# IDENTIFICATION OF PECAN WEEVILS THROUGH IMAGE PROCESSING

By

SALEH M. ASHAGHATHRA

Bachelor of Science
King Saud University
Riyadh, Saudi Arabia
1996

Master of Science
Oklahoma State University
Stillwater, Oklahoma
2003

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May 2008

IDENTIFICATION OF PECAN WEEVILS THROUGH

IMAGE PROCESSING

Dissertation Approved:

Dr. Paul Weckler

Dissertation Adviser

Dr. John Solie

Dr. Marvin Stone

Dr. Astri Wayadande

Dr. A. Gordon Emslie

Dean of the Graduate College

بسم الله الرحمن الرحيم

*In the name of Allah, the most Gracious the most Merciful*

ACKNOWLEDGMENT

My thanks and praises are due to the Almighty Allah, the most Gracious, and the most Merciful. I thank Him for all His bounties upon me. I pray for our Prophet Mohammad, (Allah’s peace and blessing be upon him), who is our role model.

I express my gratitude to my advisor, Dr. Paul Weckler, for his invaluable advice, assistance and support. I also express my deep gratitude to Dr. John Solie and Dr. Marvin Stone for their invaluable help and sincere advice. It was a memorable experience for me to work with Dr. Paul Weckler, Dr. John Solie, and Dr. Marvin Stone for my M.S. and P.hD. programs. My sincere thanks also go to Dr. Astri Wayadande for her exceptional assistance throughout this work. I feel honored to have worked with these distinguished members of the graduate faculty. I wish to thank the Insects’ Museum Technician, Mr. Don Arnold, for providing a part of the testing sample of insects.

I express my thanks to my friend Sayeed Mohammad for his help and support. I also thank my friends Dr. Othman Al-Nashwan, Dr. Mohammed Al-Saleh, Dr. Fahd Eissa, Mubashar Hassan, and Qamar Arsalan for their support.

My sincere gratitude is due to my mother, who devoted her life to help me achieve my goals. I am very grateful to my father (May Allah have mercy on him), who supported me until the last minute of his life. The support of my brothers and sisters has also helped me to overcome many odds.  I am grateful to my wife, Aisha Ashughithry, for standing next to me through out my graduate study. Her tremendous encouragement, support and the love of my kids- Nora, Jana, and Mofleh- made this task much easier.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

# CHAPTER I

# INTRODUCTION

Pecans are a native of the southeastern USA and Mexico (USDA, 2004). The nuts were originally harvested from wild trees but they have since increased in popularity such that the trees are now being extensively cultivated in the southern states, particularly in Texas and Oklahoma, and with the development of new varieties the area of cultivation is spreading farther northwards (Hill, 1983). The United States production of pecan nuts was 177,300 million pounds (USDA, 2004).

Pecans can be separated into two categories based on intensity of cultivation; *improved*, and *native and seedling* (UG, 2005). Improved pecans are produced from superior cultivars in orchard settings, while native and seedling production occurs largely from wild-harvested native stands of trees in Oklahoma and Texas. The yield, quality, and prices received are much greater for improved than native and seedling pecans (UG, 2005). Approximately, 15-24% of native pecan production comes from Oklahoma State (UG, 2005). Total pecan acreage has been estimated at over 1.4 million acres. About 60% of this is native and seedling, and 40% is improved, but the industry is moving towards greater production from improved cultivars. Due to the great value of this product, significant research has been conducted to improve its production.

More than twenty types of insects can attack the pecan tree; however, pecan weevil is one of the most destructive pests of Oklahoma pecans. It is also considered as the most serious late-season pest because it attacks the nut. The pecan weevil is the most important pest of pecans in the areas where it occurs (Harris, 1979). Nut damage is caused by adult and larva feeding, and egg laying. Starting from July through September, the adults begin emerging from the soil, where they spend 2-3 years, and feeding on the nuts. Pecan weevils mate shortly after emerging and females choose the nuts that passed the gel stage but have not hardened. Within 24 day post emergence, a female can attack 25 nuts to lay about 3 eggs in each of them (Harris, 1979). This amount of damage constitutes major damage while the amount of damage caused by adults feeding on nuts (as they feed on about 1 nut every four days) is considered minor damage (Mulder, 2004).

The present management methods for controlling pecan weevils involve detecting their emergence and then applying insecticides. Pecan weevil control requires about one to four well-timed insecticide applications (Mulder, 2004). Some integrated pest management (IPM) stations delay the first treatment until nuts have reached the gel stage of development. This is because successful pecan weevil oviposition can only occur at and after that point until shuck split (Harris, 1979). Generally, insecticide coverage of at least 20-30 days is needed for pecan weevil management.

These treatments will be economically justified in high priced, large fruited pecans if the infestation level is higher than the threshold of 500 post-emergence pecan weevil adults per hectare. The threshold for small fruited, low priced pecans is approximately 3500 pecan weevil adults per hectare. A second or even a third treatment may be needed to prevent economic damage from occurring if pecan weevils continue to emerge from the soil after an initial treatment (Harris, 1979).

There are several monitoring techniques to detect the appearance and activities of adult pecan weevil. They include inspecting dropped nuts for feeding and/or oviposition injury, and using knock down sprays, sticky bands, limb jarring, ground cover traps and assorted traps (Ree et al. 2000). Among these techniques, traps are the most commonly used method. There are different types of traps which have been utilized for monitoring weevil for example, the wire cone trap, pyramid trap, and the circle trap. Traps are placed on or under the trees with known weevil infestations. A very common trap that has been used for years is the wire cone trap. It is normally placed on the ground beneath pecan trees with a known history of pecan weevil infestations.

The number of pecan trees in an orchard block varies from 60 trees per hectare, (thinned density) to 237 trees per hectare (ultra density) (Herrera, 2000). It is recommended to use one to two traps per tree and three to five trees per orchard block (Mizell, 2003). Traps should be placed in the orchard 1 to 2 weeks before the earliest maturing varieties reach the gel stage and these traps are monitored every 2 to 3 days. Adult weevils collected in the traps should be counted and removed with each inspection (Ree et al. 2000).

For monitoring pecan weevils in a 40 hectare (100 acre) orchard, from 300 to 600 traps should be placed on 300 pecan trees. Since pecan weevil emergence varies greatly from year to year and is significantly affected by soil moisture, the initial emergence and peak population emergence can vary from orchard to orchard and tree to tree. As a result, traps have to be checked carefully during this period of time (emergence season).

Clearly, this technique of monitoring pecan weevils is labor intensive and requires very careful observation. Assuming that it would take a farmer one minute to check each trap in a 40 hectare orchard (600 traps), it would then take 10 hours to inspect all of them. That means 30 hours of work per week during the emergence period which could last for three months. The long term objective of this study is to develop an automatic monitoring system based on a wireless network imaging system. This system would detect pecan weevils as soon as they go through the imaging unit that can be incorporated inside traps. The aim of this work is the development of a recognition algorithm that can identify pecan weevils among other insects.

## 1.1 Objective

The development of a wireless network imaging system for monitoring pecan weevils in the field motivated this study. The robustness of this recognition system would replace the manual insect monitoring techniques currently in use and it would be a useful tool for pest control management, in general. The aim of this study is to develop the software part of a wireless network imaging system that can automatically identify pecan weevils in the field.

In particular, the main objective of this study is to develop a recognition algorithm that identifies pecan weevil among other insects that are naturally present in the pecan habitat by implementing several image processing techniques.

Furthermore, the software with minor modifications can be used to identify other insects.

To achieve this objective, the following tasks were undertaken:

1. Design and test an imaging system for acquiring insect’s images
2. Collect enough pecan weevils to account for all possible varieties among them and build a database of pecan weevils’ images that can be used as training data set

3. Explore the ability of different shape description methods in representing and recognizing pecan weevils
4. Develop an algorithm for identifying pecan weevils and differentiating them correctly from other insects
5. Evaluate the algorithm for robustness and speed

### 1.2 Organization of the Study

A better understanding of pecan weevils' live cycle and biology, in general, would help in designing the recognition system for them. The remaining document is organized in the following manner: Chapter II discusses in detail important aspects of a pecan weevil's life and behavior. Chapter III focuses on a literature review of the related methods used to detect insects. In this study, the ability of several recognition methods is evaluated. Chapter IV introduces these methods, with a brief mathematical background, and reviews their applications. This chapter also discusses the materials utilized in this work. The results of all individual methods are detailed and discussed in Chapter V. Based on these findings, an overall algorithm of the recognition system is proposed in Chapter VI. Finally, Chapter VII contains the major conclusions of this work and recommendations for future research.

CHAPTER II

# BIOLOGY OF PECAN WEEVIL

## 2.1 Origin and Distribution of Pecan Weevils

The pecan weevil is native to North America and can be found from New York in the east to Iowa in the west and Oklahoma, Texas and Georgia to the south (Gibson, 1969). As of 1999, pecan weevils had been found in 131 Texas counties (Ree et al., 2000). The states of Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, New Mexico, Oklahoma, South Carolina, Tennessee and Texas have also reported pecan weevil in pecan trees (Harris, 1979).

The pecan weevil, *Curculio caryae* (Horn), belongs to the order *coleoptera*, family *Curculionidae*, and subfamily *Curculioninae* (Borror et al., 1976). Twenty-seven species have undergone close scrutiny (Gibson, 1969). The family *Curculionidae* (weevil or snout beetles) is a large family with about 2500 species in North America (Borror et al., 1976).

## 2.2 Recognition Characters

*Curculio Caryae* can be recognized by its long rostrum abruptly inserted into frons and its large femoral tooth (Gibson, 1969). He described the recognition characters of pecan weevils as follows: the rostrum of the male is ¾ the length of the body and the female rostrum is as long as its body. The head is rounded, punctures elliptical and distinct, punctation extending to antennal insertion on female rostrum and to the apex though sparse, on the male rostrum. The vestiture rarely extends onto the base of female rostrum and about ½ distance to antennal insertion in male. The rostrum is abruptly inserted into the frons and slightly thickened at the base in male and female. The rostrum is longer than body, curved upward near the base, then curved down in distal ¼ and the male rostrum is usually straight to near ¾, then arcuate. It is also less than 0.6 as long as the body. The antennal insertion is about 3 mm from the base in 11-mm specimens. The antennal scape in male is as long as 5-5 ½ funicular segments whereas scape of female is as long as 4 funicular segments. The mandibles are slightly longer than the apex of rostrum. The eyes of the male are nearly circular and have about 36 facets per mm. The female eyes are flattened and have a width of about 0.66 mm. and a height of about 0.83 mm. Also, the male eye is convex and has a width of 0.66 and height of 0.75 mm. Figure 2.1 illustrates a typical male and female Pecan Weevil.

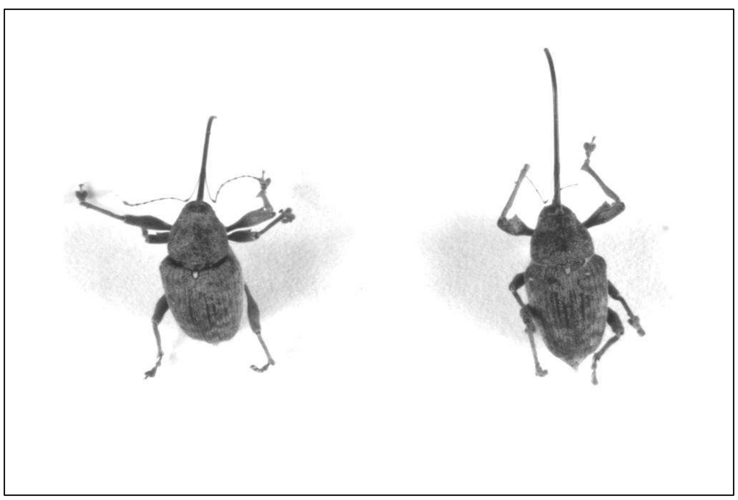


**Figure 2.1 Pecan Weevils: Male (Left) and Female (Right)**

## 2.3 General Features

Gibson (1969) also gave a general description of pecan weevils as follows: their shape is ovate-elliptical in dorsal view and elliptical in lateral view. Their length is 7.5-12 mm and width is 3.4-5.4 mm. The color of the body is dark reddish brown with the antennae and legs lighter in color. There are punctures on prothorax with or without distinct sides and the size is fairly uniform. They have a diameter of about 0.08 mm and are uniformly spaced. They are round, small, and shallow in elytral stria, and also deeper in male than in female. On metasternum, the punctures are much smaller than on prothorax and are widely set apart. The vestiture (body cover) is moderate to dense, varying from golden to grayish yellow to dark brown; with or without dark brown fasciae on elytra. The dorsum of prothorax is fairly uniformly colored and underside of body is also uniformly covered with scales of the same color or slightly paler than on dorsal surface. The squamules on prothorax are hairlike, 6-8 times longer than wide. The squamules are 4-6 times longer than wide on various areas of elytra and are variable on ventral areas of thorax. Also, most of them are 3 times longer than wide whereas on abdomen they are 4-8 times longer than wide. A few are also found on 5$^{th}$ sternum, hairlike; while those on femur are longer and are hairlike on tibia and tarsi.

## 2.4 Biology

Adult pecan weevils emerge from soil in July through September where they have spent 2 or 3 years in soil cells located 4 to 12 inches beneath the soil surface. The literature suggests that the emergence of pecan weevil coincides with the soil condition. Harris and Ring (1980) concluded after their four years of observation that pecan weevil emergence was delayed in clay soil when soil moisture was low. They also found that emergence occurs at about the same time and rate every year when soil moisture was adequate to impart sufficient friability to allow adults to emerge.

Other investigators (Gibson, 1969; Price, 1939; Hinrichs, 1965) have reached the same conclusion. The emergence of adult pecan weevil from the soil is unaffected by pecan variety (Harris, 1976). The time that adult pecan weevils emerge from the soil varies according to soil condition, season and locality (Osburn et al., 1963). According to Polles and Payne (1974), 80 to 90% of the emergence occurs between August 20$^{th}$ and September 20$^{th}$. The earliest report of adult emergence is June 15$^{th}$ (Langston, 1930) and the latest report is October 30$^{th}$ (Neel et al., 1975).

Pecan weevil larvae exhibit little horizontal movement once they have penetrated the soil, indicating that adult emergence is limited to the area covered by the canopy of the tree (Raney et al., 1970). The emergence of adults rapidly decreases as distance from the dripline of the tree increases (Tedders and Osburn, 1970). Harris (1975) found that there were no pecan weevil larvae in the soil beyond the dripline of the tree. He did not find any significant difference in the number of pecan weevil larvae from the trunk to the dripline or any significant difference in the number of pecan weevil larvae in any of the cardinal directions among native pecan trees.

Criswell et al. (1975) found that early emerging adults live longer than late emerging adults. Harris et al. (1981) reported that males live from 15.5-29.9 days and females live from 22.6-25.6 days when emergence occurs in August and September. Also, they found that the adults that were delayed in emergence until October due to drought conditions lived for shorter periods of time. Van Cleave and Harp (1971) reported average life spans for adults were 17.59 days for males and 23.05 days for females.

Upon emerging from the soil, the adults fly to pecan trees (Swingle, 1934) or crawl to the highest point nearby and then fly to the tree (Raney and Elikenbary, 1968). Female weevils tend to fly to higher parts of the pecan tree and male weevils tend to crawl to the tree more frequently than female weevils (Raney, 1969). Boethel et al. (1974) investigated in two studies (1967 and 1972) the effects of tree, height and sector on nut infestations. In both studies, they did not find any significant variation in infestation among three height levels on the trees; however, there was a significant variation in infestation among trees.

Eikenbary and Raney (1973) reported that pecan weevils moved considerably in a tree from the top of the tree to the lowest limbs and from the lower heights to the upper portion of the tree as well as around the tree. Research indicates that 77% of adults fly to the tree trunk at a height of 6 to 8 feet, 5% walk to the tree trunk and 15% fly directly to the canopy (Ree et al., 2000).

As soon as the adult pecan weevils emerge from the soil and reach the tree, they begin puncturing and feeding on pecan nuts that are in the gel stage. The feeding activity of adult weevils, both males and females, before the nuts enter the gel stage can cause nut drop (Moznette et al., 1931). This puncture is about the size of a fine needle that usually remains open (Bissell, 1935). As a result, the liquid endosperm will be released slowly producing a brown stain on shuck and shell (Brison, 1975). This type of injury may seal as the volume of the nut increases which makes it difficult to observe; whereas, punctures occurring later, when the nut is increasing in volume at lower rate, may close very little and are easier to observe (Adair, 1932). The feeding rates for males and females prior to shell hardening are low. Most weevils emerge just at or shortly after gel stage, so nut losses caused by adult feeding are small compared to those caused by egg laying (Ellis, 2007).

An average of 0.23 and 0.29 nuts per day were reported to be punctured by a male and a female weevil respectively (Calcote, 1975). Mated females punctured more nuts than virgin females while the nuts were in the water stage, indicating that mated females feed more in preparation for oviposition (Ring, 1978). After shell hardening, males only feed on the shuck and this will not cause nut drop; virgin females puncture more nuts in this state than mated females (Calcote, 1975).

### 2.5 Mating Behavior

Chemical cues are considered to be of prime importance in eliciting specific reciprocal events for male and female with visual, acoustical and tactical cues (Hartfield et al., 1982). Van Cleave and Harp (1971) indicated traps baited with live female pecan weevils captured more adults of both sexes than unbaited traps. Hartfield et al. (1982) reported that mating behavior of the pecan weevil consists of a rapid, direct orientation toward the female by the male, insertion of the median lobe, and extended copulatory period. Collins (1996) determined that pecan weevil males tapped mesothoracic legs against lateral margins of the female prior to insertion of the aedeagus and not the metathoracic legs as previously published. It was found in the same study that the male approached the female from behind, assuming a mating posture if the female was receptive, and tapped the mesothoracic leg for 3-5 sec. before inserting the aedeagus.

