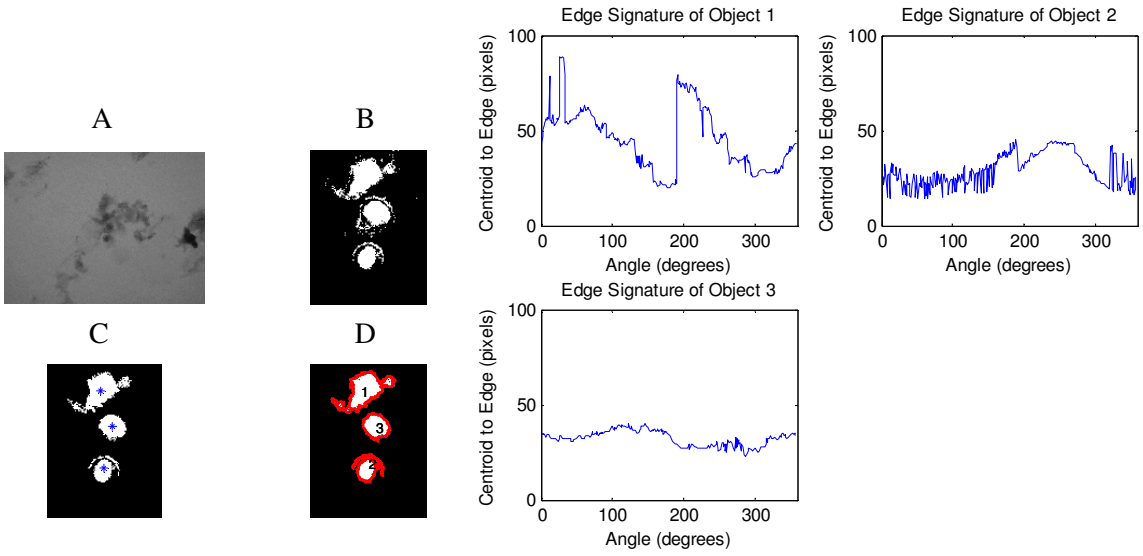

Figure 3.16 Edge Signature result of *Ascaris lumbricoides* Image #1 at a threshold of 79.

In Figure 3.16, two *Ascaris lumbricoides* parasite eggs are present. Unfortunately, MATLAB™ had identified what it believes is three parasites when only two parasites are present. The number 1 object, identified in Image D, is actually background material. Number 2 and 3 particles are the parasite eggs. By collecting enough data from the objects known to be parasite eggs, a characterization can be built in the future to eliminate the non-parasite objects. For example, the number 1 object in Figure 3.16 was included in the edge signature results, but it has centroid-to-edge distance values that vary between 20-90 pixels. Even though the *Ascaris lumbricoides* eggs do not have as much of a defined shape as the previous two examples, the centroid to edge distance is still within a range of about 20-45 pixels. The data that MATLAB™ collects would be enough to rule that this object 1 is not an egg. The results of all images will be discussed in detail in Section 4.4.

Chapter 4 Results

Several approaches were used to analyze the microscopic slide images and characterize the parasite eggs. The digital images were processed four different ways. The first approach was manual. Manual measurements were recorded by two techniques, either through the microscope eyepiece (pre-image capture) or directly after it was captured (post-image capture). The manual approach took into consideration the length and width of the parasite egg. The remaining approaches to characterize the parasite eggs involve several different types of image processing techniques. The second approach was through LabVIEW™. The previously discussed LabVIEW™ Vision Assistant™ scripts in Section 3.2.1 were used to record measurements once the parasite egg had been located within the image. LabVIEW™ has built-in measurement functions that were utilized to perform these measurements. A third approach recorded results through an analysis of the gray-scale intensity of the parasite ova cross-section that was found through a MATLAB™ code. The fourth, and final method, analyzed the parasite's edge signature which was found through another MATLAB™ code. All four methods were applied to the three parasite egg types. Thirty images per parasite type were analyzed in each of the four methods.

There are a total of fifteen individual measurements that will be statistically tested. The measurements were collected using the four approaches (Manual, LabVIEW™, Gray-scale Intensities, and Edge Signature). In the manual test, the Length and Width of each egg was measured. In the LabVIEW™ test, the Area, Equivalent Rectangular Short Side (RSS-Feret), Maximum Feret Diameter, Elongation Factor, and Perimeter were measured. In the gray-scale intensity cross-section evaluation, the (Max-Min) Intensity Difference, Mean Intensity, Number of Peaks, and Number of Peaks (5) were measured. Finally, in the edge signature test the (Max-Min) Distance Difference, Mean Distance, Number of Peaks, and Number of Peaks (5) were measured.

4.1 Manual Approach

The manual measurements were recorded by two different techniques, pre-image capture and post-image capture. This first technique will be called pre-image capture since the technique was applied before the image of the parasite egg was taken and while the prepared microscopic slide sample was still intact. The pre-image capture technique was done through the eyepiece. The eyepiece, which is also referred to as the ocular lens, comes with built-in markers. This is a type of micrometry is called ocular micrometry. The marker distances of the eyepiece are converted to micrometers by measuring a known distance on an object examined under the microscope. This measurement can be done with a

stage micrometer. The stage micrometer is placed on the microscope stage and a conversion factor is determined. The parasite egg measurements were recorded by the number of ocular lens markers. The conversion factor at a microscopic level of 100x was discovered to be 10 micrometers for one ocular mark. Thus, if the microscopic level was set to 400x, the conversion would be 2.5 micrometers for one ocular mark. Using the conversion metrics, the ocular lens marker measurements were changed to micrometers. An image of the stage micrometer taken by the microscope camera at 100x can be seen in Figure 4.1. The micrometer image was taken by laying the micrometer on the microscope and capturing the image through the microscope camera. While the prepared slide with the parasite egg sample is still on the microscope stage, the ocular micrometry manual measurements are acquired by counting the number of ocular markers needed to complete each length and width calculation. The total length of the stage micrometer is 100µm.

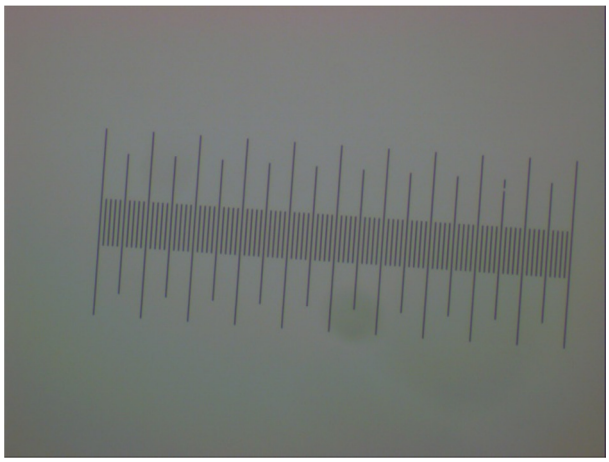

Figure 4.1 Stage micrometer used in conversion and to measure using the post-image capture technique

Some images had been taken before the ocular micrometry technique was utilized. Thus, a different manual measurement technique was applied to these images. This technique will be call post-image capture since the technique was applied after the image of the parasite egg had been captured and the prepared microscopic slide sample destroyed. The post-image capture technique was done by overlaying the captured image of the stage micrometer with the parasite ova image. Overlaying of the two images (parasite egg and micrometer) was accomplished in Microsoft PowerPoint. PowerPoint was chosen for its ability to display the images in their true pixel sizes and to remove backgrounds of images. The image of the stage micrometer was first imported into PowerPoint; and the grey background, as seen in Figure 4.1, was removed. The parasite egg image of interest was also imported to PowerPoint. An example digital image of a *Paragonimus westermani* parasite eggs is shown in

Figure 4.2. Both images maintained the proper image ratio to ensure correct distance measurements. The image of the micrometer and the parasite egg image of interest were overlaid. The measurements of length and width were directly recorded. The summary of both types of manual approach techniques (pre-image capture and post-image capture) are compared in Appendix B. The two techniques were determined to have the same outcomes based on the statistical results. In both techniques, the length and width were measured as Figure 4.3 shows.

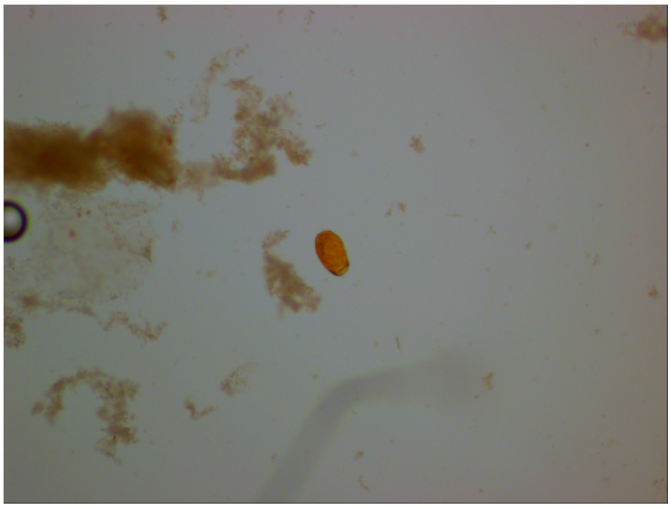

Figure 4.2 *Paragonimus westermani* example of a digital parasite egg image

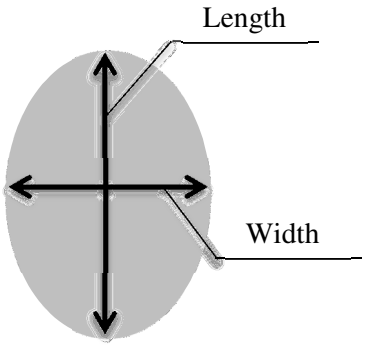

Figure 4.3 Schematic to show how the length and width measurements were taken

The manual measurements for both techniques only took into account length (longest segment) and width (perpendicular to the length) of the parasite eggs. For *Paragonimus westermani,* there were

34 samples. *Taenia* spp. had 35 samples, and *Ascaris lumbricoides* had 43 samples. The results are summarized by parasite type and can be seen below in Table 4.1, Table 4.2, and Table 4.3.

Table 4.1 *Paragonimus westermani* manual measurements (n=34)

	Mean	Median	Standard Deviation
Length (micrometers)	84.0	83.8	6.25
Width (micrometers)	50.6	50.0	3.80

Table 4.2 *Taenia* spp. manual measurements (n=35)

	Mean	Median	Standard Deviation
Length (micrometers)	39.4	40.0	3.20
Width (micrometers)	33.1	30.0	3.89

Table 4.3 *Ascaris lumbricoides* manual measurements (n=43)

	Mean	Median	Standard Deviation
Length (micrometers)	58.3	60.0	4.42
Width (micrometers)	48.0	47.5	3.16

In order to properly analyze the results, one must take into consideration the widely accepted values of width and length that are expected for the parasite eggs. These expectations, seen in Table 4.4 below, are taken from the Center for Disease Control (CDC) [3]. Gunn and Pitt [4] give the width measurement for *Ascaris lumbricoides*.

Table 4.4 Expected lengths and widths from the Center for Disease Control (CDC)

	Paragonimus westermani	*Taenia* spp.	*Ascaris lumbricoides*
Length (micrometers)	80-120	30-35	45-75
Width (micrometers)	45-70	30-35	35-50

A comparison of the gathered and expected results is plotted and compared for accuracy and relationship. In Figure 4.4, we view a box-and-whisker plot of the *Paragonimus westermani* results. The plot analyzes the data through quartiles. In the figure, the circle relates to the fourth quartile or maximum measured value. The minimum result is shown as the square. It is also referred to as the zero quartile. The median, second quartile, is represented by the triangle. The upper (third) quartile is the top of the box. It is the median of the data's upper half. The lower (second) quartile is the bottom of the box. It represents the median of the data's lower half. Therefore, the box represents the middle 50% of the data sets. The whiskers represent the 25% tails of the data on both sides of the box. The same box-and-whisker plots are used to analyze the *Taenia* spp. and *Ascaris lumbricoides* data in Figure 4.5 and Figure 4.6, respectively.

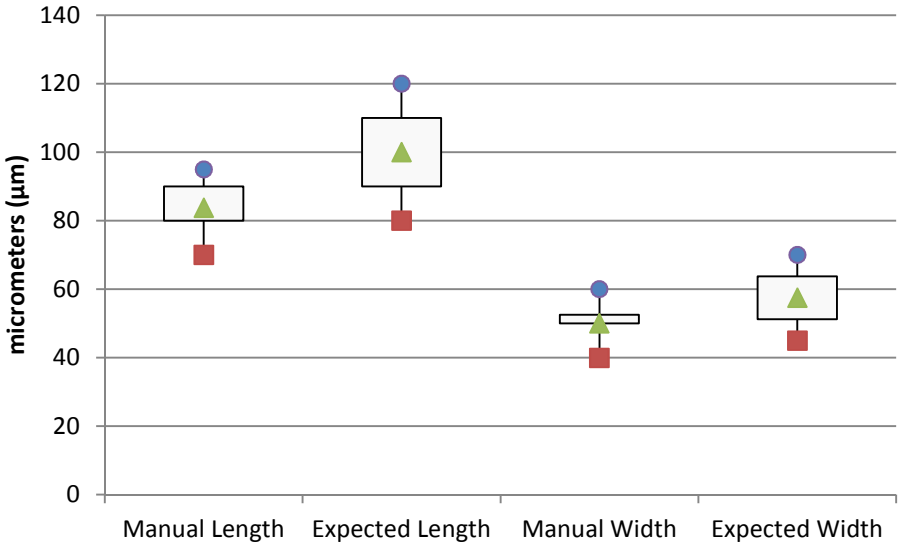

Figure 4.4 Comparison of the expected vs. manual measurements for *Paragonimus westermani.*

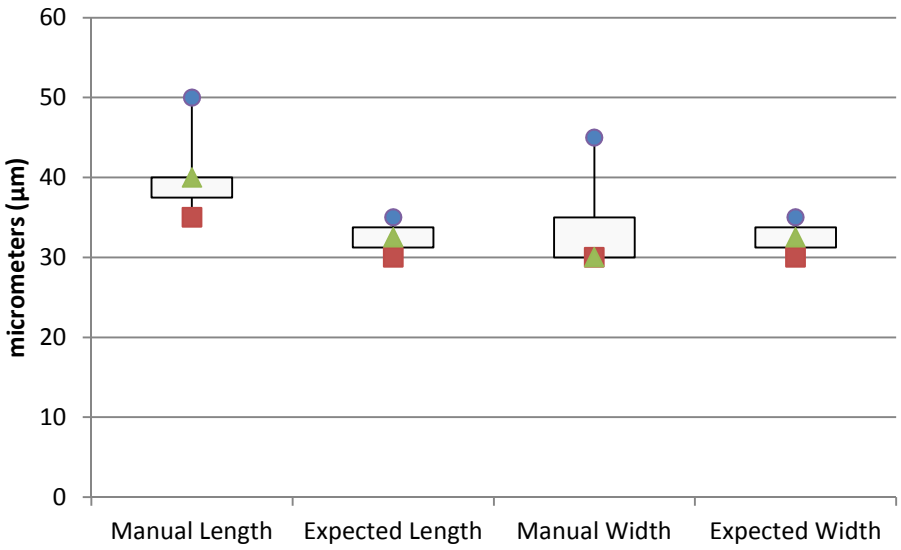

Figure 4.5 Comparison of the expected vs. manual measurements for *Taenia spp.*

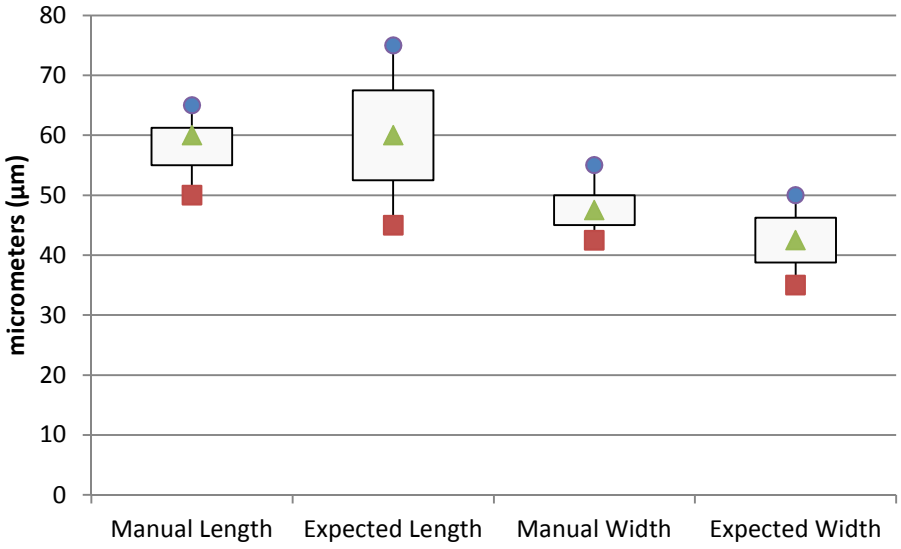

Figure 4.6 Comparison of the expected vs. manual measurements for *Ascaris lumbricoides.*

The results for all three egg measurements show that differentiation by egg types is possible. *Paragonimus westermani* has a median manual length that fell within the range of expected lengths. The minimum manual length, however, fell below the expected range by 10 micrometers. The *Paragonimus westermani* manual width resulted in the same manner; the minimum manual width fell below the expected range by 5 micrometers. The manual median of both the length and width were 16% and 13%,

respectively, below the expected median. The *Taenia* spp. results were the opposite. The manual results were larger than the expected range. For the length of *Taenia* spp., the minimum manual measurement was the only point to fall within the expected range. Overall the median manual length result was 23% larger than the expected length median. For the width of *Taenia* spp., the maximum manual measurement was 10 micrometers above the expected range. The median and minimum manual width, however, were located within the expected range. The *Ascaris lumbricoides* results were the closest to the expected values. The manual length measurements completely fell within the range of expected values. The median manual width was only 12% larger than the median expected width.

There is a measure of uncertainty based on the type of technique used to measure the lengths and widths of the parasite eggs. For the pre-image capture technique, there is an uncertainty of ±1.25 micrometers. For the post-image capture technique, there is an uncertainty of ±5 micrometers. Since the two techniques were combined in the results analysis, the uncertainty for this section will be based on ±5 micrometers.

4.2 LabVIEW™ Approach

Using the LabVIEW™ scripts discussed in Section 3.2.1, the parasite ova were found in each image. For an example the image in Figure 4.2 was processed by the LabVIEW™ script. The result is shown in Figure 4.7.

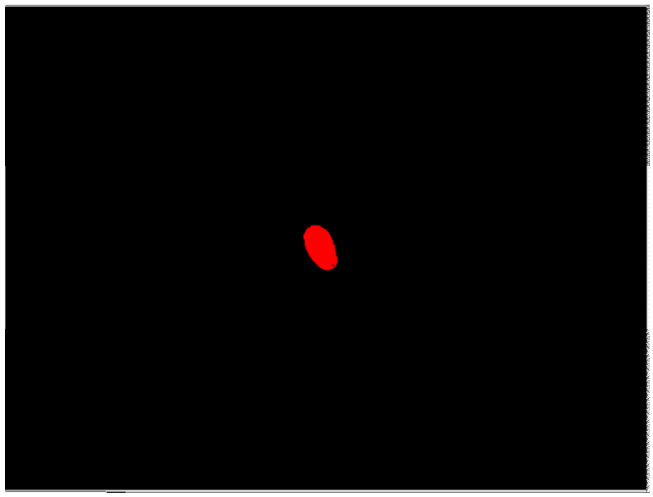

Figure 4.7 *Paragonimus westermani* digital image that has been processed by the LabVIEW™ script

LabVIEW™ has a built-in "Particle Analysis" tool that automatically measures the selected particles of interest (colored red in Figure 4.7) based on the number of pixels. Using this built-in LabVIEW™ analysis, several measurements were taken into consideration. LabVIEW™ gave the measurements for the Area of the Particle, Particle and Holes' Area, Equivalent Rectangular Short Side (Feret), Maximum Feret Diameter, Elongation Factor, and Perimeter. Before the results are summarized, it is important to understand the descriptions for these types of measurements. The Area of the Particle (A) is the number of pixels of the particle converted to the measured area (μm^2). The particle may not have been captured as a whole. When an image is set at a threshold as mentioned in Section 3.1, sometimes holes may appear in the object of interest. This is due to lighting variation either on the particle or ambient lighting in general. Thus, we rely on the Area of Particle and Holes (A_T) to retrieve the complete area. This is done by including the area of any holes that may be located within the particle. To be considered a relevant hole, the pixels of interest must be surrounded by particle pixels on all sides. Max Feret Diameter (F_d) is the distance between the two perimeter points that are located the farthest away from each other. It is also referred to as the maximum caliper length because it is measured comparable to a physical caliper. The Max Feret Diameter is equal to the Length measurement shown in Figure 4.3. RSS-Feret (RF_b) is a calculated measurement. The equation for RSS (Feret) is as follows in Equation 4.1,

$$\boldsymbol{RF_b = \frac{A_{CH}}{F_d}}, \tag{4.1}$$

where A_{CH} is the area of the convex hull and F_d is the max feret diameter. The area of the convex hull represents the area of the smallest convex polygon that can encapsulate the particle of interest. It is important to note that all measurements used the calibration explained in Section 3.2.1 to translate the results from pixels into micrometers. For more information about LabVIEW™'s Particle Measurement, refer to National Instrument's ™ *Particle Measurements* [25]. The results gathered by the LabVIEW™ scripts are listed by parasite type in Table 4.5, Table 4.6, and Table 4.7. For *Paragonimus westermani,* there were 25 samples. *Taenia* spp. had 26 samples, and *Ascaris lumbricoides* had 38 samples.

