A Thesis

Entitled

An Automated Approach of Tile Drain Detection and Extraction Utilizing High Resolution Aerial Imagery and Object-Based Image Analysis

By

Richard A. Johansen

Submitted to the Graduate Faculty as partial fulfillment of

the requirements for the Master of Arts Degree in Geography

_____
Dr. Kevin Czajkowski, Committee Chair

_____
Dr. Patrick Lawrence, Committee Member

_____
Dr. Dan Hammel, Committee Member

_____
April Ames, Committee Member

_____
Dr. Patricia R. Komuniecki, Dean
College of Graduate Studies

The University of Toledo
May 2015

An Abstract of

An Automated Approach of Tile Drain Detection and Extraction Utilizing High
Resolution Aerial Imagery and Object-Based Image Analysis

by

Richard Johansen

Submitted to the Graduate Faculty as partial fulfillment of
the requirements for the Master of Arts Degree in Geography

The University of Toledo
May 2015

Subsurface drainage is known to adversely impact the water quality and contribute to the

formation of harmful algal blooms (HABs). In early August of 2014, a HAB developed in

the western Lake Erie Basin and resulted in over 400,000 people being unable to drink

their tap water. HAB development is aided by excess nutrients from agricultural fields,

which are transported through subsurface tile and enter the watershed. Compounding the

issue, the trend has been to increase the installation of tile drains in both total extent and

density. Due to the immense area of drained fields, it is necessary to establish a cost-

effective technique to monitor tile installations and their associated impacts.

This thesis aimed at developing an automated method in order to identify subsurface tile

locations from high resolution aerial imagery by applying an object based image analysis

(OBIA) approach utilizing eCognition. This process was accomplished through a set of

algorithms and image filters, which segment and classify image objects by their spectral

and geometric characteristics. The algorithms utilized were based on the relative location

of image objects and pixels in order to maximize the robustness and transferability of the

final rule-set. These algorithms were coupled with convolution and histogram image

filters to generate results for a 10km² study area located within Clay Township in Ottawa County, Ohio.

The eCognition results were compared to previously collected tile locations from a concurrent project that applied heads-up digitizing. The heads-up digitized locations were used as a baseline for the accuracy assessment. The accuracy assessment generated a range of agreement values from 67.20% - 71.20%, and an average agreement of 69.76%. The confusion matrices calculated a range of kappa values from 0.273 - 0.416 with an overall K value of 0.382, considered fair in strength of agreement. This thesis provides a step forward in the ability to automatically identify and extract tile drains and will assist future research in subsurface drainage modeling.

Acknowledgements

Table of Contents

List of Tables

List of Figures

List of Equations

Chapter 1

1 Introduction

In August of 2014, a large harmful algal bloom (HAB) developed in the western Lake Erie Basin that resulted in over 400,000 people in the greater Toledo, Ohio area being unable to drink to their tap water (Figure 1.1). Harmful algal blooms are not uncommon in Northwest Ohio and have serious health and economic impacts on the community. In an open letter to federal officials Toledo's Mayor, Michael Collins, addressed the emergency with this statement.

"In the early morning hours of Saturday, August 2nd, until the morning of Monday, August 4th, a half million residents of Northwest Ohio and Southeast Michigan experienced the unthinkable; they were told not to consume tap water. While the water is now safe to consume, danger remains lurking off our shoreline in the form of harmful algal blooms (HABs). Eleven million people rely on Lake Erie for their water supply; millions more receive water from other bodies of water that face the same potential of being impacted by HABs."

Mayor Collins goes on to explain that these events have increased in the recent years, and calls upon the President to help protect the nearly 84% of United States' fresh water supply that rests in the Great Lakes (Collins, 2014).

Figure 1.1 Lake Erie Algal Bloom in August, 2014 (NASA, 2014)

Agricultural fertilizers and nutrient applications are speculated to play a major role in the development of the HABs, because the nutrients that are designed grow healthy crops also aid in the growth of the harmful and nuance algae. The nutrients may be carried off of the fields by means of surface or subsurface drainage, but for this thesis the main concern was subsurface drainage, also known as drainage tile or tile drains.

It is important to investigate subsurface tile because the exact extent to which these systems contribute to HABs is unknown. Secondly, monitoring tile drainage runoff is extremely difficult, due to the fact there has been little data collected on where tile have been installed, fixed, removed, and even less information on where tile have been previously installed (Fausey et al., 1995; Jaynes and James, 1987; OEPA, 2011).

This thesis investigates the history of water management with an emphasis on subsurface tile drainage, the environmental impacts associated with subsurface drainage, previous

methods of tile detection and extraction, and an attempt to develop a systematic approach that improves on previous methods of tile drain detection and extraction.

1.1 History of Water Management

For millennia mankind has had success at managing the flow of water through drainage systems. Archeologists have evidence of many civilizations controlling water flow through various practices over human history. The Greeks and Egyptians used surface drainage techniques to remove excess water from their land roughly 2500 years ago. The Romans were the first to utilize open and closed drainage systems to remove stagnate surface water, and these methods were the dominant means of drainage for more than a 1000 years (Donnan, 1976). Weaver (1964) suggested that clay tile systems, similar in style to those used today, have been utilized for over 2000 years. However, the first documented use of clay tiles was to drain excess water from the garden of the Monastery of Maubeuge in France built in 1620 (Weaver, 1964).

The practice of draining water from the land has continued to become more popular across the globe and throughout the mid-nineteenth century expanded rapidly into an industry. The first clay tile machine was invented and patented in England around 1843, and unofficially started the tile industry (Donnan, 1976). As Europeans began to migrate to America, they slowly began to implement their methods of drainage. Overall, the United States saw very little man-made drainage in its early stages, and the land that was drained was only very small and isolated plots. However, there were some important early attempts on the large scale to convert inhabitable wetlands to more functional land.

For example, the Colony of South Carolina passed an act in 1754 to drain the Cacaw

Swamp. Another attempt in 1763, by George Washington, was to drain the Dismal

Swamp that extended from Virginia to North Carolina (Beauchamp, 1987). Aside from

these attempts, few water management projects were conducted in the United States until

the 1850's.

After many years of debate and pressure from the public, congress recognized the issue

of water management and passed the Swamp Land Acts of 1849 and 1850. Collectively

they were the first federal legislation to be passed relating to draining of the land and

gave fifteen states roughly 64 million acres of land on the condition that the money

generated from the sale of the land would be used on water management projects

(Beauchamp, 1987). The management projects were essential for these states to reclaim

the swampland.

1.2 Subsurface Tile Drainage

The first material used for subsurface drainage was clay tile and were made by hand-

rolling the clay and then baking it in a kiln. This process, along with hand digging the

holes for the tile to be put in the ground, was tedious and back-breaking. There were

some advancements in technology and manufacturing during the late 1800's and early

1900's, but still was expensive and time consuming for farmers to implement. It was

estimated that even with a two man team, this method would only allow 20 to 30 feet of

clay tile to be installed a day (Fouss and Reeve, 1987).

However, a major breakthrough occurred in the 1960's with the booming plastics industry, the corrugated plastic tile (Figure 1.2). Corrugated plastic was superior to tile and other tile alternatives in many ways. First, it was incredibly light weight: 250 feet of plastic weighed only 80 pounds as compared to about 2,000 pounds of the same length and width of clay tile (Fouss and Reeve, 1987). Secondly, the polyvinyl and polyethylene plastic was incredibly strong, due to the structure of the walls, but also flexible enough to be easily bent and moved. Both the weight and the malleability of the plastic tile made it much cheaper than any of the other substitutes. The prices dropped and demand grew over the next twenty years, and by 1983 an estimated 95 percent of the subsurface drainage systems installed throughout the entire United States were corrugated plastic tile (Schwab and Fouss, 1985).

Figure 1.2 Corrugated Plastic Tile   (Niagara Block, 2014)

The advancements in plastic tubing coupled with the advancements in machinery caused a revolution in the industry (Figure 1.3). High speed plows and trenchers that are capable of digging, installing, and covering tile at a rate of 80-150 feet per minute is now common practice (Fouss and Reeve, 1987). The speed at which these machines could

operate is faster than most farmers were able to control, so another major advancement

was added to plows, a laser guided controller (Fouss et al., 1972). Agricultural drainage

practices are continually being updated and refined, and because of this, subsurface tile

systems are constantly being installed to fields in both quantity and density.

Figure 1.3 Tile Plow (Reynolds, 2014 p.4)

The reasons for installing tile drainage systems are clear and the benefits fall into two

broad categories, water management practice and land improvement technique. As a

water management tool, tile drains remove excess surface water, reduce risk, mitigate

illness and disease caused by stagnate water, and reduces runoff and erosion. The benefits

to the land are the removal of excess salts, better crop protection against pests, and

increased crop productivity (Fausey et al., 1987). Subsurface tile drainage has allowed

the Midwest region of the United States to thrive, and as a result, continue to generate

very productive and abundant crops for the entire world. Fausey et al. (1995) explain that

the Great Lakes and the Corn Belt States have over 20 million hectares of land that is

currently drained via subsurface drainage systems, and accounts for over a third of the

country's cropland (Figure 1.4).

However, over the last few decades many have begun to study the adverse impacts that this practice has on the environment and water quality. More specifically, how do subsurface drainage systems transport nutrients, nitrogen (N) and phosphorus (P), into the watershed and what are the associated impacts? (Ahiablame et al., 2011; Alexander et al., 2008; Dils and Heathwaite, 1999; (Fausey et al., 1995) Gentry et al., 2007; King et al., 2014; McDowell and Sharpley, 2001; Smith et al., 2014).

Figure 1.4 Extent of Corn Belt and Great Lakes (Holtgrieve et al., 2012 p.152)

1.3 Problem Statement and Objectives

As stated above, tile drainage systems provide many benefits both ecologically and economically, but recent studies have shown that drainage tile have negative effects as well. These adverse impacts are a result of excessive nutrient build-up on the soil, mainly nitrogen and phosphorus, that runoff directly into the watershed. The executive order of 1977 was the first attempt to monitor the impacts that subsurface drainage has on the loss

7

of wetland, and has subsequently laid the foundation of water quality and management

research (Fausey et al., 1995). For example, much of the excess nitrogen from cropland in

the Midwestern states has ended up in the Mississippi, which resulted in hypoxic

conditions in the Gulf of Mexico (Figure 1.5). The impacts are hard to quantify, because

there is a lack of prior knowledge in regards to tile quality and quantity (Jaynes and

James, 1987). Another major concern is the loss of phosphorus from cropland that is

transported into the watershed, which can wreak havoc on the environment and water

quality (Foy and Withers, 1995; McDowell and Sharpley, 2001; Sharpley et al., 2000).

Figure 1.5 Hypoxic Conditions in the Gulf of Mexico (NOAA, 2014)

A system of detecting and monitoring tile drains would allow officials to better predict

and mitigate any adverse environmental or economic impacts. More accurate data would

also assist farmers in improving their best management practices. These techniques

would allow farmers to be able to identify broken or missing tiles, which would lead to an increase in land productivity.

Previous research has been successful at determining the most likely drained areas through satellite imagery based on soil characteristics and land cover. Others studies have been able to extract tile locations through small field-size studies with relatively low accuracy (Brown, 2013; Dezső et al. 2012; Naz and Bowling, 2008; Northcott, 2000; Reynolds, 2014; Sugg, 2007). However, no research has been able to combine these methods into one approach at the scale needed. This research attempts to create an approach to automatically identify agricultural fields and extract tile drain locations through the use of high-resolution satellite imagery and object based image analysis.