## 2.6 Ovipositional Behavior

Hinricks and Thompson (1955) found that varietal susceptibility is correlated to kernel development. The period between emergence and oviposition is a minimum of 5 days (Van Cleave and Harp, 1971). It was also reported that oviposition occurs as early as 2 days post emergence but preoviposition average was 5-7 days (Criswell et al., 1975). Moznette et al. (1931) reported oviposition did not take place before the nuts' shells were hardened. The reason is that the pecan nuts will not fall after this (gel) stage. Some investigators found early rather than late maturing varieties are more susceptible to attack by the pecan weevil (Moznette et al., 1931; Swingle, 1934; Price, 1939; Osburn et al., 1966). Harris and McGlohon (1972) wrote that late maturing cultivars and varieties are usually not attacked when there are enough nuts for feeding and oviposition on earlier maturing cultivars and varieties. Harris (1976) reported that oviposition occurs as early as August 15 and latest by September 30$^{th}$. However, Harp and Van Cleave (1976) noted that this begins near the end of August.

## 2.7 Egg Laying

To deposit eggs in pecans, a female feeds through the shuck and shell to the kernel where she excavates a small cavity in the developing kernel. She turns around and, with her ovipositor, places three to four eggs per nut on the developing kernel (Ree et al., 2000). Criswell et al. (1975) reported that the average number of nuts each weevil may oviposit in was 7.8, 9.6, or 8.7. It was also found that one female could oviposit in 10-30 nuts (Swingle, 1934). In 1934, Swingle had also found that a single female could lay up to 100 eggs with an average of 25. Gill (1924) reported that the number of eggs was 7 in a single nut with the average number being 3 eggs per nut. A range of 1 to 10 eggs laid per nut was also noted by Leiby (1925). According to Moznette et al. (1931), a female oviposited 2 to 6 eggs per puncture. Generally, there is one oviposition per nut indicating that pecan weevil may shun nuts in which oviposition has already occurred (Ring, 1978). Harris (1976) stated that the regulation of the number of eggs, and thus larvae per nut, might be of selective advantage to the pecan weevil in that it may ensure enough food for each larva. The peak in egg production occurs 10-12 days post-emergence (Van Cleave and Harp, 1971).

Pecan weevil eggs are irregularly shaped; some oblong, some elliptical, and others ovate (Ring, 1978). They are clear, translucent white, with a shining surface and exceedingly delicate texture, varying in size with the length ranging from 0.027 – 0.04 inches and width of 0.02 inches (Brooks, 1910). Several investigators had reported different duration of the egg stage; Gill (1924) reported that the duration of the egg stage is 9 days. However, Moznette et al. (1931) and Bilsing (1940) found it to be one week, depending on the temperature. Swingle (1934) reported a range of 7-10 days and Harp and Van Cleave (1976) indicated that it could be from 6-14 days.

## 2.8 Larval Stage

The larvae hatch and feed in the kernel. When fully developed, larvae chew a single hole through the shuck and shell, exit the nut and drop to the ground; the time period from egg laying to larval emergence is approximately 42 days (Ree, 2000). Pecan weevil larva is a yellowish grub with a small reddish-brown head when mature and white when immature (Moznette et al., 1931). Swingle (1934) described it as white when small to creamy-white when fully-grown, measuring from ¼ to ½ of an inch in length. It was also described by Kern (1949) to be a soft, white–yellowish, cylindrical apodous body; completely chitinized head capsule and black compound eyes, ocelli and antennal structures.

Bilsing (1940) found that the larvae spent about 30 days feeding inside the nut before dropping to the ground. Aguirre Uribe (1979) studied the immature stage of pecan weevil and found that it took 7 - 8 days for eggs to hatch; 3 days for each of the 1$^{st}$ and 2$^{nd}$ instars; 4 –6 days for 3$^{rd}$ instar, and 12-25 days for the 4$^{th}$ instar (average 18.3 days). He also stated that the development time from oviposition to larval emergence averaged 33.3 days. Harp (1970) wrote that total larval feeding period was approximately 30 days in duration. Some investigators described each development stage of larvae in diameters (Kern, 1949; Harp and Van Cleave, 1976). The larval weight was found to be directly related to larval age (Aguirre Uribe, 1979).

The number of pecan weevil larvae is variable and sometimes a nut is completely packed with larvae (Bilsing, 1940). Kern (1949) indicated, in general, there would be 1-3 larvae per nut and in case that there are 5 larvae, their size will shrink because of the competition. In his study, Harris (1976) reported an average of 1.74 to 3.17 larvae per nut independent of the nut size and weight, kernel weight and infestation level. There could be several larvae per nut but only 2 or 3 will survive (Craighead, 1950; Baker, 1972).

When it is time to emerge (larval maturity), a larva bores a round hole out through the shell, then goes into the ground and pupates in preparation to pass the winter (Gossard, 1905). Hinricks and Thompson (1955) stated that usually one larva chews a circular hole (1/8 inch in diameter) through the pecan shell and/or shuck and all grubs would leave the nut by this hole. Sometimes, the pecan nuts dislodge from the tree and fall to the ground so the emerging larvae will crawl from the hole to the soil (Gossard, 1905). In some cases, larvae do not attempt to exit the nut and shrivel and die even though they are fully fed (Hamilton, 1890).

The larvae leave the nut during September and November and enter the soil to a depth of 6 or more inches (Gill, 1924). It was also reported that larval emergence lasts from late September through March with the maximum emergence occurring from October 8$^{th}$ to October 20$^{th}$ (Swingle, 1932). After the emergence, the pecan weevil larvae burrow into the soil with their abdomens (Kern, 1949). Van Cleave and Harp (1971) reported that larvae penetrate in the soil from 3-11 inches with an average of 6.5 inches.

The literature suggests that the condition of the soil correlated with the depth that larvae would burrow to. For example, the depth that a larvae burrow to in non– cultivated soil is from 1-5 inches (Moznette et al., 1931); 2-14 inches (Hinricks and Thomson, 1955); up to 8 inches (Harris, 1975) whereas in cultivated soils, the range is from 1-9 inches (Moznette et al., 1931); up to 18 inches (Hinricks and Thomson, 1955); up to 12 inches, which may suggest that pecan weevil larvae may penetrate deeper than necessary to avoid any harmful weather conditions and natural enemies (Harris, 1975). The time it takes pecan weevil larvae to burrow varied depending also on the soil condition. Chau (1949) wrote that larvae burrow 1-2 inches in 2-10 hours, and on average it takes 1 week to reach its final destination. In the soil, larvae construct (form) an earthen cell by pressing back the surrounding earth. During this period, the larvae do not feed but live on fat reserves (Kern, 1949) as they are in diapause during this stage.

## 2.9 Pupal Stage

Harp (1970) found that the duration of the pupal stage ranged from 14 to 23 days and averaged 19.1 and 20.3 days for males and females respectively in a 1967 study. He also found that the average duration of the pupal stage was 18.1 days for both sexes in a 1968 study. In the same study, Harp found approximately 90% of larvae pupate 1 year after entering the soil and the remaining pupated the following year. The pupae are white, have no coverings, and show the developing appendages of an adult (Moznette et al., 1931).

The duration of the pupal stage is 3 weeks according to several studies (Harris, 1976; Harris and McGlohon, 1972; and Swingle, 1934). Harp and Van Cleave (1976) found it to range from 14-23 days.

## 2.10 Adult Stage

After eclosion from the pupal stage, the adults remain in the soil until the following August and September (Harp, 1970). He noted that only 4% of the initial insects successfully emerged as adults with the greatest mortality occurring in the pupal and adult stages. In this way, the pecan weevil life cycle will be completed in 2-3 years (Hinricks, 1955).

Based on what has been cited in the literature, there was no relationship between the time of adult pecan weevils' emergence and the fruit stage on the host tree. Mody et al. (1976) studied the volatile components of pecan leaves and nuts and found them to contain 38 compounds. The intention of their research was to identify those constituents that could conceivably attract the pecan weevil to the leaves and the nuts. The authors stated that these constituents could also be precursors of the pecan weevil sex pheromone.

Prokopy et al. (2003) concluded in their study of odor-baited trap trees: a new approach to monitoring plum curculio (*Coleoptera: Curculionidae*), that monitoring apples on odor-baited trap trees for fresh ovipositional injury could be a useful new approach for determining need and timing of insecticide application against plum curculio in commercial orchards.

The pecan weevils can be attracted by color as well. Tedders et al. (1969) investigated the effects of color and trunk-wrap on pecan weevil catch in pyramidal traps. They found that black traps were superior to all other colors tested in attracting weevils. The study indicated that white plastic wrap was effective and easier to use than whitewash and both increased trap capture of pecan weevils. In conclusion, it appears that the pecan weevil can find the pecan tree and “sense” the odor of the fruit and leaves and find its mate; but to our knowledge, there is no evidence in the literature to test the pecan weevil’s ability to find its host tree and the maximum distance it could move to reach it.

**2.11 Summary**

Pecan weevil’s life cycle ranges from 2–3 years; most of it is underground. As soon as the adult pecan weevils emerge, they feed on pecan nuts, mate and oviposition eggs in the nuts. These eggs will hatch and larvae will be developed in 30 days. After their complete formation, the larvae would chew a hole in the nut, fall to the ground, and burrow into the soil. There it will pupate in 3 weeks, and will remain as an adult for one or two years before emerging to the pecan tree.

# CHAPTER III

# LITERATURE REVIEW

## 3.1 Insect Detection

Several methods have been used in monitoring insect's population and migration in an environment or plants. Examples of these techniques include optical and opto-electronic devices, video graph, thermal imaging, radio frequency identification, radio-telemetry, X-ray radiography, computed tomography, SODAR and SONAR (Reynolds and Riley, 2002). Monitoring the free movement of insects in the field under natural or semi-natural conditions can be done by any following procedure:

**1. Visual Method**

This approach depends on observing the insects' movements and scoring their behaviors.

**2. Night Vision Devices**

The basic idea of this technique is to amplify the available light for better vision. An objective lens focuses available light onto a photocathode which then releases electrons: the number of these electrons is greatly multiplied by some form of high voltage cascade and the resulting electron flux is used to produce an image on a phosphor screen.

### 3. Video-graphic Techniques

Riley (1993) reviewed in detail the use of video equipment to observe flying insects in the field. One major problem with the technique is maintaining a reasonable field of view whilst producing a video image distinct enough to be detectable at more than a few meters. Another common problem to most remote sensing methods is that of identifying the target. Thus, it is not surprising that outdoor flight studies have been concerned with insects either approaching traps or sources of odor plumes or other situations where identifiable species are expected to pass through a rather restricted sensing volume. The range of detection can be increased by improving the contrast between the insect and the background in particular by viewing the target against the night sky using some form of artificial illumination.

### 4. Thermal Infrared Imaging Technology

This is designed to detect objects in conditions of obscured visibility (darkness, smoke, dust, haze, etc) by utilizing the long-wave infrared (heat) radiation emitted from the objects rather than the light reflected off them.

### 5. Optical Sensors and Insect Trapping

This provides a method of recording the time of entry of insects to traps or to assign captured insects to body-size categories. This can be done by passing the insects individually through an illuminated detection volume and measuring the amount of light scattered during the transit. In the area of trapping, infrared telemetry has been used to transfer data from pheromone traps (used as piezoelectric detection mechanism) and from meteorological sensors in cotton fields to a base station computer situated in the farm office (Schouest and Miller, 1994).

### 6. Opto-Electronic Devices

Farmery (1981) developed a device where the field of view of the photomultiplier sensor intersected the illuminator beam at a selected height above ground level. Insects passing through this intercept volume could be detected and their wing-beat frequencies recorded. This device is ineffective in twilight and daytime when the luminance of the sky greatly lowers its sensitivity to insect targets. This obstacle was overcome by using a very bright xenon flash lamp working in the near infrared and a video camera equipped with a gated image intensifier which provided high-contrast images of even small flying insects against the mid-day sky (Schaefer and Bent, 1984).

Although, the above methods are applicable to insects' detection, in general; it was found that for the purposes of this work (the detection of pecan weevils in the field area) machine vision would be the most suitable detection method.

## 3.2 Detection of Insects by Machine Vision

Yu et al. (1992) worked on the identification of ichneumonid wasps using image analysis of wings. The right forewing was removed from each wasp, placed between a microscopic slide and coverslip, and then aligned for imaging. Four data sets were created for each sample; one each for vertices, veins, cells, and the whole wing. Some geometrical characteristics (length, area, orientation, etc.) for each data set were calculated. Their results showed that 100% of fifty insects were correctly classified using discriminant analysis and independent univariate comparisons of 144 characters from the data sets.

Leafhoppers (*Homoptera Cicadellidae: Draeculacephala Ball*) were identified using linear discriminant models (Dietrich et al., 1994). The incorporated system depends on a combination of qualitative external morphological features and morphometric data. First, an online interactive key was used to classify an unknown specimen in a species or species group using discrete external morphological features. Second, an image of the specimen was captured and edited for identification based on the shape. Their results showed that 89-98% of individuals may be correctly identified.

Artificial neural networks (ANNs), discriminant analysis, and k-nearest neighbors were compared in identifying fungal spores (Morgan et al., 1998). Morphometric data from spores of Pestalotiopsis species and a few species related to Truncatella and Monochaetia were used to train the ANNs. Both ANNs and statistical classifiers had similar identification success on unseen data set between 76-78% of 16 species and between 63-67% of a 19 species group.

A similar approach was applied earlier by Wilkins et al. (1999) to identify phytoplankton from cytometry data. The system was designed to identify seven freshwater and five marine phytoplankton species. Their results showed that the optimized networks and statistical methods performed similarly. They obtained a correct identification rate between 86.8% and 90.1% of data from freshwater species and between 81.3% and 84.1% of data from marine species.

Zayas and Flinn (1998) studied the detection of insects in bulk wheat samples with machine vision. Their study focused on identifying insects and body parts of *Rhyzopertha dominica* beetles in bulk wheat samples. The main objective of their work was to determine x, y coordinates of a subimage belonging to insect versus noninsect elements of the image. In combination with pattern recognition, multispectral analysis was used in their study. Their results showed that recognition rate was higher than 90% for the insects and other elements.

An optical digital system was also applied to identify five phytoplankton species (Pech-Pacheco et al., 1998). In this system, an optical filter was made for each species to be correlated with six testing images. These testing images were prepared from phytoplankton samples which may include one or more of the five species. The system was evaluated using 100 different samples and the recognition rate was 90% despite rotation, translation and scale variations.

Boddy et al. (2000) trained radial basis function artificial neural networks to discriminate between phytoplankton species based on 7 flow cytometric parameters measured on axenic cultures. The study compared the performance of two networks limited to using radially–symmetric basis functions and networks using more general oriented ellipsoidal basis functions. The results showed that the second method performed better and the overall results of correct identification were from 70-77%. The identification was very poor (<20%) when testing the system with one data set on cell growing under different lighting conditions. The best result in this study was obtained when the network was trained on a combination data set (>70%).

Identifying and counting of phytoplankton was undertaken by Embleton et al. (2003). In this study, a comparison between manual and automated counting was conducted. A combination of artificial neural networks and simple rule-based procedures were used to identify selected groups of phytoplankton. For taking the required 75 images per sample, the system took 7 minutes and between 30-40 minutes for identification and classification. The time required for the process of identifying and counting phytoplankton was similar to the manual procedure. The large number of misclassification was due to the system's inability in recognizing touched objects in the images.

The classification and identification of pollen grains was studied by France et al. (2000). The system involves two stages: first, finding the location of a pollen grain in an image taken from a slide where the pollen grains and a range of detrital materials could be differentiated; Second, classification of the pollen grains identified in the first stage into different taxonomic categories. Neural network was employed in the analysis of identifying the pollen grains. Their network had a three-layer architecture including feature extraction, pattern detection, and classification layer. Their results showed that 83% of the samples were correctly classified. However, the authors were attempting to improve the processing time, slide pre-preparation, and pollen orientation invariance in this system.

Moment invariants of butterfly wing patterns were used (White et al., 2003) to detect differences between groups of butterflies according to sex, geographical origin and culture history. Digital images of the speckled wood butterfly were used to generate moment invariant data sets which verified the differences in wing pattern. Their study showed that gray images of butterfly wing can be used to detect differences between wing surfaces even if the wing has some fading and damage. Their results suggested that the seven moment invariants would provide suitable quantitative pattern descriptors. Moments invariants could be applied even to low resolution images and still provide useful information for analysis.

A machine vision system for detection of adult beetles in wheat was devised by Ridgway et al. (2001). The system was designed to detect adult saw-toothed grain beetle (*Oryzaephilus surinamensis)* in grain. For various insect species, the reported detection rates were 89-96% for commercial samples which contained several insect species. These were all classified by the recognition stage as being insects.

Arbuckle et al. (2001) presented their Automated Bee Identification System (ABIS) where the identification was based on features extracted from their forewings. In this system, images of bees were taken by placing them manually in standard position. The venation of the wings was then identified and a set of key wing cells were determined. Images then were aligned and scaled based on the wing cells. The lengths, angles, and areas were computed. Bees' classification was achieved using Support Vector machines and Kernel Discriminant Analysis. The reported success classification rate of this system was 95% using four species. One of the restrictions of ABIS is that feature extraction algorithm includes prior expert knowledge about wing venation. Furthermore, ABIS involves a user interaction to properly place the bee in the standard wing's position.

Automatic identification of bees was also studied by Schroder et al. (2002). The identification of bees is based on characters of the fore-wing venation. In this system, images of bees would be transferred to a computer in which a user marks defined vein junction; then, junctions would be automatically connected to digitize the whole venation. Minimum training samples are 30 well-defined specimens of each sex per species. Linear and non-linear discriminant analysis methods were used in classifying species. Processing time was 5 minutes on average and the correct classification rate was found to be around 98%.

A test of pattern recognition system for identification of spiders was presented by Do et al. (1999). It was a partially automated pattern recognition system that utilized artificial neural network. For each species, between 14 and 21 individual epigyna were photographed by microscope equipped with a CCD video camera. The preparation of each specimen for imaging involved aligning the plate (containing 70% ethanol) of the epigynum to the viewing axis of the microscope. The features of spiders' species were extracted from the digital images of female genitalia using wavelet transform. Three different sized networks were assessed in discriminating a set of six species to either the genus or the species level where species represented three genera of wolf spiders (*Araneae: Lycosidae*). Their results showed that identification of specimens to the correct genus was 100% and about 81% to the correct species.