Table 4.5 *Paragonimus westermani* LabVIEW™ scripts measurements (n=25)

	Average	Median	Standard Deviation
Area of Particle and Holes (µm^2)	3540	3360	547
Max Feret Diameter (µm)	90.2	86.9	9.71
Elongation Factor (µm/µm)	2.14	2.14	0.271
RSS (Feret) (µm)	42.5	41	5.72
Perimeter (µm)	243	233	38.6

Table 4.6 *Taenia* spp. LabVIEW™ scripts measurements (n=26)

	Average	Median	Standard Deviation
Area of Particle and Holes (µm^2)	961	935	209
Max Feret Diameter (µm)	40	39.5	3.8
Elongation Factor (µm/µm)	1.557	1.54	0.0923
RSS (Feret) (µm)	25.8	25	2.94
Perimeter (µm)	117.7	114.1	13.64

Table 4.7 *Ascaris lumbricoides* LabVIEW™ scripts measurements (n=38)

	Average	Median	Standard Deviation
Area of Particle and Holes (µm^2)	1136	999	530
Max Feret Diameter (µm)	45.3	42.4	10.09
Elongation Factor (µm/µm)	1.701	1.688	0.189
RSS (Feret) (µm)	26.7	25.8	5.41
Perimeter (µm)	135.8	123.1	33.8

These LabVIEW™ approach results were added to Figure 4.4, Figure 4.5, and Figure 4.6. This will allow a comparison to be made between the manual, expected, and LabVIEW™ script results. The format will remain a box-and-whicker plot to show where the minimum, median, and maximum are located. The results are analyzed by egg type in Figure 4.8, Figure 4.9, and Figure 4.10.

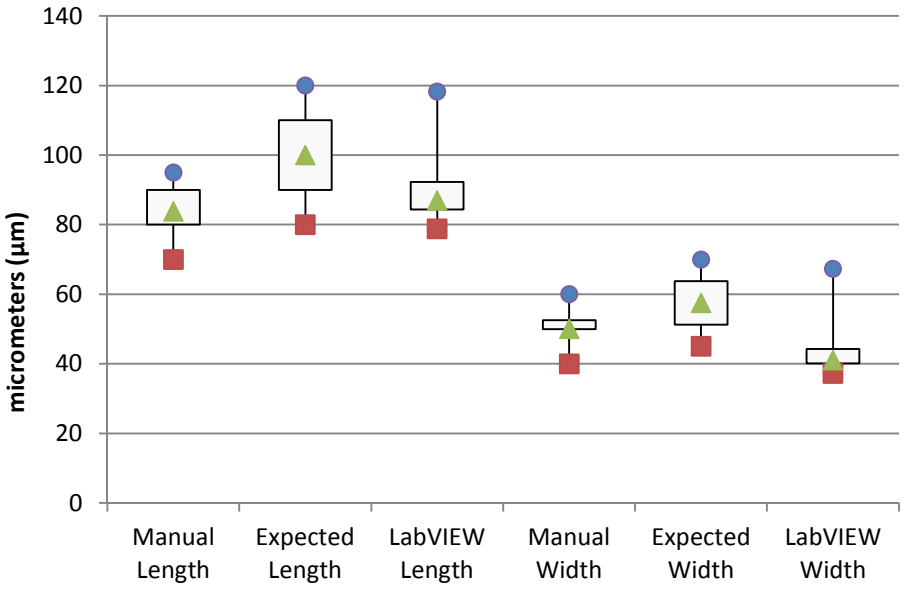

Figure 4.8 *Paragonimus westermani* comparison that includes the LabVIEW™ script results

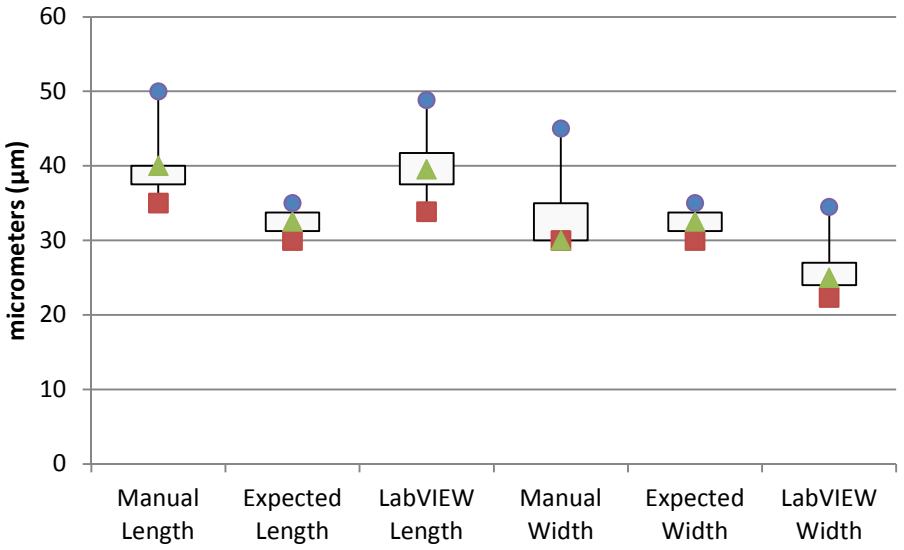

Figure 4.9 *Taenia* spp. comparison that includes the LabVIEW™ script results

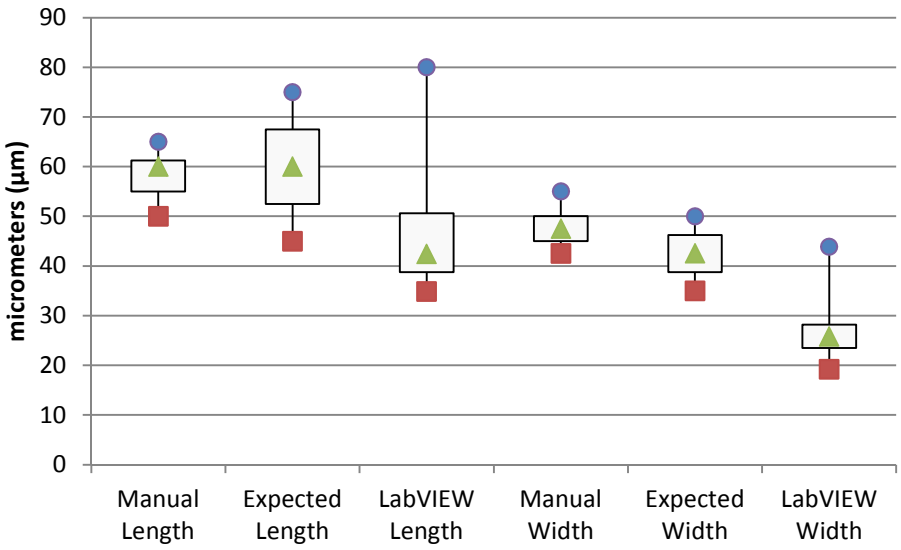

Figure 4.10 *Ascaris lumbricoides* comparison that includes the LabVIEW™ script results

The average LabVIEW™ measurements were lower than the expected measurements. *Paragonimus westermani* has a minimum LabVIEW™ length that fell below the range of expected lengths by 1.2 micrometers. The *Paragonimus westermani* LabVIEW™ width median was also lower than the range. It was 4 micrometers smaller than the expected width minimum. The LabVIEW™ medians of both the length and width were 13% and 29%, respectively, below the expected medians. The *Taenia* spp. length results were quite the opposite. The LabVIEW™ maximum length results were 13.8 micrometers larger than the expected range. For the width of *Taenia* spp., the minimum LabVIEW™ measurement was 7.7 micrometers below the expected range. The median width was also located below the expected range. The *Ascaris lumbricoides* results had the largest variance. The LabVIEW™ length median was 2.6 micrometers below the range of expected values. The median LabVIEW™ width was also 8.2 micrometers below the expected minimum width.

The measure of uncertainty is based on the calibration that was talked about in Section 3.2.1. The uncertainty is calculated to be 0.310 micrometers based on the calculations done in Appendix AError! Reference source not found.. This uncertainty is very small compared to the result values.

4.3 Gray-scale Cross-section

The parasites were also characterized by the intensity of the gray-scale in their cross-section. As discussed previously, the gray-scale intensity is a value between zero (white) and one (black). The MATLAB™ program, which was discussed in Section 3.2.3, was written to determine the gray-scale intensities across the parasite egg's cross-section. An example of the cross-section line drawn can be seen in Figure 4.11 of a *Paragonimus westermani* egg. The gray-scale intensity of the pixels along this line was measured by MATLAB™. The raw data from MATLAB™ is plotted in Figure 4.12.

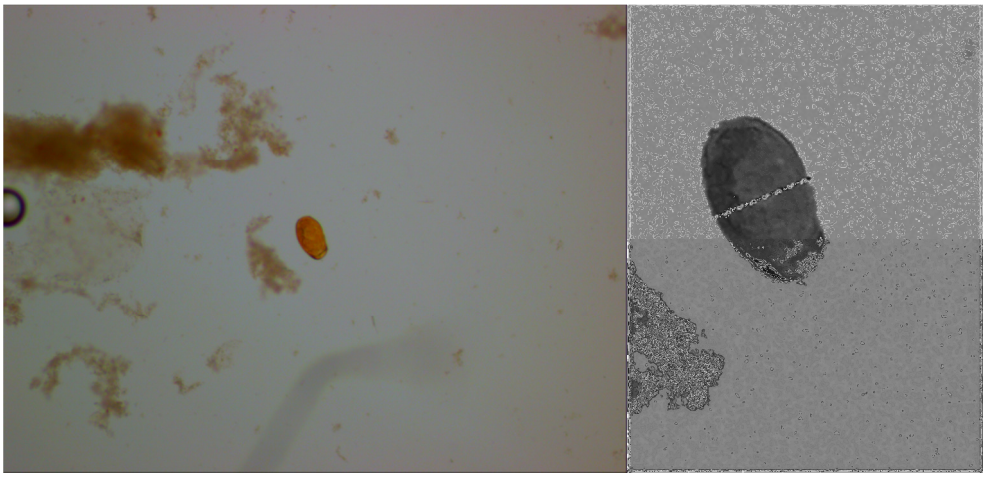

Figure 4.11 The cross-section line was inserted manually to retrieve gray-scale data

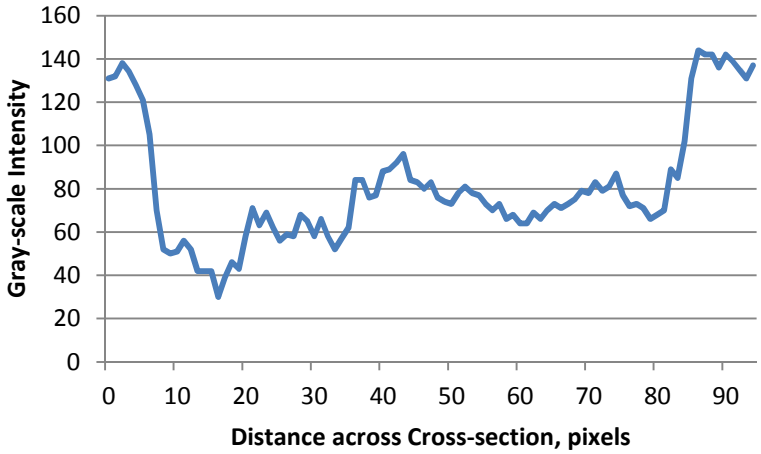

Figure 4.12 Raw data of the gray-scale intensity across *Paragonimus westermani*

Data was analyzed from the graphs. From the data the maximum intensity, minimum intensity, mean intensity, and number of peaks were found. The maximum is the highest point of intensity in the data. The minimum is the lowest point of intensity in the data. These two data points were used to calculate the (Max-Min) Intensity Difference simply by subtracting the minimum intensity from the maximum intensity. The Mean Intensity is the average of all data points. The Number of Peaks reflect the amount of peaks located along the data line. To be considered a peak, the data point to each side must be lower. The Number of Peaks (5) is a similar concept. To be considered a peak, though, the data point must be greater than the five data points to its left and right. These four measurements were used to analyze the gray-scale cross-section images. The results are categorized below in Table 4.8, Table 4.9, and Table 4.10. The MATLAB™ cross-section gray-scale program is shown in Appendix E. There were 30 samples for each parasite type.

Table 4.8 *Paragonimus westermani* cross-section gray-scale measurements (n=30)

	Median	Mean	Standard Deviation
(Max-Min) Intensity Difference	87	89	15.21
Mean Intensity	79.6	81.1	10.79
Number of Peaks	22	21.8	4.38
Number of Peaks (5)	10	10.5	1.737

Table 4.9 *Taenia* spp. cross-section gray-scale measurements (n=30)

	Median	Mean	Standard Deviation
(Max-Min) Intensity Difference	74.5	75.2	10.5
Mean Intensity	79.4	80.1	6.97
Number of Peaks	14.5	14.13	3.17
Number of Peaks (5)	7	7.37	1.326

Table 4.10 *Ascaris lumbricoides* cross-section gray-scale measurements (n=30)

	Median	Mean	Standard Deviation
(Max-Min) Intensity Difference	69	69.3	14.67
Mean Intensity	93.7	95	10.25
Number of Peaks	20	20.4	3.72
Number of Peaks (5)	10	10.03	1.65

A box-and-whisker format is used in Figure 4.13, Figure 4.14, Figure 4.15, and Figure 4.16 to compare the gray-scale cross-section results among the three egg types.

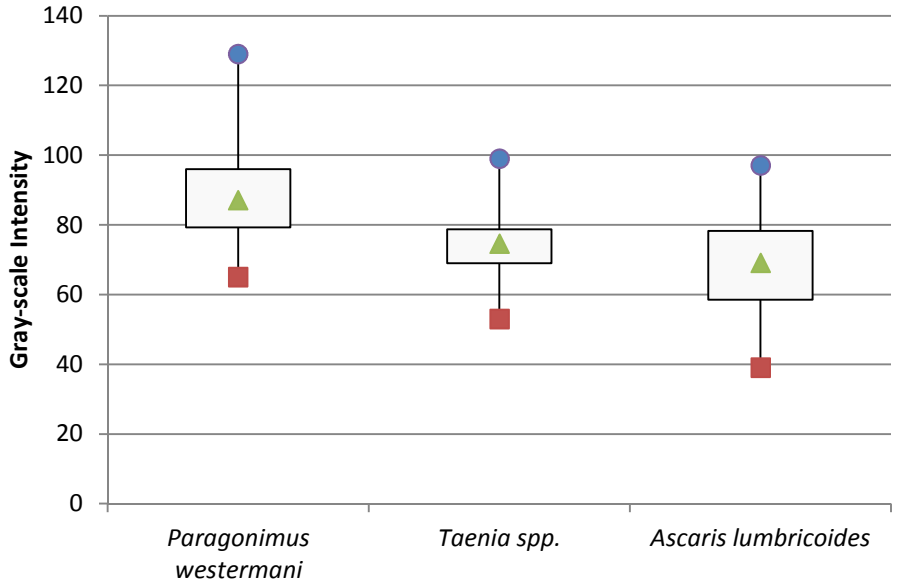

Figure 4.13 (Max-Min) Intensity Difference results are compared by parasite type

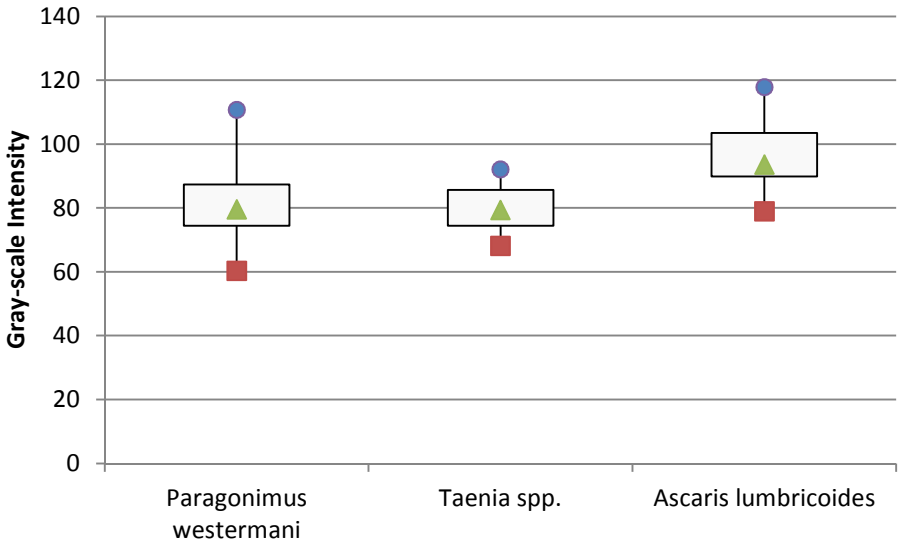

Figure 4.14 Mean intensity results are compared by parasite type

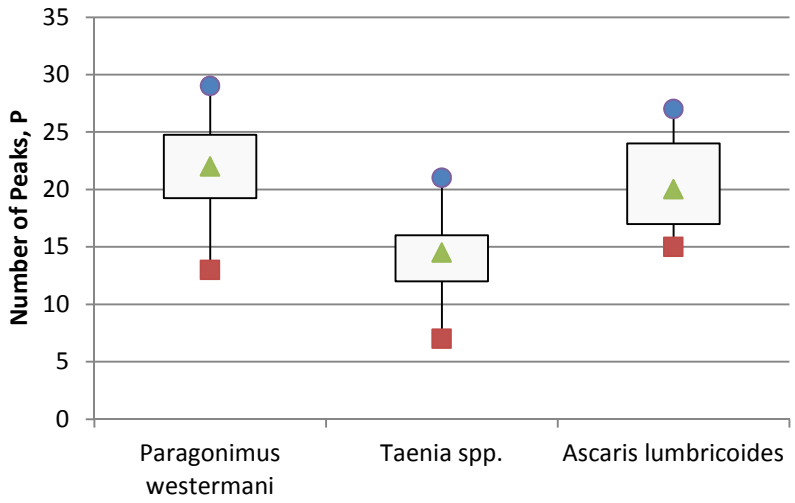

Figure 4.15 Number of Peaks results are compared by parasite type

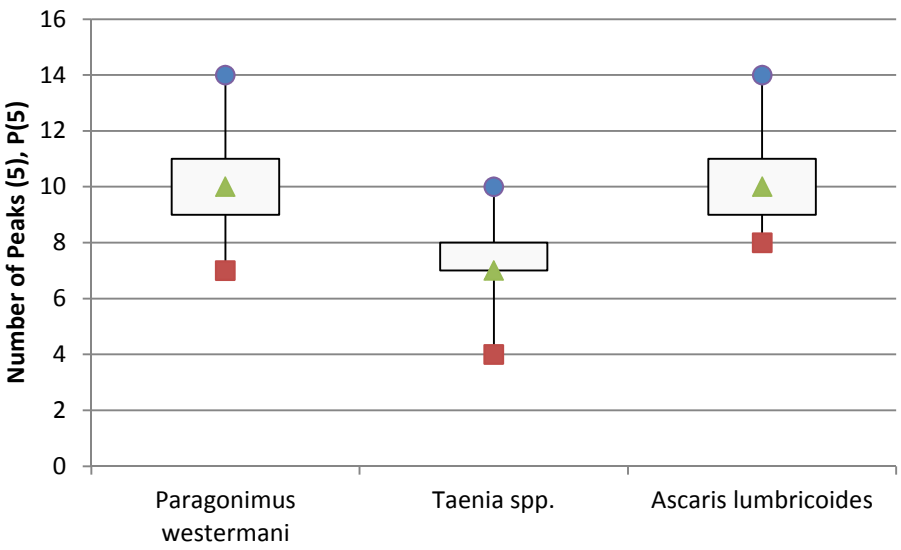

Figure 4.16 Number of Peaks (5) results are compared by parasite type

The gray-scale cross-section data differs from previous data that was analyzed. The previous manual and LabVIEW™ approaches had expected results that were compared to the measurements. Gray-scale cross-section data does not have any expected results to be compared to. Since the intensity

values determine the shade (white to black) of the pixels, we are able to determine some results from the data. A higher mean or median intensity value is essential due to darker pixels. The darker pixels could be a result of thicker egg edges or a complex, dark egg interior. The opposite would be true for a low mean or median intensity value. A low value would be due to thin egg edges and a simple, light egg interior. The number of peaks would be higher if the interior of the egg was complex and varying from light to dark. The gray-scale cross-section is also based on the lighting in the original digital image. Although the lighting did not change drastically, the current analyzed images did not have a process for controlling the lighting conditions. The lighting is dependent upon the number of lights, types of light, and the adjustment of the microscope. If the lighting hardware is set up and the microscope settings are not changed, the lighting conditions will be repeatable.