Objectives:

1. Develop a working algorithm in the OBIA software eCognition.

2. Apply the algorithm in order to accurately and cost effectively extract tile drain data.

Chapter 2

2 Literature Review

Subsurface drainage systems have a long history and can result in monumental change to the physical landscape, as seen in the transformation of The Great Black Swamp to highly productive farmland in roughly a century (Kaatz, 1955). More recently, research has focused on the environmental implications of tile drains and drainage practices, but the lack of geographic information regarding the tile make any predictions about the extent of environmental degradation incredibly difficult. The early method of tile detection was through tile probing, which is an in-situ process where an individual tests, with a metal probe, for the tile location in a field. Tile probing is time intensive, physically demanding, and only practical on field-sized sites. The advancements in computing and remote sensing techniques have created a new avenue for detecting tiles via satellite imagery and high-resolution aerial photographs.

2.1 The Great Black Swamp

One particularly harsh swampland located in Northwestern Ohio was called The Great Black Swamp (Figure 2.1). The Great Black swamp covered an area of 1,500 square miles extending from the Western Lake Erie Basin well into Indiana (Kaatz, 1955). Kaatz

describes the area as "one continuous region of standing water or so wet as to ooze water when walked upon in all seasons except the very driest". Naturally, the areas with the most efficient natural drainage were developed first, but this left large a portion of NW Ohio undeveloped through much of the nineteenth century. Efforts to drain the area with ditches and furrows were employed, but were not effective enough to remove all the excess water. Subsurface tile drainage systems were a necessity, and the tile industry that recently migrated to the US began to thrive in the late 1800's. The Black Swamp region saw a surge of tile manufacturing facilities by the end of the 1870's totaling over fifty in the region alone (Kaatz, 1955). The tile drain demand continued and over only a few decades converted nearly one million acres of unusable swampland to one of the most productive agricultural lands in the world (Fausey et al., 1995).
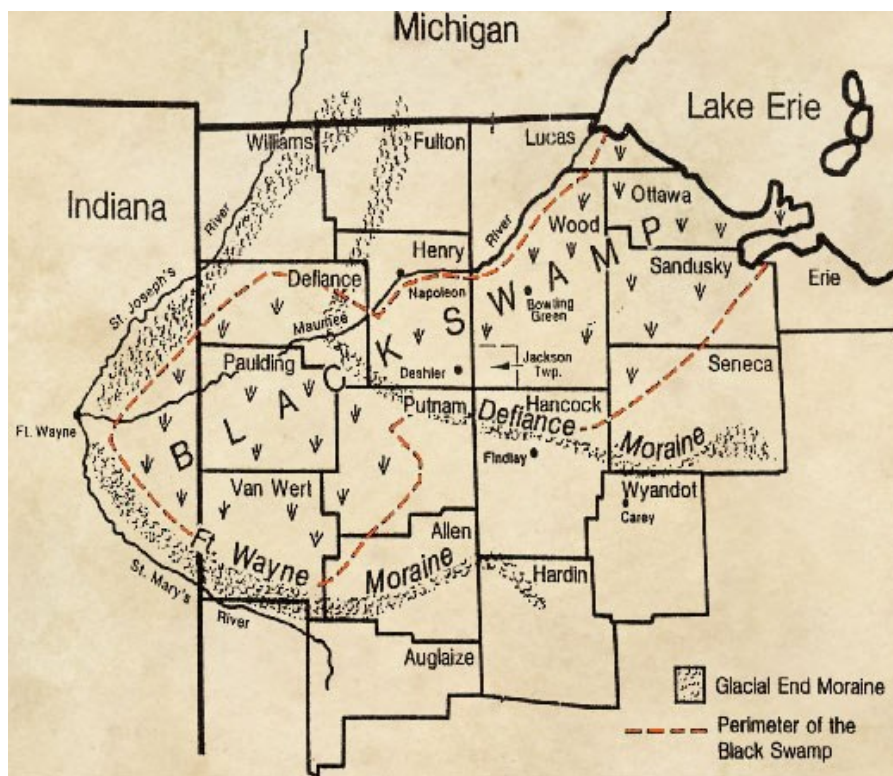

Figure 2.1 The Great Black Swamp (The Black Swamp Conservancy, 2014)

2.2 Environmental Impacts

There has been a great deal of research conducted on the transportation of nutrients originating from agricultural fields that ultimately affect water quality. Since subsurface drainage has become the dominant means of removing excess water from cropland, recent environmental studies have begun to focus on the impacts drainage tile pose to water quality.

An early investigation focused on subsurface tile drainage was conducted by Kladivko et al. (1991). This research examined the influence tile spacing had on the overall water quality, measured through water runoff, nutrient loading, and pesticide transport. This research concluded that narrower spacing (6 meters) resulted in a lower water quality than tiles with larger spacing (12 and 24 meters). However, trends show that farmers are installing tile at increased quantity and density, which could result in even more stress to the ecosystem and poorer water quality (OEPA, 2011).

Dils and Heathwaite (1999) investigated the role of subsurface tile with regards to the transport of phosphorus (P). The P lost through subsurface tile represented a very small percentage of the total P needed for healthy crops, so from a farmer's economic perspective this has a minimal impact on their harvest or budget. However, this small percentage may be detrimental to the health of the watershed. The authors also made the conclusion that subsurface tile systems create networks that connect vast areas of agricultural land and aggregate all of the phosphorus lost into the same watershed. The authors acknowledge that there is a lack of locational information regarding the large tile networks and the impacts at the network level are difficult to quantify.

McDowell and Sharpley (2001) studied and quantified the amount of phosphorus lost from varying agricultural soil types. These soils have built-up concentrations of P from continued applications of fertilizers, and the high concentrations of P eventually percolate through the soils and are transported through tile drains. The main focus was on the concentration levels of fertilizer applications and the impact of different soil characteristics impact on P lost through drainage.

Gentry et al. (2007) examined the transport of phosphorus in agricultural areas by subsurface tile throughout three study areas in East-Central Illinois, and investigated how precipitation events impact the flow rate of subsurface tile and subsequently the transportation of P. The authors concluded that during extreme precipitation events tile played less of a role than during dry years. This is due to the surface runoff that occurs during a large storm, which quickly drains off of cropland. However, the P lost due to the surface runoff will likely flow into a drainage ditch, stream, or other water bodies, and eventually work its way into the watershed. This study also observed that after the first flush, the initial surface runoff of a precipitation event, P concentrations were found to remain high, unlike nitrogen (N) or herbicides. This means that P may remain in soils and concentrate over a long period of time and exemplifies the complexity of tile drainage effects on differing nutrients, chemical compounds, and their transportation.

Recent studies have looked explicitly at the impacts of P loss through tile drainage and the potential contribution to harmful algal blooms (HABs). Smith et al. (2014) generated results that have contradicted previous findings, in regards to the influence of P loading from tile drainage. In agricultural fields within the St. Joseph River Watershed they calculated that tile drains produce between 25%-80% of the total P lost. This number is

much higher than what was previously thought, which highlights the importance and influence subsurface tile drains may have on P transportation. King et al. (2014) conducted a study over an eight year period to investigate a watershed in Central Ohio and found that P concentrations significantly exceeded EPA proposed recommendations for the study area. These findings demonstrate that the high phosphorus concentrations are likely to play a major role in the development of HABs in Lake Erie.

The Ohio Lake Erie Phosphorus Task Force Report concluded that at the current estimated P totals of 0.6 to 1.1 kg/ha for cropland surrounding western Lake Erie Basin conditions are capable of producing algal blooms (Ohio Phosphorus Task Force, 2013). It is also recommended that spring P loading be decreased by 40% in order to reduce the degree and frequency of HABs.

In a report by the National Center for Water Quality Research (NCWQR) (2011), where it is estimated 7% of the total P that enters Lake Erie was transported from the Maumee River from point source pollution. This means that nearly 93% of the total P that is deposited in Lake Erie is from non-point source pollution such as farm fields.

Another survey conducted by the Great Lakes Protection Fund (2010) investigated 657 agricultural fields with an average size of 25.2 acres in the Sandusky Watershed. This survey gives insight into the actual numbers of fields containing subsurface drainage. Nearly 91.5% of the fields surveyed contained some degree of subsurface tile coupled with 74.6% of field edges within 1000 feet of a watercourse (field ditch, stream, or river). The combined efforts of farm field surveys, research on nutrient transportation, and knowledge of HABs production, has laid the foundation for better water quality

management. However, without the knowledge of tile locations, average spacing, or density it is difficult to calculate the contributions to algal blooms as well as enforce any regulations at the watershed level, so a method of tile identification and extraction must be created.

2.3 Previous Methodologies in Tile Detection

Verma et al. (1996) attempted to map tile lines within Vermilion County, Illinois using color infrared (CIR) aerial photographs. Tile drains remove excess water from soil directly above the tiles faster than the soil in-between tiles, which creates spectral differences between the wet and dry soils that can be seen with the naked eye (Hoffer, 1972). However, the aerial photographs must be taken during ideal conditions in order to generate the most effective results. The ideal conditions are two to three days after a rainfall of 2.54cm (Verma et al., 1996). The photographs were scanned and geo-referenced using IDRISI in order to allow for geographically precise data manipulation. Then through applying various band combinations, the aerial photographs were analyzed to demonstrate each combination's effectiveness. The bands utilized for this study were bands 1, 2, and 3, representing blue, green, and red. The best results were obtained through the product of band 2 and band 3, which was considered "excellent, and the quotient of band 2 and band 3, which was considered "good". A classification scheme was produced using the combination, band 2 multiplied by band 3, and the study area was broken in three classes, wet, moist, and drained. An accuracy assessment was conducted using the tile probing method. Results showed that main lines were detected almost every

time on the first attempt. However, lateral lines were more difficult to detect due to their

smaller sizes. This research demonstrated successfully that CIR photographs could be

utilized in the detection of tile drains, but lack quantitative rigor in the results and

accuracy assessment.

Northcott et al. (2000) took a similar approach using CIR imagery and GIS layers (Soil

properties and hydrologic features) to delineate subsurface tile systems in East Central

Illinois. This process included scanning and geo-referencing aerial photographs and then

overlaying the GIS layers to determine most likely areas where tiles would occur. The

research utilized a technique called heads-up digitizing, which a user manually "draws"

lines on a computer screen within the viewing window where the drains are located. The

digitizing process creates shapefiles for each of the individual tile lines while maintaining

their geographic properties, which are essential for any data analysis. No formal

accuracy assessment was undertaken for this research. The results produced a map of tile

drains within a comparatively large study area. However, this process requires manually

scanning through imagery to detect the variations in soil reflectance, and is too time-

consuming for larger study areas.

Sugg (2007) investigated the extent of subsurface tile drainage to fill the void in data that

has been spotty at best. This research shows that large scale attempts to collect

information about tile drains have been conducted via census surveys by the government

but have been eliminated resulting in a data gap. Sugg states "no truly comprehensive

information on the status of agricultural drainage has been published since the

aforementioned 1987 USDA report" (Sugg, 2007, p.2). This study examined eighteen

states across the United States to generate what areas would most likely contain subsurface tiles and what areas could benefit the most with the installation of new tile drains. The process was undertaken by overlaying the land cover classification from the 1992 National Land Cover Dataset (NLCD), soil information from USDA's State Soil Geographic Database (STATSGO), and county-level Soil Survey Geographic Database (SSURGO). The results were generated specifically through the five poorest drained soil types and row crop data, but the methodology makes two assumptions. First, that subsurface drainage is the only drainage practice used in these locations, and secondly that subsurface tile systems are not in locations with better natural drainage.