A study of automating the identification of insects by Weeks et al. (1997) described a semi-automated digital imaging system to discriminate five species of *Ichneumonidae*. The wings of the specimens were used for distinguishing purposes. The algorithm followed in this study had three major steps; first, training step in which features of wings were extracted for each insect; second, principal components were employed to represent the morphology of wings of each individual insect; third, correlation among the species based on their characteristics principal components was performed to evaluate the likelihood of a new image to the trained sets. The overall result of this method was 94% when the system was tested on 175 images of wings of the five *Ichneumonids*.

In their second study, Weeks et al. (1999) considered the identification of wasps using principal component associative memories. Fifty specimens from each of five species from two genera of pimpline ichneumonid wasps collected from Costa Rica were included in their study. For each wasp, the right forewing was removed and mounted in Canada balsam on a microscope slide with cover slip. Images were then acquired from a CCD camera placed on the microscope. The wings were oriented by adjusting the microscope slide to have the anterior margin in parallel with the x-axis. The principal component method was used for identification where the images of wings were rearranged into column vectors consisting of concatenated rows of pixel intensities. The system employed the differences between a pair of reconstructed images produced when unknown image is included in and then excluded from the training set encoded by the associative memory. The reconstructed images were correlated using a non-parametric statistical correlation metric given by Kendall's method. The results of this approach showed that 86% of the species were correctly identified.

Identification of live moths (Macrolepidoptera) was explored by Watson et al. (2003). Their system is known as Digital Automated Identification System (DAISY). Thirty-five species were used as training images for the system. The system requires the user to align the forewings of each moth for image capture and segmentation. Identification method of species depends on its correlation value with an optimal linear combination of the principal components of each class. Their results showed that correct identifications were heavily affected by the accuracy of forewings alignment. Correct classification was about 83% whereas the best result was 100% and the worst was 35%. The second version of the DAISY system adopted another method beside pattern correlation known as the normalized vector difference (NVD) algorithm which requires including adequate samples for each species. The classification algorithm was based on n-tuple classifier (NNC) and plastic self organizing maps (PSOM).

The limitations of morphometric features for the identification of black-lip Pearl Oyster larvae were evaluated by Paugam et al. (2006). The goal of their study was to determine the most significant morphological identification measures in identifying black-lip Pearl Oyster larvae to be able to distinguish them from three other related species. The software (Optilab, Grafetec, France) was used to automatically number retrained areas, processed from larvae images, and perform a series of 42 measurements like coordinates, optical density, shape parameters, etc. The principal component analysis was used to determine the best descriptors among the 42 data sets which turned out to be six parameters (ellipse ratio, elongation factor, compactness factor, moment of inertia, type factor). Statistical analysis of the descriptors showed that 77% of the black-lip larvae were correctly identified.

Mayo et al. (2007) continued their investigation of automatically identifying live moths. In this study, the accuracy rate was 85% without manual specification of region of interest like their previous system, Digital Automated Identification System (DAISY). A dataset of 774 images of individual live moths belonging to 35 different species was built. After extracting the feature vectors of each image, a toolkit (collection of machine learning algorithms for data mining tasks), WEKA, (The University of Waikato, San Francisco) was used to classify the moths by species. About 11,300 numeric features were extracted from each image at multiple points during processing. These features were a combination of global image and also local image statistics, obtained by centering a grid of total size 600×600 pixels over the centroid of the moth. This mask was then subdivided into 400 square patches, each patch around 30×30 pixels in size. For each area, the mean, minimum, maximum, and standard deviation of the pixel values was calculated and added to the feature vector. However, this approach may not be as effective for smaller species that occupy less space in the image.

A group of twelve researchers (Larios et. al., 2007) have worked for the development of an automated approach to identify stonefly larvae. Their project includes designing and building a mechanical device that can transport insects through the field of a microscope and automatically photograph them. The stoneflies were imaged by an apparatus that manipulates the specimens into the field view of a microscope. The classification process included: identification of regions of interest, representation of these regions as SIFT vectors, classification of SIFT vectors into learning descriptors, formation of histogram of detected descriptors, and classification of the descriptor histogram via state-of-the-art ensemble classification algorithm.

The authors (Larios et. al., 2007) tested three region descriptors on each image: Hessian-affine detector (Mikolajczyk and Schmid, 2004), Kadir entropy detector (Kadir et al., 2004), and the authors' detector principal curvature-based region (PCBR). The results of this study showed that the best classification was obtained when using a combination of all three detectors. The successful rate for four-class was 82% where it was 95% for three-class accuracy. However, at this stage, the system is not completely automated because the specimen is inserted manually into the acrylic to be pumped through the tube. Also, when photographing the specimen at different views, human operator has to make a decision about which of the images is suitable for recognition process. There are some issues that the authors were considering to improve their system. These include the mechanical apparatus and the software (Larios et. al., 2007).

The identification of bivalve larvae with image analysis using discriminant analysis was studied by Hendriks et al. (2005). Two methods were applied for recognition; first, compiling species-specific dimensions and second, compiling the shape parameters (contour of the larvae shell) from the available dataset. The first method failed to provide reliable results whereas the second one when applied to large larvae (length > 150µm) using discriminant analysis showed better results. The technique could identify up to 74% of the large larvae correctly.

### 3.3 Analysis of Previous Work

Among the closely related studies to this project, Digital Automated Identification System (DAISY) (Watson et al., 2003), Automated Bee Identification System (ABIS) (Arbuckle et al., 2001), Species Identification Automated and web Accessible (SPIDA) (Do et al., 1999), and the Automated Insect Identification through Concatenated Histograms of Local Appearance (AIICHLA) (Larrios et al., 2007), were very significant in the field. However, these systems have some limitations and may not be applicable for identifying all insects. The target group that DAISY was designed to identify is Ophioninae (*Hymenoptera: Ichneumonidae*). For accurate classification, the system requires that insects are aligned for capturing their image. In other words, the system may not be applicable for field application where no human interaction is preferred. Furthermore, for insects that are closely related and similar in shape, large number of training images would be required especially with the random n-tuple classifier (NCC) used in this system. However, the system could be a good tool for routine identification of a targeted group of insects.

ABIS system was designed specifically to identify bees based on differences of their forewings. It requires user interaction for aligning the specie's wing before capturing its image. Also, the system is limited to species with membranous wings as the algorithm depends on a specific set of characters of the wing venation for identification. In the SPIDA-web system, manual manipulation of spider specimen is required for proper image acquisition. User interaction is also required for region selection and preprocessing of images. The AIICHLA system is specifically designed to identify stonefly larvae which live in water. An operator has to make sure that the larvae are in the standard orientation for properly capturing their images.

Based on the cited literature of insects' identification systems, no fully automated system for identifying insects in the field have been developed thus far. Furthermore, to our knowledge, no recognition system has been designed specifically for identifying pecan weevils. The absence of such a system motivated the research of this project.

CHAPTER IV

# MATERIALS AND METHODS

## Recognition Methods

### Introduction

The shape of an object is an important feature for certain image recognition application. In general, there are two criteria for representing the shape of an object. First, the shape descriptors should be sufficiently accurate so that they uniquely represent that shape. Second, they should be broad enough to be insensitive to minor variations among objects of the same type. This applies, in particular, to biological objects (as is the case in this study), since they are irregular objects.

The shape of objects can be represented by different methods. These methods are generally classified under two major categories of shape representation, the boundary-based and region-based methods. Boundary-based representations utilize only the information of the shape boundary whereas the region-based techniques consider the internal and external details of the shape.

The boundary-based method can represent the boundary as a whole, and a feature vector derived from the whole boundary is used to describe the shape. The other approach of representing the boundary is by segmenting them into primitives using certain criterion. Techniques like Shape Signature, Fourier Descriptors, Wavelet Descriptors, Chain Code, and Gaussian are examples of this approach.

The region based method has been widely implemented in many applications because it is usually simple to acquire and is descriptive sufficiently. In this technique, all pixels within a shape region are taken into account to obtain the shape representation. An advantage of this technique is that these pixels can be utilized to describe non-connected and disjoint shapes. However, region-based representations do not emphasize boundary features, which could be very crucial for some recognition applications. Examples of some common region-based methods include the Area, Holes, Euler number, Geometrical moments, Zernike moments, Pseudo-Zernike moments, and Legendre moments. Figure 4.1 presents the major classification of these shape representation methods.

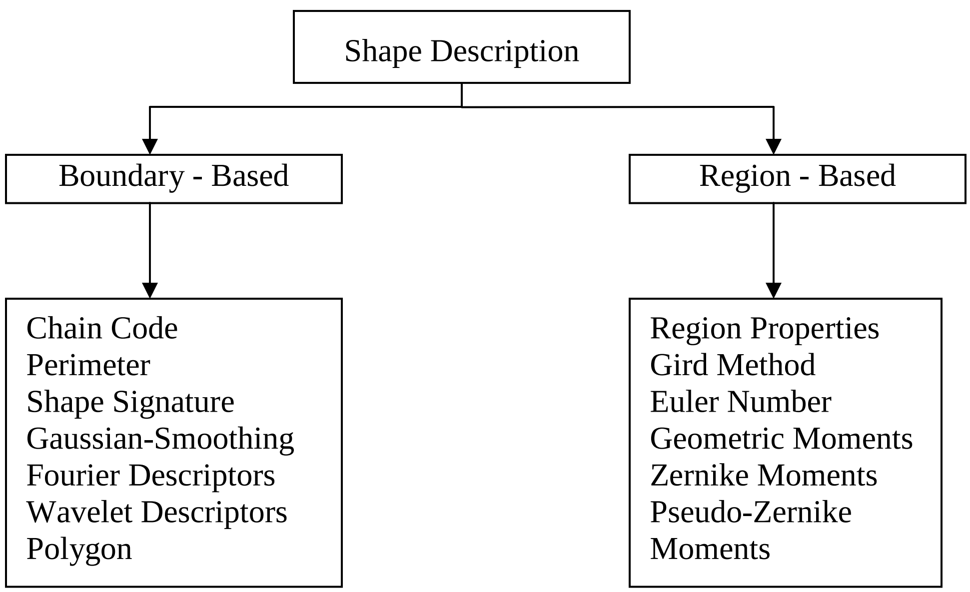



**Figure 4.1 Taxonomy of Some Shape Description Methods**

Boundary–based methods depend on extracting the boundary information of object’s shape. However, sometimes, this information may not be available or accurate due to variety of reasons.  On the other hand, Region-based methods are not heavily dependent on shape boundary information (as Boundary-based method) but they do not represent local features of the shape. Therefore, both classes of shape description are necessary for general object recognition application.

In this study, methods from both types of shape descriptors were used. Fourier descriptors and String matching methods were implemented as boundary-based method, whereas, Geometric moments, Zernike moments, and Region properties were selected from Region-based method. In addition to these methods, the Normalized cross-correlation method was also employed in this application. These selected methods are reviewed and discussed in the next section of this chapter. For each method, a brief introduction is given followed by the mathematical background and the algorithm used for that method.

## 4.1 Normalized Cross- Correlation Based Template Matching

### Introduction and Review

Correlation-based template matching is a standard way of finding a match of a template in a given image. The template could be a part of an object or the whole object that need to be found in an image. The task of correlation-based matching is extended to find the location of the template in the tested image. The correlation calculates the similarity between the template and the area in the input image. A large value of the correlation indicates a high similarity. In this method, no pre- or post- processing is required. The computational complexity of correlation-based method depends on the size of the template and the image. However, when the searching area (image) is large, it is expected that the searching process will take longer time. Moreover, the computational cost will be high in case of searching large databases.

Goshtasby et al. (1984) used two-stage template matching with cross correlation as the similarity measure. In a human face recognition task, Brunelli and Poggio (1993) compared two algorithms, one based on the computation of a set of geometrical features, and the second based on correlation template matching. The results showed that perfect recognition was achieved when using template matching where only 90% correct recognition was obtained when using the other method.

Correlation-based template matching is also applicable to binary images. Wallace (1988) applied boundary correlation to geometrically matched rigid object-model contours with image contours. The normalized form of correlation has been computed in the spatial domain because there is no simple and efficient frequency domain expression (Lewis 1995).  Darrell et al. (1996) selected normalized correlation to be the similarity measure between an image and a set of spatial view models to obtain real-time performance in hand and face gestures recognition application.

Choi and Kim (2002) used correlation as a first stage in a rotation and illumination invariance template matching method. In this step, matching candidates are selected using the vector sum of circular projections of the sub-image, which is low in computational cost. In the second stage, matching is performed only on these selected candidates using Zernike moments. Farag et al. (2004) developed an algorithm based on a combination of Normalized cross-correlation template matching and Bayesian post-classification to detect and recognize lung cancer during image mass screening of spiral computer tomographic chest scan.

In a real-time tracking system of a sequence of images, Wang et al. (2005) studied the use of template matching.  Kim et al. (2003) used a combination of three correlation template matching- sum of absolute difference (SAD), sum of squared difference (SSD), and maximum absolute differences (MAD). Kohandani et al. (2006) proposed a template edge matching algorithm that utilized the edge information to perform the template matching. Their results showed that the proposed algorithm has higher error in finding the template location but performed four times faster on average than the method based on original algorithm.

**Matching by Correlation**

The template will be denoted as $t(x, y)$ of size $Q \times L$ that is to be matched with an image $f(x, y)$ of size $M \times N$ where the size of the template should be less than or equal to the size of the image. The sum of squared differences (SSD) is a similarity measure widely used in computer vision. In a gray level image, differences of the sum squared of each corresponding template and input image pixel is taken as an indication of the similarity between the template and the searched area of the image (Storring and Moeslund, 1997).

$$SSD(x, y) = \sum_{n=1}^{N}\sum_{m=1}^{M}\left[f(x+n, y+m) - t(n, m)\right]^2 \tag{4.1}$$

The cross-correlation can be derived from the SSD as

$$SSD(x, y) = \sum_{n=1}^{N}\sum_{m=1}^{M} f(x+n, y+m)^2 + \sum_{n=1}^{N}\sum_{m=1}^{M} t(n, m)^2 - 2\cdot\sum_{n=1}^{N}\sum_{m=1}^{M} f(x+n, y+m)^2 \cdot t(n, m) \tag{4.2}$$

In this equation, the energy of the searched area and the template are represented by the first and second terms respectively. The last term is the cross correlation (CC) which forms the correlation between the image and the template. The value of the CC ranges from zero (no match), to [N.M. $255^2$] the maximum value. The need for normalizing the cross correlation (NCC) term appeared since the energy of the different searched area in an image is not usually constant (Storring and Moeslund, 1997). The CC can be normalized as follows:

$$NCC(x,y)=\frac{\sum_{n=1}^{N}\sum_{m=1}^{M} f(x+n,y+m)^2\cdot t(n,m)}{\sqrt{\sum_{n=1}^{N}\sum_{m=1}^{M} f^2(x+n,y+m)\sum_{n=1}^{N}\sum_{m=1}^{M} t^2(n,m)}} \tag{4.3}$$

As Equation 4.3 shows, the normalization is done by dividing the CC with the square root of the energy of the searching area and the template. The range of the NCC is between [0, 1] where zero indicates no match and 1 for identical match.

In this study, NCC was used with a simple algorithm to identify pecan weevils among other insects. First, the program reads the gray level input image and the image of pecan weevil stored in the database. Then, the input image will be treated as a template and normalized cross correlation performed between this template and the database images one by one. If the value of the correlation is greater than the experimentally determined threshold (0.75), then the input image will be recognized as a pecan weevil.

## 4.2 Geometric Moment Invariants

### Introduction and Review

Hu (1962) presented a theory of two-dimensional moment invariants for planar geometric figures based on the work of the 19$^{th}$ century mathematicians Boole, Cayley and Sylvester. In this theory, a set of invariants based on combinations of regular moments using algebraic invariants were derived. These seven moments are invariant to translation, scale, and orientation.  He also implemented these moments to recognize the Latin alphabetic characters.

Many studies on deriving different types of invariants and using invariant moments for object recognition have been carried out since the introduction of Hu's theory. Ezer et al. (1994) compared the performance of moment invariants and Fourier descriptors in recognizing planar shapes. They found that moment invariants are affected more under changes in size. They concluded that recognition was better when using the silhouette of the shape. With a significantly lower error rate, Dudani et al. (1977) implemented moment invariants for the automatic recognition of six aircraft types using 132 images of them. Wong et al. (1995) described and implemented an automatic approach for the generation of moment invariants. These new generated moments were used for the recognition of English alphabets.

Abu Mostafa and Psaltis (1984) investigated some aspects of information loss, suppression, and redundancy encountered in moment invariants. They described the behavior of moment invariants in the presence of additive noise. Goshtasby (1985) described rotationally invariant template matching using normalized invariant moments. In this technique, the zeroeth-order moment was used first to determine the likely match positions, and the second and third-order moments were later used to determine the best match position among the likely ones.

Flusser and Suk (1994) used affine moment invariants, derived by them, as the features for recognizing handwritten characters independent of their size, slant and other variations. To estimate the object pose in an image, Mukundan and Ramakrishnan (1996) used geometric moments as image feature descriptors to establish the correspondence between the object and image space transformations. They (1993) also used a moment based method for determining the three-dimensional rotational parameters of a rigid body. Second- and third-order moment invariants were used to construct the feature vector for the scale and orientation.

Sluzek (1988) presented a new approach based on the moment invariants to match partially occluded contours. Sluzek (1995) also used moment invariants for the identification and inspection of 2-D shapes. Using family of shapes created by occluding the object by circles located in the object's center of area, it was possible to create from a single moment invariant many shape descriptors. In 2005, Sluzek combined template matching and moments into a method of designing detectors for various image patterns. The method used a circular window of the size related to the object, then, for each location of the window, moment features were used to determine the best match to the actual content of the window.