4.4 Edge Signature

The final characterization is referred to as the edge signature. The signature data was generated and analyzed through a MATLAB™ code. The parasite egg image in Figure 4.17 was processed as explained in Section 3.2.4. The edge signature is plotted linearly, as seen in Figure 4.18.

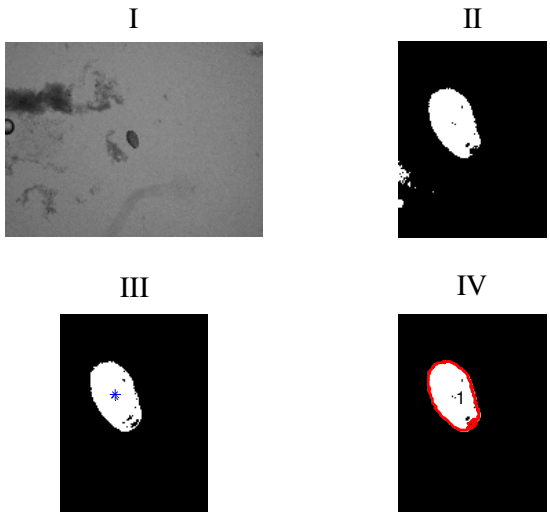

Figure 4.17 MATLAB™ code found the edge signature of a *Paragonimus westermani* egg

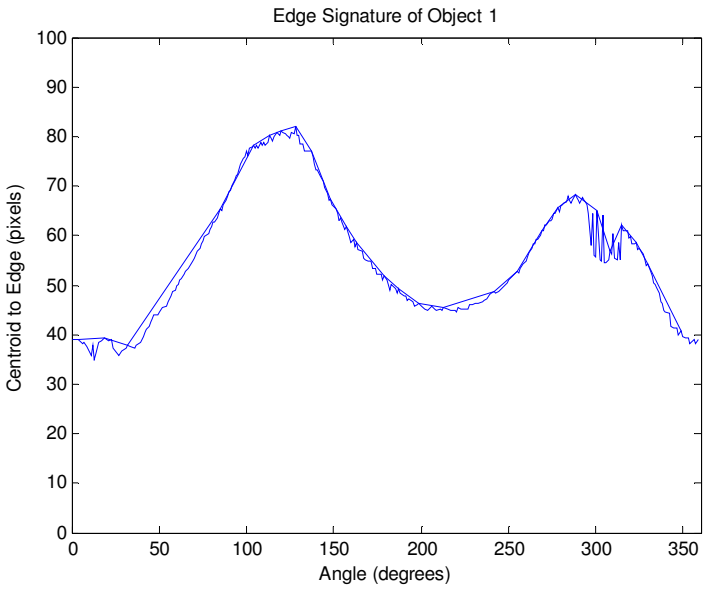

Figure 4.18 Raw MATLAB™ edge signature result of the *Paragonimus westermani* egg

Since the data format is similar to the gray-scale cross-section data format, the same four metrics for measurement will be used. For *Paragonimus westermani,* there were 28 samples. *Taenia* spp. had 28 samples, and *Ascaris lumbricoides* had 36 samples. The summary of results can be seen below in Table 4.11, Table 4.12, and Table 4.13 below.

Table 4.11 *Paragonimus westermani* parasite edge signature results (n=28)

	Median	Mean	Standard Deviation
(Max-Min) Distance Difference (μm)	35.8	38.3	10.71
Mean Distance (μm)	51.8	52	3.22
Number of Peaks	67.5	67.1	7.84
Number of Peaks (5)	31	30.8	2.69

Table 4.12 *Taenia* spp. parasite edge signature results (n=28)

	Median	Mean	Standard Deviation
(Max-Min) Distance Difference (μm)	8.4	9.86	4.17
Mean Distance (μm)	28.4	29.2	3.38
Number of Peaks	50	49.8	6.63
Number of Peaks (5)	20	20.2	2.78

Table 4.13 *Ascaris lumbricoides* parasite edge signature results (n=36)

	Median	Mean	Standard Deviation
(Max-Min) Distance Difference (µm)	20.9	23.7	12.25
Mean Distance (µm)	33	34	6.76
Number of Peaks	62.5	65	13.69
Number of Peaks (5)	27.5	27.4	5.07

A box-and-whisker format in Figure 4.19, Figure 4.20, Figure 4.21, and Figure 4.22 is used to represent the edge signature data for all three parasite eggs.

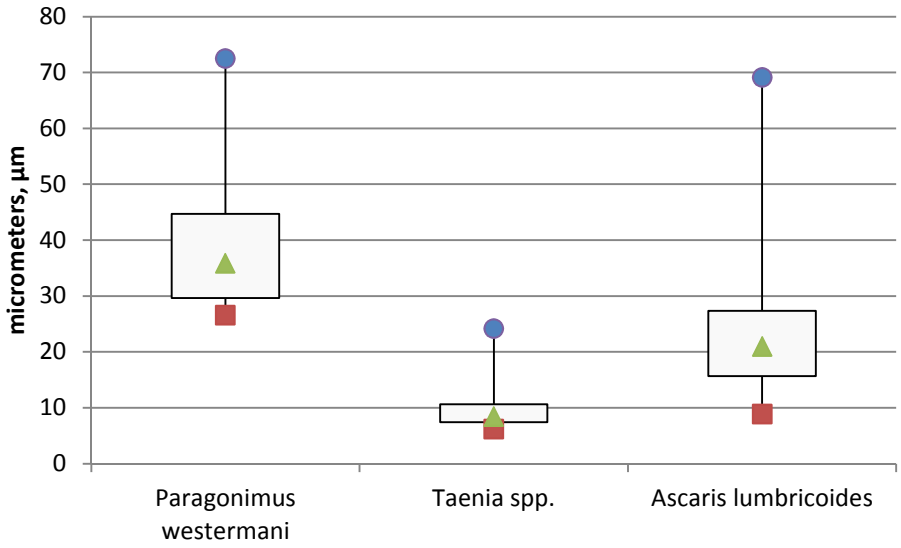

Figure 4.19 (Max-Min) Distance Difference of the parasite edge signatures

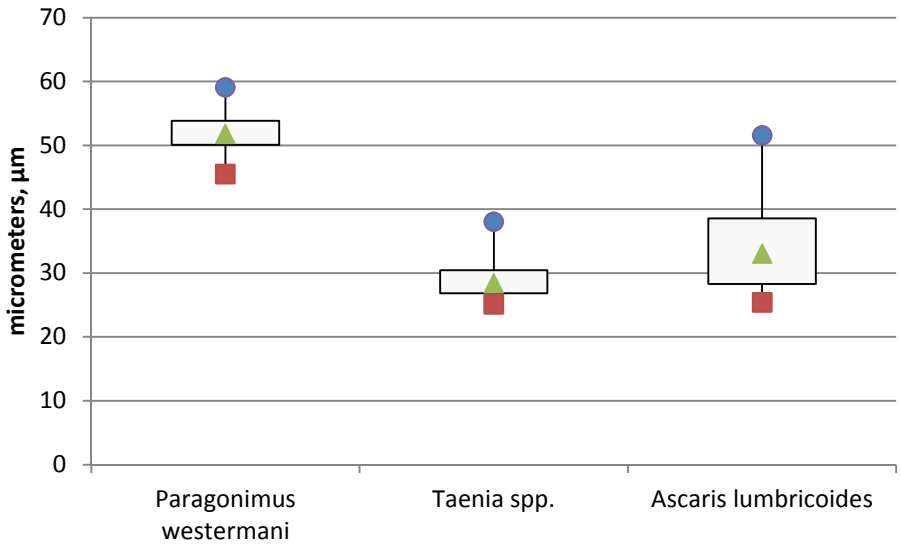

Figure 4.20 Mean Distance of the parasite edge signatures

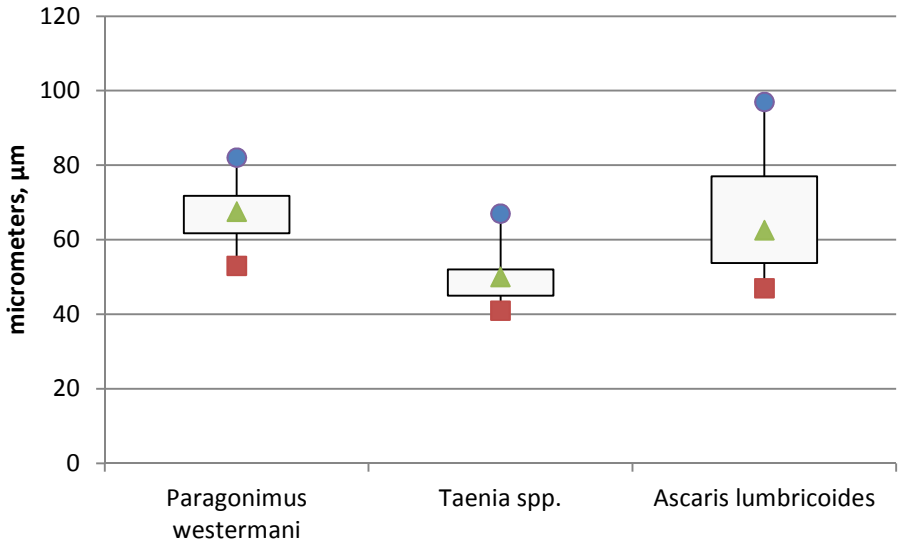

Figure 4.21 Number of Peaks in the parasite edge signatures

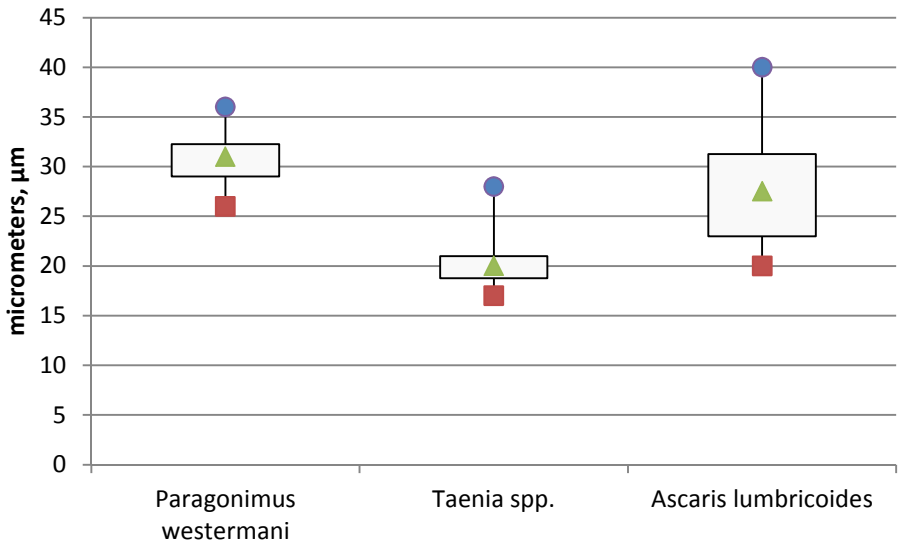

Figure 4.22 Number of Peaks (5) in the parasite edge signatures

The edge signature results are similar to the gray-scale cross-section data in that, there are no expected results to which the achieved results can be compared. The edge signature is able to give information about the size and shape of the parasite egg. The (Max-Min) Difference is proportional to the roundness of the parasite egg. The higher the (Max-Min) Difference the more oval the parasite egg is in shape. The smaller the (Max-Min) Difference the more circular the parasite egg is in shape. A higher mean distance value is due to a larger parasite egg. The smaller mean distance value is from a small parasite egg. If a uniform parasite egg edge was discovered, the number of peaks would be low. The number of peaks would be higher if the edge of the parasite egg had some unaccounted for pixels, due to thresholding error. If threshold value is not correctly calculated, some of the parasite edge can be removed. These missing pixels can account for a larger number of peaks in the data, a smaller (Max-Min) Difference, and a smaller mean distance value.

Chapter 5 Analysis of Results

In the previous chapter, the results of fifteen measurements based on four different approaches were presented. Through the use of statistical tests, one can determine which of these measurements are effective at differentiating between eggs of different parasites. In this chapter, statistical methods will be used to process the data discussed in Chapter 4. Two types of statistical tests will be used in series: single factor global Analysis of Variance (ANOVA) test and Multiple Comparison tests. There are three potential outcomes once statistical tests are applied to the data. First, a given measurement may not be able to distinguish between any of the three parasite eggs. Thus, all three measurements are deemed to be statistically similar. Second, the measurement may only be able to distinguish between some, but not all, of the parasite eggs. Two parasite eggs could be classified as significantly different, but the third parasite egg may be too similar to differentiate. This type of statistical result will still be useful and could be used in conjunction with other measurements. By using two or more measurement results in conjunction, one may be able to differentiate between all three parasite eggs. Finally, there will be measurement results where statistical results show that the eggs of all three species of parasites can be differentiated. These results will be the most helpful and effective at differentiating between the eggs.

5.1 Statistical Results

The single factor Analysis of Variance (ANOVA) test was used to compare the eggs of three species based on the results from fifteen measurements discussed in Chapter 4. In order for an ANOVA test to be conducted, there must be at least three variables. These variables are represented by the three parasite eggs. The term "single factor" from the single factor ANOVA test alludes that the parasite eggs will be analyzed by one factor (measurement) at a time. The single factor ANOVA test does not require the number of samples per factor (measurement) to be the same for each egg species. The single factor ANOVA is separated into the following two parts: global test and multiple comparison test.

The first part of the ANOVA test is called the single factor global ANOVA test. The single factor global ANOVA test determines if all of the mean measurement values are equal. The p-value, which represents probability, is a resultant of the global ANOVA test. If the resulting p-value is less than 0.0001, there is a low probability that the mean measurements of the parasite egg are equal. This also determines that at least one pairing of parasite eggs is statistically different. If the global ANOVA test has results that are deemed statistically different (p-value < 0.0001), then separate multiple comparison tests are run to determine which parasite eggs for that particular factor are statistically different. If the results are not statistically significant, the mean measurements of the parasite egg are

determined to be equal. No further analysis would be performed where the mean measurements are equal. For all fifteen measurements, the single factor global ANOVA test was demonstrated to be statistically significant.

The multiple comparison tests take a statistically significant global test and determine which pairs within the three egg types are significantly different. Multiple comparison tests compare the mean values of one factor (measurement) for all three parasite eggs to determine if the means between pairs are equal. When the results show that the parasite eggs have equal means, it is demonstrated that the measurements of eggs of the parasites are not significantly different. When the multiple comparison test results show that the mean values are not equal, the results are proven statistically different. Statistically different results are needed to differentiate between parasite egg types. The single factor ANOVA test results, both global tests and multiple comparison tests, will be presented and summarized.

5.1.1 ANOVA Global Test

The single factor global ANOVA test is used when there are several groups present that are represented by single factor. The single factor represents one particular measurement, which are represented by the mean value. The global test is used to determine if the mean values between groups are equal or unequal for all of the groups. Specifically, in this case, the groups represent the three parasite egg types. Statistically, these are referred to as the null and alternative hypothesis. The null hypothesis for the three parasite eggs states,

$$\boldsymbol{H_o}: \boldsymbol{\mu_1 = \mu_2 = \mu_3}\,, \tag{5.1}$$

where the mean values (μ) are set equal to each other. If the null hypothesis is proven true, the result signifies that all mean measurements between parasite eggs are statistically the same for that particular measurement. The alternative hypothesis for the three parasite eggs states,

$$\boldsymbol{H_a}: \boldsymbol{\mu_a \neq \mu_b}\,, \tag{5.2}$$

where the mean measurement values, μ_a and μ_b, are not equal to each other. The mean measurement values could be from any two of the parasite eggs. The alternative hypothesis does not guarantee that all parasites pairs will meet the "not equal to" condition. It can only signify statistical difference between the groups in at least one case.

In order to run the global ANOVA test, one measurement (factor) will be selected. There are a total of fifteen individual measurements that will be statistically tested. The measurements were collected in the four approaches (Manual, LabVIEW™, Gray-scale Intensities, and Edge Signature). In the manual approach, the Length and Width of each egg was evaluated. In the LabVIEW™ approach, the Area, RSS (Feret), Maximum Feret Diameter, Elongation Factor, and Perimeter were captured. The

gray-scale intensity cross-section approach evaluated the (Max-Min) Intensity Difference, Mean Intensity, Number of Peaks, and Number of Peaks (5). Finally, in the edge signature approach the (Max-Min) Distance Difference, Mean Distance, Number of Peaks, and Number of Peaks (5) were analyzed.

For each of these measurements, the single-factor global ANOVA test is run. The results are calculated in Excel Spreadsheet application (Microsoft™, Redmond, WA) by the single factor ANOVA test, which is found in Excel's data analysis tools. The global test gives the following results: mean values, variance (σ^2), sum of squares between-group variability (SSB), sum of squares within-group variability (SSW), degree of freedom between groups (df_b), degree of freedom within groups (df_w), mean square between groups (MSB), mean square within groups (MSW), test statistic (F), probability ($p - value$), and critical test statistic (F_{crit}). The variance is calculated for each group of parasites by the equation,

$$\boldsymbol{\sigma^2 = \frac{\sum_i (y_i - \bar{y})^2}{n - 1}}, \tag{5.3}$$

where y_i is the value of the ith sample, $\bar{y}$ is the mean, and n is the number of samples. The variance is used to calculate the standard deviation. Standard deviation (σ) is the square root of variance. The sum of squares between-group variability (SSB) is a value that represents the between-sample variability. The equation for SSB is,

$$\mathbf{SSB} = \sum_{i=1}^{t} \boldsymbol{n_i (\bar{y}_i - \bar{y}_{..})^2}, \tag{5.4}$$

where t is the number of groups, n_i is the number of samples, $\bar{y}_i$ is the sample mean, $\bar{y}_{..}$ is the overall mean. The SSB value compares the variability of the sample mean to the overall mean. The overall mean is represented by the equation,

$$\bar{\boldsymbol{y}}_{..} = \frac{\sum_{i=1}^{n_i} \sum_{j=1}^{n_j} \mathbf{y_{ij}}}{\boldsymbol{n_T}}, \tag{5.5}$$

where y_{ij} is the value of the jth sample of group i and n_T is the total number of samples. The sum of squares within-group variability (SSW) is a value that represents the within-sample variability. The equation for SSW is,

$$\mathbf{SSW} = \sum_{i=1}^{t} \sum_{j=1}^{n_i} \boldsymbol{(y_{ij} - \bar{y}_{i\cdot})^2}, \tag{5.6}$$

where y_{ij} is the value of the jth sample of group i and $\bar{y}_i$ is the sample mean for group i. The SSW finds the variability of y_{ij} with respect to the mean $\bar{y}_i$. The total sum of squares, TSS, can be calculated from the results of SSB and SSW. The equation for TSS is,

$$\mathbf{TSS} = \sum_{i=1}^{t}\sum_{j=1}^{n_i}(\boldsymbol{y_{ij}} - \bar{\boldsymbol{y}}_{..})^2 = \mathbf{SSB} + \mathbf{SSW}, \quad 5.7$$

where the within-sample sum of squares is added to the between-sample sum of squares. The total sum of squares gives a result that shows the variability both between and within-sample variance. Other results from the ANOVA tests are the degrees of freedom. There is a calculated degree of freedom between samples, within samples, and for the total. The equation to calculate the degree of freedom between samples (df_b) is,

$$\boldsymbol{df_b} = \mathbf{t} - \mathbf{1}\,, \quad 5.8$$

where t is the number of groups. More specifically, there are three groups of parasite eggs. The equation to calculate the degree of freedom within samples (df_w) is,

$$\boldsymbol{df_w} = \boldsymbol{n_T} - \mathbf{t}\,, \quad 5.9$$

where n_T is the total number of samples for all three groups of parasites. The degrees of freedom of the total (df_t) is calculated by,

$$\boldsymbol{df_t} = \boldsymbol{n_T} - \mathbf{1} = \boldsymbol{df_b} + \boldsymbol{df_w}\,, \quad 5.10$$

where the degrees of freedom between and within samples can be added together. Other results of the ANOVA test are the mean squares. The mean square between samples (MSB) is calculated by the equation,

$$\mathbf{MSB} = \frac{\mathbf{SSB}}{\boldsymbol{t} - \mathbf{1}} = \frac{\mathbf{SSB}}{\boldsymbol{df_b}}, \quad 5.11$$

where the sum of squares between samples is divided by the degree of freedom between samples. The mean square within samples (MSW) is calculated by the equation,

$$\mathbf{MSW} = \frac{\mathbf{SSW}}{\boldsymbol{n_o} - \boldsymbol{t}} = \frac{\mathbf{SSW}}{\boldsymbol{df_w}}, \quad 5.12$$

where the sum of squares within samples is divided by the degree of freedom within samples. The main purpose for calculating the mean squares, for the scope of this project, is to calculate the test statistic. The test statistic (F) equation is,

$$\mathbf{F} = \frac{\mathbf{MSB}}{\mathbf{MSW}}, \quad 5.13$$

where the mean square between samples is divided by the mean square within samples. The test statistic is used to compare the two mean square values and can be thought of as a "signal to noise" ratio. The value is used to determine if the null hypothesis is accepted or rejected.