Brown (2013) extended Sugg's method by obtaining and comparing eight multi-scale and multi-temporal GIS layers over Central Minnesota. The study compiled land use data from the NLCD (2001) and NASS (2008), Soil and hydrology properties from SSURGO, and slope information from the United States Geologic Survey (USGS) to predict most likely locations that contain artificial tile drainage.

Naz and Bowling (2008) developed an automated process of detecting tile drainage from remote sensing data in Tippecanoe County, Indiana. The research applied a decision-tree classification (DTC) methodology to examine various GIS layers, including land cover, soil drainage properties, and slope. The imagery was then processed using a filter technique, spatial convolution, in order to divide the image and to enhance the edges between high and low frequency areas. The results of the filters were then classified into two classes, tile and non-tile, based on the density slice classification. An accuracy assessment was conducted with previously known tile locations within the study area. The Hough transformation produced the most accurate results and the least amount of

discontinuity between tile lines, which in turn reduces the amount of user-time spent connecting tile lines. Another study by Naz et al. (2009) utilized the same methodology of DTC and image processing tools to carry out research in another study area in West Central Indiana. The replicability of this study is important to demonstrate that techniques for tile detection can be implemented in various geographic locations.

Thompson (2010) continued the work of others within this field through the use of National Agricultural Imagery Program (NAIP) data and performing an unsupervised classification. Next, the author applied an edge detection technique to identify sharp contrast in the data in order to detect tile locations. Finally, he applied the heads-up digitizing process to generate shapefiles of the tile drain locations within Wood County, Ohio. The shapefiles were compared to geo-referenced tile drain blue prints to measure accuracy. A comparable methodology was used by Andrade (2013) to investigate the ability of mapping tile drain locations in the Eagle Creek Watershed in Iowa. This research also took into account 10-day rain averages to determine likely locations of tile drains, because the ability to utilize remotely sensed images is highly dependent on precipitation amounts just prior to the collection of the imagery (Andrade, 2013).

2.4 Object Based Image Analysis & eCognition

Over the last couple decades a transition has been occurring in the field of remote sensing. Publications such as *What's Wrong with Pixels* by Blaschke and Strolbe (2001), sum up the dissatisfaction with the limitations of pixel-based image analysis on very high resolution imagery. This new phenomenon started with emergence of easily accessible

high resolution images, and is most effective on imagery with spatial resolutions under five meters (Hay and Castilla, 2006). The main processing tool of object based image analysis (OBIA) is segmentation, a technique that divides an image into sections or segments similarly to the way the human brain would process an image. A segment is a region that is homogenous in at least one characteristic (Blaschke, 2009).

The main advantages of OBIA over pixel-based:

1. The segmentation process most resembles the way humans organize and comprehend images.
2. Image objects are less affected by the modifiable aerial unit problem (MAUP) than their pixel-based counterpart.
3. OBIA utilizes object features (shape, texture, and spatial relations) beyond just the spectral properties.
4. This technique dramatically reduces the computational load placed on an operating system.
5. OBIA software has become more powerful and more affordable (eCognition). (Hay and Castilla, 2006)

This is not a comprehensive list of advantages, but broadly covers the key points important for this research. Since the early 2000's the number of scientific publications using OBIA has increased dramatically in a wide variety of disciplines (Blaschke, 2009). Krause et al. (2004) applied OBIA on aerial photographs to detect changes in both time and scale on mangroves in Northern Brazil. OBIA software has been successful in mapping hydrologic soil properties demonstrated by Corbane et al. (2008). Thomas et al.

(2003) successfully extracted land-cover/ land-use (LC/LU) data from very high resolution imagery, less than one meter, and estimated storm-water runoff in Scottsdale, Arizona. One of the most important aspects of OBIA methodology is the ability to develop rule-sets that may be transferable. This transferability was demonstrated by Schopfer and Moller (2006) and Walker and Blaschke (2008), and in theory allows a standardized set of rules or parameters to be developed and applied across different spatial and temporal scales. However, OBIA techniques at this point are still highly dependent on the dataset. These are only a handful of examples of the work being done with the advancements with OBIA. For a comprehensive list of examples see Blaschke (2009).

At the beginning of the OBIA craze there was no single best software in order to accomplish the sophisticated computational tasks needed to process high resolution imagery. The software previously available was not user-friendly and was expensive. However, in late 2000 Definiens Imaging GmbH developed software, eCognition, to address the growing demand for OBIA and the ability to analyze geometric properties and spatial relations of pixels (Flanders et al. 2003; Trimble, 2014). eCognition has proven to be a popular software for conducting classification trees and rule-based research as demonstrated by Flanders et al. (2003) who conducted research on cut block delineation in Calgary, Canada to Dezső et al. (2012) who applied multiple segmentation techniques with eCognition to delineate land use classifications.

There are two broad segmentation methodologies, merge-based or cut-based. In merge-based, or the bottom-up approach, the user starts at the pixel level and through parameters, generated by rule-sets, aggregates the pixels into groups of similar

characteristics. This grouping process is continued until the classification is complete. Dezső et al. (2012) demonstrated the flexibility of the software and the many different techniques available to achieve the same result. Their research applied the sequential linking, best merge algorithm, and graph based merge to effectively show a variety of bottom-up techniques. The second process is called cut-based, or top-down approach, which as expected, works in the reverse of the cut-based method. The cut-based method starts with the entire image as one segment and then through iterations continually subsets the image into many smaller segments until the image is classified appropriately. Three techniques available that utilize this approach are minimum mean cut algorithm, minimum ratio cut algorithm, and normalized cut algorithm (Dezső et al. 2012).

Reynolds (2013) used the eCognition software to develop rule-sets in order to automatically detect tile drains systems over five independent fields using high resolution aerial photography of Northwest Ohio. Following the detection, an extraction technique was implemented within the eCognition software to transfer the data from raster to a vector format. The method was successful in the overall goal of detecting and extracting tile drain locations. However, there are two issues with this research that must be addressed: the study only investigated an individual field at a time, and the rule-sets were changed due to localized variations. The rule-sets developed in this study are extremely difficult to transfer to different geographic locations and scales. A more rigorous set of parameters must be established that can accommodate for local variation and spatial differences.

eCognition has been well established as an excellent tool to accomplish OBIA research, as seen through previous studies and the rapid growth of peer-reviewed journal article

publications over the last twenty years (Beck et al. 2013; Blaschke, 2009; Definiens, 2009; Dezső et al. 2012; Flanders et al. 2003; Reynolds, 2014). With the newly released eCognition software and easily accessible high resolution imagery, it is possible to fill the void in both knowledge and methodology by developing a rigorous technique of automatically detecting and extracting tile drains through high resolution aerial photography and satellite imagery.

Chapter 3

3 Methodology

3.1 Study Area

This research attempted to develop a replicable rule-set for extracting field locations, containing tile drainage systems, as well as the underground tile drainage structure. The study area was a subset of Clay Township, which lies within Ottawa County, Ohio (Figure 3.1). This area was chosen to correspond to a concurrent research project, under the supervision of Dr. Kevin Czajkowski, Mapping Drain Tile and Modeling Agricultural Contribution to Nonpoint Source Pollution in the Western Lake Erie Basin at The University of Toledo. Another important set of characteristics of this study area are the proximity to Lake Erie and that the landscape is dominated by agriculture fields. These two features are crucial for this research because they connect the impact of agricultural runoff and harmful algal blooms. Finally, the large variations contained in this study area make for a more robust rule-set that is able to be reproduced in other geographical locations.

Figure 3.1 Study Area in Clay Township, Ottawa County, Ohio

The image above shows the extent of the study area chosen for this research. The first

subset map is of the state of Ohio and the grey portion is Ottawa County (bottom left).

The second subset map is of Ottawa County with the grey portion containing Clay

Township (bottom middle). The final subset map displays the extent of Clay Township

and contains the previously collected hand-digitized tile drain information for this area

shown in red. This map contains the outline of the study area for this research (bottom

right). The main image is the OSIP high resolution aerial image of the study area, shown

in true color at a scale of 1:30,000 (top). This imagery was utilized throughout the

development of the algorithm process.

3.2 Imagery

High-resolution imagery available for analogous research is available freely at The Ohio

Statewide Imaging Program (OSIP) and the National Agricultural Imagery Program

(NAIP). Images from OSIP are available in both red, green, blue (RGB) images and color

infrared (CIR) at one foot and one meter resolution respectively. However, for this

research, the best results were generated from RGB imagery from OSIP (figure 3.2). A

concern with OSIP imagery is the lack of metadata. There is no information contained in

the image to describe the time and conditions of when the imagery was captured. This is

critical for identifying tile drainage since, as stated in section 2.1, the conditions needed

to detect tile are uncommon. It is known that the imagery has been collected during the

leaf-off period, from March to April, and the majority of fields are bare soil with the

exception of winter crops.
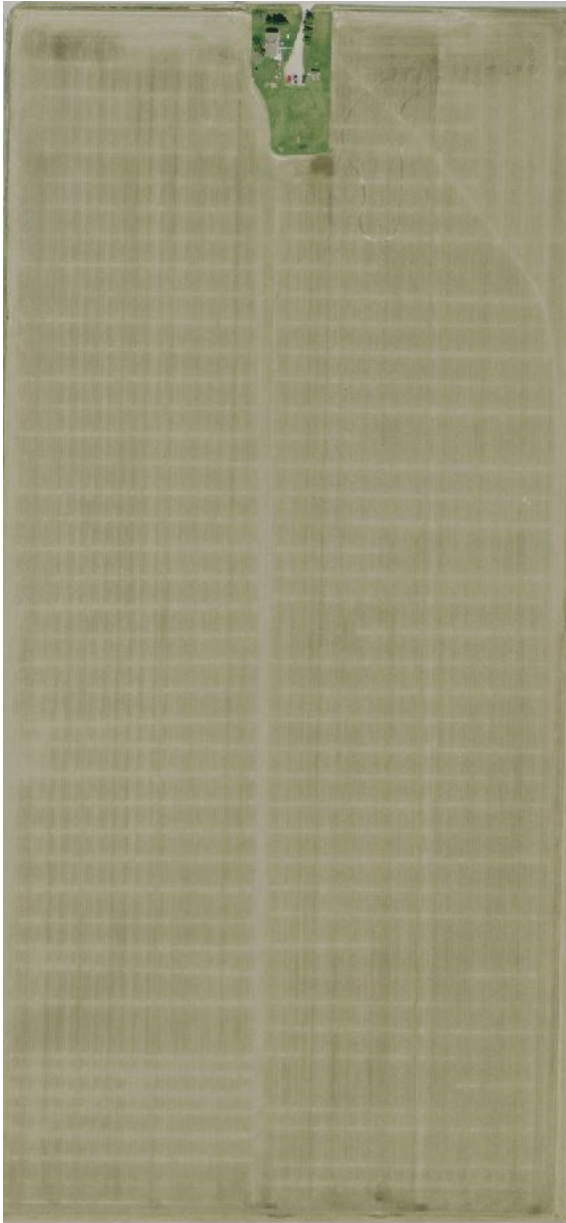
Figure 3.2 Example of OSIP Imagery with RGB Bands displayed as true color

The images examined contain three bands red (R), green (G), and blue (B), and each of

the bands are imported as a separate layers within eCognition. The software is capable of

analyzing characteristics of each layer, individually or any combination of the three, in

order to extract the pertinent information needed.