El-Khaly and Sid- Ahmed (1990) used moment invariants for the recognition of Arabic characters and their isolation from the printed text. Tsirikolias and Metzios (1993) presented a set of moments which were normalized with respect to standard deviation and used them for optical character recognition. Maitra (1979) provided some improvement in moment invariants to make them invariant to the changes in contrast. Hupkens and De Clippeleir (1995) applied normalization to moment invariants to overcome the intensity invariant of noisy image. Wang and Healey (1998) developed a method for recognizing color texture under various illumination conditions and geometric configurations based on Zernike moments.

Zhang and Lu (2002) used geometrical moments in image retrieval application and concluded that they perform very well on similarity transformed and affinely transformed contour-based shapes. However, their performance was not very effective in the case of distorted contour-based and scaled shapes.

### Object Recognition by Geometrical Moment Invariants

For robust and reliable recognition results, it is desirable to utilize methods which are invariant to translation, scale, orientation, and rotation. One of these methods that have been widely implemented in image pattern recognition and image classification is moment invariants. Hu (1962) defined seven descriptors which computed from central moments through order three that are independent to object translation, scale and orientation. Translation invariance is obtained by computing moments that are normalized with respect to the center of gravity. The size invariance can be computed from algebraic invariants. From the second and third order values of the normalized central moments, a set of seven invariant moments can be computed which are independent of rotation (Sarfraz, 2006). Given a two-dimensional image (Gonzalez and Woods, 2001), $f(x, y)$, the moments of order $(p+q)$ are defined as:

$$M_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x, y)\, dx\, dy \tag{4.4}$$

where p, q = 0,1,2,…,∞.

The central moments are defined as:

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})^p (y-\bar{y})^q f(x, y)\, dx\, dy \tag{4.5}$$

where $\bar{x} = \frac{m_{10}}{m_{00}}$ *and* $\bar{y} = \frac{m_{01}}{m_{00}}$

For a digital image, Equation (4.5) becomes:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \tag{4.6}$$

The central moments of order up to 3 are:

$$\begin{aligned} \mu_{00} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^0 f(x, y) \\ &= \sum_x \sum_y f(x, y) \\ &= m_{00} \end{aligned} \tag{4.7a}$$

$$\begin{aligned} \mu_{10} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^0 f(x, y) \\ &= m_{10} - \frac{m_{10}}{m_{00}}(m_{00}) \\ &= 0 \end{aligned} \tag{4.7b}$$

$$\begin{aligned} \mu_{01} &= \sum_x \sum_y (x - \bar{x})^0 (y - \bar{y})^1 f(x, y) \\ &= m_{01} - \frac{m_{01}}{m_{00}}(m_{00}) \\ &= 0 \end{aligned} \tag{4.7c}$$

$$\begin{aligned} \mu_{11} &= \sum_x \sum_y (x - \bar{x})^1 (y - \bar{y})^1 f(x, y) \\ &= m_{11} - \frac{m_{10}\ m_{01}}{m_{00}} \\ &= m_{11} - \bar{x} m_{01} = m_{11} - \bar{y} m_{10} \end{aligned} \tag{4.7d}$$

$$\begin{aligned} \mu_{20} &= \sum_x \sum_y (x - \bar{x})^2 (y - \bar{y})^0 f(x, y) \\ &= m_{20} - \frac{2m_{10}^2}{m_{00}} + \frac{m_{10}^2}{m_{00}} \\ &= m_{20} - \frac{m_{10}^2}{m_{00}} \\ &= m_{20} - \bar{x} m_{10} \end{aligned} \tag{4.7e}$$

$$\mu_{02} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^2 f(x,y)$$
$$= m_{02} - \frac{m_{01}^2}{m_{00}} \tag{4.7f}$$
$$= m_{02} - \bar{y} m_{01}$$

$$\mu_{21} = \sum_x \sum_y (x-\bar{x})^2 (y-\bar{y})^1 f(x,y)$$
$$= m_{21} - 2\bar{x} m_{11} - \bar{y} m_{20} + 2\bar{x}^2 m_{01} \tag{4.7g}$$

$$\mu_{12} = \sum_x \sum_y (x-\bar{x})^1 (y-\bar{y})^2 f(x,y)$$
$$= m_{12} - 2\bar{y} m_{11} - \bar{x} m_{02} + 2\bar{y}^2 m_{10} \tag{4.7h}$$

$$\mu_{30} = \sum_x \sum_y (x-\bar{x})^3 (y-\bar{y})^0 f(x,y)$$
$$= m_{30} - 3\bar{x} m_{20} + 2\bar{x}^2 m_{10} \tag{4.7i}$$

$$\mu_{03} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^3 f(x,y)$$
$$= m_{03} - 3\bar{y} m_{02} + 2\bar{y}^2 m_{01} \tag{4.7j}$$

The normalized central moments $(\eta_{pq})$ can be defined as:

$$\eta_{pq} = \frac{\mu^{pq}}{\mu_{00}^{\gamma}} \qquad \text{where} \qquad \gamma = \frac{p+q}{2} + 1 \tag{4.8}$$

The seven geometrical moment invariants derived by Hu (1962) from the second and the third order moments are:

$$\phi_1 = \eta_{20} + \eta_{02} \tag{4.9}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \tag{4.10}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{21} - 3\eta_{03})^2 \tag{4.11}$$

$$\phi_4 = (\eta_{30} + 3\eta_{12})^2 + (\eta_{21} + 3\eta_{03})^2 \tag{4.12}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12}) + (\eta_{30} + 3\eta_{12})^2 \left[ (\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[ 3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \tag{4.13}$$

$$\phi_6 = (\eta_{20} - 3\eta_{02}) \left[ (\eta_{30} + 3\eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] + 4\eta_{11}(3\eta_{30} - \eta_{12})(\eta_{21} + \eta_{03}) \tag{4.14}$$

$$\phi_7 = (3\eta_{21} - 3\eta_{03}) + (\eta_{30} + 3\eta_{12}) \left[ (\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 \right] \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[ 3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2 \right] \tag{4.15}$$

**Similarity Measure**

The idea of template matching is to measure the similarity between an input image and a database of known shapes. When comparing two images, say G and H, then, two sets of values $g(i)$ and $h(i)$, will be produced by the two images. The similarity (distance) between them can be measured as $d(i) = g(i) - h(i)$. The smaller the distance is, the closer the two shapes are to each other and vice versa. If the distance, $d(i) = 0$, then the two shapes are identical. In measuring similarity, it is always desired to represent the result by a single value instead of a set of values like in $d(i)$. This can be done by treating $d(i)$ as a vector in multi-dimensional space where the length of this vector represents the distance between the two compared images (Sarfraz, 2006). The value of the distance can be obtained from the square root of the sum of the squares of the elements of $d(i)$.

To calculate the similarity degree of the corresponding moment invariants of an input insect's image and the database of pecan weevils' images, Euclidean Distance (ED) was utilized as the classifier measure. The ED's equation can be written as:

$$D = \sqrt{\sum_{i=1}^{n} (g(i) - h(i))^2} \tag{4.16}$$

For a given input (insect's image), these seven moment invariants were calculated and compared with the ones of each pecan weevil image stored in the database. Euclidean similarity measure was used for comparison. The algorithm used in this study to measure the seven moment invariants has closely followed the algorithms presented in Gonzalez et al. (2004) and Sarfraz (2006). The algorithm was designed to execute the following steps:

1. Load the input gray level image
2. Convert the raw image to a binary
3. Label the image, remove all objects except the largest (insect's body)
4. Extracted object in the last step gets loaded as a new image
5. Boundaries of the image will be detected
6. Calculate the seven moments
7. Normalize the moments by dividing them with the first one to eliminate the scale effect
8. Compare the obtained moments of the tested image with all moments of the pecan weevil images using Euclidean distance
9. Input image with least Euclidean distance is recognized as pecan weevil

## 4.3 Zernike Moments

### Introduction and Review

Geometrical moments are the most common and simplest moments; but in many cases they do not represent the image features efficiently (Zhenjiang, 2000). Zernike moments perform better than other methods despite the fact that they are more complicated than geometrical moments. Zernike moments were first introduced by

Teague (1980) who defined them based on Zernike polynomials. These moments are orthogonal to each other which allow them to represent the properties of images with no redundancy or overlap of information between the moments. Zernike moments have been used in many recognition applications and their performances have been compared with other methods. Different combinations of Zernike moments and some recognition methods have been studied as well. Hwang and Kim (2006) proposed an approach that computes Zernike moments from a digital image eight times faster than an existing method.

The literature strongly suggested the superiority of Zernike moments in many identification tasks over geometric moment invariants (Hu's moment). Liao and Pawlak (1996) addressed the issue of increasing the accuracy and efficiency of moment invariants and proposed some techniques for their improvement. These approaches were used for image reconstruction from the orthogonal Legendre moments. Zernike moments are the most desirable shape descriptors among other moments (Zhang et al., 2004).

Teh (1988) compared the performances of geometrical moments, Legendre moments, Zernike moment, pseudo-Zernike moments, rotational moments, and complex moments. Their results showed that Zernike moments were better than other types of moments in terms of information redundancy and overall performance. Khotanzad and Hong (1990) introduced a set of invariants of features based on Zernike moments. These features were tested using clean and noisy images from a 26-class character data set and a 4-class lake data set. Belkasim et al. (1991) studied the effectiveness of different moment invariants in the recognition of handwritten numerals and aircrafts. The best overall performance was obtained when applying their method of deriving Zernike moments along with normalized scheme even with some additive noise.

In retrieval application, Kim and Kim (2000) used Zernike moments to match images of large database under various transformations and similarity–based retrieval. Their results showed that Zernike moment can be used as an effective descriptor of global shape retrieval. Zhenjiang (2000) utilized the properties of Zernike moments to define the method of roundness measurement, which was used to analyze the roundness of rose flower shapes. Shen and Ip (1999) compared the performance of several moment invariants method. The correct classification achieved by Zernike moments was 98.7% compared to 75.3% for Li's moments.

Lin and Chou (2003) studied the first 36 Zernike moments and found the dependence relations between them. They applied their proposed compact representation of Zernike moments on image retrieval application which turned out to be faster than original representation. Park and Kim (2004) used Zernike moment as a region–based classifier in their proposed shape–based image retrieval method. They concluded that using Zernike moment beside Fourier descriptors provided better results than other similarity methods. In a comparison of 2-D moment-based description techniques, Di Ruberto and Morgera (2005) found that Zernike moments gave the best result compared to the others.

**Object Recognition by Zernike Moments**

Zernike moment descriptor has the properties of rotation invariance, robustness to noise, expression efficiency, fast computation and multi-level representation for describing the various shapes of patterns (Kim and Kim, 2000). In many comparison studies of moments based methods (Teh and Chin, 1988; Lin and Chou, 2003; Belkasim et al., 1991; Zhang and Lu, 2004; Park and Kim, 2004; Ezer, 1994; Padilla-Vivanco et al., 2007 and Liao and Pawlak, 1996) Zernike moments outperformed the others methods, especially the geometrical moments.

Zernike introduced a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle of $x^2 + y^2 = 1$. The computation of Zernike moments from an input image consists of three steps: computation of radial polynomials, computation of Zernike basis function, and computation of Zernike moments by projecting the image on to the basis function (Hwang and Kim, 2006). The form of these polynomials is as follows:

$$V_{nm}(x, y) = Vnm\left(\rho, \theta\right) = R_{nm}\left(\rho\right)\exp\left(jm\theta\right) \tag{4.14}$$

where $n = 0,1,2,3,....\infty$ n is called “order”, m is a positive and negative integer (known as “repetition”) with constraint that $n - |m|$ is even and $|m| \leq n$, $\rho$ is the length of vector from origin to $(x, y)$ pixel, $\theta$ is the angle between vector $\rho$ and $x$ axis in counter-clockwise direction, $R_{nm}$ is the radial polynomial defined as:

$$R_{nm} = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s\left[(n-s)!\right]\rho^{n-2s}}{s!\left(\frac{n+|m|}{2} - s\right)!\left(\frac{n-|m|}{2} - s\right)!} \tag{4.17}$$

These polynomials are orthogonal and satisfy the orthogonal properties for the same repletion,

$$\int\int_{x^2+y^2\leq 1} \left[V_{nm}(x, y)\right] V_{pq} \quad (x, y)dxdy = \frac{\pi}{n+1}\delta_{np}\delta_{mq} \tag{4.18}$$

where $\delta = \begin{cases} 1 & (a = b) \\ 0 & (\text{otherwise}) \end{cases}$

The Zernike moments of order $n$ with repetition $m$ for a continuous image function $f(x,y)$ outside the unit circle is:

$$A_{mn} = \frac{n+1}{\pi} \iint_{x^2+y^2 \le 1} f(x,y)\,[V_{nm}(\rho,\theta)]\,dxdy \tag{4.19}$$

In Equation 4.17, the integral can be replaced by summations (since all the images are digital) as follows:

$$A_{mn} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y)\,[V_{nm}(\rho,\theta)] \qquad \text{where } x^2 + y^2 \le 1 \tag{4.20}$$

The Zernike moments are computed for an image by considering the center of the image as the origin and the pixel coordinates are mapped to the range of the unit circle. The computation will not include pixels outside the unit circle. The orthogonality implies no redundancy or overlap of information between the moments with different orders and repetitions (Hwang and Kim, 2006). In this case, each moment will be a unique and independent representation to a given image.

**Similarity Measure**

For this method, the degree of similarity was evaluated by calculating the Euclidean Distance of the corresponding Zernike moments of an input insect's image and the database of pecan weevils' images:

$$D = \sqrt{\sum_{i=1}^{n} \left(g(i) - h(i)\right)^2} \tag{4.21}$$

where $g(i)$ and $h(i)$ are the i$^{th}$ corresponding Zernike moment for the unknown image and the pecan weevil image respectively.

### Algorithm

In this study, an algorithm presented by Chang (2000) at the University of Texas was adopted for computing Zernike moments. The algorithm executes the following steps:

1. Load the input (gray level) image
2. Convert the raw image to a binary
3. Label the image, removes all objects except the largest (insect's body)
4. Extracted object in the last step gets loaded as a new image
5. Normalize the coordinates assuming the center is the origin
6. Compute the unit disk mask
7. Mask the pixels which are lying inside or on unit circle by clipping the input image matrix regions into a vector
8. Compute Zernike polynomial over unit circle at a given order n and repetition m which will result in the magnitude of the matrix and the phase

## 4.4 Fourier Descriptors

### Introduction and Review

Fourier descriptors (FDs) method is one of the boundary-based shape descriptors techniques and has been implemented for classification of different types of objects. FDs can provide characteristics of an object that uniquely represent its shape (Sarfraz, 2006). Their main advantage is invariance to translation, rotation and scaling of the observed object. Invariant FDs which were first introduced by Granlund (1972), described the rotational symmetry of a shape. He tested the method in classifying hand printed letters which resulted in 98% correct classification.

Many FD methods have been developed for shape analysis, character recognition, shape coding, shape classification, and shape retrieval (Zhang and Lu, 2002). Persoon and Fu (1986) used FDs in character recognition and machine parts recognition. Rui et al. (1999) proposed a modified Fourier descriptor to describe and compare closed planar curves which account for the effects of spatial discretization of shapes. Kauppinen et al. (1995) used FDs in 2D shape classification and found that it performed better than autogressive modeling method especially in the presence of noise. Chellappa and Bagdazian (1984) described a transform coding scheme for image boundaries using Fourier computations. Reeves et al. (1988) used FDs in recognizing a three-dimensional object from a two-dimensional image where they performed almost the same as the standard moments. Park and Kim (2004) concluded that image retrieval success rate increased significantly when FDs were used.

Zhang and Lu (2002) proposed a generic Fourier descriptor that derived by applying two-dimensional Fourier transform on a polar-raster sampled shape image. Their experimental results showed that the proposed method performed better than some existing contour-based and region-based shape descriptors. In image retrieval application, Zhang and Lu (2002) studied the performance of several FDs that have been derived from different signatures. Their results showed that shape retrieval using FDs derived from centroid distance signature is significantly better than that using FDs derived from the other three signatures. In another comparative study, Zhang and Lu (2002) found that FDs performed better than the curvature scale space method in terms of robustness, low computation, hierarchical representation, retrieval performance.

**Object Recognition by Fourier Descriptors**

Fourier descriptors are produced by the Fourier Transformation which represents the shape in the frequency domain. The lower frequency descriptors store the general information of the shape of the shape and the higher frequency the smaller details (Sarfraz, 2006). Therefore, the lower frequency components of the Fourier descriptors are sufficient for general shape description.

The boundary of a shape consists of $K$ points in the $xy-$plane. Tracing once around the boundary from an arbitrary starting point $(x_0, y_0)$, in the counterclockwise direction, at a constant speed produces a sequence of coordinate pairs $(x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots, (x_{K-1}, y_{K-1})$. For representing traversal at a constant speed, it is necessary to interpolate equidistant points around the boundary. The boundary can be represented as the sequence of coordinates $s(k) = [x(k), y(k)]$, for $k = 0, 1, 2, \ldots, K-1$. The coordinate pair of shape boundary can be described as a complex number as:

$$s(k) = x(k) + jy(k) \tag{4.22}$$

where $j = \sqrt{-1}$ . This representation changed the problem from two-dimensional to one-dimensional case.

The discrete Fourier transform of Equation 4.18 is:

$$a(u) = \frac{1}{K}\sum_{k=0}^{K-1} s(k)e^{-j2\pi uk/K} \tag{4.23}$$

for $u = 0, 1, 2, \ldots, K-1$ and the complex coefficients $a(u)$ are known as Fourier descriptors of the boundary. The inverse Fourier transform of Equation 4.19 is:

$$s(k) = \frac{1}{K}\sum_{u=0}^{K-1} a(k) e^{-j2\pi uk/K} \tag{4.24}$$

where ‘k’ is the number of points in the boundary and ‘s’ is the featured value from Fourier descriptors for object recognition and representation. As mentioned earlier, high frequency components account for fine detail and low frequency components determine global shape, therefore, not all Fourier descriptors are required for general object recognition. Instead, only the first P coefficients should be used. In this case, Equation 4.20 can be rewritten as:

$$\hat{s}(k) = \sum_{u=0}^{P-1} a(u) e^{j2\pi uk/K} \tag{4.25}$$

As a result, smaller the P value is, the finer details would be lost on the boundary. On the other hand, the fewer components we include in the calculations, the faster the algorithm will be.