In Table 5.1, the global test results for the manual approach are displayed. In the table, "σ" stands for standard deviation. The P-value is the statistical term that proves significance. If the P-value is less than 0.001, the groups are statistically different from each other. If the P-value is greater than

0.001, the groups cannot be distinguished between each other. In Table 5.1, both P-values are significantly less than 0.001. This indicates that, for each measurement type, at least one pair of the parasite types is significantly different. Significant difference indicates that the means of the parasite egg measurements are not equal. Once significant difference is idicated, the parasite eggs can be differentiated and identified based on the raw data results. The LabVIEW™, gray-scale, and edge signature statistical results can also be seen in Table 5.2, Table 5.3, and Table 5.4 respectively, for all of the single-factor ANOVA tests, α=0.05. The Greek letter "α" represents the significance level of the test. This significance level, α, is the probability that the null hypothesis will be reject when, in fact, it should be accepted.

Table 5.1 The manual approach measurements were found to be significant.

	Paragonimus westermani n=34		*Taenia* spp. n=35		*Ascaris lumbricoides* n=43		
	Mean	(± σ)	Mean	(± σ)	Mean	(± σ)	*P-value*
Length (µm)	84.0	(6.25)	39.4	(3.20)	58.3	(4.42)	6.10E-65
Width (µm)	50.6	(3.80)	33.1	(3.89)	48.0	(3.16)	1.004E-40

Table 5.2 The LabVIEW™ test measurements were all found to be significant.

	Paragonimus westermani n=25		*Taenia* spp. n=26		*Ascaris lumbricoides n=38*		
	Mean	(± σ)	Mean	(± σ)	Mean	(± σ)	*P-value*
Area of Particle and Holes (µm^2)	3430	(581)	961	(209)	1136	(530)	1.13E-34
Max Feret Diameter (µm)	90.2	(9.71)	40	(3.80)	45.3	(10.09)	7.40E-38
Elongation Factor (µm/µm)	2.14	(0.271)	1.557	(0.0923)	1.701	(0.1890)	1.50E-17
RSS (Feret) (µm)	42.5	(5.72)	25.8	(2.94)	26.7	(5.41)	9.38E-23
Perimeter (µm)	243	(38.6)	117.7	(13.64)	135.8	(33.8)	2.49E-26

Table 5.3 The gray-scale cross-section results were all found to be significant.

	Paragonimus westermani n=30		*Taenia* spp. n=30		*Ascaris lumbricoides n=30*		
	Mean	(± σ)	Mean	(± σ)	Mean	(± σ)	*P-value*
(Max-Min) Intensity Difference	89	(15.21)	75.2	(10.50)	69.3	(14.67)	7.67E-07
Mean Intensity	81.1	(10.79)	80.1	(6.97)	95	(10.25)	9.96E-09
Number of Peaks	21.8	(4.67)	14.13	(3.76)	20.4	(4.52)	7.79E-12
Number of Peaks (5)	10.5	(1.737)	7.37	(1.326)	10.03	(1.650)	1.04E-11

Table 5.4 The edge signature results all proved to be statistically significant.

	Paragonimus westermani n=28		*Taenia* spp. n=28		*Ascaris lumbricoides* n=36		
	Mean	(± σ)	Mean	(± σ)	Mean	(± σ)	*P-value*
(Max-Min) Distance Difference	38.3	(10.71)	9.86	(4.17)	23.7	(12.25)	1.19E-16
Mean Distance	52	(3.22)	29.2	(3.38)	34	(6.76)	6.33E-31
Number of Peaks	67.1	(7.84)	49.8	(6.63)	65	(13.69)	3.61E-09
Number of Peaks (5)	30.8	(2.69)	20.2	(2.78)	27.4	(5.07)	1.81E-16

All of the individual factors (measurements) have a p-value less than 0.001. This indicates that for each factor at least one parasite egg type is statistically different. In Table 5.5, the values for the test statistic, F, and the critical test statistic, F_{crit} , are compared. The critical test statistic is found in the t-Test tables [26]. The test statistic comparison is another way to see if the null hypothesis is rejected of accepted. If the F value is greater than F_{crit} , the at least one pair of values is statistically different. If F is less than F_{crit} , all pairs all statistically the same. These results are another way to determine statistical significance between parasite egg measurements.

Table 5.5 Comparison of F and F_{crit} of all measurement factors determines statistical differences

	$p-value$	F	F_{crit}
Manual Approach			
Length	6.1 E-65	767	3.08
Width	1.004 E-40	243	3.08
LabVIEW™ Approach			
Area	1.130 E-34	222	3.10
Elongation	1.499 E-17	62.9	3.10
Max Feret	7.40 E-38	271	3.10
Perimeter	2.49 E-26	126.4	3.10
RSS	9.38 E-23	96.9	3.10
Gray-scale Cross-section			
(Max-Min) Difference	7.67 E-07	16.6	3.10
Mean Intensity	9.96 E-09	22.9	3.10
Number of Peaks	7.79 E-12	34.8	3.10
Number of Peaks (5)	1.037 E-11	34.3	3.10
Edge Signature			
(Max-Min) Difference	1.192 E-16	56.9	3.10
Mean Distance	6.33 E-31	167.8	3.10
Number of Peaks	3.61 E-09	24.4	3.10
Number of Peaks (5)	1.806 E-16	56.0	3.10

In this section, the single factor global ANOVA test was explained. All relating equations were discussed in detail. The single factor global ANOVA test was then completed through Excel, and the results were analyzed. In all measurement cases, the parasite eggs between types are shown to be statistically different. This was confirmed by the P-value analysis and the comparison between F and F_{crit} . The next step of the single factor ANOVA test is a multiple comparison test. This statistical comparison between the parasite egg measurement means will determine which parasite egg types can be differentiated.

5.1.2 Multiple Comparisons

In all of the single factor global ANOVA test results, it was determined that every measurement was, in fact, significant. Thus, multiple comparisons are required to gain additional details of which parasite egg measurements are statistically significant. A multiple comparison test is only applied to the results in the global test that are deemed to be statistically different. It takes the measurement that was

determined to be statistically different in the global test and now compares two parasite's mean measurement values to see which ones are statistically different. For example, since there are three parasite egg types being compared, there will be three comparisons made in each multiple comparison test. One multiple comparison test will be run for each type of measurement, in this case the number is 15. Therefore, the multiple comparison tests will be run a total of 45 times. There are several ways to compute the multiple comparison testing. The following are several types of multiple comparison tests: Scheffe, Bonferroni, Tukey, SNK, and LSD. The Tukey method was chosen to complete the multiple comparison tests because the numbers of samples within the parasite groups are not the same, and the Tukey method allows for this.

There are several inputs and outputs that are involved in Tukey's multiple comparison testing. The inputs include the following: mean values for each group (μ); number of subjects for each group (n_i); number of groups, which are the three parasite types; mean square within groups (MSW), which is an output of the global ANOVA test; and the level of significance (α), which is 0.05. The output is the Q-test statistic (Q) and the Q-critical value (Q_{crit}). The equation used in Tukey's multiple comparison tests to find the Q-test statistic is the following,

$$\mathbf{Q} = \frac{\boldsymbol{\mu_1 - \mu_2}}{\sqrt{\mathbf{2}}} \sqrt{\boldsymbol{s_W}^2 \left(\frac{\mathbf{1}}{\boldsymbol{n_1}} + \frac{\mathbf{1}}{\boldsymbol{n_2}}\right)}, \tag{5.14}$$

where μ_1 is the mean of the measurement for parasite 1, μ_2 is the mean of the measurement for parasite 2, ${s_W}^2$ is the mean square within, n_1 is the number of samples for parasite 1, and n_2 is the number of samples for parasite 2. The multiple comparison tests are done by calculating the Q-test statistic so that it can be compared to the Q-critical value, which is located in the Studentized range table [26]. If the Q-critical value is larger, the null hypothesis is accepted. If the Q-critical value is smaller, the null hypothesis is rejected. The null hypothesis states that $\mu_1 = \mu_2$.

To find the Q-critical value in the studentized range table, the degree of freedom (df) and number of groups must be known. The degree of freedom equation is,

$$\boldsymbol{df} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{k}} \boldsymbol{n_i} - \boldsymbol{k}, \tag{5.15}$$

where n_i is the number of samples for parasite type "i" and k is the number of groups. For all calculations in this experiment, $k = 3$. The number of samples for each parasite type fluctuated by the measurement type. This fluctuation may cause the Q-critical value to vary slightly. The values for the number of samples by measurement types are seen in Table 5.6.

Table 5.6 The number of samples varied by measurement type and parasite.

	Manual	LabVIEW™	Cross-section gray-scale	Edge Signature
Paragonimus westermani (n_p)	34	25	30	28
Taenia spp. (n_t)	35	26	30	28
Ascaris lumbricoides (n_a)	43	38	30	36

Since each of the measurements was determined to be significant in the single factor global ANOVA testing, as shown in Section 5.1.1, the multiple comparison tests will be applied to all measurements. The Manual multiple comparison test results can be seen in Table 5.7 below. The LabVIEW™, gray-scale cross-section, and edge signature multiple comparison test results follow in Table 5.8, Table 5.9, and Table 5.10.

Table 5.7 Manual data results in the multiple comparison tests

	Decision Rule: Reject Ho if the Q-test ≥ Q-critical value α=0.05				
Length	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	55.2 ≥ 3.37	TRUE	Yes
	2	Ho: μp = μa	33.6 ≥ 3.37	TRUE	Yes
	3	Ho: μt = μa	24.8 ≥ 3.37	TRUE	Yes
Width	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	28.4 ≥ 3.37	TRUE	Yes
	2	Ho: μp = μa	4.43 ≥ 3.37	TRUE	Yes
	3	Ho: μt = μa	25.9 ≥ 3.37	TRUE	Yes

Table 5.8 LabVIEW™ data results in the multiple comparison tests

	Decision Rule: Reject Ho if the Q-test ≥ Q-critical value α=0.05				
Area	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	26.13 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	26.98 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	2.08 ≥ 3.38	FALSE	No
Elongation Ratio	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	15.11 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	12.67 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	4.16 ≥ 3.38	TRUE	Yes
Max Feret Diameter	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	29.43 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	29.24 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	3.50 ≥ 3.38	TRUE	Yes
Perimeter	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	20.49 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	19.50 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	3.30 ≥ 3.38	FALSE	No
RSS (Feret)	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	17.13 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	17.96 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	1.09 ≥ 3.38	FALSE	No

Table 5.9 Gray-scale cross-section data results in the multiple comparison tests

	Decision Rule: Reject Ho if the Q-test ≥ Q-critical value α=0.05				
(Max-Min) Difference	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	5.56 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	7.95 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	2.39 ≥ 3.38	FALSE	No
Mean	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	0.56 ≥ 3.38	FALSE	No
	2	Ho: μp = μa	8.00 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	8.56 ≥ 3.38	TRUE	Yes
Number of Peaks	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	11.08 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	2.02 ≥ 3.38	FALSE	No
	3	Ho: μt = μa	9.06 ≥ 3.38	TRUE	Yes
Number of Peaks (5)	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	10.86 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	1.62 ≥ 3.38	FALSE	No
	3	Ho: μt = μa	9.24 ≥ 3.38	TRUE	Yes

Table 5.10 Edge signature data results in the multiple comparison tests

	Decision Rule: Reject Ho if the Q-test ≥ Q-critical value α=0.05				
(Max-Min) Difference	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	15.09 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	8.28 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	7.85 ≥ 3.38	TRUE	Yes
Mean	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	24.35 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	20.57 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	5.47 ≥ 3.38	TRUE	Yes
Number of Peaks	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	8.92 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	1.14 ≥ 3.38	FALSE	No
	3	Ho: μt = μa	8.39 ≥ 3.38	TRUE	Yes
Number of Peaks (5)	Pair	Null Hypothesis	Is the following inequality true?		Reject Ho?
	1	Ho: μp = μt	14.62 ≥ 3.38	TRUE	Yes
	2	Ho: μp = μa	5.06 ≥ 3.38	TRUE	Yes
	3	Ho: μt = μa	10.56 ≥ 3.38	TRUE	Yes

The results of the above multiple comparison tests are now analyzed with the main goal of accurately distinguishing between the parasite types. It is important to note which multiple comparison tests failed to reject the null hypotheses. Tukey's multiple comparison test results show that the mean measurement value for the majority of the parasites is significantly different. There were no manual measurement comparisons that failed. Three of the LabVIEW™ multiple comparison tests failed. They are as follows: Area (*Taenia* spp. vs. *Ascaris lumbricoides*), Perimeter (*Taenia* spp. vs. *Ascaris lumbricoides*), and RSS (Feret) (*Taenia* spp. vs. *Ascaris lumbricoides*). Four of the gray-scale cross-section multiple comparison tests failed. They are as follows: (Max-Min) Intensity Difference (*Taenia* spp. vs. *Ascaris lumbricoides*), Mean Intensity (*Paragonimus westermani* vs. *Taenia* spp.), Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*), and Number of Peaks (5) (*Paragonimus westermani* vs. *Ascaris lumbricoides*). Only one of the edge signature multiple comparison tests failed, Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*).

5.2 Summary

In this chapter the single factor ANOVA test was explained, demonstrated, and discussed in two parts: the single-factor global ANOVA test and the multiple comparison test. The global test results showed that for all measurements at least one pairing of parasite egg types would result in a statistical difference of value. From these results, the multiple comparison tests were run to determine which parasite eggs within one specific measurement failed to be statistically different. There were no manual measurement comparisons that failed to be statistically significant. Three of the LabVIEW™ multiple comparison tests failed. They are as follows: Area (*Taenia* spp. vs. *Ascaris lumbricoides*), Perimeter (*Taenia* spp. vs. *Ascaris lumbricoides*), and RSS (Feret) (*Taenia* spp. vs. *Ascaris lumbricoides*). Four of the gray-scale cross-section multiple comparison tests failed. They are as follows: (Max-Min) Intensity Difference (*Taenia* spp. vs. *Ascaris lumbricoides*), Mean Intensity (*Paragonimus westermani* vs. *Taenia* spp.), Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*), and Number of Peaks (5) (*Paragonimus westermani* vs. *Ascaris lumbricoides*). Only one of the edge signature's multiple comparison tests failed, Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*). Failure to reject the null hypothesis determines that the means of the measurements are considered equal. Thus, the failed tests cannot be used to distinguish between parasite types. There were 36 comparisons out of the 44 that did not fail to reject the null hypothesis. The measurements where no null hypothesis failures occurred are the following: manual Length, manual Width, LabVIEW™ Elongation Ratio, LabVIEW™ Max Feret Diameter, edge signature (Max-Min) Difference, edge signature Mean, and edge signature Number of Peaks (5). These measurements may be used when attempting to differentiate between parasite eggs from the examined species.

Chapter 6 Discussion and Conclusions

As a final summary, the main contributions of this research will be highlighted. There are several motivations behind this project: to empower more individuals with the tools to diagnose parasitic disease and to provide relief to those suffering from parasitic diseases. The current method of diagnosing a parasite infection is to collect and analyze a fecal sample. Due to the slide preparation techniques, thorough slide analysis, and knowledge of parasite types required; detection has been most effective if the slide is analyzed by a trained individual looking at the microscope slides of the fecal sample with a high-quality microscope. These motivations and current detection methods led to several research goals for this project.

The goal of this research is to forge a path that can eventually lead to an automatic detection and identification method for parasite eggs. This would enable untrained volunteers to provide a helping hand to trained specialist, perform parasite detection themselves, and ultimately improve healthcare for underserved populations. This would also enable the trained specialist to be quicker and more efficient at the sample analysis. In order to accomplish this, there is a need to develop methods for the automatic detection of parasite eggs and the differentiation of their species using various measurements. Therefore, there is a need to understand the image processing steps in order to differentiate the parasite egg from the background material in the digital image. In order to analyze various methods and measurements for parasite detection and determination in digital images, eggs from three parasite species were photographed and captured by a microscope camera. For the three parasite eggs selected, computer programs were created that calculate quantitative descriptors to aid in detection and differentiation between the three species. Through statistical analysis, the efficiency of these descriptors in distinguishing eggs of the three parasites was determined. These statistical results as well as the ability to locate the parasite egg in the image will be summarized for the reader. A glance back at the research steps taken to get to this point will be recalled.

6.1 Overview of Completed Work

In the preparation for this research, current methods for parasite detection and previous research in automatic parasite egg detection were studied. The current method of diagnosing an intestinal parasite infection is to collect and analyze a fecal sample. The fecal sample may be prepared for analysis through several different methods including the use of saline, sugar solution, or a fixative such as Formalin. Once the fecal suspension is created, a microscope slide is prepared. The entire slide must then be analyzed using a microscope. Due to the complexity of the steps required to prepare a fecal sample

suspension and the vast differences between parasite egg types; historically, detection has been most effective if the slide is analyzed by a trained individual examining the microscope slides of the fecal sample with a high-quality microscope.

Several similar research projects on automatic parasite egg detection were found and used as a basis for this work. Three themes were discovered in the previous research on automatic parasite egg detection and identification. They are the following: Hardware and Software, Feature Detection, and Classification. Pre-captured images in JPEG files from Kansas State University were utilized in research done by Dogantekin et. al [5]. For the research contained in this paper, digital images were taken under a typical lab setting with a light-microscope. Sommer [10] used a 400x magnification setting on digital images, even though normal slide analysis magnification is 100x. Therefore, the parasite egg was the sole focus of the image. No background material, in which to differentiate against, was present. In this work, background material was included in the image so that the parasite egg was not the only object in the view. This helped develop the detection codes to operate in a more real-world situation. As in previous research [5-7], MATLAB™ was utilized for processing. Programs for separation code, cross-section gray-scale analysis, and edge signature code were created by the author in MATLAB™. Unlike any previous works, LabVIEW™ Vision Assistant™ was utilized to develop image processing scripts for locating and characterizing the parasite eggs.

Several of the feature detection methods found in previous research were studied and built upon. In Yang, et. al. [7], morphological features were extracted through the use of median filtering, binary thresholding, and segmentation. All three of these processing techniques are used in the research of this paper. Yang's feature extraction method plots angular trajectory of the parasite egg. In this work, the same technique was implemented in MATLAB™ as the Edge Signature. The data that Sommer [10] collected from parasite eggs was for quantification purposes only; there was no attempt made to identify the parasite egg species based on the results. Yang, et. al. [7] discovered methods to discriminate between the parasite eggs and debris in the slide. A goal of the research of this paper was to not only collect quantitative characterization data from parasite eggs, but also statistically demonstrate that this data can distinguish between types of parasite eggs and any background material. To help with the added background material, a method was adopted from Liu, et. al. [13] to separate overlapping or touching objects. Liu, et. al. [13] specifically used this approach to separate liver cells that were touching and overlapping. Since the parasite eggs and fecal matter in the digital images also may be touching or overlapping, a similar technique was applied through the development of the MATLAB™ separation code.

Classification uses stored results of previous parasite egg measurements to identify future parasite eggs. This was the third and final main topic found in the previous research. Yang et. al [7] used the results from the feature extraction measurements to identify the species of the egg. Artificial neural networks were then used to train the computer to identify eggs in future images. Nugent et. al [9] used active learning of the feature extraction results so that future images are processed more successfully. The initial steps for creating a classification method have been applied in this work. Feature extraction results were captured and analyzed statistically. The statistical analysis shows that a classification method, if implemented for future parasite egg results, would be successful in the determination of parasite egg species. Future research would need to be completed to put a classification system into action.

For this work, digital microscopic images were collected for three parasite eggs. Slides were created and analyzed using the sedimentation technique from the three different parasite egg suspensions. The fecal matter sample created the background material within which the parasite eggs would be distinguished. Images were captured at 100x total magnification using the autofocus feature on the camera. The resulting digital images are all 2048 x 1536 pixels.

Before any image processing can take place, calibration must occur. The microscope was calibrated, which allowed for the manual measurement of lengths and widths of the parasite eggs through the eyepiece. At 100x magnification, there were 10 microns for every ocular division in the microscope eyepiece. There was also a calibration of the digital images by using a micrometer and the point coordinate calibration method. Calibration of the digital images allowed for real-life measurements to be calculated from the pixel measurements within an image.

Once the images were calibrated, several processes conditioned the image for parasite identification. Image enhancements were performed through color plane extraction, filtering, and thresholding. In order to transform a color image to a gray-scale image, one color plane must be removed. The analysis of one color plane, also known as color plane extraction, will result in a gray scale image. A color plane (red, blue, or green) was chosen to maximize the contrast in the image for better results. Filtering steps manipulate the gray-scale image in order to separate and define objects. Segmentation extracts features from the image such as edges or blobs. Convolution is a type of filtering that sharpens edges based on kernel size.