3.3 Software

3.3.1 ArcGIS

ArcGIS is an ESRI product package that contains the software ArcMap 10.2, which was utilized throughout this project. This was the software needed to complete the hand-digitized mapping of the previously collected tile locations from the research project Mapping Drain Tile and Modeling Agricultural Contribution to Non-point Source Pollution in the Western Lake Erie Basin. ArcMap 10.2 was also the main component used in conducting the accuracy assessment, in which the results from the eCognition rule-set were validated with the previously collected data.

3.3.2 eCognition

eCognition is a software currently owned by Trimble Navigation Ltd, and is essential for this object-based research. In chapter 2, the importance of OBIA was highlighted and many examples of how this technology has been applied were cited. The process uses various segmentation and classification techniques to detect spatial, spectral, and geometric differences between pixels and image objects. The software also allows the user to develop rule-sets that run linearly from one algorithm to the next. This is important for non-experts and makes for a more seamless and user friendly interface. Another important aspect of eCognition is the ability to export the data in a vector format, which is easily transferable to ArcMap and other GIS software.

3.4 Hand-Digitized Lines

Dr. Czajkowski's project mapped an estimated 2,000 agricultural fields that contained tile drains, by utilizing the heads-up digitizing approach on high-resolution aerial imagery within ArcMap 10.2. This process was completed by manually scanning through images to identify fields containing tile and drawing individual polylines for each tile visible in every field surveyed. The tile lines were observed by the contrast between wet and dry soil, which typically follows one of two patterns, checkerboard or fishbone (Figure 3.3-3.4). However, fields may also contain random patterns that increase the difficulty of tile detection (Figure 3.5). Once a field has been confirmed to contain tile the technician used the edit toolbox to hand draw the tile lines in the viewing window. The hand-digitizing method is relatively simple when tile are clearly displayed on the viewing window, but the ideal conditions are rare and results may contain errors.
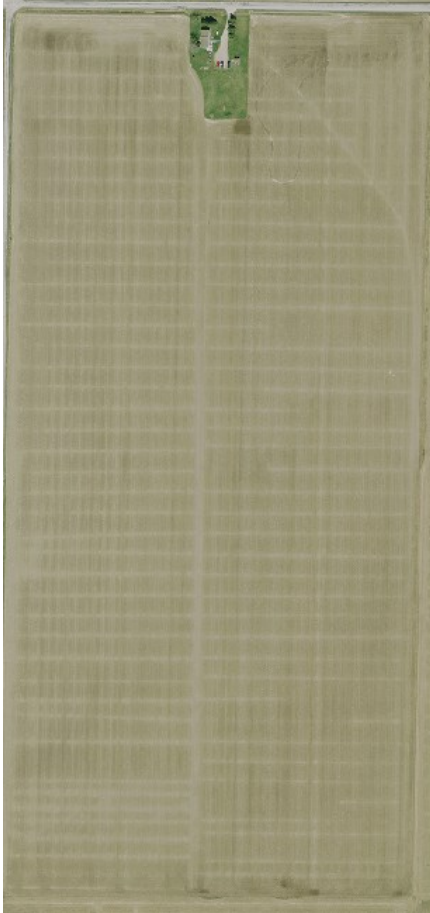
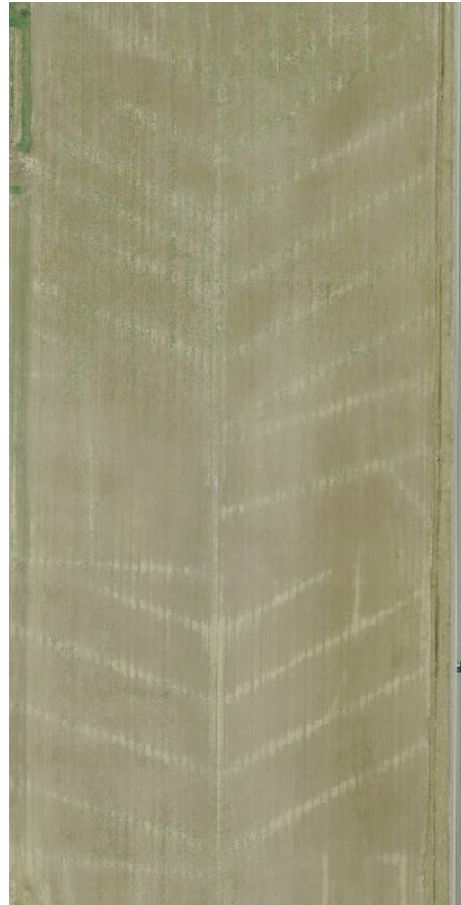Figure 3.3 Checkerboard Tile Pattern

Figure 3.4 Fishbone Tile Pattern

Figure 3.5 Random Tile Pattern

The two major errors that occurred during the hand digitizing process were caused by fields that contained plow lines or fields exhibiting minimal variations in soil brightness. Plow lines are lines that are created after a tractor or combine crosses a field, and the markings left behind are nearly identical to the patterns of tile drains. One way to distinguish between plow lines and tile lines is the spacing between lines, because tile drains are normally spaced much further apart than the lines generated from farming equipment. However, when spacing of tile and plow lines are similar, around 10 meters, it is nearly impossible to differentiate. Another source of error occurred when a field contained minimal spectral variation. The human eye is unable to distinguish between areas of nearly identical brightness, and the lack of variation may result in misplaced or missing tiles.

In order to overcome these errors, multi-temporal imagery was acquired to help determine difficult fields. The users compared multiple images and made use of color-infrared imagery to draw tiles on a single field, while maintaining the integrity of the geographic information of the line. After all fields were hand-digitized they were checked for accuracy. This was accomplished by each field being validated by two separate technicians to ensure accuracy. The fields were judged on the total percentage of tile mapped per field. The project mapped roughly 2,000 fields, 500 per county, and each county was required to have an accuracy of at least 95%. Due to the high level of accuracy of this project, this dataset was used as a baseline in comparing the results of this thesis in the accuracy assessment.

3.5 Developing a Rule-Set

The goal of this research was to develop a transferable and efficient rule-set, or

algorithm, that successfully identifies agricultural fields that contain tile drainage

systems, as well as the tile drainage network on each field. This was completed by

identifying contrasting pixels of wet and dry soil that are indicative of tile drains. The

process generated results that were the combination of prior knowledge and experimental

style trial and error. However, the eCognition software contains over 130 algorithms that

can investigate a multitude of object features characteristics; for a comprehensive list see

Trimble (2012). Each algorithm contains a set of parameters that can be altered which

affects the end result. That being understood, there are many routes to generate nearly

identical results. This research aimed at developing a rule-set that was both efficient and

accurate. Efficiency maybe overlooked in some research, but is essential when analyzing

imagery measured in the gigabytes. Processing power and time is a major limitation

associated with research using high-resolution imagery, and both of these were

considered when the study area and the algorithms were chosen.

3.5.1 Initial Segmentation

The process begins with importing the high-resolution imagery, where each band is

automatically given the names layer 1, layer 2, and layer 3, which correspond to red,

green, and blue, respectively. For this project, the names were left in their default position

for simplicity. This is done under the modify project window, where the user can create a

subset, add thematic layers, and manually edit pixel size. Once the image has been

loaded, the process of developing a rule-set can begin. The first step needed for any rule-

set is an initial segmentation, which aggregates individual pixels into groups based on a

scale parameter and their homogeneity. The primary segmentation used was a multi-

resolution segmentation, which is designed to minimize the average heterogeneity of the

objects at a given scale. This algorithm converts the image from over 110,000,000 pixels

to 142 image objects, or homogeneous group of pixels, and makes the image much easier

to process. Each image object is shown with a blue outline in figure 3.6.
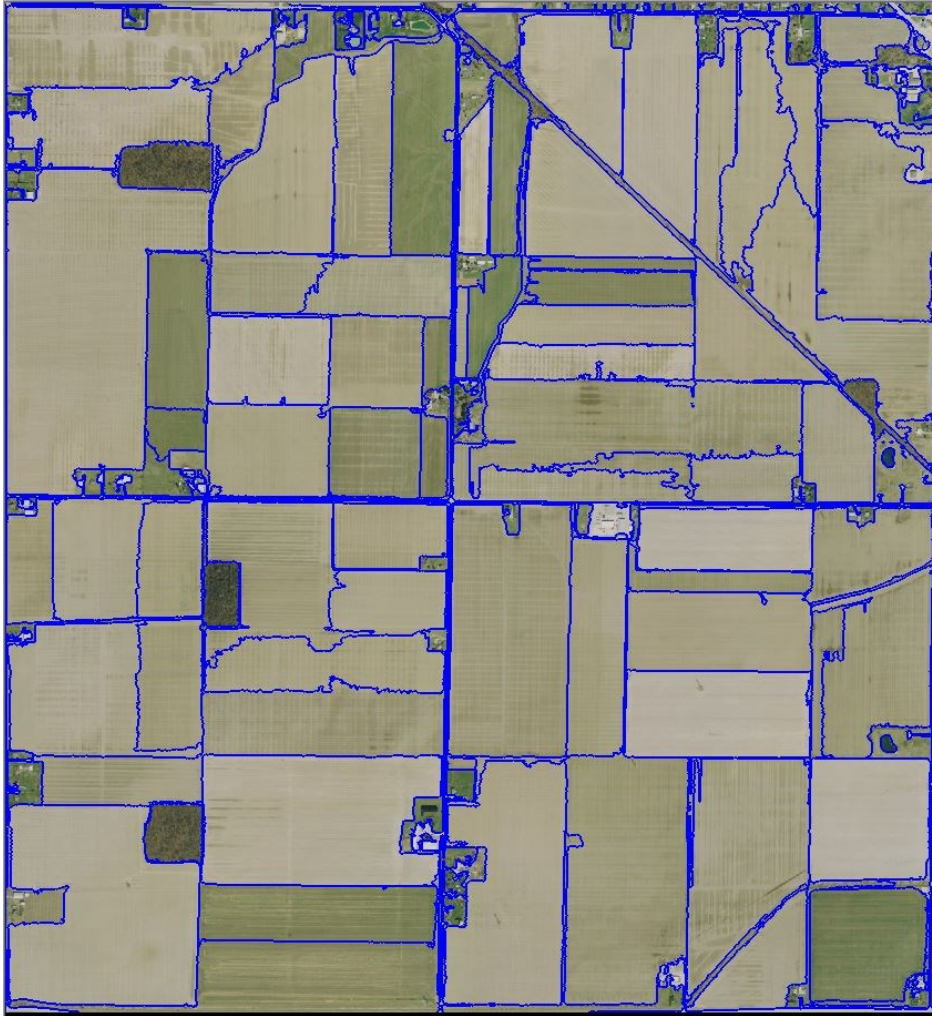
Figure 3.6 Initial Multi-Resolution Segmentation

3.5.2 Brightness Interval Classification

The next step is to classify each of these 142 objects into a brightness interval, a range of

brightness values, in order to separate out agricultural fields from non-agricultural fields.