The Fourier descriptors are not invariant directly to the geometrical changes like rotation, translation, scale, and starting point of processing. However, applying specific transformation to each individual variant would make Fourier descriptors insensitive to that particular change. The following set of equations present the Fourier descriptors for a boundary sequence that experience changes in rotation, translation, scaling, and changes in starting point (Gonzalez and Woods, 2001):

Rotation, $$a(u) = \frac{1}{K}\sum_{u=0}^{K-1} s(k) e^{j\theta} e^{-j2\pi uk/K} \tag{4.26}$$

Translation, $$a(u) = \left(\frac{1}{K}\sum_{u=0}^{K-1} s(k) e^{-j2\pi uk/K}\right) + \Delta_{xy} \tag{4.27}$$

where $\Delta_{xy}$ is the added constant displacement to all coordinates in the boundary $s(k)$ which can be presented as:

$$s(k) = [x(k) + \Delta x] + j[y(k) + \Delta y]$$

Scaling, $$a(u) = \left( \frac{1}{K} \sum_{u=0}^{K-1} s(k) e^{-j2\pi uk/K} \right) .\alpha \tag{4.28}$$

Starting Point, $$a(u) = \frac{1}{K} \sum_{u=0}^{K-1} s(k) e^{-j2\pi uk_0/K} \tag{4.29}$$

In Equation 4.28, the boundary sequence is defined as: $s(k) = x(k - k_0) + jy(k - k_0)$, which indicates that the starting point of the sequence changes from $k = 0$ to $k = k_0$.

**Similarity Measure**

Euclidean Distance (ED) was implemented in this method as well, as a classifier to measure the similarity degree of the corresponding Fourier descriptors of an input insect image and the database of pecan weevil's images. The ED's equation can be written as:

$$D = \sqrt{\sum_{i=1}^{n} (g(i) - h(i))^2} \tag{4.30}$$

Using Fourier descriptors, an acquired image is recognized as pecan weevil when the value of ED is less than or equal to 1.0 (experimentally determined threshold).

**Algorithm**

A simple and fast algorithm was implemented in identifying pecan weevils among other insects. This algorithm follows closely the methods presented by Sarfraz (2006) and Gonzalez et al. (2004) with some modifications.

1. Load all images
2. Convert the loaded images to binary images
3. Label connected components in binary image
4. Include the main object and eliminate the rest
5. Create a binary image containing only the main object
6. Find the boundary of the image
7. Convert the complex pairs to 1-D vector of descriptors
8. Select a number that could represent the shape of the insect (450 Fourier Descriptors )
9. Normalize the selected Fourier descriptors by dividing each element of the 450 descriptors by the first element (alpha)
10. Measure the similarity between a given image and the training set by using Euclidean distance
11. An input image Euclidean distance less than or equal the experimentally determined threshold of 1.0 will be recognized as pecan weevil

## 4.5 String Matching

### Introduction and Review

String matching is one of the boundary-based descriptors in which the boundary of a shape can be represented by a string. Strings are one-dimensional structures representing the boundary of two-dimensional shape. This representation requires an appropriate method for reducing the two-dimensional relations to one-dimensional form (Gonzalez and Woods, 2001). The fundamental concept of using strings as descriptors is to extract the connected line segments from the shape to be recognized. The approach implemented in this study is to track the contour of an insect and code the result with segments of specified direction (angles).

Sze and Yang (1981) proposed two similarity measures between strings. Wu and Wang (1999) proposed a two-stage string matching method to recognize two-dimensional objects. First, global string matching is conducted; then, in the second stage, the local dissimilarity measure is computed. Wolfson (1990) introduced two algorithms to find the longest common sub-curve of two 2-D curves. In 2001, Wu introduced a feature for two-dimensional object recognition named as “the reciprocal of compactness of triangles formed by two adjacent dominant String matching techniques for two-dimensional objects”. The author concluded based on some experimental results that no parameters need to be set in the recognition process.

Lee and Lee (1998) used a two-stage String matching method in a scheme to inspect two-dimensional objects for dimensional and shape verification in industrial environment. Their method first determines the invariant starting point for boundary tracing and locating the corner points of a curved object for polygon. In the second stage, their method utilizes the feature string for each tested object to find the exact correspondence to one of several model objects.

Pavlidis (1979) introduced a contour matching algorithm. Pavlidis and Ali (1979) used a syntactic shape analyzer whose output is a description of the contour as a sequence of arcs, each with a set of attributes. Sze and Yang (1981) proposed a variation of the previous study. Gdalyahu and Weinshall (1999) proposed an algorithm for matching curves under substantial deformations and arbitrary large scaling and rigid transformations and applied the algorithm to silhouette matching. In their design, they constructed a syntactic representation for both curves and an edit transformation.

**Matching by Strings**

In this method, the boundary of an insect is represented by a string which is generated by coding the interior angles of the polygons. Then, strings were generated from a given angle array by quantizing the angles into $45^\circ$ increments which produced strings whose elements were numbers between 1 and 8 (Gonzalez et al., 2004) with 1 increment. Table 4.1 presents this relationship. For an input image of unknown insect and pecan weevil, the two boundaries can be coded into strings denoted $a_1\ a_2,...,a_n$ and $b_1\ b_2,...,b_n$ respectively. If $\alpha$ represents the number of matches between the two strings, and the match takes place in the $k^{th}$ location, then the number of unmatched symbols can be described as: $\beta = \max(|a|, |b|) - \alpha$ where $|a|$ and $|b|$ are the length of the string representing the unknown insect and the pecan weevil images respectively. In this case, the value of $\beta$ is equal to zero if the two images are identical.

**Table 4.1 Designated Integers for Each Angle's Range**

| Angle Range | Symbol Representing the Range |
|---|---|
| $0^\circ \le \theta < 45^\circ$ | 1 |
| $45^\circ \le \theta < 90^\circ$ | 2 |
| $90^\circ \le \theta < 135^\circ$ | 3 |
| $135^\circ \le \theta < 180^\circ$ | 4 |
| $180^\circ \le \theta < 225^\circ$ | 5 |
| $225^\circ \le \theta < 270^\circ$ | 6 |
| $270^\circ \le \theta < 315^\circ$ | 7 |
| $315^\circ \le \theta < 360^\circ$ | 8 |

**Similarity Measure**

Even though there are many definitions of string similarity, a simple measure between strings was implemented in this study which can be represented by the following ratio:

$$D = \frac{\alpha}{\beta} = \frac{\alpha}{\max(|a|, |b|) - \alpha} \tag{4.31}$$

The value of $D$ is equal to zero when none of the symbols in $a$ (unknown insect's image) and $b$ (pecan weevil's image) is matched; $D$ is equal to infinite when the two images are identically matched. In String matching, a tested image is recognized as pecan weevil if the D value is greater than or equal to the value (1.0) of the threshold.

**Algorithm**

1. Load all images
2. Convert the loaded images to binary images
3. Label connected components in binary image
4. Include the main object and eliminate the rest
5. Create a binary image containing only the main object
6. Compute the minimum perimeter polygon (by using the Matlab function *minperpoly*) and form it in a string
7. Compute the internal polygon angles of the string by coding the interior angles of the polygon
8. Form a string from the angles array by quantizing the angles into $45^\circ$ increments in which the elements of this string are integers from 1 to 8 representing the angles' increments.
9. Convert the yielded string from a string of integers to a character string

10. Measure the similarity between two strings using Matlab function *strsimilarity*

### 4.6 Regional Properties Descriptors

While the aim of this study is to identify pecan weevils among other insects, it is also desired to keep such a system as simple as possible. A regional property is one of the approaches among regional descriptors as it deals with the region(s) of the image instead of its boundary. It is a simple method for describing important properties of image regions such as, the area, centroid, and orientation. Although there are many insects that are very close to pecan weevils in terms of shape description, one important feature can be utilized to distinguish pecan weevils from other insects. This feature is the pecan weevil's rostrum. It has been previously mentioned (Chapter II) that pecan weevil can be recognized by its long rostrum which is ¾ the length of the male's body and as long as the female's body.

As pecan weevil is not the only insect that has a rostrum, hence, utilizing this feature alone (major-axis length) may not be very effective. Therefore, this feature was related to other features in order to form a unique representation of pecan weevils. The area, major-axis length, and minor-axis length were used to describe pecan weevils in this project. The area of the selected region is defined as the number of pixels in that region. The major-axis length can be defined as the length (in pixels) of the major axis of the ellipse that has the same second moments as the region. Finally, the minor-axis length is the length (in pixels) of the major axis of the ellipse that has the same second moments as the region (Gonzalez et al., 2004).

**Similarity Measure**

Euclidean Distance (Equation 4.29) was utilized to measure the similarity between the testing images and the database of pecan weevils' image. In this method, an input image is recognized as pecan weevil if the ED value is greater than or equal to the experimentally determined threshold which was found to be 1.0.

**Algorithm**

1. Load all images
2. Convert the gray level image to double Precision Floating Point form
3. Clean the image from noise by applying low pass filter
4. Define the boundaries of the image to eliminate the zero-padding artifacts around the edge of the image
5. Convert the resulted images to binary images
6. Label connected components in binary image
7. Include the main object and eliminate the rest (object which its area is >1000 pixels)
8. Create a binary image containing only the main object
9. For each image, compute the area, major-axis length, and minor-axis length of the object and form the three components in one vector
10. Measure the similarity between an input image and the database of pecan weevil's images
11. Recognize the input image as pecan weevil image if the ED is $\leq 1.0$

## 4.7 Materials

### Collection of Insects

Traps were set up for pecan weevils at different locations in Stillwater, Oklahoma. The other source of insects' samples was the Entomology Museum at Oklahoma State University. Over 205 pecan weevils were collected by both sources and these included both males and females. The collected weevils varied in their size, color, age. Furthermore, about 27 other types of insects were also collected to be part of the testing set. These insects are normally present in the pecan habitat. Some insects had 2, 3, 4, or 5 replicate samples and nine insects have only one sample. The names of insects used in the testing data set and their number of replicates are presented in Table 4.2.

**Table 4.2 Insects Used for Testing the Algorithm**

| Family | Insect | Number of Replicates |
|---|---|---|
| Pentatomidae | Apis Mellifera L | 4 |
| Pentatomidae | Brochymena Guadripustulata (Fab) | 5 |
| Acridiae | Chortophaga Viridifasciata (Deg) | 4 |
| Buprestidae | Chrysobothris Femorata (Oliv) | 5 |
| Carabidae | Coleoptera Carabidae | 1 |
| Elateidae | Acrosterunum Hilaris (Say) | 5 |
| Elateridae | Condoerus Lividus (Deg) | 5 |
| Reduviidae | Hemiptera | 1 |
| Microlepidutera | Hyphantria Cunea (Drury) | 4 |
| Alydidae | Leptoglossus Opposites (Say) | 2 |
| Cercopidae | Lepyronia Gibbosa (Ball) | 5 |
| Membracidae | Metealfa Pruinosa (Say) | 4 |
| Lepidoptea | Plathypena Scabra (Fab) | 5 |
| Formicidae | Tomostethus Multicinctus (Rohwer) | 4 |
| Curculionidae | Pantomorus Pallidus (Horn) | 5 |
| Curculionidae | Naupactus Leucoloma (Boh) | 5 |
| Curculionidae | Cyrtepistomus Castaneus (Roolofs) | 2 |
| Curculionidae | Compsus Auricephalus (Say) | 3 |
| Curculionidae | Conotrachelus Elegans (Say) | 5 |

Furthermore, each insect used in this study is shown in Figure 4.2 where the last row of that image shows pecan weevils.

**Figure 4.2 Sample of Insects used as the Testing Set**

**Image Acquisition**

In template-based application, training set of image should be a real representative of the targeted object or shape. Even though traps were checked regularly, few pecan weevils were found alive. Experiments showed that those live weevils die in short time when kept in cages. Moreover, it is very hard to position live weevils appropriately for imaging without causing some damage to their bodies or losing them since they can fly. As a result, live collected insects were set to “sleep” by placing them in a refrigerator at 40 $^{o}$F for 60 minutes.

Muscles of pecan weevil shrink and pull the legs in shortly after they die. This generally results in all six legs of pecan weevil remaining close to the body or sometimes touching the insect's abdomen. Since this system is designed to identify live pecan weevils in the field, images of them in such positions would not simulate the natural appearance of the insects in the field. Therefore, preserved insects' parts (legs and antenna) were stretched out to so that they would appear similar to the position in live insects. In order to achieve good results without losing these fragile parts, some careful pre-processing steps were also undertaken to prepare the collected insects for imaging so that they would appear like live insects.

The first pre-processing step was to put the insects in a humidifying chamber for 10 days. The humidified environment helps in making the insects' parts (legs and antenna) more flexible for stretching them out such that they are closer to their normal position. The second step was to align each insect at the camera view for imaging. All insects were approximately placed at a reference position and orientation. Images of insects were then acquired with the image acquisition system described below.

**The imaging system**

The imaging system consists of an AVT F-145B CCD black and white camera which is an IEEE 1394 SXGA+ camera. It is equipped with a 1.45 megapixel 2/3" progressive CCD sensor. This camera was manufactured by Allied Vision Technologies GmbH 2003, Stadtroda, Germany. Images of insects were of the size of 335×285 pixels. The lighting system is an Aristo MS1417 (Aristo Grid Lamp Products Inc. Washington D.C) model which consists of two parts- lamp housing and power pack. The lamp house (Figure 4.3) is 43.18×35.56 x 7.62 centimeters (l x b x h) respectively and is equipped with a cold cathode grid lamp. As suggested in the literature, diffused light chamber would enhance edge detection and body reflection. Therefore, a diffused light chamber was designed and fabricated in the departmental workshop.

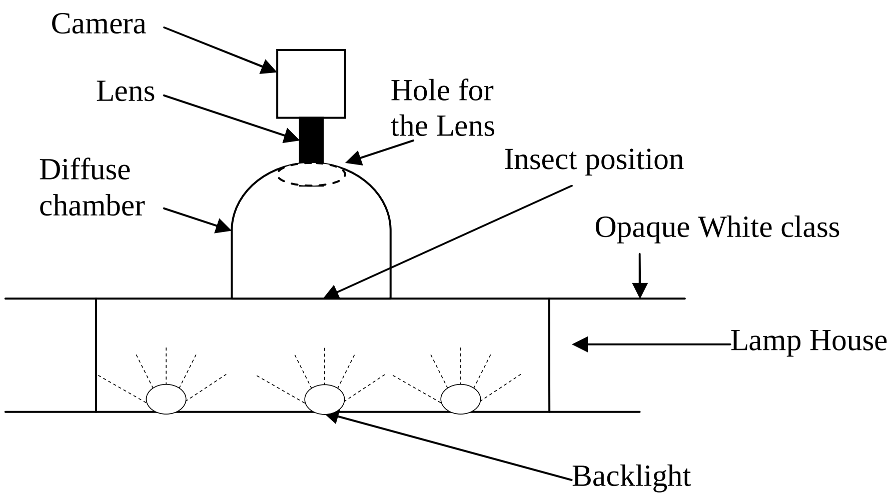



**Figure 4.3 Imaging System**

This tool helped in reducing the specular reflection from external light sources. It is 45.72 centimeters in length, 23.8 centimeters in width, and 12.7 centimeters in height. The chamber has an opening of 3 inch radius to allow the lens to go through the chamber. An opaque white-class cover (0.3175 centimeter thick) is used on top of the lighting box. A Dell Optiplex GX745, Pentium® D, 3.4 GHz CPU was used in this project. MATLAB® (R2006a) image processing software was utilized to conduct these experiments.

# CHAPTER V

# RESULTS AND DISCUSSION

## Introduction

The ultimate goal of this study is to develop a simple yet robust algorithm that can be used to identify pecan weevils among other insects. This algorithm is anticipated to serve in a system that employs a wireless imaging technique to detect targeted insects in the field.

In this chapter, individual experiment results for each method are illustrated and discussed for reliability and performance speed. After discussing all methods, a comparison study between individual methods is presented. These discussions are followed by a comparison of the performance of all methods.

## 5.1 Normalized Cross- Correlation Based Template Matching

Correlation-based method is a standard method in object recognition and has been utilized in many applications for its simplicity and reliability. This technique finds the “best” match and its location of an input shape (template) in a given image. The degree of similarity between a template and an image ranges from 1.0 for identical match and 0.0 for no match. A higher correlation value indicates a large similarity between compared images and vice versa.

In spite of the fact that in this technique images do not require any pre- or post-processing, it showed higher positive identification rate among most other methods. The issue of computational complexity of correlation-based method is not a major concern in this project because of the limited number of pecan weevils in the training set. Moreover, since this method will serve as part of a system that would serve in a wireless imaging network in the field, time of recognition is not as critical as identifying the weevils correctly.

Generally, processing time can be a very significant factor in selecting a recognition method over others if the designed system is anticipated to serve in a production line (industrial application). In this project, however, identification time is important but not critical because instantaneous user action is not necessary. Traditionally, growers would check pecan weevil traps every three to five days during the insect's emergence season (July-September). When the searching area (image) is large, it is expected that the searching process will take longer time. Moreover, images used in this project are resized to $114 \times 134$ to ensure reasonable processing speed.

Threshold at which insects are recognized as pecan weevil was experimentally determined using 204 pecan weevils. In this test, Normalized cross-correlation was performed between each image of this data set and the rest of the pecan weevils. This step showed the maximum possible correlation value of each individual weevil which helped in choosing an appropriate overall threshold.