Thresholding was applied to the newly-formed gray-scale images to separate background and foreground pixels. Otsu's global approach to thresholding assumes that a bimodal grayscale-intensity distribution is present within the image. Another form of thresholding, adaptive thresholding operates under the pretense that the background is non-uniform and does not fall under the bimodal distribution.

Both global and adaptive thresholding methods, are applied to the digital images within this project. All pixels that are located below the threshold are now considered background pixels. All pixels that are located at or above the threshold are foreground pixels, also known as pixels of interest. The new image created by implementing the threshold is referred to as a binary image.

Morphological operators enhanced the binary image to complete the final step of gathering data for parasite characterization. Morphological operators function by manipulating the shapes of objects in a binary image. There are four main types of operations: dilation, erosion, closing, and opening [18, 23]. Dilation enlarges the foreground pixels. Erosion is a morphological operator that reduces the foreground objects at its edge. Closing is a combination of a dilation operation followed by an erosion operation. Opening is a combination of an erosion operation followed by a dilation operation.

One manual approach and four automated approaches located the parasite eggs and gathered parasite characterization data. The manual approach used manual measurements of the parasite eggs within the digital images. The following are the four automated approaches: LabVIEW™ Vision Assistant™ scripts, MATLAB™ separation code, MATLAB™ cross-section grayscale analysis, and MATLAB™ edge signature analysis. Three LabVIEW™ scripts, one for each parasite egg type, were written within LabVIEW™'s Vision Assistant™. These scripts are summarized in Figure 6.1. Several types of image processing techniques were applied. Through these image processing steps, the location of the parasite egg is determined.

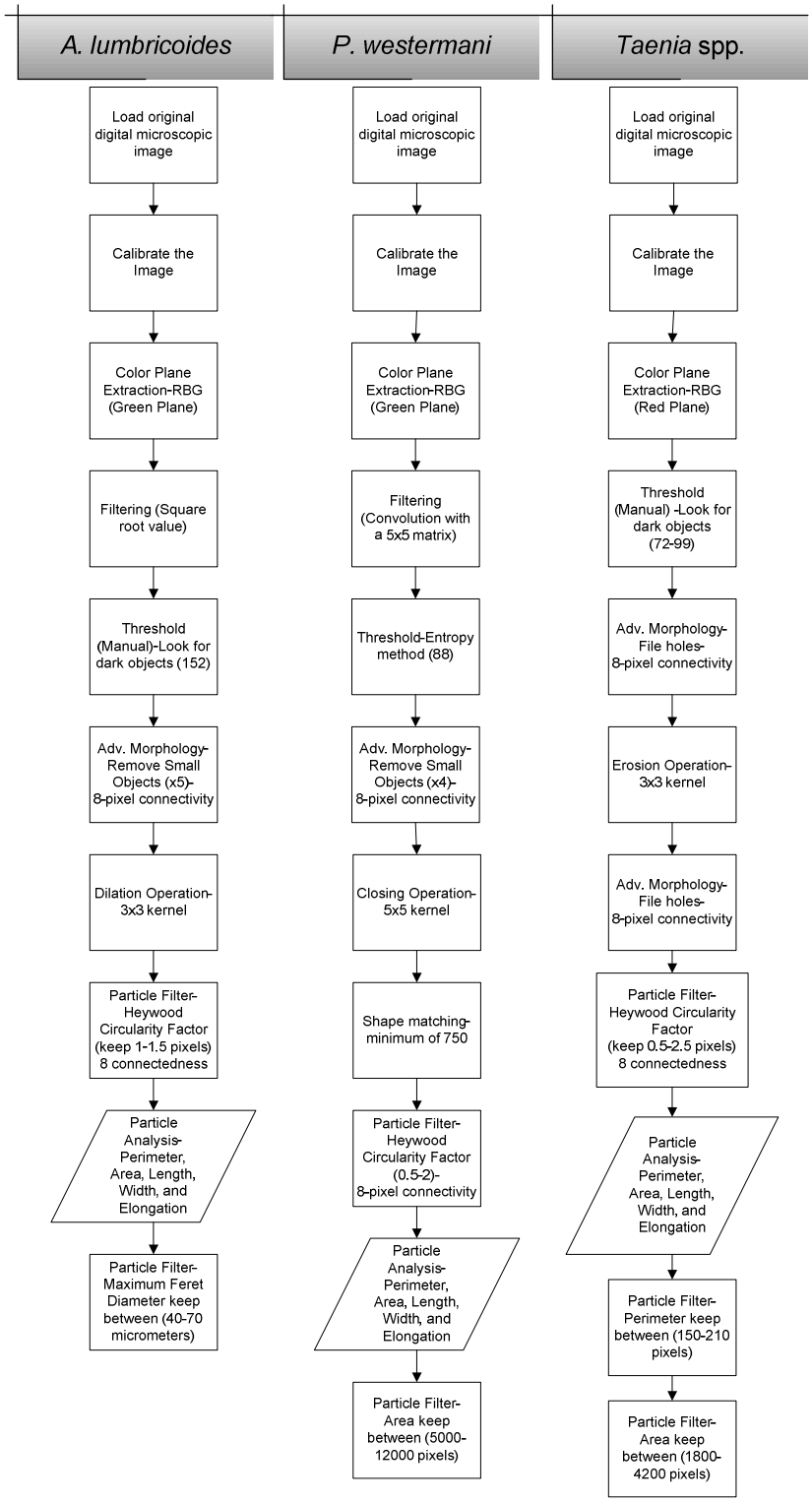

Figure 6.1 Flow diagram of the LabVIEW™ scripts for the three parasite types

A MATLAB™ separation code was also utilized to separate the parasite eggs from the background matter and any touching objects. Overlapping of both parasite and background particles may be present. Minor overlapping can be reduced with separation and filtering. The success of separating overlapping images relies on the following two steps: finding the interior cores of the foreground objects and dilating the objects to find their linking points. Another automated approach employs MATLAB™ to collect data about the cross-section grayscale intensities of the parasite eggs. Because the interior and shell of parasite eggs can be distinctively different, this program was implemented to better characterize the different parasite egg types. The final MATLAB™ approach, edge signature, provides a way to capture the shape of an object quantitatively. The distance between the boundary and the centroid creates the edge signature plot.

Each of the five approaches collected data, in the form of quantifiers, which can be utilized in locating and distinguishing between parasite egg types. Out of the 44 quantifiers in the manual and automated approaches, 36 quantifiers proved to be statistically significant in characterizing the parasite eggs. These 36 measurements, along with the summary of data collected in this research, are used as the basis for parasite egg differentiation.

6.2 Results Summary

There were four approaches to analyze the digital images and characterize the parasite eggs. The first approach was manual. Manual measurements were recorded by one of the following techniques: through the microscope eyepiece (pre-image capture) or after the digital image was captured (post-image capture). The two techniques had the same outcome based on the statistical results. The length and width measurements for all three eggs show that differentiation by egg types is feasible. *Paragonimus westermani* eggs have a median manual length that fell within the range of expected lengths. The minimum manual length, however, fell below the expected range by 10 micrometers. The *Paragonimus westermani* manual width resulted in the same manner; the minimum manual width fell below the expected range by 5 micrometers. The manual median of both the length and width were 16% and 13%, respectively, below the expected median. The *Taenia* spp. egg results were quite the opposite. The manual results were larger than the expected range. For the length of *Taenia* spp., the minimum manual measurement was the only point to fall within the expected range. Overall the median manual length result was 23% larger than the expected length median. For the width of *Taenia* spp., the maximum manual measurement was 10 micrometers above the expected range. The median and minimum manual width, however, were located within the expected range. The *Ascaris lumbricoides*

egg results were the most successful by comparison. The manual length measurements completely fell within the range of expected values. The median manual width was only 12% larger than the median expected width.

The remaining three approaches to characterize the parasite eggs involve several different types of image processing techniques. Once the parasite egg had been located within the image, LabVIEW™ Vision Assistant™ scripts used built-in measurement functions to record sizes. The LabVIEW™ measurements were consistently lower than expected values. The next approach analyzed the gray-scale intensity of the parasite ova cross-section in MATLAB™. The final method examined the parasite's edge signature in MATLAB™. Thirty images per parasite type were analyzed using each of the four methods.

Fifteen individual measurements were statistically tested. The measurements were collected using the four approaches (Manual, LabVIEW™, Gray-scale Intensities, and Edge Signature). In the manual test, the Length and Width of each egg was measured. In the LabVIEW™ test, the Area, RSS (Feret), Maximum Feret Diameter, Elongation Factor, and Perimeter were measured. In the gray-scale intensity cross-section evaluation, the (Max-Min) Intensity Difference, Mean Intensity, Number of Peaks, and Number of Peaks (5) were measured. Finally, in the edge signature test the (Max-Min) Distance Difference, Mean Distance, Number of Peaks, and Number of Peaks (5) were measured.

The statistical analysis began with single factor global Analysis of Variance (ANOVA) test and multiple comparison tests. The global test results showed that for all measurements at least one pairing of parasite egg types would result in a statistical difference of value. For all fifteen measurements, the single factor global ANOVA test proved to be statistically significant. From these results, the multiple comparison tests, using Tukey's method, were performed. A multiple comparison test is only applied to the results in the global test that are deemed to be statistically different. This test will determine which parasite eggs within one specific measurement fail to be statistically different. A summary of the multiple comparison results is shown in Table 6.1.

Table 6.1 Summary of the multiple comparison test results

	Significant Difference?	*P. westermani* vs. *Taenia* spp.	*P. westermani* vs. *A. lumbricoides*	*A. lumbricoides* vs. *Taenia* spp.
	Manual			
Length	Yes	✓	✓	✓
Width	Yes	✓	✓	✓
	LabVIEW™ scripts			
Area	Yes	✓	✓	-
Elongation Ratio	Yes	✓	✓	✓
Max Feret Diameter	Yes	✓	✓	✓
Perimeter	Yes	✓	✓	-
RSS (Feret)	Yes	✓	✓	-
	Gray-scale Cross-section			
(Max-Min) Difference	Yes	✓	✓	-
Mean	Yes	-	✓	✓
Number of Peaks	Yes	✓	-	✓
Number of Peaks (5)	Yes	✓	-	✓
	Edge Signature			
(Max-Min) Difference	Yes	✓	✓	✓
Mean	Yes	✓	✓	✓
Number of Peaks	Yes	✓	-	✓
Number of Peaks (5)	Yes	✓	✓	✓

All of the manual measurement comparisons are statistically significant. Three of the LabVIEW™ multiple comparison tests failed to be statistically significant. They are as follows: Area (*Taenia* spp. vs. *Ascaris lumbricoides*), Perimeter (*Taenia* spp. vs. *Ascaris lumbricoides*), and RSS (Feret) (*Taenia* spp. vs. *Ascaris lumbricoides*). Four of the gray-scale cross-section multiple comparison tests failed. They are as follows: (Max-Min) Intensity Difference (*Taenia* spp. vs. *Ascaris lumbricoides*), Mean Intensity (*Paragonimus westermani* vs. *Taenia* spp.), Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*), and Number of Peaks (5) (*Paragonimus westermani* vs. *Ascaris lumbricoides*). Only one of the edge signature's multiple comparison tests failed to be statistically significant, Number of Peaks (*Paragonimus westermani* vs. *Ascaris lumbricoides*). These failed tests cannot be used to distinguish between parasite types. There were 36 comparisons out of the 44 that did not fail to reject the null hypothesis. The measurements where no null hypothesis rejections occurred are the following: manual Length, manual Width, LabVIEW™ Elongation Ratio, LabVIEW™ Max Feret Diameter, edge signature (Max-Min) Difference, edge signature Mean, and edge signature Number of Peaks (5).

6.3 Recommendations

The manual Length, manual Width, LabVIEW™ Elongation Ratio, LabVIEW™ Max Feret Diameter, edge signature (Max-Min) Difference, edge signature Mean, and edge signature Number of Peaks (5) measurements should be the first used when attempting to differentiate between parasite eggs from the three examined species.

There are many opportunities to complete more research on this topic. First, a method to achieve uniform digital images must be developed. The quality of the images is rooted in the lighting, slide preparation, and camera features which corresponds to the ability of the computer programs to locate parasite eggs. While the program may be able to locate a parasite egg, the parasite egg may be misidentified if the image quality is poor.

Although a general image processing procedure was created, future research should fine-tune and qualify the best image processing methods. Since the most critical part of the image processing is thresholding, special attention should be given to improve this step. Periodically it was difficult for the program to locate the outer shell of the *Ascaris lumbroicodes* egg. The software would only detect the inner portion of the parasite egg, which is similar in size to the *Taenia* spp. egg. Thus, some measurements between these two eggs could not show a statistical difference. The resulting false positive identifications, as well as false negatives, need to be improved. Refinement of the image processing techniques will improve measurement quality. The final image processing procedure needs to combine all three parasite egg species. Once a combined parasite image processing tool is created, field tests should be performed.

A database of measurement results should be created for eggs of several intestinal parasites to provide a larger data sample. The expansion of the database and decision tree will require more digital images. The addition of more parasite eggs to the database would test the differentiation ability of the measurement parameters. This database will be the basis for differentiation between parasite types in the automated parasite detection codes. A decision tree based on the statistical results of the data can be built to automate processing. A single flow decision matrix would link all identifiers and parasites. In order to begin the transformation of this research into a commercial product, the different approaches taken for locating the parasite eggs must be integrated. A critical goal for this effort must be the minimization of computations and the ability for high through-put. The integration of approaches from MATLAB™ and LabVIEW™ Vision Assistant™ should take place within LabVIEW™.

Once an integrated program is created, it could be packaged in a desirable form to consumers. Mobile phone based applications would place the product into the hands of the consumer as a portable

device. This would be especially valuable for field use. Untrained users would likely be accustomed to the mobile phone application form. Within the walls of a medical facility, the software could be integrated with motorized, translation microscope stages. The moveable microscope stages would allow for hands-free parasite detection. Only the suspected parasite eggs gathered from the automated scan of the microscope slide would be presented to the trained staff. This would focus the staff's attention on positive identifications rather than screening the entire slide to detect eggs. The process would become further automated, but the professional opinion would still be required.

6.4 Final Conclusion

LabVIEW™ Vision Assistant™ and MATLAB™ provided valuable image processing tools. This initial research has set the stage for the software structure, image processing technique, and the measurements that should be used for automated parasite detection. The image processing techniques have proven successful at differentiating parasite eggs from background material. Forty-four separate measurements were analyzed through four different approaches. Thirty-six of the measurements were statistically significant in the determination of the parasite eggs identities. These measurements have demonstrated that it is feasible to develop an automated parasite detection algorithm through image processing.

References

[1] 2013, "World Health Organization (WHO)," http://www.who.int/en/.

[2] Bovik, A. C., 2009, The essential guide to image processing, Academic Press, Burlington, Mass.; London.

[3] 2010, "DPDx - CDC Parasitology Diagnostic Web Site," Web Page, http://dpd.cdc.gov/dpdx.

[4] Gunn, A., and Pitt, S. J., 2012, Parasitology: An Integrated Approach, Wiley.

[5] Dogantekin, E., Yilmaz, M., Dogantekin, A., Avci, E., and Sengur, A., 2008, "A robust technique based on invariant moments – ANFIS for recognition of human parasite eggs in microscopic images," Expert Systems with Applications, 35(3), pp. 728-738.

[6] Avci, D., and Varol, A., 2009, "An expert diagnosis system for classification of human parasite eggs based on multi-class SVM," Expert Syst. Appl., 36(1), pp. 43-48.

[7] Yoon Seok, Y., Duck Kun, P., Hee Chan, K., Min-Ho, C., and Jong-Yil, C., 2001, "Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network," Biomedical Engineering, IEEE Transactions on, 48(6), pp. 718-730.

[8] Sommer, C., 1996, "Digital image analysis and identification of eggs from bovine parasitic nematodes," Journal of Helminthology, 70(02), pp. 143-151.

[9] Nugent, C., Cunningham, P., and Kirwan, P., 2006, "Using active learning to annotate microscope images of parasite eggs," Artif. Intell. Rev., 26(1-2), pp. 63-73.

[10] Sommer, C., 1998, "Quantitative characterization of texture used for identification of eggs of bovine parasitic nematodes," Journal of Helminthology, 72(02), pp. 179-182.

[11] Broyles, D. A., 2002, "Rapid shape characterization of crushed stone by PC-based digital image processing," Master of Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

[12] Sommer, C., 1998, "Quantitative characterization, classification and reconstruction of oocyst shapes of Eimeria species from cattle," Parasitology, 116(01), pp. 21-28.

[13] Liu, B., Fang, X., Wang, W., and Zheng, Z., "Automatic separation of overlapping objects," Proc. Proceedings of the 4th World Congress on Intelligent Control and Automation, pp. 2901-2905 vol.2904.

[14] "Olympus ", http://www.olympusamerica.com.

[15] "WARD'S Natural Science," http://wardsci.com/.

[16] "National Instruments' LabVIEW," http://www.ni.com/labview/.

[17] "Mathworks' MATLAB," http://www.mathworks.com/products/matlab/.

[18] Shapiro, L. G., and Stockman, G. C., 2001, Computer vision, Prentice Hall, Upper Saddle River, NJ.

[19] Otsu, N., 1979, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, 9(1), pp. 62-66.

[20] Castleman, K. R., 1996, Digital image processing, Prentice Hall, Englewood Cliffs, N.J.

[21] Fisher, R. B. D.-H., K.; Fitzgibbon, A.; Robertson, C.; Trucco, E., 2005, Dictionary of Computer Vision and Image Processing, John Wiley & Sons

[22] Bovik, A. C., 2000, Handbook of image and video processing, Academic Press, San Diego.

[23] Davies, E. R., 2005, Machine vision : theory, algorithms, practicalities, Elsevier, Amsterdam; Boston.

[24] Bingham, L. X., Fang; Weizhi, Wang; Zhiyong, Zheng, 2002, "Automatic Separation of Overlapping Objects," Proceedings of the 4th World Congress on Intelligent Control and Automation, pp. 2901-2905.

[25] 2011, "Particle Measurements - NI Vision 2011 Concepts Help - National Instruments," Web page, http://zone.ni.com/reference/en-XX/help/372916L-01/nivisionconcepts/particle_measurements/.

[26] Ott, L., and Longnecker, M., 2010, An Introduction to Statistical Methods And Data Analysis, Brooks/Cole Cengage Learning.

Appendices

Appendix A: LabVIEW™ Vision Assistant™ Calibration

The calibration tool chosen within LabVIEW™ Vision Assistant™ was a simple calibration. When the simple calibration is used in microscopic inspections, the results are referred to as coarse measurements. Simple calibrations measure the pixel distance between two points of a known length, also called the point coordinates method. Figure A.1 represents the calibration process.

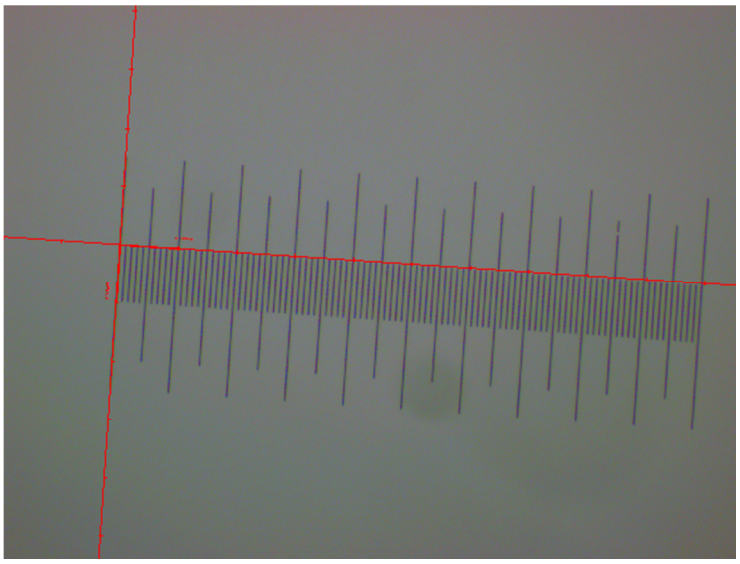

Figure A.1 Calibration process within the LabVIEW™ script results in uncertainties

The red lines are the axes that were manually placed as an overlay for the stage micrometer image. In was important that the axes line up with the micrometer measurements. The axes are lined up by choosing two points on the image- Point 1 and Point 2. In Figure A.1, the points chosen were at each end of the micrometer, which is a known 100 micrometers in length. The calibrations are calculated, and they are automatically completed within LabVIEW™. A mathematical representation of the calibration is shown in Equation A1.1, A1.2, and A1.3.

Example:

Point 1: (364, 659)

Point 2: (525, 670)

Distance of 100 micrometers.