The brightness values are generated through the calculation below (Equation 3.1):

$$\bar{c}(v) = \frac{1}{w^B} \sum_{k=1}^{K} w_k^B \bar{c}_k(v)\Big|$$

- w_k^B is the brightness weight of image layer k with $w_k^B = \{ {0 \atop 1}$
- K is the number of image layers k used for calculation
- w^B is the sum of brightness weights of all image layers k used for calculation with $w^B = \sum_{k=1}^{K} w_k^B$
- $\bar{c}_k(v)$ is mean intensity of image layer k of image object v
- c_k^{min} is the darkest possible intensity value of image layer k
- c_k^{min} is the brightest possible intensity value of image layer k.|

Equation 3.1 Brightness Algorithm (Trimble, 2014 p. 236)

The brightness intervals created are as follows: $\geq$200, $\geq$190, $\geq$180, $\geq$170, $\geq$160, $\geq$150, $\geq$140, $\geq$130, and $\leq$ 130. It is important to run the classification algorithm in this order because the classifications are based on unclassified pixels. Using a top-down approach effectively classifies all image objects without overlap or reclassification. Each brightness interval represents a range of brightness values, and for simplicity, the ranges are called by the name of the brightness interval, which represent all values that fall within that range. For example, brightness interval 190 represents any brightness value from 190-199, and brightness interval 180 represents any brightness value from 180-189, et cetera. After each brightness interval has been classified, the copy map tool is applied to create a new map for each brightness interval. This step allows for the reclassification of each interval to be undertaken individually, and assists in the processing time required to perform more complex tasks.

3.5.3 Image Filtering

After the brightness intervals are separated, the next step is to apply a series of image filters to each interval. The first image filter is a histogram filter contained under the Layer Normalization algorithm's menu. The histogram process normalizes the image by stretching the values to incorporate the entire pixel value range based on the histogram of the selected image (Figure 3.7). This process essentially highlights the differences between pixels by reducing the possible range of values.

Figure 3.7 Histogram Filter (Trimble, 2014 p. 135-136)

Another image filter utilized in this research is the Convolution Filter, which is designed to minimize the local variation. This is completed by selecting a kernel size, a specified group of pixels in the shape of a square, and then recalculates all the pixels sampled to the average resulting in a smoother image. For this research, the algorithm applied was the Gaussian Blur, which expands the simple convolution filter to generate the new pixel value instead of the average (Equation 3.2). The Gaussian Blur is a popular blurring function that helps eliminate noise and reduce processing time (Figure 3.8).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

(Where σ is the standard deviation of the distribution.)

Equation 3.2 Convolution Filter Algorithm (Trimble, 2014 p. 133)

Figure 3.8 Gaussian Blur (PublicWiki, 2010)

The final filter was applying the convolution filter on the histogram layer. This

maximized the difference between each pixel within the image, and then smoothed those

variations out over a 9x9 kernel. The image filters are essential for this research because

it highlights the contrasts between the wet and dry soils at the pixel level, but then

reduces the tiny variations between pixels and greatly improves the processing ability.

The image filter processes is below shown in figure 3.9-3.11. These figures demonstrate

the importance of each filter and show how each affects the appearance of field tile.
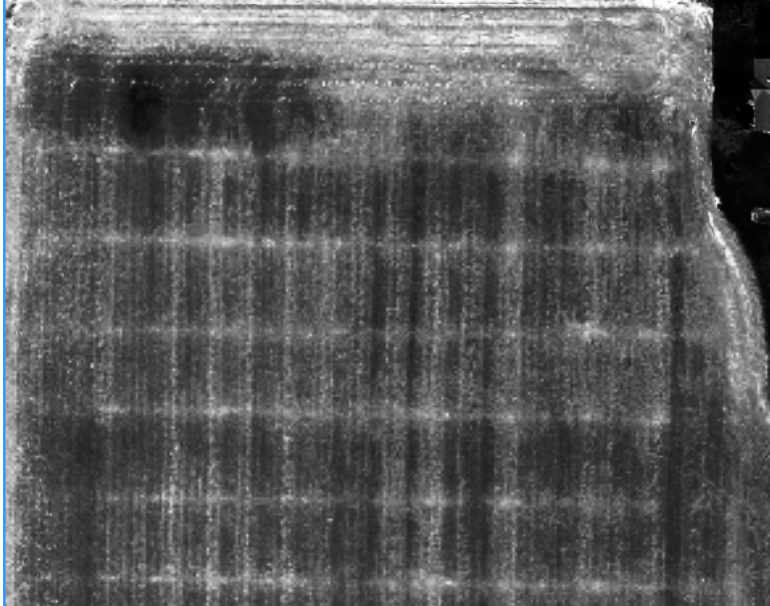
Figure 3.9 Histogram Filter

Figure 3.10 Convolution Filter

Figure 3.11 Convolution Filter on Histogram

3.5.4 Re-segmentation of the Brightness Intervals

The next series of steps involves re-segmenting each brightness interval based on the same multi-resolution segmentation parameters as the original except the scale and layer weights. Each class was re-segmented at the parameter scale of 10, in order to ensure that each brightness interval re-segmented independently (figure 3.12). Another difference on this segmentation is that the image filters are now included in the layer weights. A closer look at one field helps demonstrate the true size of the new segmentation. These are displayed at the field scale (figure 3.13) and tile scale (figure 3.14). These small image objects are the building blocks for the new scheme and through their specific properties are able to classify the image into two classes, Tile or Non-Tile.

Figure 3.12 Re-Segmentation of the Study Area

Figure 3.13 Re-Segmentation at Field Scale

Figure 3.14 Re-Segmentation at Tile Scale

3.5.5 Tile Classification

The final step in the segmentation and classification section is to classify each brightness

interval's newly segmented image objects into one of the two classes. This process was

the most difficult, because the image objects have an insurmountable amount of localized

variation. The soil profile above the tile has different spectral and geometric

characteristics in every field, and many fields have dramatic variations that can occur

within an area less than 100 feet. Campbell and Wynne (2011, p.19) wrote an analogy

that exemplifies the obstacle within this research. The authors discuss the issue of trying

41

to develop a spectral signature for corn. Throughout the life cycle of corn it has many

spectral signatures, and coupled with different soil types, shading, time the image was

collected, etc. it can be nearly impossible to generate a single spectral signature for any

one phenomena. In order to overcome this dilemma, this research made use of a pair of

algorithms that examine image object characteristics relative to their neighbors. This

results in a methodology that can be transferred over geographic locations and through

time series imagery.

This process was accomplished by a small series of steps. The first step was to remove

any object that was larger than 10,000 pixels; this eliminated all fields that were not

associated with the current brightness interval. The next step was to run the assign class

tool on the convolution-histogram layer based on the values of the Border Contrast

algorithm (Equation 3.3). This pixel-based algorithm examines the edge pixels of an

image object and calculates the difference between neighboring pixels.

The border contrast is defined as the mean value of the pixel edge contrasts for all pixel
edges in $E(v)$:

$$bc = \frac{1}{n} \times \sum_{[e=(p,q) \text{ in } E(v)]} c_{k(q)} - c_{k(p)}$$

with $n = \#E(v)$

- Let v be an image object,
- Let $E(v)$ = {set of all pixel pairs (p,q) : p is in $P_{v_{outer}}$ and q is in $P_{v_{inner}}$ and q is in $N_{6(p)}$ } be the set of all pixel edges of v
- For e in $E(v)$ the pixel edge contrast in layer k is contrast(e) : $c_{k(q)} - c_{k(p)}$

Equation 3.3 Border Contrast Algorithm (Trimble, 2014 p. 244)

The output of this algorithm is used to determine which class each image object is to be placed. Any image object that is less than 0 or greater than 10 is considered Non-Tile. The unclassified image objects are then processed by the next algorithm, Contrast to Neighbor Pixels (Equation 3.4). The Contrast to Neighbor Pixel algorithm continues the process of classifying Non-Tile objects by calculating the mean difference in contrast to the surrounding area of a given size.

$$1000 \times \left(1 - \frac{\bar{c}_k \left(B_v(d) - P_v \right)}{1 + \bar{c}_k (P_v)} \right)$$

- $B_v(d)$ is the extended bounding box of an image object v with distance d with $B_v(d)$ equal to $\{(x,y,z) : x_{min}(v) - d \le x \le x_{max}(v) + d \,,\, y_{min}(v) - d \le y \le y_{max}(v) + d \,,\, z_{min}(v) - d \le z \le z_{max}(v) + d\}$
- P_v is the set of pixel/voxels of an image object v
- $\bar{c}_k$ is the mean intensity of image layer k.

Equation 3.4 Contrast to Neighbor Pixels (Trimble, 2014 p. 245)

The results of this algorithm are then used to classify any unclassified object that has a value of less than 10 or greater than 125 as Non-Tile. The final step in this classification is to convert all remaining unclassified image objects to Tile. This series is done for each brightness interval independently (figure 3.15). This image shows a completely classified field with image objects classified as Tile in blue and Non-Tile Objects in grey. A clear systematic pattern of vertical and horizontal tile lines is visible.

Figure 3.15 Rule-Set Final Results displaying tile lines in blue and non-tile areas in gray.

3.5.6 Merge & Export Shapefiles

The final step in the rule-set is to merge all of the tile image objects into one image object. This step dramatically reduces the processing time by converting the more than 100,000 image objects into one. An example of the full rule-set, for brightness interval 130, is shown in figure 3.16. Once this has been completed for each brightness interval the merged image object is exported into a vector format, as smoothed polygons, suitable for ArcMap 10.2, and where the accuracy assessment can be conducted.

Figure 3.16 Complete Rule-Set for Brightness Interval 130

3.6 Accuracy Assessment

An accuracy assessment was conducted to evaluate the effectiveness of the eCognition rule-set through two methods, image object identification and tile verification. The first method was to determine the number of correctly identified image objects for each brightness interval. This process was conducted with a simple visual test, using the high-resolution aerial imagery, and the results were classified as either correctly identified or incorrectly identified. These results allow for a simple method to calculate the algorithm's ability to detect tiled fields in each brightness interval.