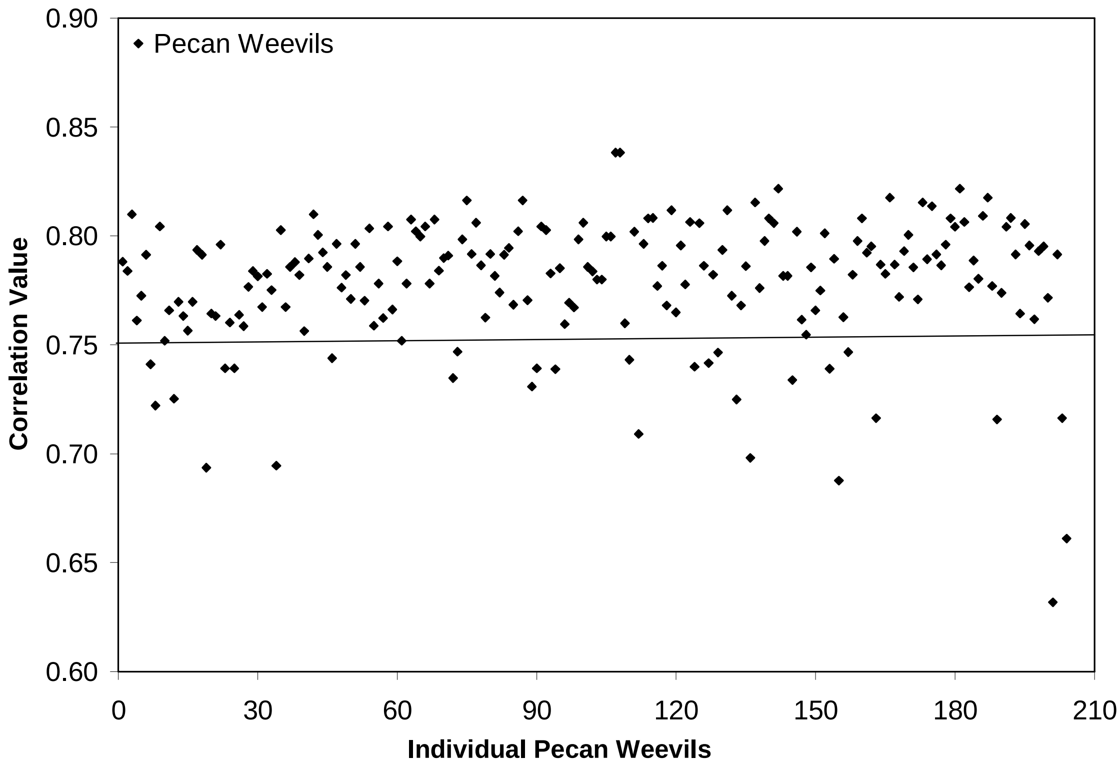



**Figure 5.1 Threshold of Recognizing Pecan Weevils Using Normalized Cross-Correlation**

Figure 5.1 illustrates that 86% of the pecan weevils have at least one match that is greater than or equal to a correlation value of 0.75. As a result, the threshold was set to be a correlation value of 0.75 or greater. In other words, an insect would be recognized as pecan weevil using Normalized cross-correlation method if its correlation value with an individual pecan weevil (training set) is greater than or equal to 0.75.

The method was then tested using two types of data sets; the first one consisted of 30 pecan weevils that were randomly selected from a group of 200 pecan weevils. The second group set is a group of 19 different insects (74 insects) that are naturally present in the pecan habitat (Table 4.2). The results of this experiment showed that the average correlation value of pecan weevils was 0.79 which is above the threshold of recognition (0.75). Using the testing set, twenty seven pecan weevils out of thirty were positively recognized. On the other hand, when correlating the second group of insects with the training set, seventy non-pecan weevil insects out of seventy four were correctly classified.

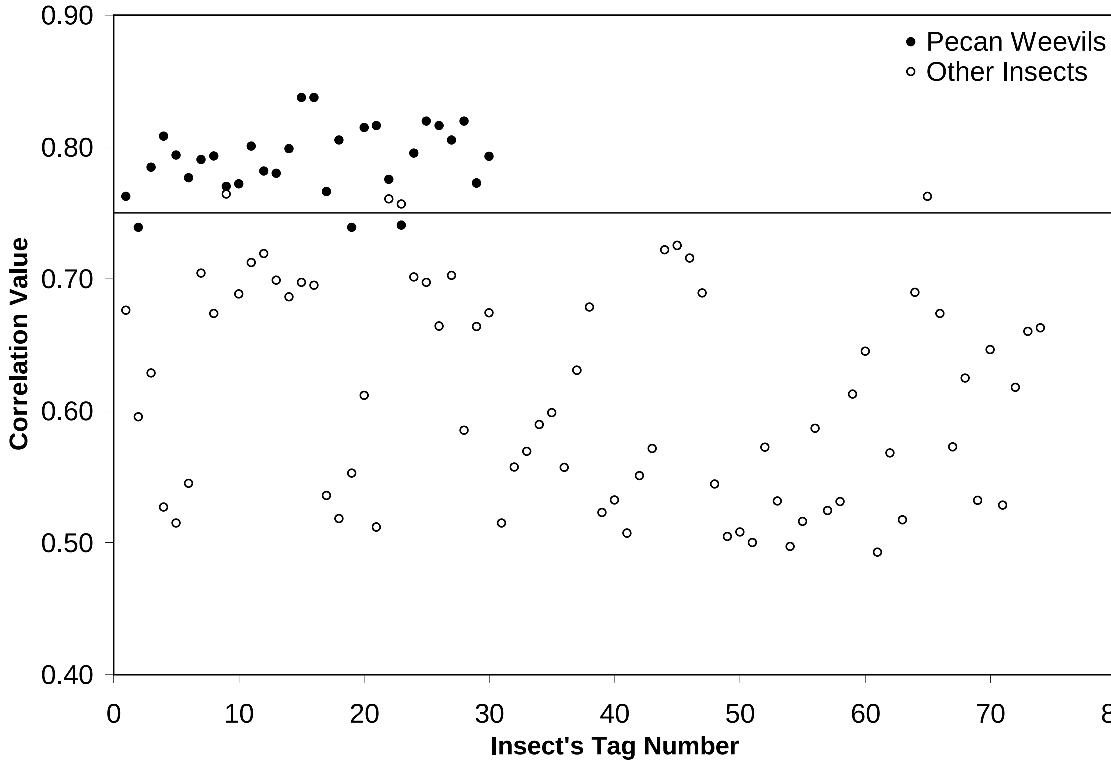



**Figure 5.2 Recognition Results for Pecan Weevils and Other Insects Using Normalized Cross-Correlation Method**

Figure 5.2 illustrates the results of using Normalized cross-correlation method to identify pecan weevils among other insects. In this figure, pecan weevils are represented by the solid circles while the other insects are represented with hollow circles. Clearly, it can be noticed that this method can distinguish pecan weevils from other insects. First, 90% of the pecan weevils are above the experimentally determined threshold of 0.75. Moreover, the three pecan weevils which fall below the threshold line were very close (0.74) to the passing criteria and not significantly away from being correctly distinguished.

Figure 5.2 shows that 95% of the non-pecan weevils were correctly classified. The 5% error represents the misclassification of four insects that are very close in shape and size to the pecan weevil. Figure 5.3 shows, on the first row, the four insects that were misclassified and on the second row, an example of some pecan weevils that are somewhat similar in shape and size to the above insects. This similarity is not surprising because all these four insects are in fact weevils that share similar body shape and size.

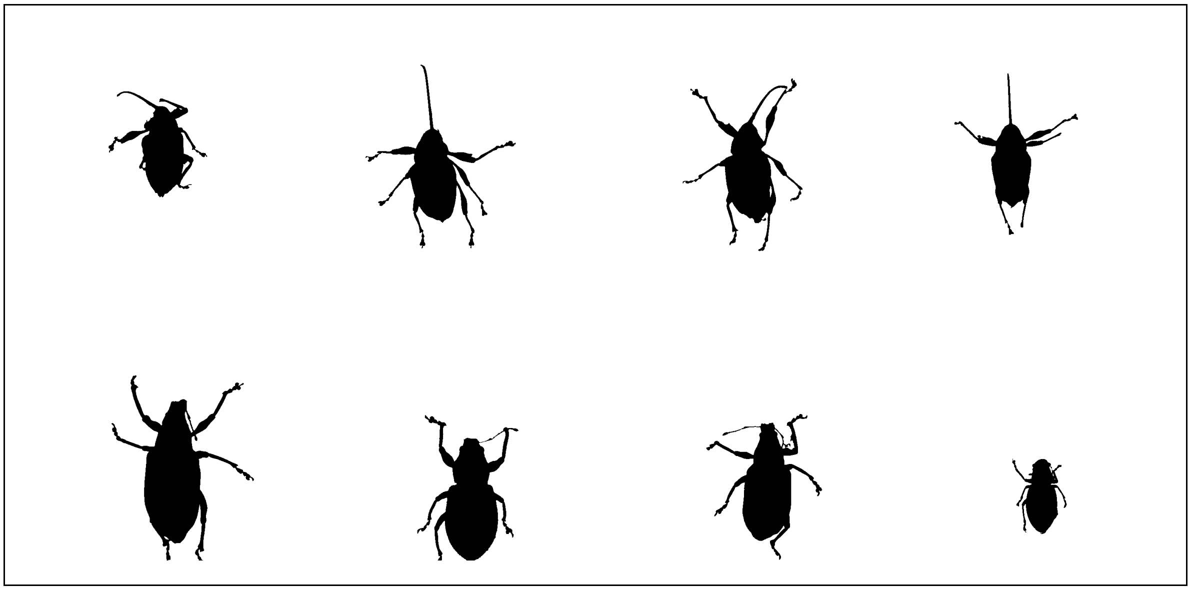


**Figure 5.3 Misclassified Insects (First Line) and Actual Pecan Weevils (Second row) Similar to the Misclassified Insects**

Table 5.1 summarizes the performance of the Normalized cross-correlation method in recognizing pecan weevils. It can be noticed that the correlation value of 60% of non-pecan weevil insects is less than or equal to 0.65, which is significantly less than the threshold of positive recognition (0.75). The results strongly suggest that the method would be a good tool in identifying pecan weevils, in general. Furthermore, it would facilitate correct recognition among insects which are quite similar to pecan weevils.

**Table 5.1 Performance Summary of Normalized Cross-Correlation Method**

| Average Correlation Value | Pecan Weevils | Other Insects |
|---|---|---|
| Ratio of insects with correlation value $\leq 0.50$ | 0% | 2.7 % |
| Ratio of insects with correlation value $\leq 0.55$ | 0% | 31% |
| Ratio of insects with correlation value $\leq 0.65$ | 0% | 60% |
| Ratio of insects with correlation value $\geq 0.75$ | 90% | 5.4% |
| Type I Error | 10% | - |
| Type II Error | 5% | - |

The training set of 205 pecan weevils was chosen to account for all possible variation among pecan weevils in terms of size and parts' orientation. The average time for recognizing unknown insects would take 22 sec., on average. In other words, it would take about 22 sec. to correlate an input image with all 205 pecan weevils' images to determine if the correlation value of any of them is greater than or equal to the threshold. This time is not what would be expected if the input image is in fact a pecan weevil. That is because it would be recognized at least once before correlating that image with the whole training set. The results had shown that pecan weevils which have been positively recognized have 5.9 matches on average. This result indicates that performance time using this method is within acceptable range especially when considering the large template used in this study.

Normalized cross-correlation method has been used by itself in many recognition applications but in this project, this method would not ensure 100% positive recognition. Furthermore, correlation technique is known to be sensitive to variation in rotation which may negatively affect the recognition rate and reduce the reliability of the system. The following method is implemented for two reasons: first, to use its properties in describing and recognizing pecan weevils and second, to help in correcting for any rotational invariance of the input images.

### 5.2 Region Properties Method

Insects are the earth's most varied organisms and almost three-quarters of all animals (Daly et al, 1998). In body size, most insects are 1 to 10 mm in length. In this project where many recognition methods are implemented, it is desired to adopt a method that would act as a "filter" which only allows insects to proceed through recognition steps if they are matching pecan weevils in certain criteria. Such procedure would significantly reduce the amount of data that the recognition algorithm has to handle and increase the system's efficiency. The method used for achieving this goal is known as the Region properties method, which depends on evaluating some geometrical measurement of a shape or object. These descriptors would then be utilized to represent that particular silhouette in any identification or recognition problem.

Region properties method provides several descriptors that can be used in the area of image recognition; for example, the *area*, *centroid*, *orientation*, *Euler number*, among others. In this study, three measurements (descriptors) were adopted to represent pecan weevils and these included area, major axis length, and minor axes length. The major axes length is defined as the length (in pixels) of the major axis of the ellipse that has the same second moments as the region. The minor axis can be related to the length (in pixels) of the minor axis of the ellipse that has the same second moments as the region length. The term area here refers to the number of pixels in a given region (Gonzalez et al, 2004).

Pecan weevils belong to the superfamily *Curculionidea* which is the largest group (65,000 species) of order *Coleoptera*. They, as members of this group, are most specialized in having rostrums which are used in preparing oviposition holes as well as in feeding. Therefore, a vector that has the three selected Region properties descriptors was formed for each insect as a unique representation. The significant point, here, is the major axis length of the insects. This is true because in case of pecan weevils, it is almost always the case that the major axis of their body is in fact the length of their rostrum plus the length of the body (head and abdomen). This general rule has some exceptions especially when an image of pecan weevil was taken while its rostrum was reoriented or broken.

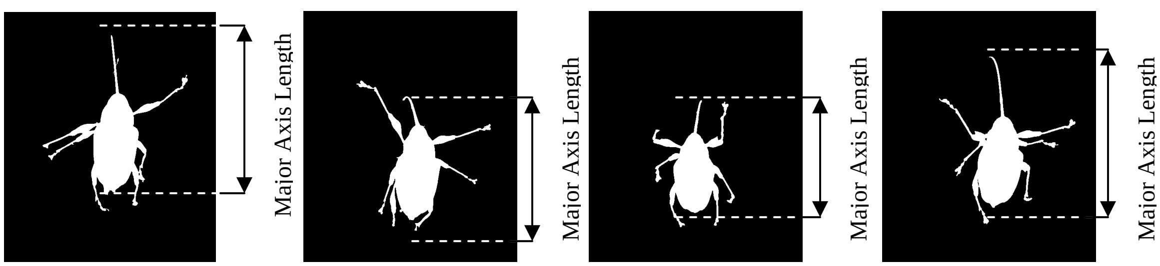



**Figure 5.4 Samples of Pecan Weevils with the Major Axis Length of their Bodies**

Figure 5.4 illustrates the measurement of pecan weevils' major axis length. This distinctive characteristic of pecan weevil (rostrum) was utilized by relating it to two more geometrical measurements formed in one vector. The idea behind combining these three components in one single vector is to make it more specific to the target insect. In other words, it was hypothesized that an insect with such major and minor axis and area is most likely to be a pecan weevil. This is so because other insects could have longer axis but smaller area or vice versa. Before conducting the test, threshold was first experimentally determined.

Figure 5.5 presents the results of applying the Region properties method to the training data set of 205 pecan weevils. The results showed that 80% of the pecan weevils have at least one positive match with minimum Euclidean distance less than 1.0. The average of number of positive matches for this group (threshold $\leq 1.0$) was three matches per insect. Although 90% of the pecan weevils were positively identified below a threshold of 2.0, it was preferred to go for a more biased threshold of 1.0.

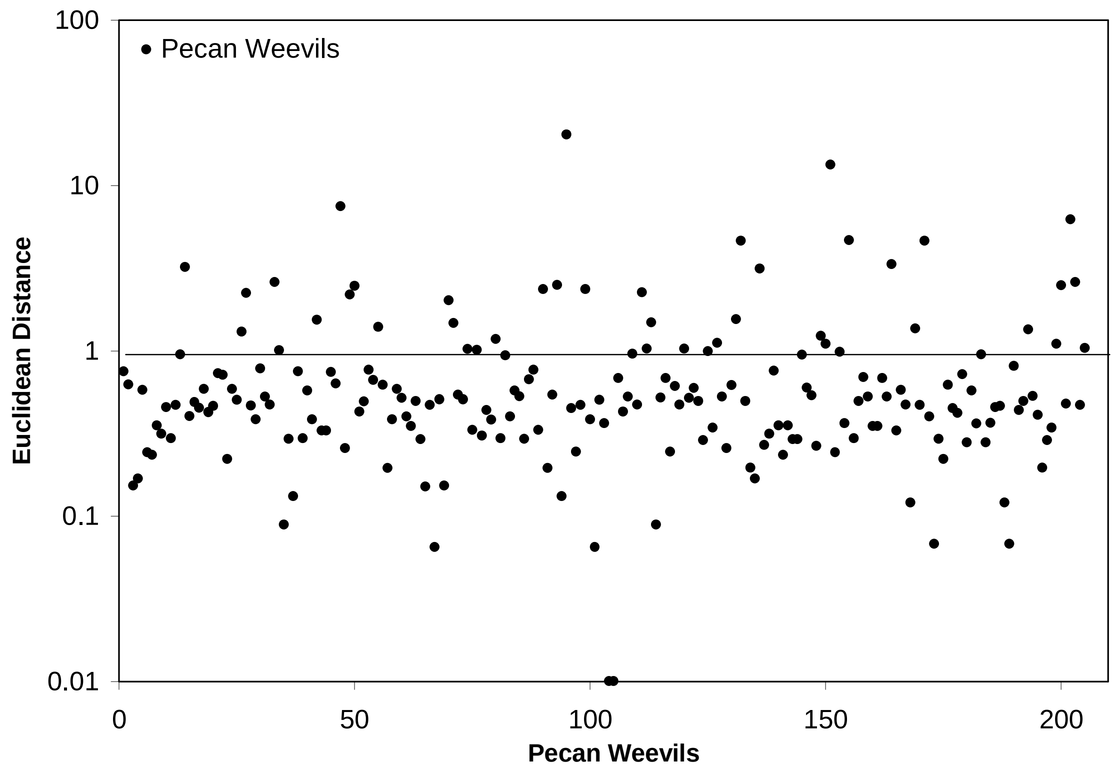



**Figure 5.5 Threshold of Recognizing Pecan Weevils Using Region Properties Method**

With this criterion of recognition, two experiments were conducted. The first one was done on a randomly selected group of thirty pecan weevils. This testing set is the same sample that was used to test the performance of all other recognition methods. In the other experiment, the Region properties method was applied to a data set consisting of 74 insects (non-pecan weevils). Using Euclidean distance for measuring similarity, the results of first experiment showed that 27 pecan weevils out of 30 were positively identified. This 90% successful recognition rate was achieved with threshold of less than or equal to 1.0. On the other hand, all non-pecan weevil insects group were correctly classified except five samples. In other words, more than 93% of the non-pecan weevil insects were accurately identified.