The distance between the points is calculated in the following equation,

$$\mathbf{D} = \sqrt{\boldsymbol{dx^2 + dy^2}}, \qquad \textbf{A1.1}$$

where dx is the distance between the x-pixels and $\boldsymbol{dy}$ is the distance between the y-pixels. Their equations are as follows,

$$\boldsymbol{dx = |x_1 - x_2|,} \quad \textbf{A1.2}$$

$$\boldsymbol{dy = |y_1 - y_2|,} \quad \textbf{A1.3}$$

where the x values are the two x-values of the points, and the $\boldsymbol{y}$ values are the two x-values of the points. For this application, the distance found in Equation A1.1 is 161.4 pixels. Since these pixels represented 100 micrometers in real-world length, the digital image is calibrated in LabVIEW™ to have 0.620 micrometers per pixel.

Half of the smallest reference length represents the uncertainty of the calculation. In the case of this research, the smallest reference length can either be one pixel or one micrometer from Figure A.1. In both cases, the uncertainty of the LabVIEW™ Vision Assistant™ calibration is 0.310 micrometers.

Appendix B: Measurement uncertainty calculations

The manual measurements were recorded by two different techniques, pre-image capture and post-image capture. This first technique will be called pre-image capture since the technique was applied before the image of the parasite egg was taken and while the prepared microscopic slide sample was still intact. The pre-image capture technique was done through the eyepiece. The eyepiece, which is also referred to as the ocular lens, comes with built-in markers. This is a type of micrometry is called ocular micrometry. The marker distances of the eyepiece are converted to micrometers by measuring a known distance. Half of the smallest reference length represents the uncertainty of the calculation. In the case of this pre-image capture, the smallest reference length is one ocular marker. The conversion factor at a microscopic level of 100x was discovered to be 10μ for one ocular mark. The uncertainty of the manual pre-image capture technique is 5μ.

Some images had been taken before the ocular micrometry technique was discovered. Thus, a different manual measurement technique was applied to these images. This technique will be call post-image capture since the technique was applied after the image of the parasite egg had been captured and the prepared microscopic slide sample destroyed. The post-image capture technique was done by overlaying the captured image of the stage micrometer with the parasite ova image. Overlaying of the two images (parasite egg and micrometer) was accomplished in Microsoft PowerPoint. PowerPoint was chosen for its ability to display the images in their true pixel sizes and to remove backgrounds of images. The image of the stage micrometer slide and the parasite egg image of interest were overlaid. The measurements of length and width were directly recorded. Half of the smallest reference length represents the uncertainty of the calculation. In the case of this pre-image capture, the smallest reference length is 1μ in distance. The uncertainty of the manual pre-image capture technique is 0.5μ.

Appendix C: Image results from MATLAB™ Separation Code

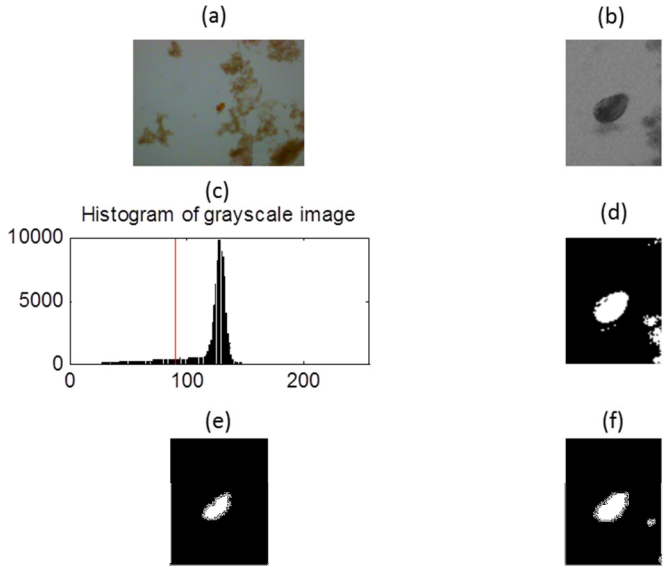

Figure C.1 Pre-processing for MATLAB™ Separation Code *Paragonimus westermani* Image #15

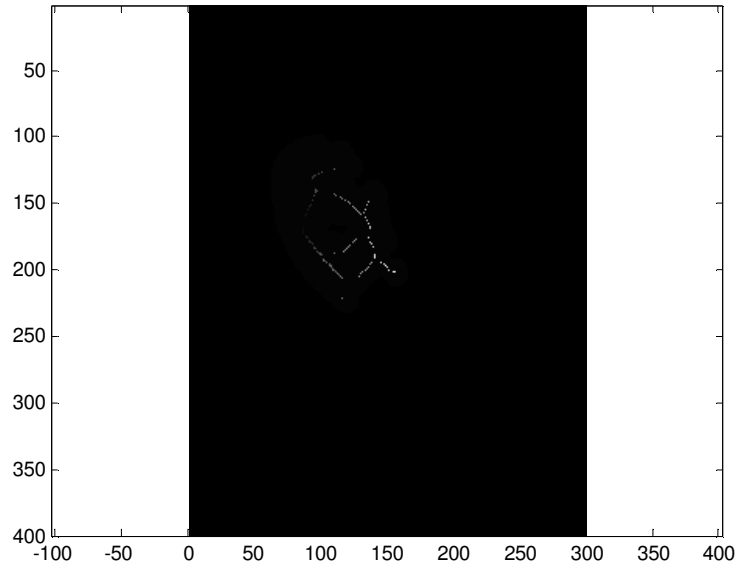

Figure C.2 *Paragonimus westermani* Image #15 result from MATLAB™ Separation Code

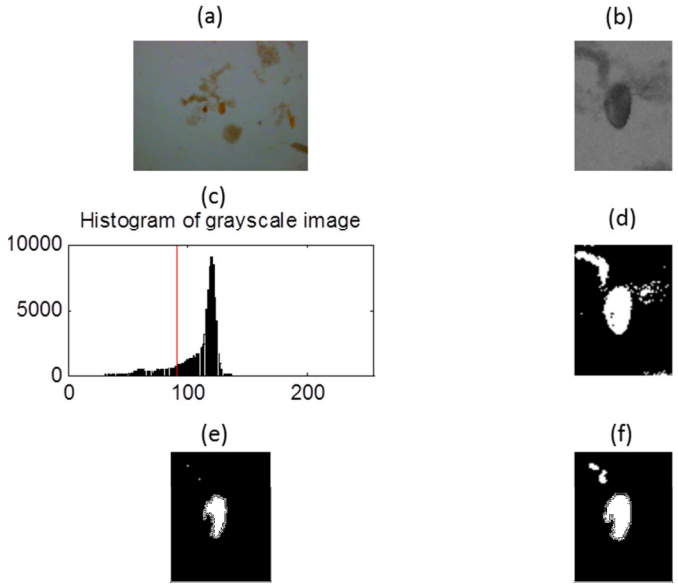

Figure C.3 Pre-processing for MATLAB™ Separation Code *Paragonimus westermani* Image #30

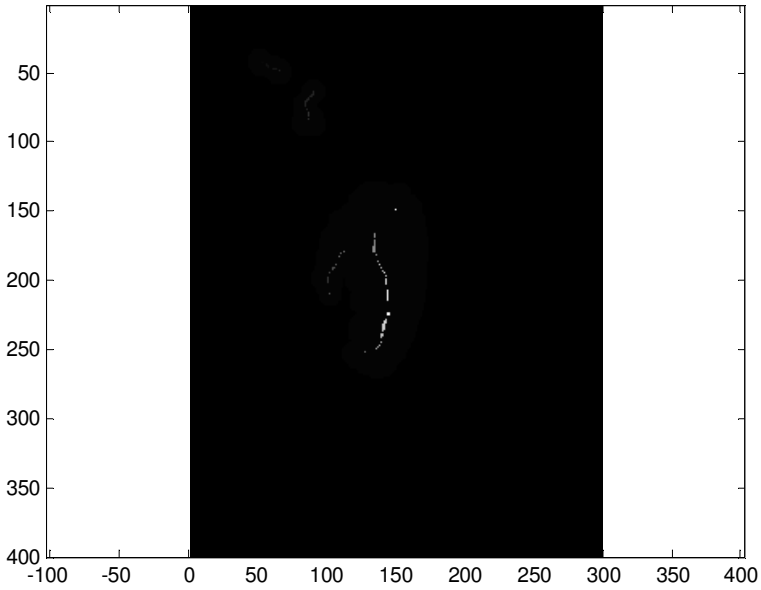

Figure C.4 *Paragonimus westermani* Image #30 result from MATLAB™ Separation Code

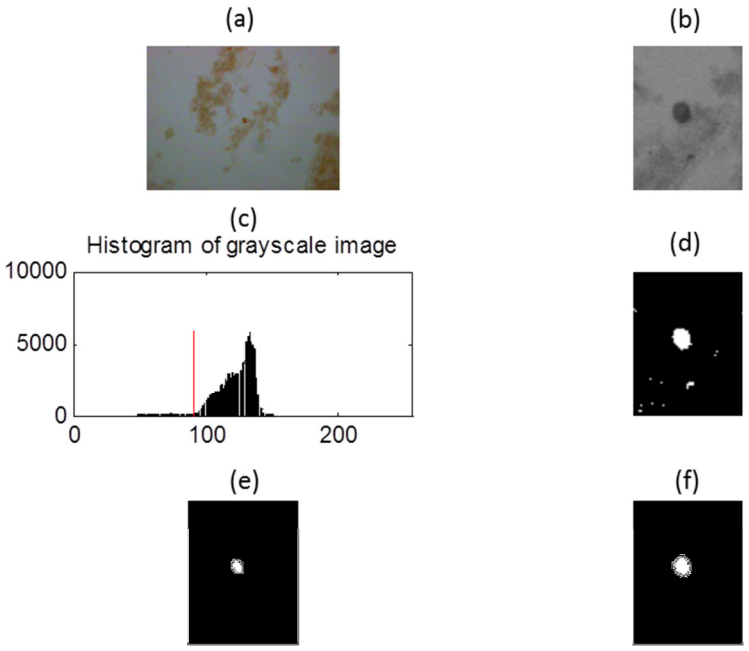

Figure C.5 Pre-processing for MATLAB™ Separation Code *Taenia* Image #1

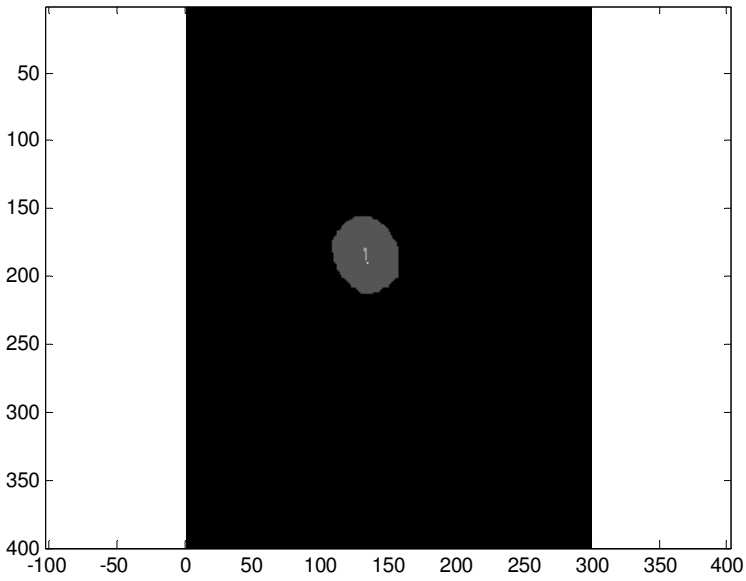

Figure C.6 *Taenia* Image #1 result from MATLAB™ Separation Code

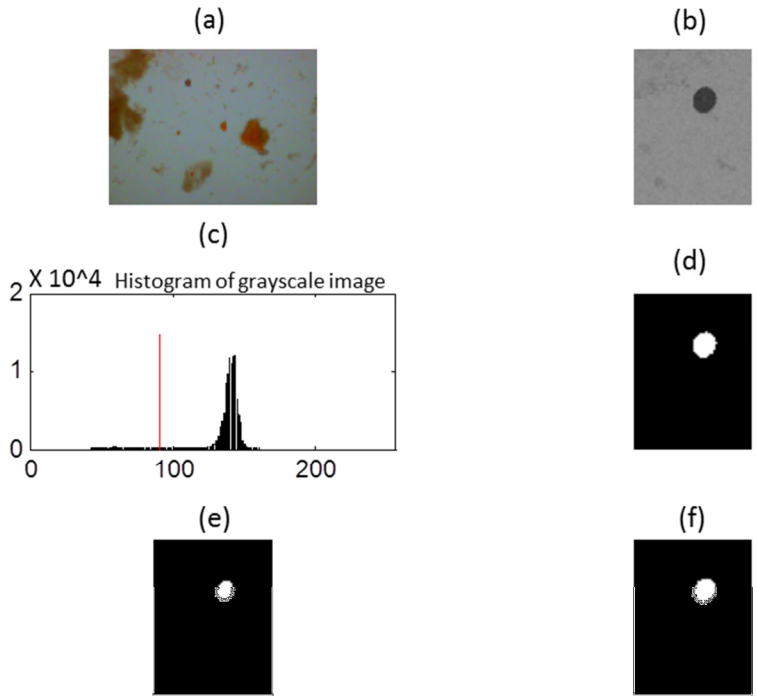

Figure C.7 Pre-processing for MATLAB™ Separation Code *Taenia* Image #15

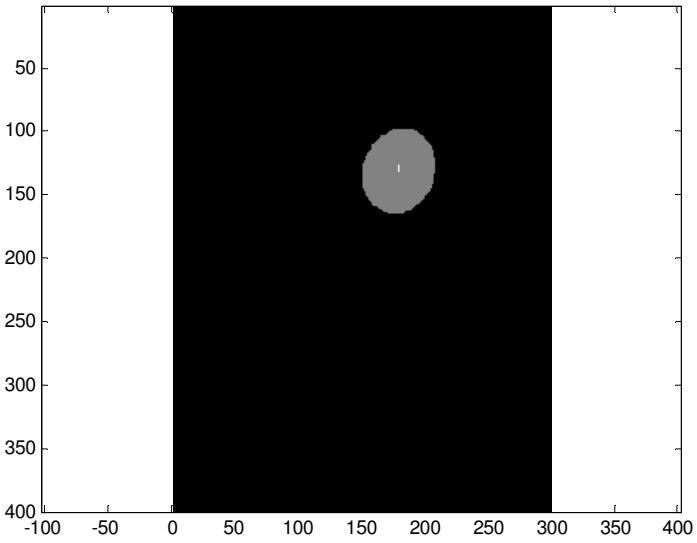

Figure C.8 *Taenia* Image #15 result from MATLAB™ Separation Code

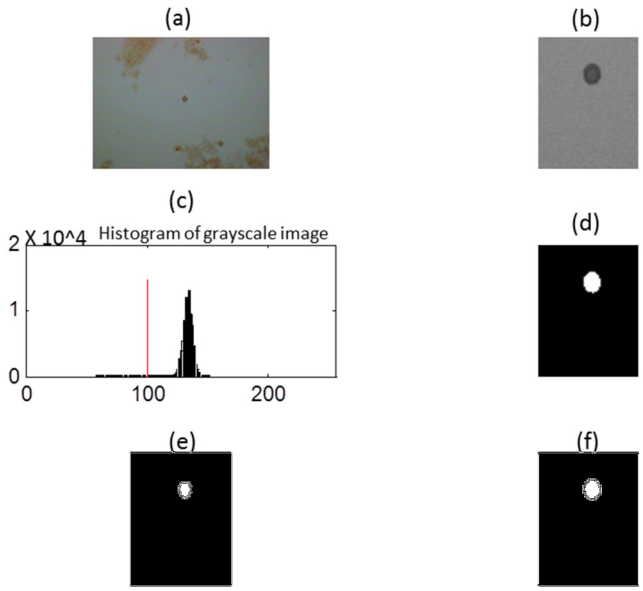

Figure C.9 Pre-processing for MATLAB™ Separation Code *Taenia* Image #30

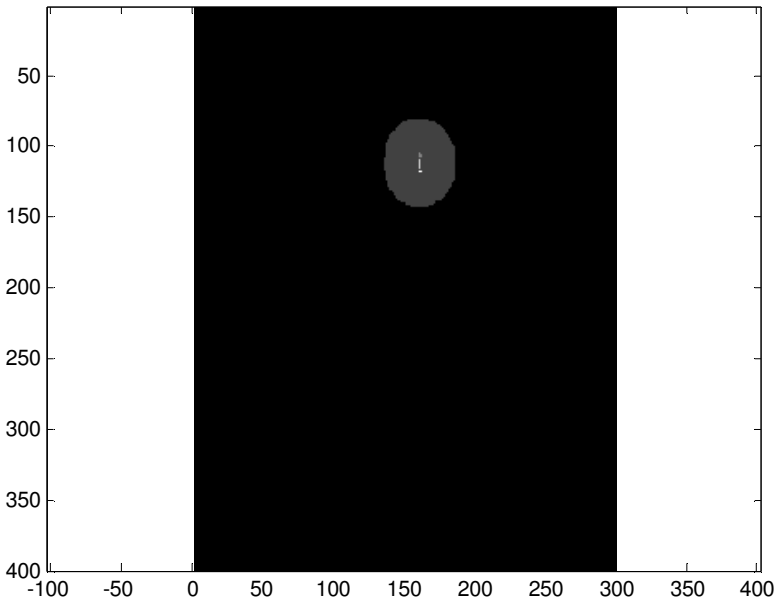

Figure C.10 *Taenia* Image #30 result from MATLAB™ Separation Code

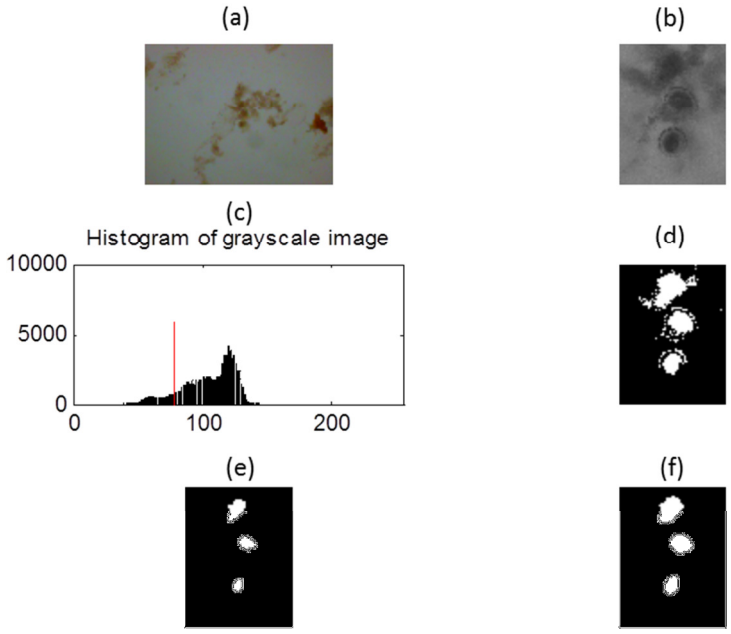

Figure C.11 Pre-processing for MATLAB™ Separation Code *Ascaris lumbricoides* Image #1

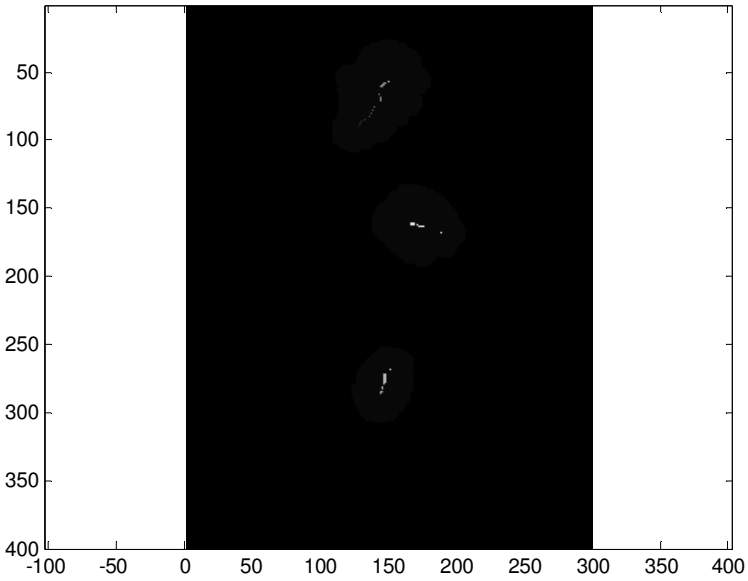

Figure C.12 *Ascaris lumbricoides* Image #1 result from MATLAB™ Separation Code

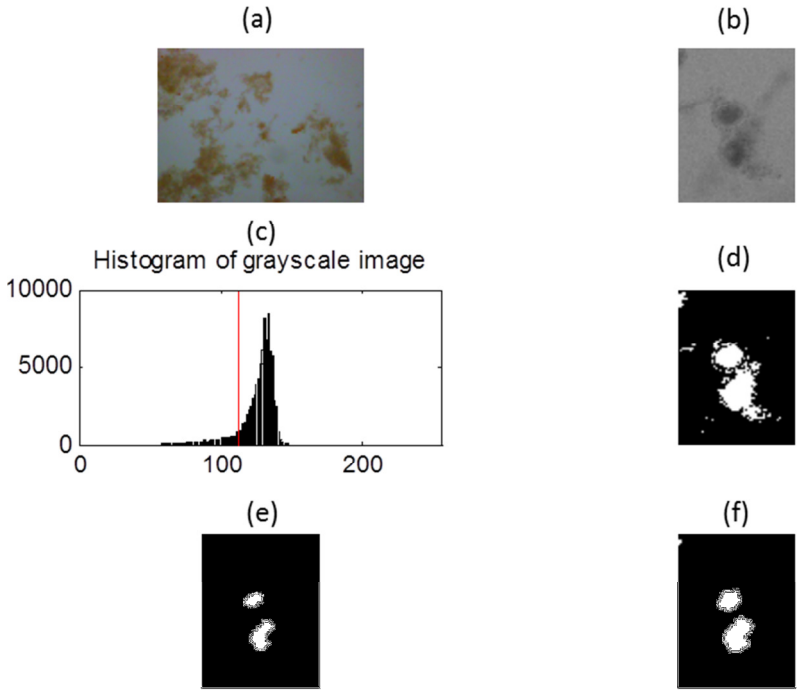

Figure C.13 Pre-processing for MATLAB™ Separation Code *Ascaris lumbricoides* Image #15