The tile verification approach looked at the generated shapefiles from the eCognition rule-set and compared them to the previously collected hand-digitized ti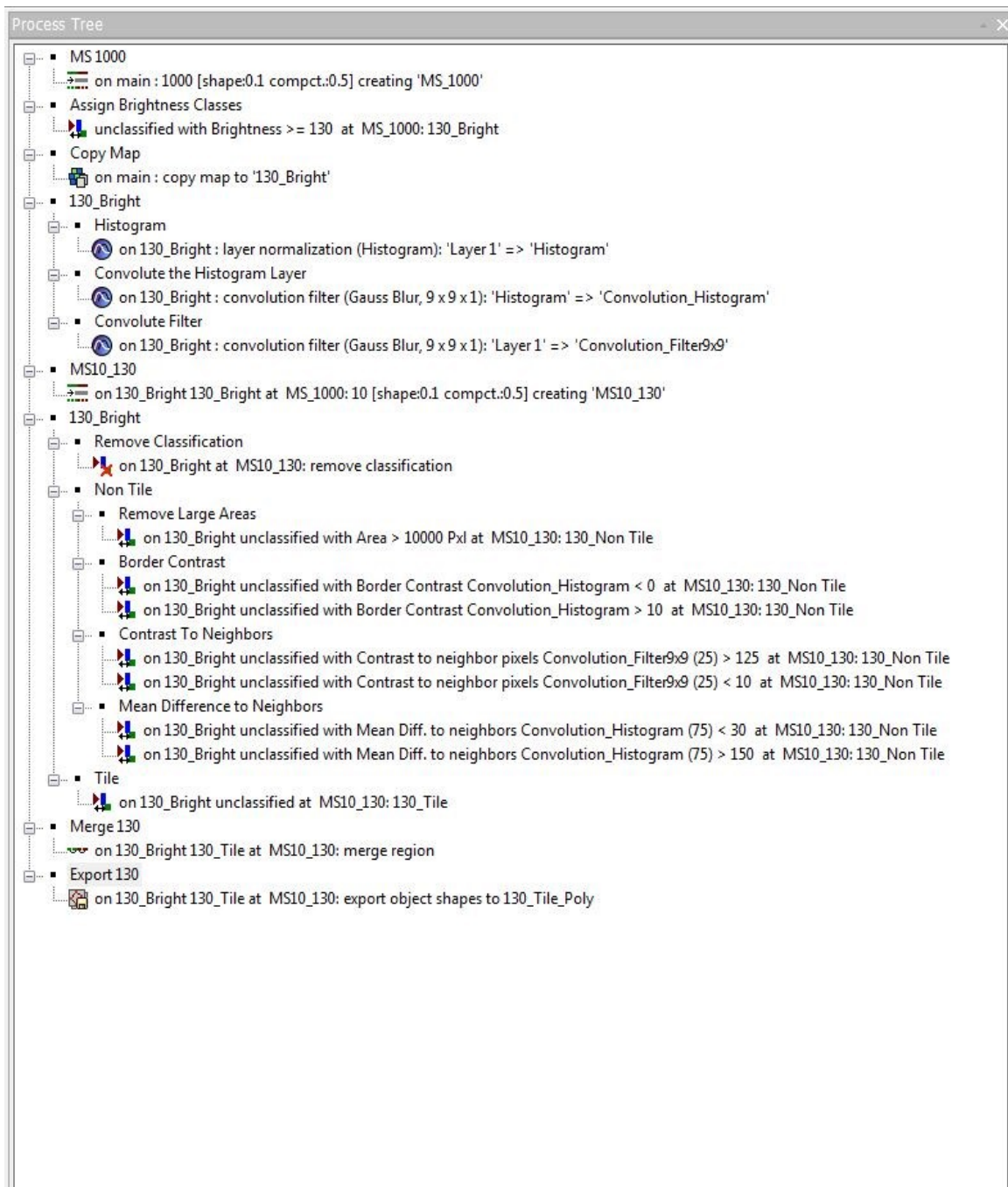le locations. A buffer of ten feet is added around the hand-digitized polylines, because the rule-set was designed to detect the contrast in soil characteristics, not the subsurface tile lines themselves. This method of buffering was successfully demonstrated in Reynolds (2014). The buffered tile lines are clipped, using ArcMap 10.2, into their associated brightness interval, which allows for each brightness interval to be evaluated. For each brightness interval, which contains both the eCognition generated polygons and the hand-digitized buffered polygons; a set of 250 random points was created by applying the Create Random Points tool. A minimum boundary of 10 feet was determined per point, which means two points cannot be located within 10 feet of each other (Figure 3.17). This

method was adopted because some brightness intervals may only contain small total

areas, and more points are unlikely to generate better results. The random points were

only plotted within the image objects that contain both sets of tile line information. This

is done because not all fields within the study area were hand-digitized.

Figure 3.17 Random assessment points with ten foot buffer

The randomly generated points are evaluated and grouped into one of four classifications.

1. Positive – points identified by both the hand-digitized and eCognition methods as Tile.

2. Negative – points identified by both the hand-digitized and eCognition methods as Non-Tile.

3. False Positive – points identified by the eCognition method as Tile, but the hand-digitized method identified as Non-Tile.

4. False Negative – points identified by the eCognition method as Non-Tile, but the hand-digitized method identified as Tile.

A confusion matrix was created for each brightness interval as well as the entire study area. Next the Total Percent Agreement and Total Percent Error were calculated, by adding the sum of the Positive and the sum of the Negative points, and adding the sum of the False Positive and the False Negative points, respectively. Finally, the kappa statistic is calculated utilizing the Total Percent Agreement and the Total Percent Error (Equation 3.5).

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)},$$

Equation 3.5 Kappa Statistic   (Corpus Linguistic Methods, 2013)

Where Pr(a) is the relative observed agreement and Pr(e) is the expected agreement based on the hypothetical probability. Results can range from 0-1, with k=0 meaning there is no agreement and k=1 there is complete agreement.

Chapter 4

4 Results

4. 1 Image Object Identification

The results from the image object identification method demonstrated the usefulness of each brightness interval to detect image objects that contain tile drains. These results were simply derived from the multi-resolution segmentation algorithm using the scale parameter of 1000 applied to each brightness interval independently. The initial segmentation of the study area, roughly 10.22km², was divided into 142 image objects (Figure 4.1). A more detailed look into the classification of these image objects is displayed in Table 4.1.

Figure 4.1 Brightness Interval Classification

These results highlight the importance of brightness within the tile identification process. The vegetation brightness interval contains the largest number of image objects, but this is due to many localized variations in residential areas, such as yards, urban parks, small forested areas, etc. The 13 image objects containing tile are a result of the subdivision of four agricultural fields and one residential garden (Figure 4.2). For this reason, any image object with a brightness value under 130 was considered vegetation and was excluded in the final results. This rule-set is capable of accommodating vegetated areas, but will reduce overall accuracy dramatically.

Figure 4.2 Example of Vegetation Error

The three most important brightness intervals were 140, 150, and 160, which were dominated by image objects containing tile at 90.48%, 88.89%, and 87.50% respectively. The range 140-160 outlines the brightness values that most accurately identified agricultural fields and subsequently tiled areas. These brightness values contained 56 image objects, and 50 of them contained tiled fields. The three brightness intervals made up roughly 73.50% of the total tiled image objects, and are vital for the rest of this process.

The next series of brightness intervals which included 170, 180, and 190, began to see a decline in the accuracy of agricultural field detection. There were a total of 13 image objects in these three intervals, 9 in 170, 4 in 180, and 0 in 190. Of the 13 image objects

only one contained tile and was located in the brightness interval 170. These results show

that image objects above the 170 interval tend to be too bright to be agricultural fields

and are likely impervious surfaces. However, it should be noted that the algorithm did not

detect any image objects with brightness values exceeding 190.

The brightness algorithm generates similar results to the well-known normalized

difference vegetation index (NDVI), but unlike NDVI, the brightness algorithm does not

require the infrared band. One drawback of the brightness algorithm is that values are

highly dependent on the scale parameter, because scale is essentially the resolution of

each image object. This relationship means that the larger the scale parameter, the smaller

the variation, and the smaller scale parameter, the larger the variation.

Table 4.1 Image Object Identification

Brightness Interval	Image Object Containing Tile	Percent Containing Tile %	Image Object Not Containing Tile	Percent Not Containing Tile %	Total Number of Image Objects
Vegetation(<130)	13	20.63%	50	79.37%	63
130	4	40%	6	60%	10
140	19	90.48%	2	9.52%	21
150	24	88.89%	3	11.11%	27
160	7	87.50%	1	12.50%	8
170	1	11.11%	8	88.89%	9
180	0	0	4	100%	4
190	0	0	0	0	0
Bright Objects(>200)	0	0	0	0	0
				Total	142

4.2 Results at the Individual Field Level

The overall results for this study were pleasing because they eliminated the need for the

user to manipulate the rule-set for every field in order to make it transferable, as was the

case for Reynolds (2014). The three sample fields shown below allow for an in-depth

discussion of the rule-set's ability to detect tile drains at the field level (Figure 4.3-4.5).

Figure 4.3 Field Results Sample Field #1 original image (left) classified (right)

The image above demonstrates the strength of the algorithm to be able to detect tile

drains in soil that exhibit minimal brightness variation. Low variations in agricultural

fields are a common occurrence, but a successful rule-set must be able to accommodate

this concern. Albeit, there are errors contained in the image, most notably edge effects

and discontinuous lines, but the tile lines are clearly defined and visible to the naked eye.

The algorithm even allows for the distinction of tile lines within areas that are extremely

blurred, as seen in the eastern and the southwestern portions of the image.

54

Figure 4.4 Field Results Sample Field #2 original image (left) classified (right)

Sample field #2 highlights the threshold of the rule-set's capability. The errors associated with this image are caused by overlapping issues including minimal variation in brightness, edge effects, and crop rows. The image does not exhibit a clear distinction between wet and dry soil except for the very middle of the image, where the algorithm correctly identified tile. One source of error is that the majority of the image is covered with diagonal crop rows, as seen in a northeast to southwest pattern. Image filtering has greatly reduced this error, but it is an inherit issue that still may cause confusion in both the automated algorithm and a human technician. However, the rule-set is still able to successfully detect tile lines for much of the field even under these difficult conditions.
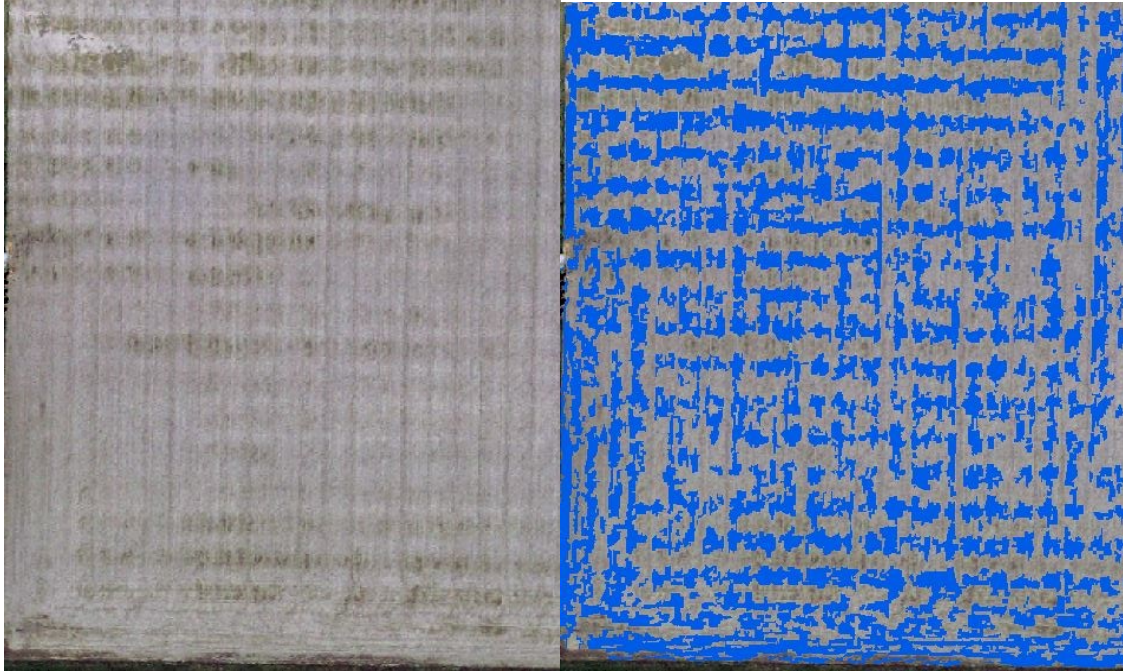
Figure 4.5 Field Results Sample Field #3 original image (left) classified (right)

The third field demonstrated the strength of the rule to correctly identify a field, because there are little edge effect or crop row errors. This image was taken under nearly perfect conditions to display the full extent of the tile drainage system. The only major error was the omission of a part of a large vertical tile and some smaller horizontal lines in the southeastern portion of the image.

These three sample fields represent the typical conditions and results generated from this rule-set when applied to high-resolution imagery. The tile line locations are successfully extracted even under imperfect or complex conditions, without any additional manipulation to the algorithm. The results generated give a clear picture of the extent,

spacing, and density of tile drains within a field, and demonstrated the ability of being

reproduced over a large study area.

4.3 Tile Verification

The eCognition results were validated through a comparison to the hand-digitized tile line

locations (Figure 4.6-4.7). The previously collected hand-digitized tile locations were

validated by two separate technicians with a combined accuracy of 95% or greater. This

highly accurate dataset allowed for the baseline to assess the eCognition rule-set.