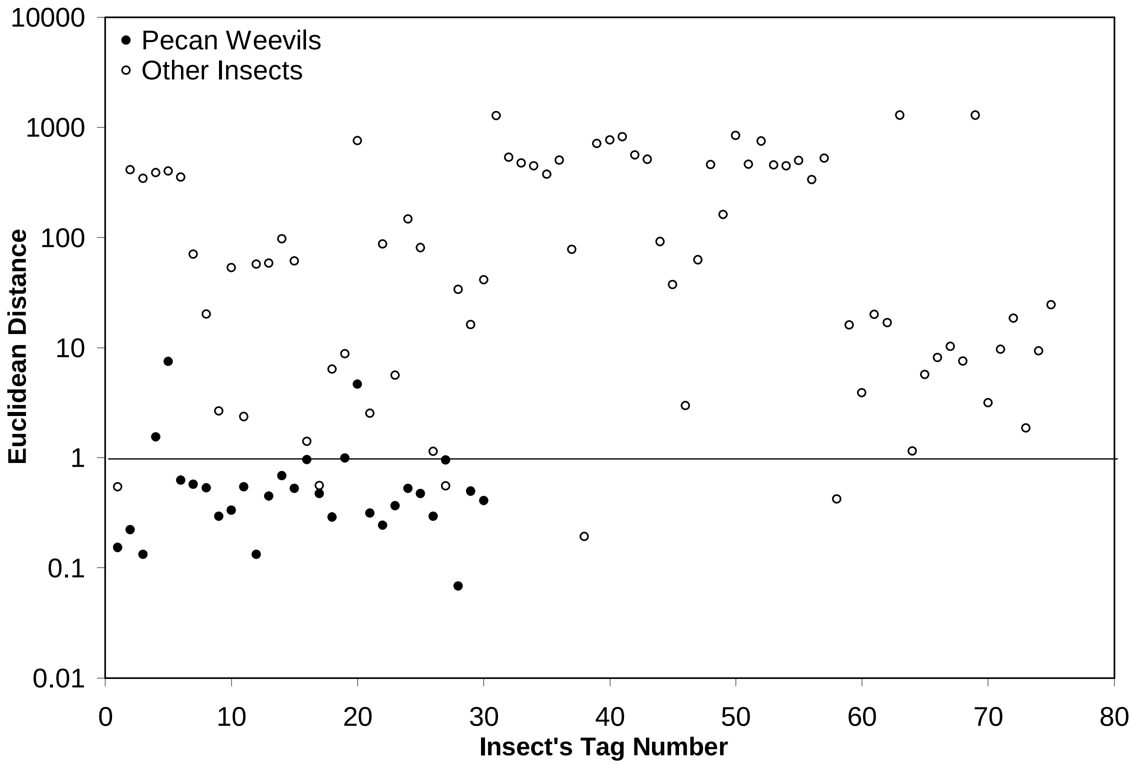



**Figure 5.6 Recognition Results for Pecan Weevils and Other Insects Using Region Properties Method**

Figure 5.6 shows the significance of the Region properties method wherein non-pecan weevil insects are clearly distinguished from pecan weevils. The results indicated that the Euclidean distance value of 40% of the non-pecan weevil insects was at least 100.0 whereas the recognition criteria was an Euclidean distance value of less than or equal to 1.0. It is worth mentioning here that about 70% of the non-pecan weevil insects have an Euclidean distance value of greater than or equal to 10.0. These encouraging results suggest that Region properties method would be a useful tool in the pecan weevils' recognition system.

Furthermore, Region properties method can be used to correct for any rotational invariants in pecan weevils' images. One of the parameters that Matlab function *regionprops* provides is the orientation, which is the angle (in degrees) between the x-axis and the major axis of the ellipse that has the same second moments as the region (Gonzalez et al, 2004). As illustrated in Figure 5.4, the major axis for pecan weevil is the length of its abdomen and rostrum. As a result, the orientation of pecan weevil's image can be corrected for rotation invariants prior to a rotational invariant method. This step would increase the reliability and efficiency of such a system.

### 5.3 String Matching

String matching is classified as region-based descriptors in which the two-dimensional boundary of a shape is represented by one-dimensional string. This is done by extracting the connected line segments from the shape and coding the interior angles of the polygons. The resulting angle array for that shape would then be quantized into $45^{\circ}$ increments which produced strings whose elements were numbers between 1 and 8 with 1 increment (Gonzalez et al, 2004).

Figure 5.7 shows two images of insects and their polygon representation. The string of angles is then used in the recognition process as a unique representative for each individual insect. For measuring the degree of similarity between pecan weevils and any given insect, a simple measure between strings was implemented in this study which was represented in Equation 4.26. The degree vary from D equal to zero when the input image of ($a$) and any image of ($b$) pecan weevils do not match, and infinite when the two images are identically matched**.**

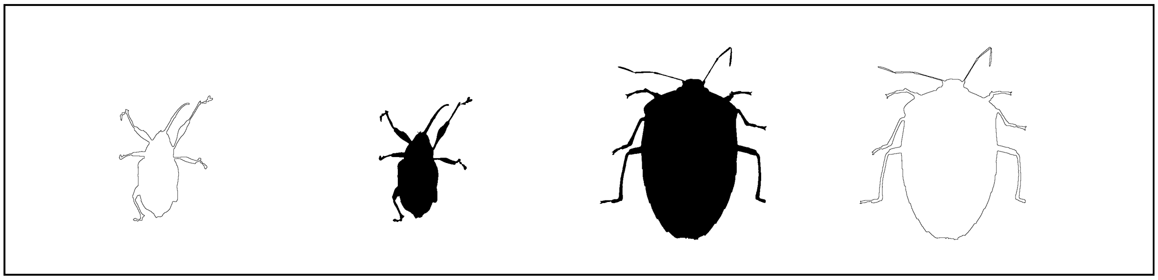


**Figure 5.7 Original Images of Pecan Weevils with their Polygons**

In this method, a tested image is recognized as pecan weevil if the D value is greater than or equal to 1.0. This threshold was determined based on the maximum similarity value of the 205 pecan weevils when compared to each other. Figure 5.8 presents the results of measuring the degree of similarity among the training data set using String matching method.  It can be noticed that the maximum similarity values of more than 74% of the pecan weevils' population were greater than or equal to 1.0. The result also showed that each weevil which has a similarity value above the threshold has at least 3.75 weevil matches, on average.

The testing set of insects which includes 75 non-pecan weevil insects and 30 pecan weevils were used in this experiment as it has been used throughout this study. Figure 5.9 presents the results of using String matching method to identify pecan weevils.

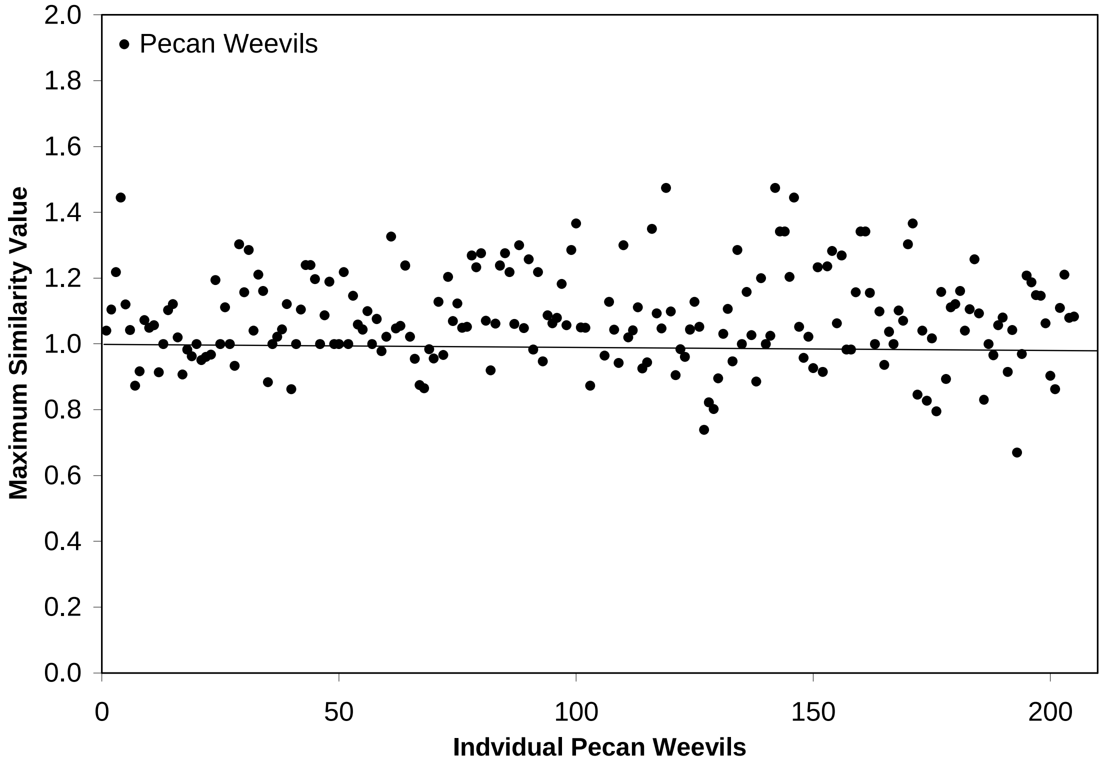



**Figure 5.8 Threshold for Recognizing Pecan Weevils Using String Matching Method**

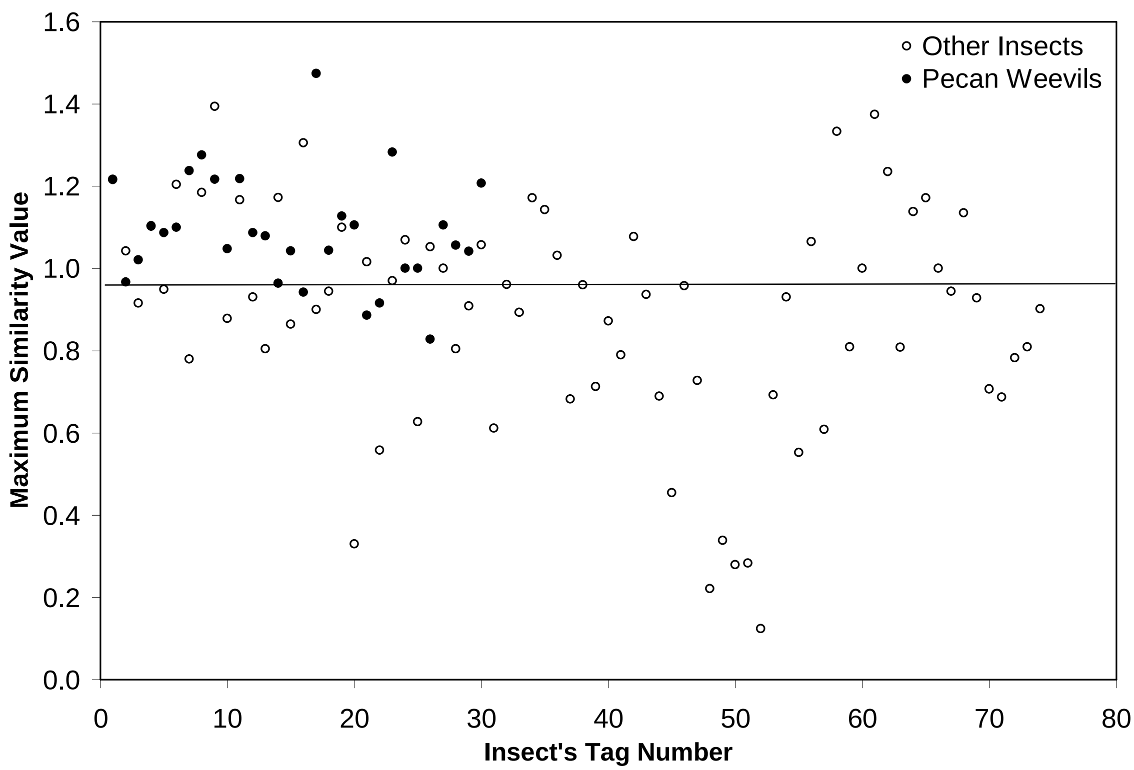



**Figure 5.9 Recognition Results for Pecan Weevils and Other Insects Using String Matching Method**

It can be noticed that 80% of the pecan weevils were correctly classified whereas this ratio was 62% for the non-pecan weevil insects.

This result is based on a selected threshold and it can vary as the threshold moves up or down the scale. Table 5.2 summarizes the possibilities of correct identification ratio based on different thresholds. As evident from Table 5.2, changing the threshold will increase the correct identification of one group and reduce it for the other one.

**Table 5.2 Recognition Ratio at Different Threshold Values**

| **Group** | **Recognition Rates at Different Threshold Values** | | | |
|---|---|---|---|---|
| | 0.80 | 0.90 | 0.95 | 1.00 |
| Pecan Weevils | 100% | 93% | 87% | 80% |
| Other Insects | 30% | 42% | 57% | 62% |

Among the misclassified insects, there were 13 weevils which are somewhat similar to pecan weevils. This is about 46% of the total number of the misclassified insects. The limited number of descriptor (angles represented by digits from 1 to 8) which represent the boundary of a given shape may partly explain the relatively high type II error when using this method. This appears to be true especially when coding the boundary of objects that do not follow some linear changes such as an insect's body. Nevertheless, String matching method can significantly contribute to a pecan weevil recognition system.

### 5.4 Geometric Moment Invariants

The recognition system was tested by using two sets of data- the pecan weevil and other insects testing groups. In this experiment, the seven moments were calculated first for the whole training set of 205 pecan weevils for the purpose of defining a recognition threshold. Then, moments were calculated for the testing data. During this processing, it was noticed that the value of the seventh moment for some images was zero. This is because these Hu Moment invariants use central moments to achieve invariance to changes in shift. As a result, the odd orders of these central moments give value of zero for images with symmetry which may affect the ability of the recognition system in identifying and classifying such images.

It was found that 30% of the images in the training set and 27% in the testing set used in this study were associated with this symmetry. The symmetry occurs when an image has symmetry in the $x$ and or $y$ directions and at centroid. This problem is caused by the use of the image centroid in the calculation of the central moments (Palaniappan et al., 1999). This reduces the number of descriptors that can be used in classifying images where it is always desirable to have more features for successful classification especially when working with noisy images.

In spite of the symmetry problem in some images, the first experiment was conducted using a training set of 147 pecan weevils. This was the number of pecan weevils whose images did not suffer from the symmetry problem and thus all their seven moments could be calculated. This data set is about 30% less than the standard number of pecan weevils used for training the system. This reduction of training images is expected to have some impact on the recognition rate of this method. Figure 5.10 presents the results of measuring the similarity of pecan weevils to each other. The results showed that 80% of the pecan weevils have a similarity degree (Euclidean distance) less than or equal to 0.21. This threshold was chosen to be the recognition criterion for the following tests.

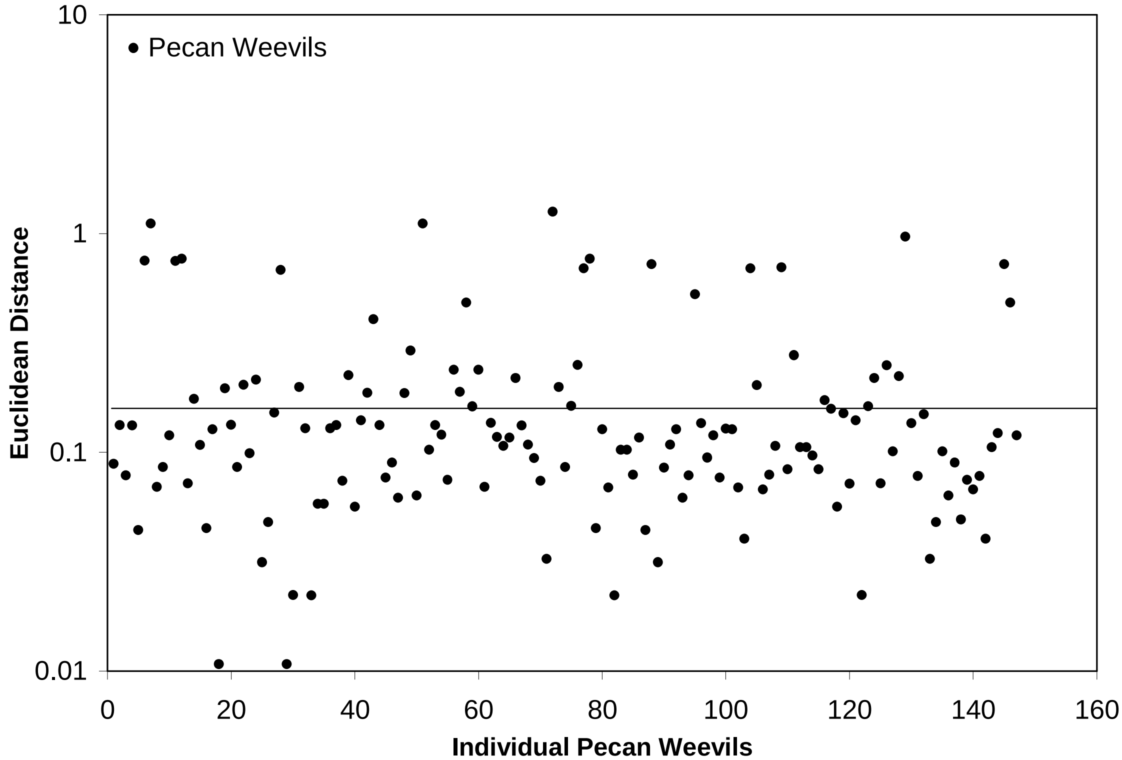



**Figure 5.10 Threshold for Recognizing Pecan Weevils Using Geometrical Moment Invariants Method**

The seven geometrical moments were used first for identifying the 30 randomly selected samples of pecan weevils. Results of this test showed that 26 pecan weevils (87%) were correctly classified. The second experiment was conducted using 54 non-pecan weevil insects. These insects were the only insects that had a non-zero seventh moment. In other words, 27 images of this group have the problem of symmetry in which the seventh moment of each image in this group was zero and hence they could not be used in this test. The results of this experiment are presented in Figure 5.11 where it can be noticed that 63% of the non-pecan weevil insects were correctly classified. Although this result is not very promising, one can observe that the correctly identified images were significantly distinguished from pecan weevils. Here, about 39% of non-pecan weevils have similarity value of greater than 1.0 whereas the threshold value was 0.21.