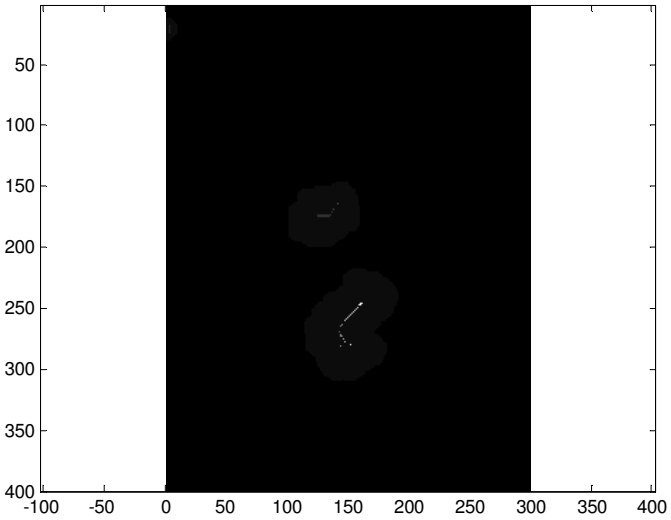

Figure C.14 *Ascaris lumbricoides* Image #15 result from MATLAB™ Separation Code

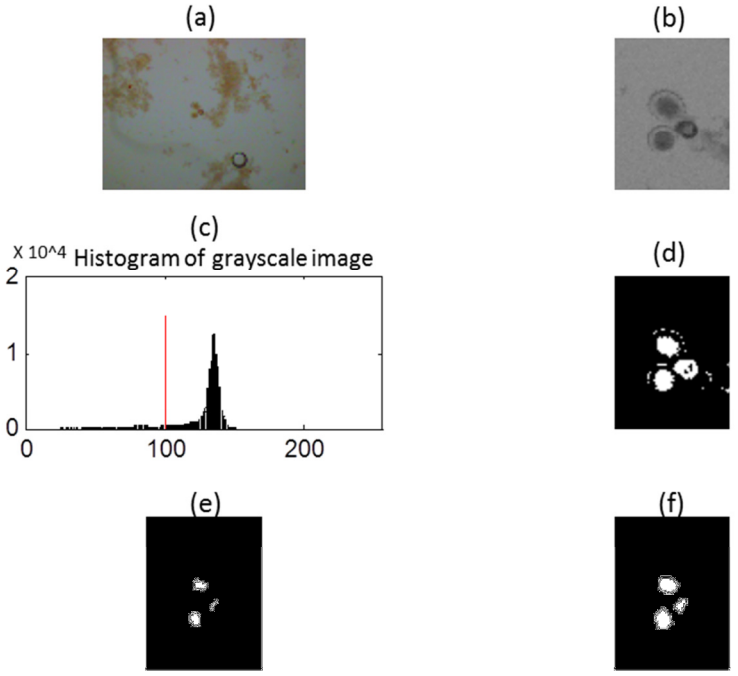

Figure C.15 Pre-processing for MATLAB™ Separation Code *Ascaris lumbricoides* Image #30

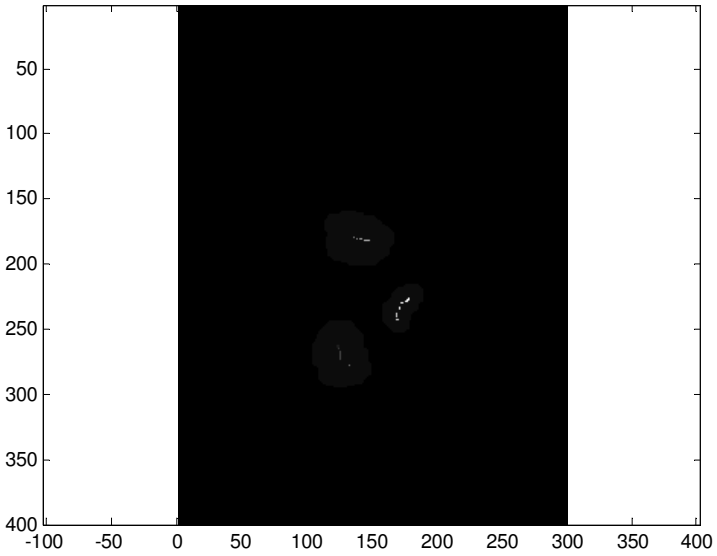

Figure C.16 *Ascaris lumbricoides* Image #30 result from MATLAB™ Separation Code

Appendix D: Separation MATLAB™ code

```
Separation LEO 2

clc
clear all
close all
format compact
format short e

ParasiteImage = imread('100xtaenia30.bmp');
subplot(3, 2, 1);
imshow(ParasiteImage);

greenPlane = ParasiteImage(:, :, 2);

if ParasiteImage == imread('100xAL30.bmp')
    Subsection=greenPlane(500:900,800:1100);
elseif ParasiteImage == imread('100xAL24.bmp')
    Subsection=greenPlane(200:600,600:900);
elseif ParasiteImage == imread('100xAL25.bmp')
    Subsection=greenPlane(300:700,900:1200);
elseif ParasiteImage == imread('100xAL27.bmp')
    Subsection=greenPlane(200:600,400:700);
elseif ParasiteImage == imread('100xAL29.bmp')
    Subsection=greenPlane(300:700,800:1100);
elseif ParasiteImage == imread('100xtaenia15.bmp')
    Subsection=greenPlane(200:600,600:900);
elseif ParasiteImage == imread('100xtaenia9.bmp')
    Subsection=greenPlane(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xtaenia16.bmp')
     Subsection=greenPlane(600:1000,800:1100);
elseif ParasiteImage == imread('100xtaenia19.bmp')
     Subsection=greenPlane(300:700,900:1200);
elseif ParasiteImage == imread('100xtaenia20.bmp')
     Subsection=greenPlane(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xpwestermani16.bmp')
    Subsection=greenPlane(700:1100,1200:1500);
elseif ParasiteImage == imread('100xpwestermani17.bmp')
    Subsection=greenPlane(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani19.bmp')
     Subsection=greenPlane(600:1000,800:1100);
elseif ParasiteImage == imread('100xpwestermani20.bmp')
     Subsection=greenPlane(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani22.bmp')
     Subsection=greenPlane(600:1000,1200:1500);
elseif ParasiteImage == imread('100xpwestermani24.bmp')
     Subsection=greenPlane(1000:1400,1300:1600);
elseif ParasiteImage == imread('100xpwestermani29.bmp')
     Subsection=greenPlane(600:1000,700:1000);
else
Subsection=greenPlane(600:1000,900:1200);
```

```
end

subplot(3, 2, 2);
imshow(Subsection);

%% Make histogram

[pixelCount grayLevels] = imhist(Subsection);
subplot(3, 2, 3);
bar(pixelCount);
title('Histogram of grayscale image');
xlim([0 grayLevels(end)]);

thresholdValue = 100;

hold on;
maxYValue = ylim;
hStemLines = stem(thresholdValue, maxYValue(2), 'r');
children = get(hStemLines, 'children');
set(children(2),'visible', 'off');

binaryImage = Subsection < thresholdValue;
subplot(3, 2, 4);
imshow(binaryImage);


%% Open the image (Erode then Dilate)
se2 = strel('disk',9);
ErodedBinary = imerode(binaryImage,se2);

subplot(3,2,5);
imshow(ErodedBinary)

se1=strel('disk',8);
DilateBinary = imdilate(ErodedBinary,se1);

subplot(3,2,6);
imshow(DilateBinary)

%% The first scan is completed starting in the top left of the image and
%  works it's way across the columns and then down the rows.

[x,y]=size(DilateBinary);
FirstScan=int8(DilateBinary);
for i=2:1:x
    for j=2:1:y
        if FirstScan(i,j)==0;
            FirstScan(i,j)=0;
        elseif FirstScan(i,j)==1;
            z=[FirstScan(i,j-1)+1, FirstScan(i-1,j)+1];
            FirstScan(i,j)=min(z);
        end
```

```
    end
end

%% The second scan is then implemented. It goes the opposite the first scan
%  by starting in the lower right corner and working its way across the
%  columns and up the rows.

SecondScan=FirstScan;
for i=(x-1):-1:1
    for j=(y-1):-1:1
        if SecondScan(i,j)==0;
            SecondScan(i,j)=0;
        elseif SecondScan(i,j)~=0;
            d=[SecondScan(i+1,j)+1, SecondScan(i,j+1)+1, SecondScan(i,j)];
            SecondScan(i,j)=min(d);
        end

    end
end

%% The next step is to find the core of the objects.
%

CoreObjects=SecondScan;
for i=2:1:(x-1)
    for j=2:1:(y-1)
        f=[SecondScan(i-1,j-1),SecondScan(i-1,j), SecondScan(i+1,j-
1),SecondScan(i,j-1), SecondScan(i+1,j),SecondScan(i-1,j+1),
SecondScan(i,j+1), SecondScan(i+1,j+1)];
        if SecondScan(i,j)>=max(f);
            CoreObjects(i,j)=SecondScan(i,j);
        else CoreObjects(i,j)=1;
        end
    end
end

%% Since the rows and columns on the outmost edges are not correctly
%  we will set those to ones.

Edges=CoreObjects(2:(x-1),2:(y-1));
[x,y]=size(Edges);

%% We find the regional maximum and label the core area

Max=imregionalmax(Edges);
Connect = bwconncomp(Max);
labeled = labelmatrix(Connect);
RGB_label = label2rgb(labeled, @gray, 'b', 'noshuffle');
figure()
title('Colormap of the Labeled Connected Objects')
imshow(RGB_label,'InitialMagnification','fit');
```

```
%% The labeled objects need to be reassigned a value k (k=2:1:cc.NumObjects)

Relabel=labeled;
for i=1:1:x
    for j=1:1:y
        if labeled(i,j)>0;
            Relabel(i,j)=labeled(i,j)+1;
        else Relabel(i,j)=0;
        end
    end
end

%% The rest of the object area is then re-marked

ReMark=Relabel;
for i=1:1:x
    for j=1:1:y
        if labeled(i,j)==0;
            if Edges(i,j)==0;
                ReMark(i,j)=0;
            else ReMark(i,j)=1;
            end
        end
    end
end

figure()
[r1,c1] = size(ReMark);                          %# Get the matrix size
imagesc((1:c1)+0.5,(1:r1)+0.5,ReMark);           %# Plot the image
colormap(gray);                                %# Use a gray colormap
axis equal                                     %# Make axes grid sizes equal

% % Separation of any remaining overlap of images occurs.
[Separation2]=Separator(Connect,ReMark,x,y);

figure()
title('Final Separation Objects Found')
[r1,c1] = size(Separation2);                          %# Get the matrix size
imagesc((1:c1)+0.5,(1:r1)+0.5,Separation2);           %# Plot the image
colormap(gray);                                       %# Use a gray colormap
axis equal
%
```

Appendix E: Image results from MATLAB™ Cross-section Gray-scale Analysis

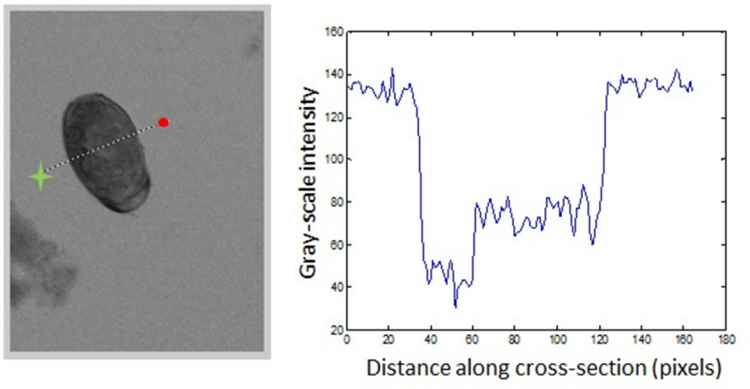

Figure E.1 *Paragonimus westermani* Image #1 cross-section gray-scale analysis results

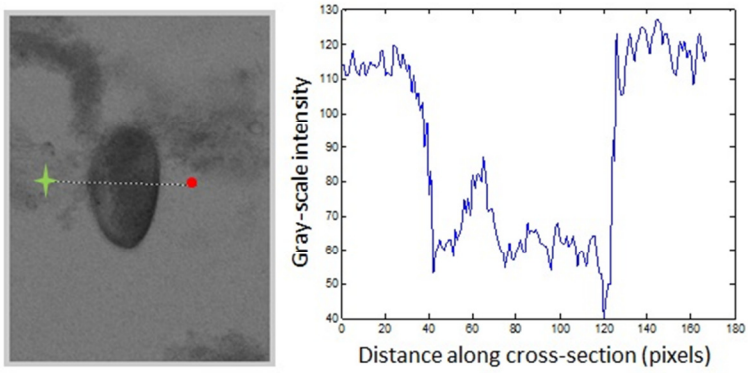

Figure E.2 *Paragonimus westermani* Image #30 cross-section gray-scale analysis results

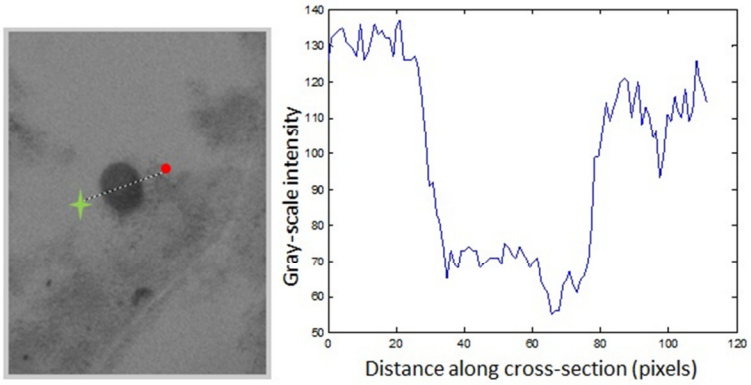

Figure E.3 *Taenia* spp. Image #1 cross-section gray-scale analysis results

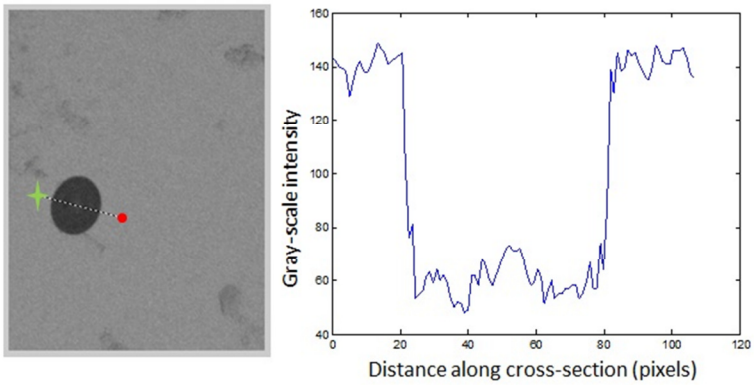

Figure E.4 *Taenia* spp. Image #15 cross-section gray-scale analysis results

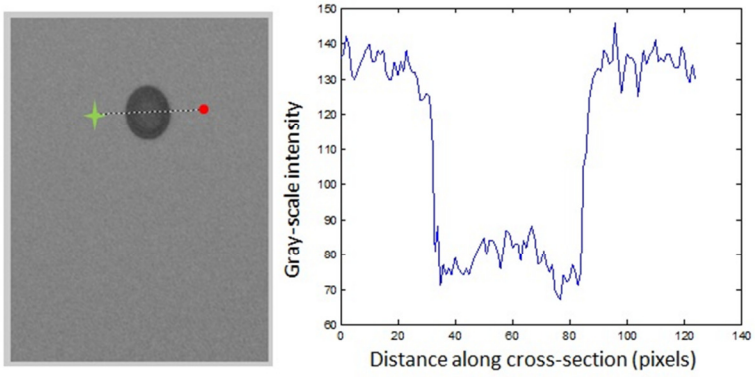

Figure E.5 *Taenia* spp. Image #30 cross-section gray-scale analysis results

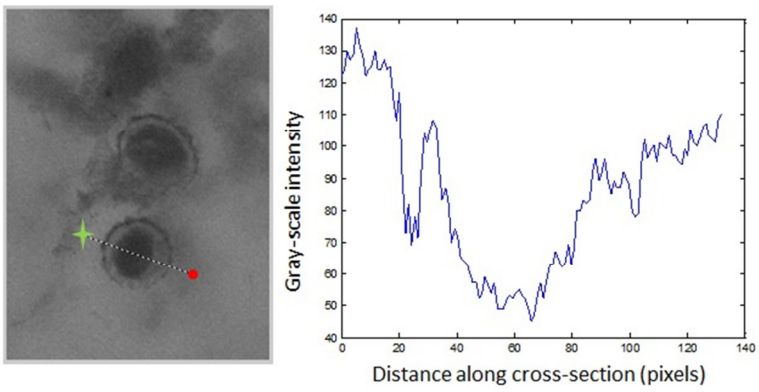

Figure E.6 *Ascaris lumbricoides* Image #1A cross-section gray-scale analysis results

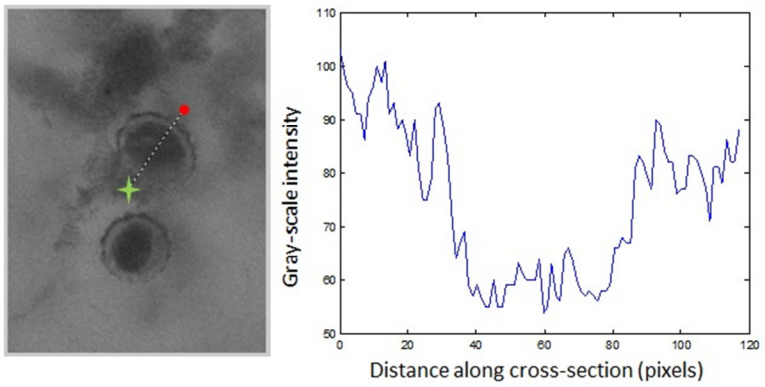

Figure E.7 *Ascaris lumbricoides* Image #1B cross-section gray-scale analysis results

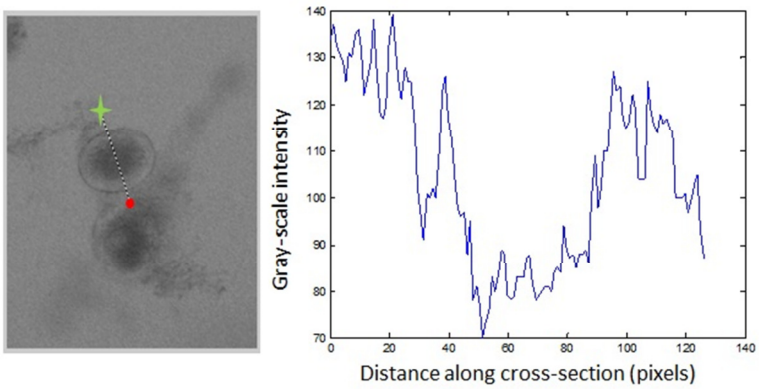

Figure E.8 *Ascaris lumbricoides* Image #15A cross-section gray-scale analysis results

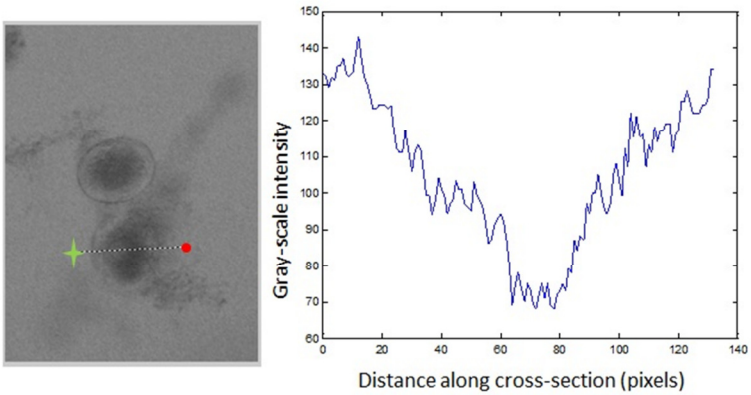

Figure E.9 *Ascaris lumbricoides* Image #15B cross-section gray-scale analysis results