Figure 4.6 Hand Digitized Lines for Study Area

Figure 4.7 Rule-Set Results for Study Area

The tile verification was completed within ArcMap 10.2, through a process of clipping

and dissolving both sets of shapefiles. This was completed in order to generate areas that

contain both sets of data encompassed within an appropriate brightness interval (Figure

4.8). The accuracy assessment was conducted by generating 250 random points per

brightness interval at a minimum distance of ten feet apart. Each point was then visually

identified as one of the four classes, positive, negative, false positive, false negative.

Figure 4.8 Accuracy Assessment Locations

The brightness intervals show little variation between the percentage agreements, with a range of 67.20% to 71.20%. The small range is important because it demonstrates how the rule-set is able to successfully identify tiled fields throughout a complex image without sacrificing the accuracy of one interval for another. The overall percentage agreement of all of the brightness intervals was 69.76%.

Table 4.2 Brightness Interval Agreement

Brightness Interval	Total Percent Agreement %	Total Percent Error %
130	71.20%	28.80%
140	70%	30%
150	70.80%	29.20%
160	69.60%	30.40%
170	67.20%	32.80%
Average	69.76%	30.24%

The most common false positive error was classifying the edge of a field as tile (Figure

4.9). The underestimation of tile was 12.72%, which was classified as false negative. The

major reason for this error is the assumptions made by technicians when conducting the

hand-digitized lines. Technicians connected tile lines across entire fields even if they are

not visible to the naked eye. This is logical, because based on prior knowledge most

fields exhibit a "natural" pattern. However, the rule-set is incapable of making these

logical assumptions and can only classify an image object as tile if it falls within the

specified parameters.

Figure 4.9 Example of Edge Effects

The data showed more variation when examining the kappa statistic (K) for each

brightness interval independently shown in tables 4.3 through 4.9. The brightness interval

with the least strength of agreement was 130, with a kappa value of 0.273 (Table 4.3).

This brightness interval contains brightness values that border the vegetation interval and

may explain the relatively low accuracy.

Table 4.3 Confusion Matrix – Brightness Interval 130

		Brightness Interval 130			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	32	37	69	27.60%
	Not-Tile	35	146	181	72.40%
	Total	67	183	250	
	Percentage	26.80%	73.20%		71.20%
	Kappa=	0.273			
	Strength of Agreement=	Fair			

The brightness interval 140 revealed a sizable increase in accuracy from the previous

brightness interval to a kappa value of 0.36 (Table 4.4). The strength of agreement is still

only considered fair, but includes a much larger proportion of agricultural land than

brightness interval 130.

Table 4.4 Confusion Matrix – Brightness 140

		Brightness Interval 140			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	54	50	104	41.60%
	Not-Tile	25	121	146	58.40%
	Total	79	171	250	
	Percentage	31.60%	68.40%		70.00%
	Kappa=	0.36			
	Strength of Agreement=	Fair			

The kappa statistic continues to improve in brightness interval 150, with a K value of 0.416 and strength of agreement considered moderate (Table 4.5). This is the most accurate brightness interval with 177 of the 250 random points in agreement. This interval lies in the middle of the suggested range of brightness values, 140-160, for detecting tile drained agricultural fields.

Table 4.5 Confusion Matrix – Brightness 150

		Brightness Interval 150			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	80	45	125	50.00%
	Not-Tile	28	97	125	50.00%
	Total	108	142	250	
	Percentage	43.20%	56.80%		70.80%
	Kappa=	0.416			
	Strength of Agreement=	Moderate			

Brightness interval 160 displays a slight decrease in accuracy with a kappa value of 0.393

and strength of agreement considered fair (Table 4.6). This interval seems to be at the

high end of the range of brightness values most accurate for detecting tile as determined

by this study.

Table 4.6 Confusion Matrix – Brightness 160

		Brightness Interval 160			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	79	47	126	50.40%
	Not-Tile	29	95	124	49.60%
	Total	108	142	250	
	Percentage	43.20%	56.80%		69.60%
	Kappa=	0.393			
	Strength of Agreement=	Fair			

The last interval that exhibited any tiled image objects, brightness interval 170, starts to

illustrate the potential threshold of brightness values to accurately detect tile (Table 4.7).

Brightness interval 170 had a kappa value of 0.338 and fair strength of agreement. As

noted earlier, any image object above this range, >170, was considered non-tile without

any extra investigation.

Table 4.7 Confusion Matrix – Brightness 170

		Brightness Interval 170			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	96	40	136	54.40%
	Not-Tile	42	72	114	45.60%
	Total	138	112	250	
	Percentage	55.20%	44.80%		67.20%
	Kappa=	0.338			
	Strength of Agreement=	Fair			

The final table is the overall accuracy of the study area, which is the aggregate of all brightness intervals (Table 4.8). The total number of correctly identified random points was 872 out of 1250, which resulted in a kappa statistic of 0.382 and the strength of agreement was considered fair.

Table 4.8 Confusion Matrix – Study Area

		Study Area			
		Hand Digitized			
		Tile	Not-Tile	Total	Percentage
eCognition	Tile	341	219	560	44.80%
	Not-Tile	159	531	690	55.20%
	Total	500	750	1250	
	Percentage	40.00%	60.00%		69.76%
	Kappa=	0.382			
	Strength of Agreement=	Fair			

Table 4.9 illustrates the total agreement percentage and total error percentage for each brightness interval. The overall accuracy for each brightness interval is nearly identical, with the range of percent agreement from 67.20%, in interval 170, to 71.20% in interval 130. The average agreement for the entire study area is 69.76%. The overestimation of tile for the entire study area was 17.52%, classified as false positive, and was the result of many areas within image objects with very high contrasts between edges not caused by tile.

Table 4.9 Tile Verification

Tile Verification									
Brightness Interval	Positive	Percent Positive %	Negative	Percent Negative %	False Positive	Percent False Positive %	False Negative	Percent False Negative %	Total
130	32	12.80%	146	58.40%	37	14.80%	35	14.00%	250
140	54	21.60%	121	48.40%	50	20.00%	25	10.00%	250
150	80	32.00%	97	38.80%	45	18.00%	28	11.20%	250
160	79	31.60%	95	38.00%	47	18.80%	29	11.60%	250
170	96	38.40%	72	28.80%	40	16.00%	42	16.80%	250
Average	68.2	27.28%	106.2	42.48%	43.8	17.52%	31.8	12.72%	

For the complexity of the data utilized and the relatively simple methodology applied these results are attractive, because they represent a robust and transferable rule-set. Similar results have been acquired with field-sized study areas and more complicated rule-sets. Previous attempts required the rule-sets to be manipulated in order to be transferable to different study areas.

Chapter 5

5 Conclusions

Overall, this research provided a major step forward in the ability to identify and extract

tile drains; generating a total percent agreement of 69.76% with a fair ranking kappa

statistic of 0.382. The results were grouped into two separate processes, image object

identification and tile verification, which combined, produced the final results. The first

process was the image object identification, which automatically characterized large

homogeneous areas based on their associated brightness interval. One issue that was

encountered during this research was the error propagation caused by vegetation. This

research used techniques that worked well on bare soil only, and as a result any area with

vegetation, or a brightness value under 130, was considered Non-Tile. One possible

solution is to use multi-temporal imagery and run the rule-set images that were collected

during different growing seasons. However, this technique was unable to be employed in

this study. This limitation is acknowledged and further research must be done to

investigate identifying tiled fields with vegetation cover or low brightness values.

It is concluded that the majority of the tiled image objects fall within the brightness range

of 140-160. This insight into the spectral properties of agricultural fields is important for

large scale projects, where processing time and computing power are limited. Using this

range of values would allow for the most effective route in order to identify agricultural

fields. Another conclusion drawn from this study is that the brightness value of 180 or greater is considered a threshold for agricultural fields. This means all fields above this threshold may be considered Non-Tile without any further segmentation or classification. However, this has not been tested on other study areas and should only be used as a guideline.

The tile verification process re-segmented each brightness interval into a series of much smaller image objects and then reclassified each image object as Tile or Non-Tile. A critical step was applying the image filters, Histogram and Convolution, on the small image objects. The combination of these sets of image filters provided the means necessary to enhance the pertinent contrasts between image objects plus smooth out the erroneous noise. Without these two filters, this methodology would have created negligible results, because the variations between image objects without these filters are too vast for the algorithms to produce accurate results at the scale studied.

After a lengthy series of test trials with various combinations of algorithms and parameters, it was decided to only apply the two most effective algorithms. The two algorithms, Border Contrast and Contrast to Neighbor Pixels, were determined to be the most efficient in detecting the contrast between soil brightness values that are indicative of tiled fields. The final rule-set was excellent at identifying small contrasts in soil, and even some undetectable to the human eye.

These results demonstrated a fair strength of agreement, which is acceptable for the experimental nature of this study. However, there are areas where improvements are needed. The main sources of error are caused by the rule-set's inability to distinguish

between different types of soil contrasts. For example, the contrasts between the edge of a field and cropped areas, called edge effect, are nearly identical to the contrast between the tiled portion and non-tiled portion within a field. This same error may also occur when crop rows are visible, which exhibit similar patterns as tile lines. Additional parameters added to the rule-set may be necessary to mitigate these errors, but might dramatically increase the processing time. An unforeseen limitation of this research was the inability to convert the final output into useable polylines. Polylines would have allowed for a more direct comparison to the hand-digitized lines, which would have likely been more accurate.

However, the most important aspect of this rule-set is that it holds true throughout the study area, and no modifications are needed to be made to the rule-set for any individual field. This thesis provides a method to extract the density and extent of tiled fields with much less effort and user-time required. Ultimately, this thesis provides the means in which others will be able to model nutrient transport through subsurface tile drainage over larger areas more effectively and efficiently.

Chapter 6

6 Future Research

Throughout this thesis there were a few key issues that arose, which would benefit

greatly with further investigation. The first issue is the software eCognition itself, which

has a steep learning curve. New editions of eCognition such as version 9, may be more

user friendly and may be accompanied with new algorithms that would greatly benefit

future research. Version 9 is available now, but was unavailable at the time this research

was conducted. Further study is needed to comprehensively experiment with all

algorithms, parameters, and image filtering techniques.

More investigation is needed to expand the work accomplished by this thesis, with

special focus on refining the brightness intervals. The brightness intervals for this study

were only guidelines based upon using intervals of ten. Much can be learned from

determining the exact brightness ranges that are produced by agricultural fields.

The availability of the imagery is still a major concern because the ideal conditions

needed for tile identification are rare. This issue is likely to be solved in the near future

with the advancements in remote sensing techniques and decreasing cost. Another avenue

that should be investigated is the use of unmanned aerial vehicles (UAVs), or drones, to

collect accurate high-resolution imagery taken under the conditions of the user's

choosing. Finally, alternate sensors that utilize wavelengths outside of the visible or near

infrared should be applied in order to investigate their effectiveness.

References

Agri Drain Corporation. (2014). *Equipment & Tools.* Retrieved from Agri Drain
Products: http://ecommerce.agridrain.com/509-1318-thickbox/5-heavy-duty-tile-
probe.jpg

Ahiablame, L. M., Chaubey, I., Smith, D. R., & Engel, B. A. (2011). Effect of tile effuent
on nutrient concentration and retention efficiency in agriculutral drainage ditches .
Agricultural Water Management, 1271-1279.

Alexander, R. B., Smith, R. A., Schwarz, G. E., Boyer, E. W., Nolan, J. V., & Brakebill,
J. W. (2008). Differences in Phosporus and Nitrogen Delivery to The Gulf of
Mexico from the Mississippi River Basin. *Environmental Science Technology*,
822-830.

Andrade, C. (2013). An Exploratory Study on Heads Up Photo Interpretation of Aerial
Photography as a Method for Mapping Drainage Tile. *Papers in Resource
Analysis at Saint Mary's University of Minnesota University Services*.

Beauchamp, K. H. (1987). A History of Drainage and Drainage Methods. In U. ERS,
Farm Drainage in the United States; History, Status, and Prospects. (pp. 13-28).
Washington, D.C.

Beck, M. W., Vondracek, B., & Vinje, J. (2013). Semi-automated analysis of high-
resolution aerial images to quantify docks in glacial lakes. *ISPRS Journal of
Photogrammetry and Remote Sensing*, 60-69.

Blaschke, T. (2009). Object Based Image Analysis for Remote Sensing. *ISPRS Journal of
Photogrammetry and Remote Sensing*.

Blascke, T., & Strobl, J. (2001). What's wrong with pixels? Some recent developments
interfacing remote sensing and GIS. *GIS - Zeitschrift Fur
Geoinformationssysteme*, 14 (6) 12-17.

Brown, D. C. (2013). Applying a Model to Predict the Location of Land Drained by
Subsurface Drainage Systems in Central Minnesota. *Papers in Resource Analysis
at Saint Mary's University of Minnesota University Services*.

Campbell, J. B., & Wynne, R. H. (2011). *Introduction to Remote Sensing.* New York: The Guilford Press.

Collins, D. M. (2014, August 11). Open Letter to Federal Officeholders. Toledo, Ohio, USA: Toledo Mayor's Office.

Corbane, C., Raclot, D., Jacob, F., Albergel, J., & Andieux, P. (2008). Remote Sensing of Soil Surface Characteristics from a Multiscale Classification Approach. *Catena 75*, 308-318.

Corpus Linguistic Methods. (2013, December 24). *A Practical Introduction with R and Python*. Retrieved from Cohens Kappa: https://corpuslinguisticmethods.files.wordpress.com/2014/01/cohenskappa.jpg?w=316

Definiens. (2009). *Definiens eCognition Developer 8, Reference Book* . Munich, Germany: Definiens.

Dezső , B., Fekete, I., Gera, D., Giachetta, R., & László, I. (2012). Object-Based Image Analysis in Remote Sensing Applications Using Various Segmenation Techniques. *Annales Univ. Sci. Budapest*, 103-120.

Dils, R. M., & Heathwaite, A. L. (1999). The Controversial Role of Tile Drainage in Phosphorus Export from Agricultural Land. *Water Science Technology*, 55-61.

Donnan, W. W. (1976). An Overview of Drainage Worldwide. *Proceedings of the Third National ASAE Drainage Symposium.* St. Joseph Michigan.

Fausey, N. R., Brown, L. C., & Belche, H. W. (1995). Drainage and Water Quality in Great Lakes and Cornbelt States. *Journal of Irrigation and Drainage Engineering*, 283-288.

Fausey, N. R., Doering, E. J., & Palmer, M. L. (1987). Purposes and Benefits of Drainage. In U. ERS, *Farm Drainage in the United States; History, Status, and Prospects.* (pp. 48-51). Washington, D.C.

Flanders, D., Hall-Beyer, M., & Pereverzoff, J. (2003). Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. *Can. Journal of Remote Sensing*, 441-452.

Fouss, J. L., & Reeve, R. C. (1987). Advances in Drainage Technology: 1955-85. In U. ERS, *Farm Drainage in the United States; History, Status, and Prospects.* (pp. 30-45). Washington, D.C.

Fouss, J. L., Fausey, N. R., & Reeve, R. C. (1972). Draintube Plows: Their Operation and Laser Grade Control. *Proceedings of Second National ASAE Drainage Symposium*, (pp. 39-42). St. Joseph Michigan.

Foy, R. H., & Withers, P. J. (1995). The Contribution of agricultural phosphorus to eutrophication. *Proceeding No. 365.* Peterborough, England: The Fertilizer Society.

Gentry, L. E., David, M. B., Royer, T. V., Mitchell, C. A., & Starks, K. M. (2007). Phosphorus Transport Pathways to Streams in Tile-Drained Agricultural Watersheds. *Journal of Environmental Quality*, 408-415.

Great Lakes Protection Fund. (2010). *Phosphorus stratification in relation to agricultrual production practices and the soil resource base in the Sandusky River Watershed.* Tiffin, Ohio: Heidelberg University.

Hay, G. J., & Castilla, G. (2006). *Object-Based Image Analysis.* Calgary, AB, Canada: Foothills Facility for Remote Sensing and GIScience.

Hoffer, R. M. (1972). *Agricultural and Forest Resource Survey from Space.* West Lafayette, Indiana: LARS/Purdue University.

Holtgrieve, D. G., Hardwick, S. W., & Shelley, F. M. (2012, January 12). The Geography of North America: Environment, Culture, Economy.

Jaynes, D. B., & James, D. E. (1987). *The Extent of Farm Drainage in the United States.* Ames, IA: US Department of Agriculture Economic Research Service.

Kaatz, M. R. (1955, March). The Black Swamp: A Study in Historical Geography. *Annals of the Association of American Geographers*.

King, K. W., Williams, M. R., & Fausey, N. R. (2014). Contributions of Systematic Tile Drainage to Watershed-Scale Phosphorus Transport. *Journal of Environmental Quality: Phosphorus fate, Management, and Modeling in Artificially Drained Systems*.

Kladivko, E. J., Van Scoyoc, G. E., Monke, E. J., Oates, K. M., & Pask, W. (1991). Pesticide and nutrient movement into subsurface tile drains on silt loam soil in Indiana. *Journal of Environmental Quality*, 20(1) 264-270.

Krause, G., Bock, M., Weiers, S., & Braun, G. (2004). Mapping land-cover and mangrove structures with remote sensing techniques: A contribution to a synoptic GIS in support of coastal management in North Brazil. *Environmental Management 34(3), 429-440.*, 34(3), 429-440.

McDowell, R. W., & Sharpley, A. N. (2001). Approximating Phosphorus Release from Soils to Surface Runoff and Subsurface Drainage. *Journal of Environmental Quality*, 30:508-520.

National Aeronautics and Space Adminstration. (2014). *Goddard Space Flight Center's Images.* Retrieved from NASA Earth Observatory: http://eoimages.gsfc.nasa.gov/images/imagerecords/84000/84125/LakeErieBloom_amo_2014215.jpg

National Center of Water Quality Research. (2011). *Water Quality News and Notes: Dissolved Phosphorus From Cropland Runoff.* Tiffin, Ohio: Heidelberg University.

National Oceanic and Atmospheric Administration. (2014). *Hypoxia Watch.* Retrieved from NOAA: National Coastal Data Development Center: http://www.ncddc.noaa.gov/hypoxia/

Naz, B. S., & Bowling, L. C. (2008). AUTOMATED IDENTIFICATION OF TILE LINES FROM REMOTELY SENSED DATA. *Transactions of the ASABE*, 1937-1950.

Naz, B. S., Ale, S., & Bowling, L. C. (2009). Detecting subsurface drainage systems and esitmating drain spacing in intensively managed agricultural landscapes. *Agricultural Water Management*, 627-637.

Niagara Block. (2014). *Drain Tile.* Retrieved from Niagara Block: http://www.niagarablock.com/application/webroot/userfiles/image/F-DrainTile%201001.JPG

Northcott, W. J., Verma, A. K., & Cooke, R. A. (2000). Mapping Subsurface Drainage Systems using Remote Sensing and GIS. Milwaukee, WI: ASAE/CSAE-SCGR Annual International Meeting.

Ohio Environmental Protection Agency. (2011). *Ohio Lake Erie Phosphorus Task Force Report II.* Toledo, Ohio: Ohio Environmental Protection Agency.

Ohio Geographically Reference Information Program. (2011). *OSIP Data for Ohio by County.* Retrieved from OSIP Data: http://gis3.oit.ohio.gov/geodatadownload/osip.aspx

Ohio History Connection. (2014). *Black Swamp Map*. Retrieved from Ohio History Central: http://www.ohiohistorycentral.org/images/a/ae/300x300xBlack_Swamp_map.jpg.pagespeed.ic.5UZtr6gPK6.jpg

Ohio Phosphorus Task Force. (2013). *Ohio Lake Erie Phosphorus Task Force II Final Report*. Columbus, Ohio: Ohio Department of Agriculture, Ohio Department of Natural Resources, Ohio Environmental Protection Agency.

PublicWiki. (2010, December 13). *ITK Examples Baseline Smoothing* . Retrieved from Discrete Gaussian Image Filter: http://itk.org/Wiki/images/2/28/ITK_Examples_Baseline_Smoothing_TestDiscret eGaussianImageFilter.png

Reynolds, E. (2014, August). An Automated Method of Identifying the Location of Agricultural Field Drainage Tiles. Toledo, Ohio.

Schopfer, E., & Moller, M. S. (2006). Comparing metropolitan areas - Transferable object-based image analysis approach. *Photogrammetrie, Fernerkundung, Geoinformation*, 277-286.

Schwab, G. O., & Fouss, J. L. (1985). Plastic Drain Tubing: Successor to Shale Tile. *Agricultural Engineering*, 23-26.

Sharpley, A. N., Foy, B., & Withers, P. (2000). Practical and Innovative measures for the control of agricultural phosphorus losses to water: An Overview. *Journal of Environmental Quality*, 29:1-10.

Smith, D. R., King, K. W., Johnson, L., Francesconi, W., Richards, P., Baker, D., & Sharpley, A. N. (2014). Surface Runoff and Tile Drainage Transport of Phosphorus in the Midwestern United States. *Journal of Environmental Quality: Phosporus fate, Management, and Modeling in Artificially Drained Systems*.

Sugg, Z. (2007). *Can GIS Lead to Better Estimates of Subsurface Drainage Extent?* Washington, DC: World Resources Institue.

The Black Swamp Conservancy. (2014). *Historic Map*. Retrieved from The Black Swamp Conservancy: http://www.blackswamp.org/clientuploads/historic_map.jpg

Thomas, N., Hendrix, C., & Congalton, R. G. (2003). A comparison of urban mapping methods using high-resolution digital imagery. *Photogrammetric Engineering & Remote Sensing*, 963-972.

Thompson, J. (2010). Identifying Subsurface Tile Drainage Systems Utilizing Remote Sensing Techniques. https://etd.ohiolink.edu/ap/10?206094541742112::NO:10:P10_ETD_SUBID:781 84.

Trimble. (2014). *eCognition Developer Reference Book.* Munich, Germany. Retrieved from http://www.ecognition.com/

United States Department of Agriculture Economic Research Service. (1987). *Farm Drainage in the United States; History, Status, and Prospects.* Washington, DC.

Verma, A. K., Cooke, R. A., & Wendte, L. (1996). Mapping subsurface drainge systems with color infrared aerial photographs. *GIS and Water Resources.* Ft. Lauderdale, FL: American Water Resource Association's 32nd Annual Conference and Symposium.

Walker, J. S., & Blaschke, T. (2008). Object-based landcover classification for the Phoenix metropolitan area. *International Journal of Remote Sensing*, 2021-2040.

Weaver, M. M. (1964). *History of Tile Drainage.* Waterloo, New York: M. M. Weaver.