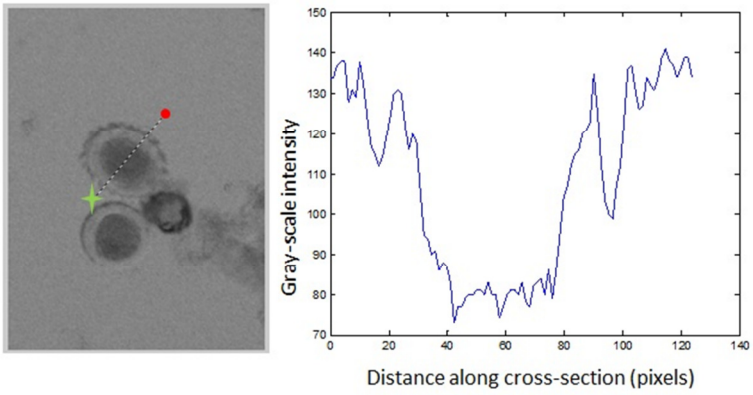

Figure E.10 *Ascaris lumbricoides* Image #30A cross-section gray-scale analysis results

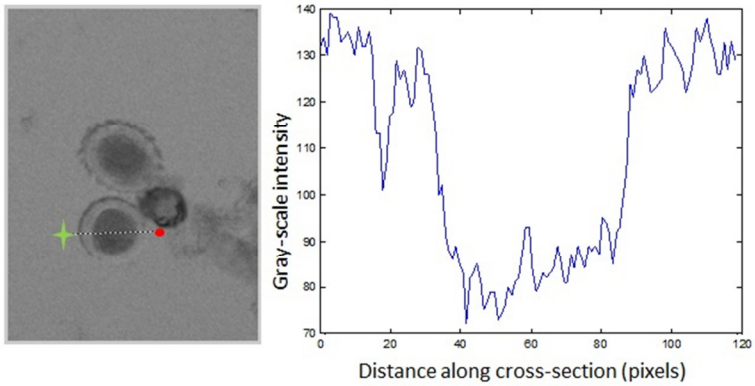

Figure E.11 *Ascaris lumbricoides* Image #30B cross-section gray-scale analysis results

Appendix F: Gray-scale intensity MATLAB™ code

Raw Cross Section Data

```
%% RawCrossSectionData.m
% This MATLAB™ program calls images and with help from the user determines
% the gray-scale intensities across the cross-section of the parasite eggs.
%

clc
clear all
close all
format compact
format short e

%The image of parasite egg is called.
ParasiteImage = imread('100xpwestermani15.bmp');
subplot(2, 2, 1);
imshow(ParasiteImage);

%The green plane is removed. This converts the image to gray-scale.
greenPlane = ParasiteImage(:, :, 2);

% In the images below, the parasite eggs were not centered. The image is
% cropped to reveal the parasite eggs.
if ParasiteImage == imread('100xAL30.bmp')
    Subsection1=greenPlane(500:900,800:1100);
elseif ParasiteImage == imread('100xAL24.bmp')
    Subsection1=greenPlane(200:600,600:900);
elseif ParasiteImage == imread('100xAL25.bmp')
    Subsection1=greenPlane(300:700,900:1200);
elseif ParasiteImage == imread('100xAL27.bmp')
    Subsection1=greenPlane(200:600,400:700);
elseif ParasiteImage == imread('100xAL29.bmp')
    Subsection1=greenPlane(300:700,800:1100);
elseif ParasiteImage == imread('100xtaenia15.bmp')
    Subsection1=greenPlane(200:600,600:900);
elseif ParasiteImage == imread('100xtaenia9.bmp')
    Subsection1=greenPlane(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xtaenia16.bmp')
     Subsection1=greenPlane(600:1000,800:1100);
elseif ParasiteImage == imread('100xtaenia19.bmp')
     Subsection1=greenPlane(300:700,900:1200);
elseif ParasiteImage == imread('100xtaenia20.bmp')
     Subsection1=greenPlane(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xpwestermani16.bmp')
    Subsection1=greenPlane(700:1100,1200:1500);
elseif ParasiteImage == imread('100xpwestermani17.bmp')
    Subsection1=greenPlane(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani19.bmp')
     Subsection1=greenPlane(600:1000,800:1100);
elseif ParasiteImage == imread('100xpwestermani20.bmp')
```

```
     Subsection1=greenPlane(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani22.bmp')
     Subsection1=greenPlane(600:1000,1200:1500);
elseif ParasiteImage == imread('100xpwestermani24.bmp')
     Subsection1=greenPlane(1000:1400,1300:1600);
elseif ParasiteImage == imread('100xpwestermani29.bmp')
     Subsection1=greenPlane(600:1000,700:1000);
else
Subsection1=greenPlane(600:1000,900:1200);
end

subplot(2, 2, 2);
imshow(Subsection1);

%The "improfile" command allows the user to click on the image to begin
%drawing a line. The user then drags a line across the mid section of the
%egg and click again to determine the end of the line.
line=improfile

%The following lines are drawing information out of the raw data.
N=length(line);
Ts=1;
fs=1/Ts;
Tmax=(N-1)*Ts;
T= [0:Ts:Tmax];
subplot(2, 2, 3);
plot(T,line);
xlabel('Number of Pixels')
ylabel('Gray-Scale Intenstiy')

[pks1,locs]=findpeaks(line,'minpeakdistance',1);
[pks2,locs]=findpeaks(line,'minpeakdistance',5);

NumberofPeaks=length(pks1)
NumberofPeaks5=length(pks2)
```

Appendix G: Image results from MATLAB™ Edge Signature

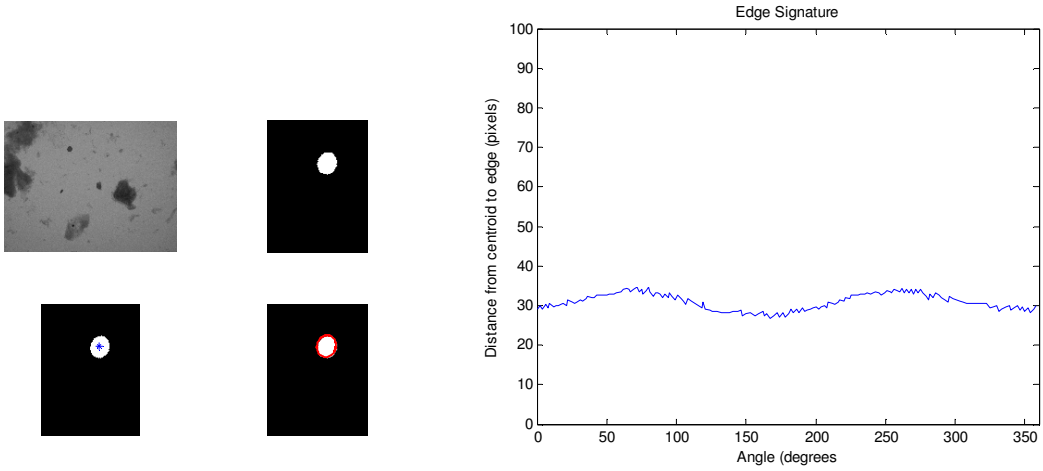

Figure G.1 Edge Signature result of *Taenia* spp. Image #15 at a threshold of 90.

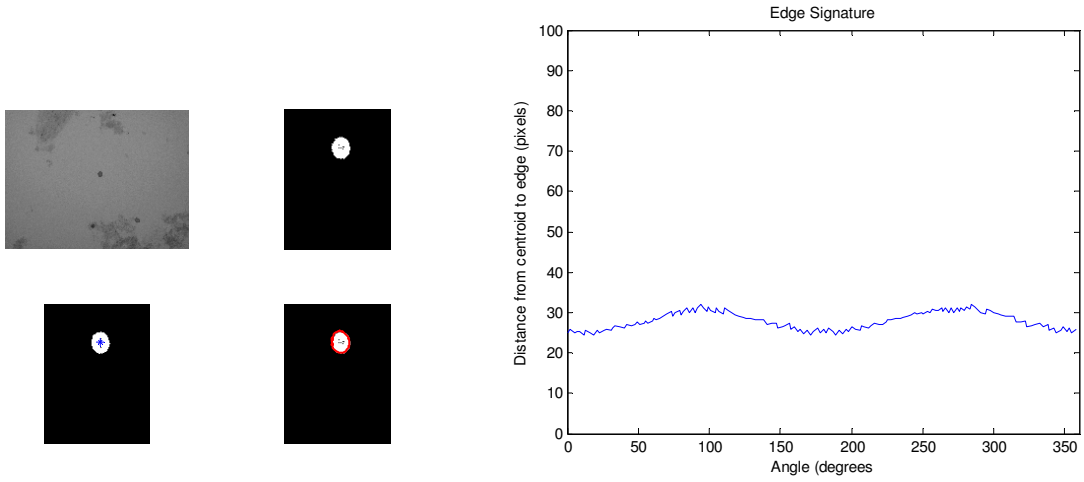

Figure G.2 Edge Signature result of *Taenia* spp. Image #30 at a threshold of 90.

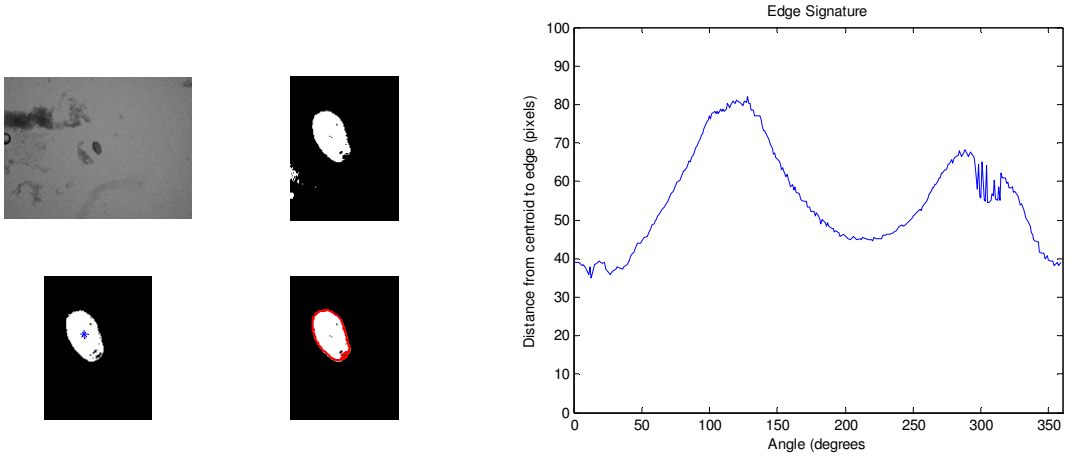

Figure G.3 Edge Signature result of *Paragonimus westermani* Image #15 at a threshold of 90.

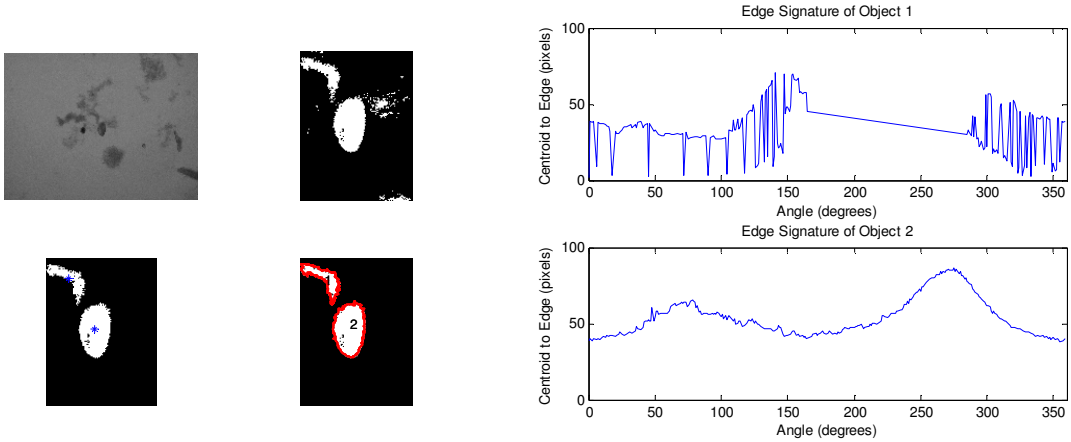

Figure G.4 Edge Signature result of *Paragonimus westermani* Image #30 at a threshold of 90.

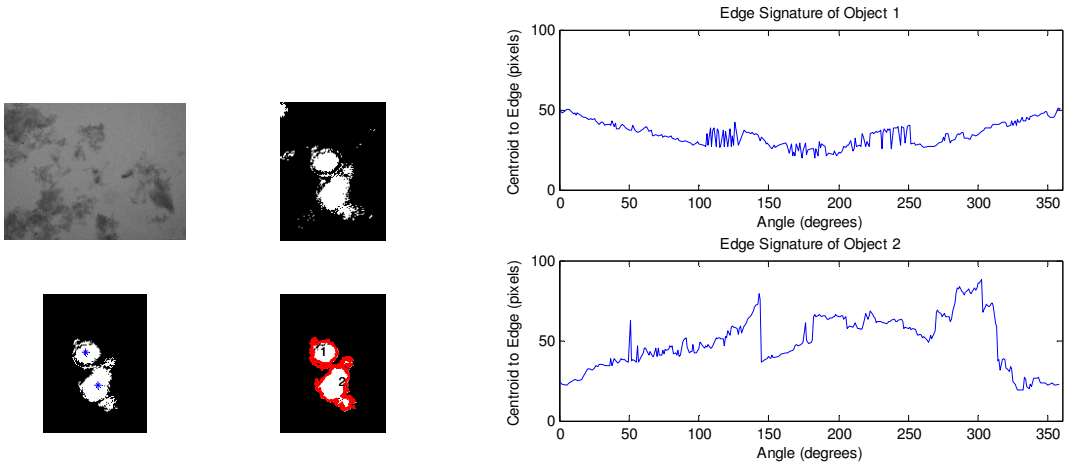

Figure G.5 Edge Signature result of *Ascaris lumbricoides* Image #15 at a threshold of 113.

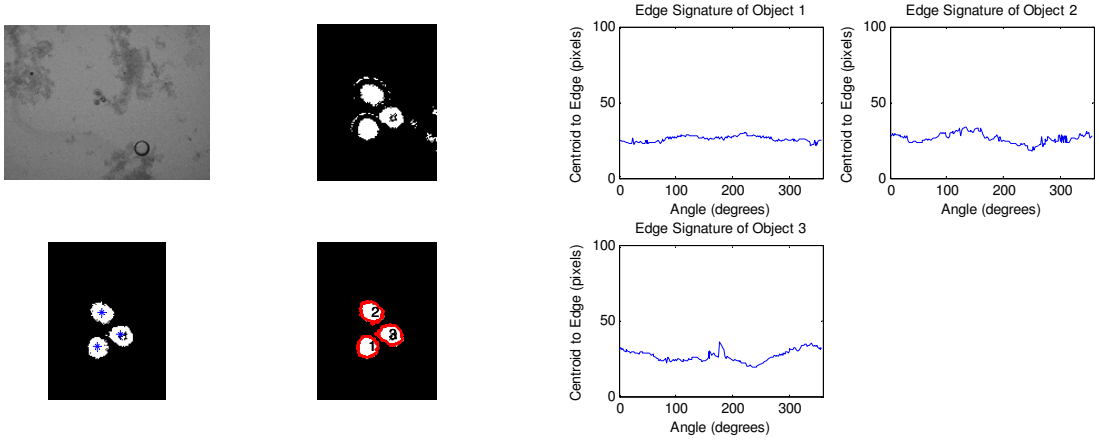

Figure G.6 Edge Signature result of *Ascaris lumbricoides* Image #30 at a threshold of 100.

Appendix H: Edge signature MATLAB™ code

```
%% CentroidEdgeProfile.m
% This MATLAB™ program calls images and determines the parasite egg profile,
% also known as the signature. The signature, distance from the center to
the outer edge, is plotted linearly against
% the angle of measurement around the egg (0-360 degrees).

clc
clear all
close all
format compact
format short e

%The image is read in, converted to gray-scale, and subsectioned.
%The conversion to gray-scale is done by extracting the green color plane.
ParasiteImage = imread('100xAL30.bmp');
greenPlane = ParasiteImage(:, :, 2);
subplot(2, 2, 1);
imshow(greenPlane);

%% A manual "fixed" global threshold is performed to convert to binary.
thresholdValue = 100;
binaryImage1= greenPlane < thresholdValue;

% In the images below, the parasite eggs were not centered. The image is
% cropped to reveal the parasite eggs.
if ParasiteImage == imread('100xAL30.bmp')
    Subsection1=binaryImage1(500:900,800:1100);
elseif ParasiteImage == imread('100xAL24.bmp')
    Subsection1=binaryImage1(200:600,600:900);
elseif ParasiteImage == imread('100xAL25.bmp')
    Subsection1=binaryImage1(300:700,900:1200);
elseif ParasiteImage == imread('100xAL27.bmp')
    Subsection1=binaryImage1(200:600,400:700);
elseif ParasiteImage == imread('100xAL29.bmp')
    Subsection1=binaryImage1(300:700,800:1100);
elseif ParasiteImage == imread('100xtaenia15.bmp')
    Subsection1=binaryImage1(200:600,600:900);
elseif ParasiteImage == imread('100xtaenia9.bmp')
    Subsection1=binaryImage1(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xtaenia16.bmp')
     Subsection1=binaryImage1(600:1000,800:1100);
elseif ParasiteImage == imread('100xtaenia19.bmp')
     Subsection1=binaryImage1(300:700,900:1200);
elseif ParasiteImage == imread('100xtaenia20.bmp')
     Subsection1=binaryImage1(800:1200, 1200:1500);
elseif ParasiteImage == imread('100xpwestermani16.bmp')
    Subsection1=binaryImage1(700:1100,1200:1500);
elseif ParasiteImage == imread('100xpwestermani17.bmp')
    Subsection1=binaryImage1(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani19.bmp')
     Subsection1=binaryImage1(600:1000,800:1100);
```

```
elseif ParasiteImage == imread('100xpwestermani20.bmp')
     Subsection1=binaryImage1(500:900,700:1000);
elseif ParasiteImage == imread('100xpwestermani22.bmp')
     Subsection1=binaryImage1(600:1000,1200:1500);
elseif ParasiteImage == imread('100xpwestermani24.bmp')
     Subsection1=binaryImage1(1000:1400,1300:1600);
elseif ParasiteImage == imread('100xpwestermani29.bmp')
     Subsection1=binaryImage1(600:1000,700:1000);
else
Subsection1=binaryImage1(600:1000,900:1200);
end
subplot(2, 2, 2);
imshow(Subsection1);

%%
CC = bwconncomp(Subsection1);
numPixels = cellfun(@numel,CC.PixelIdxList);

for k=1:1:length(numPixels)

    [biggest(k)] = numPixels(k);
    if biggest(k) < 2000
        Subsection1(CC.PixelIdxList{k}) = 0;
    else
        Subsection1(CC.PixelIdxList{k}) = 1;
    end
end


%% Calculate the centroids of the objects
s = regionprops(Subsection1, 'Centroid');
centroids = cat(1, s.Centroid);

subplot(2, 2, 3);
imshow(Subsection1)
hold on
plot(centroids(:,1), centroids(:,2), 'b*')
hold off


%% Show object boundaries

[B,L,N,A] = bwboundaries(Subsection1);
subplot(2, 2, 4); imshow(Subsection1); hold on;


for cnt=1:N,
    for k=1:length(B),
    if(~sum(A(k,:)))
       boundary = B{cnt};
       plot(boundary(:,2),boundary(:,1),'r','LineWidth',2);
       text(mean(boundary(:,2)),mean(boundary(:,1)),num2str(cnt));
```

```
    end
    end
end

%% The next step is to obtain the signature of each object that has been
located.
subplotrow = ceil(sqrt(N));
for cnt=1:N;
    boundary=B{cnt};

[st, angle, x0, y0]=signature(boundary);
figure(2)
subplot(subplotrow,round(N/subplotrow),cnt);
plot(angle,st);
title(['Edge Signature of Object ', num2str(cnt)])
xlabel('Angle (degrees)')
ylabel ('Centroid to Edge (pixels)')
axis([0,360,0,100])

%Find the numerical values of results
[pks1,locs]=findpeaks(st,'minpeakdistance',1);
[pks2,locs]=findpeaks(st,'minpeakdistance',5);
hold on;
%plot(angle(locs),pks1);
plot(angle(locs),pks2);

format short;

NumberofPeaks=length(pks1)
PeaksPerDegree=length(pks1)/360

NumberofPeaks5=length(pks2)
PeaksPerDegree5=length(pks2)/360

MeanDistance=mean(st)

MAX=max(st);
MIN=min(st);

DifferenceMaxMin=MAX-MIN

end
```