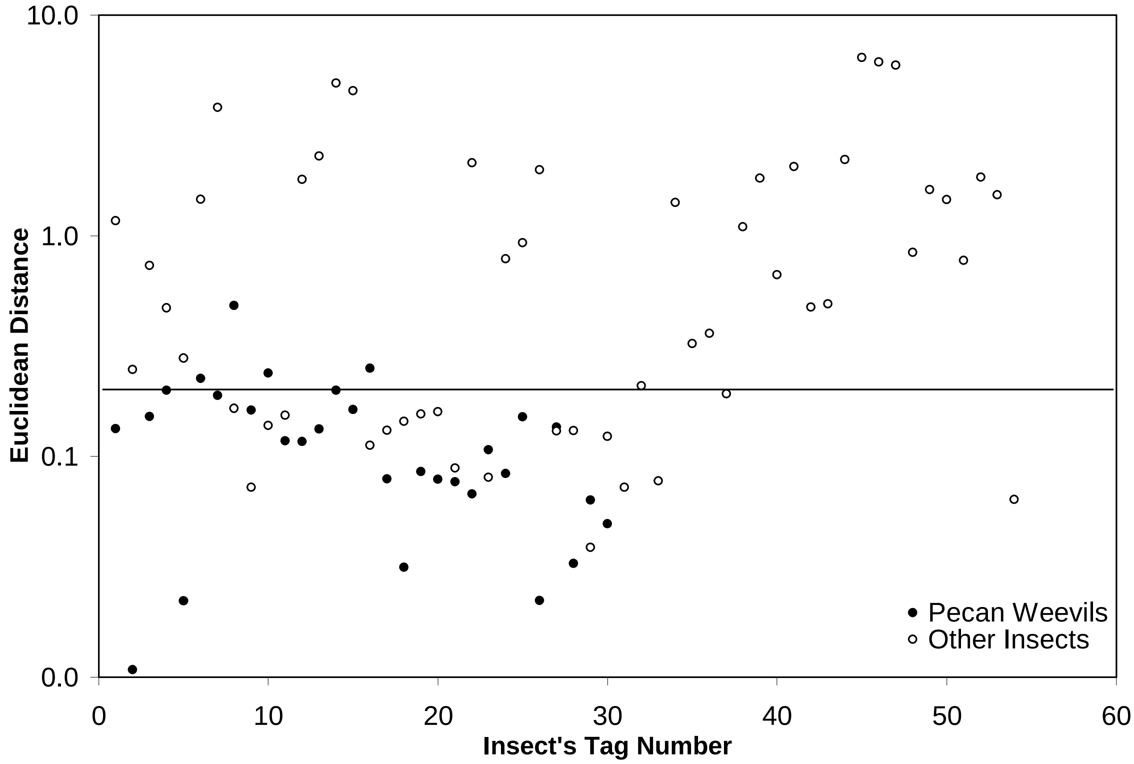



**Figure 5.11 Recognition Results for Pecan Weevils and Other Insects Using the Seven Geometrical Moment Invariants Method**

The issue of symmetry in about 30% of the training and testing data sets constrains the number of moments to be six instead of seven. One obvious result of this reduction is an even lower recognition rate as less number of descriptors will be used to represent images. But on the other hand, no images from the training and testing data sets will be discarded because of the symmetry problem. Figure 5.12 presented the threshold of recognition using the first six moments. Using the 205 images of pecan weevils (training set), results showed that 80% of them were at a Euclidean distance of less than or equal to 0.06 from each other. This threshold was adopted in testing the method on the two data sets.

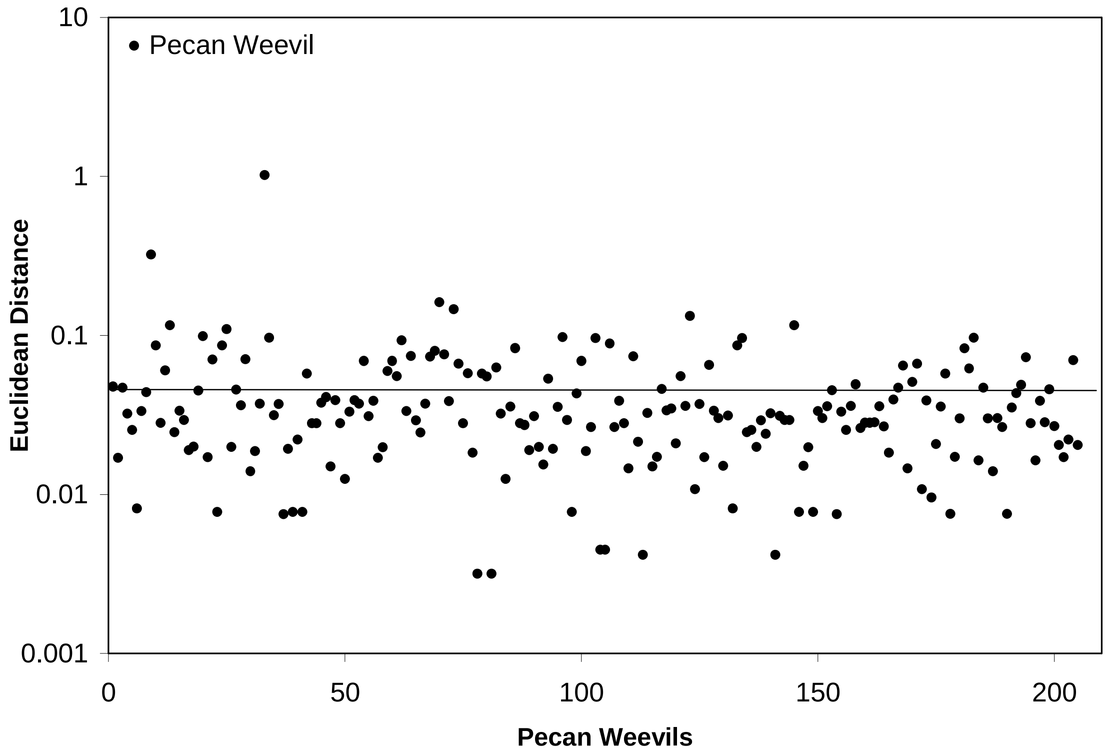



**Figure 5.12 Threshold for Recognizing Pecan Weevils Using Geometrical Moment Invariants Method (Six moments)**

As illustrated in Figure 5.13, results of applying the first six moments of the geometrical moment invariants method showed that 87% of the pecan weevils were correctly classified. This successful recognition rate is the same when the first seven moments were used for recognition. Therefore, one can conclude that there was no difference in recognition rate of pecan weevil images when using either the first seven or six moments of the geometrical moment invariants. For the non-pecan weevil testing set, results indicated that 47% of this group was correctly classified. As the number of samples in this test included all images (74 insects), this relatively low recognition rate cannot be compared with the previous experiment where only 54 insects could be included. However, when comparing the same 54 images which were used in both experiments, the method that uses 6 moments failed to recognize six images (11%) of the testing set which were correctly recognized using the seven moments approach.

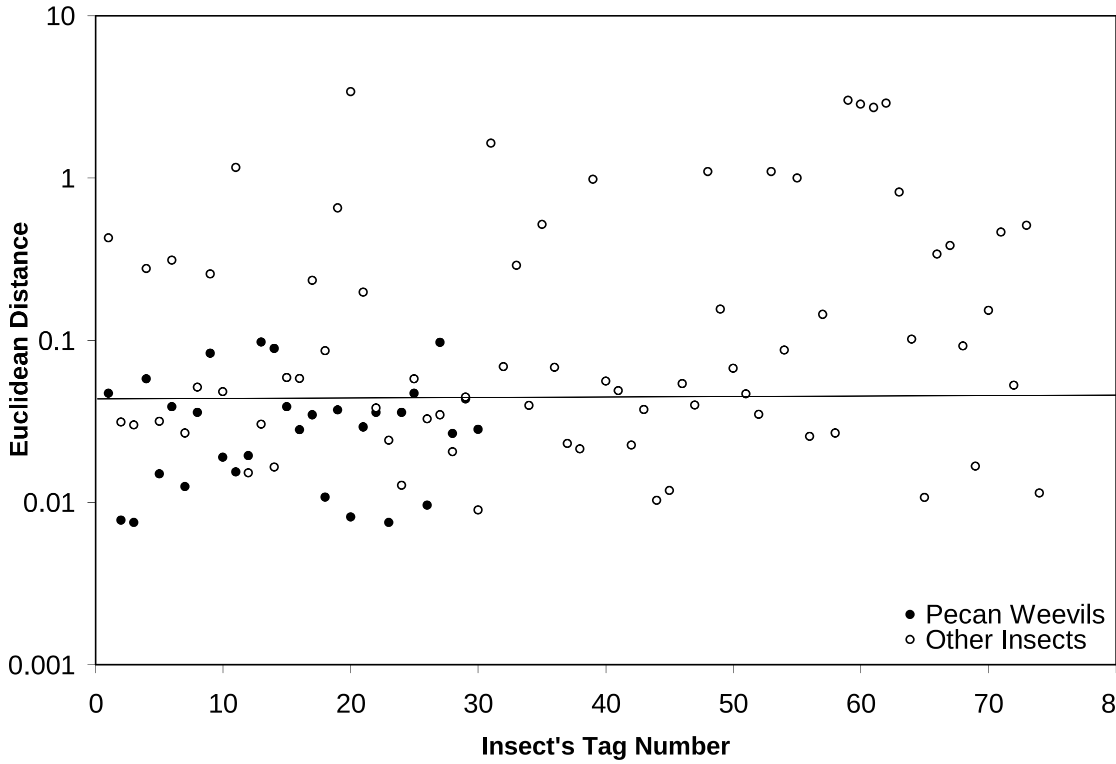



**Figure 5.13 Recognition Results for Pecan Weevils and Other Insects Using the Six Geometrical Moment Invariants Method**

These results were not surprisingly different than what would be expected especially after reducing the number of moments (eliminating the seventh moment) used in recognition. At any rate, if geometrical moment invariants method is to be used in such a recognition system, 53% of non-pecan weevil insects are expected to be misclassified. This higher error rate motivated the exploration of an alternative moments-based method. Zernike moments has the potential of performing better than the geometrical moment invariants, since it has been shown in the literature (for example, Teh, 1988) to outperform other methods including the geometrical moment invariants.

### 5.5 Zernike Moments

Although Geometrical moment is the most common and simplest moment, results showed that it was not performing efficiently in recognizing pecan weevils in this study. Zernike moment method is known to be more complicated than the geometrical moment. Also, at higher moment orders, Zernike Moment takes longer time in calculating its features; however, it generally outperforms the geometrical moment method. Zernike moment introduces a set of complex polynomials which form a complete orthogonal set over the interior of the unit circle.

The orthogonality property enables one to separate out the individual contribution of each order moment to the reconstruction process (Khotanzad and Hong, 1990). Although higher order moments carry finer details of an image, they are also more susceptible to noise. Therefore, experiments were carried out with different orders of Zernike moments to determine the optimal order for our problem. This yielded five different orders of moments. These orders were 3, 5, 10, 20, and 30 and these were applied to the testing data sets for evaluating their performance. Threshold of recognition was performed for each individual order using the training data of 205 pecan weevils.

Figure 5.14 presents the threshold result for Zernike moment of order 30 in which 80% of the training set were similar to each other at Euclidean distance value of less than or equal to 13.8. This criterion was used in the first experiment where 30 pecan weevils and 74 other insects (testing sets) were tested for recognition. The results of this test showed that 90% of the pecan weevil group was correctly classified. The successful recognition in the non-pecan weevil testing set was 73%. When comparing these results with that of geometrical moment invariants, a significant improvement of recognition rate

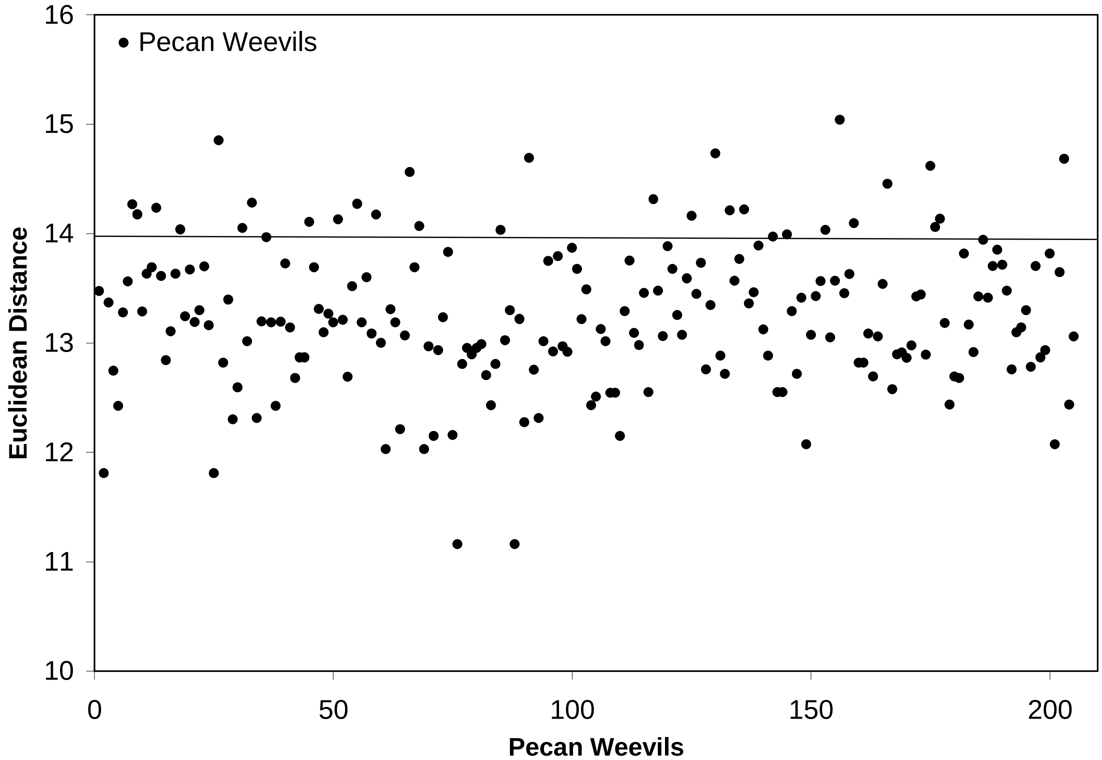



**Figure 5.14 Threshold for Recognizing Pecan Weevils Using Zernike Moment of Order 30**

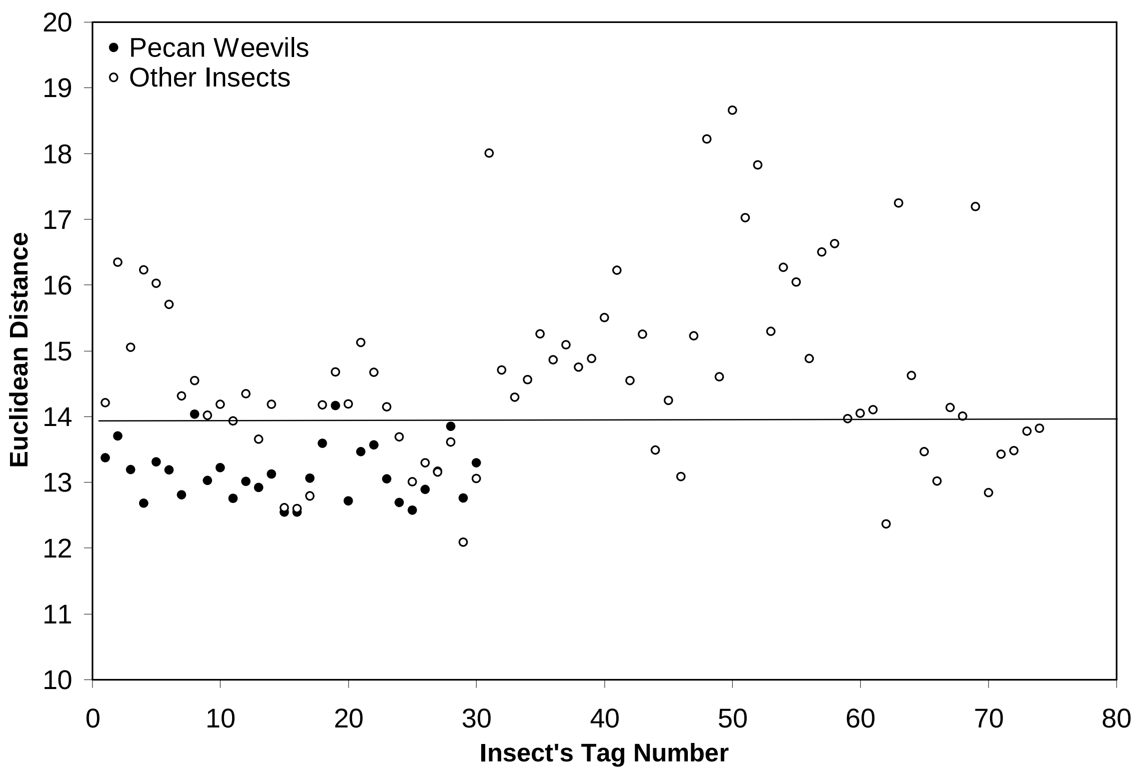



**Figure 5.15 Recognition Results for Pecan Weevils and Other Insects Using the Zernike Moments Method of Order 30**

can be noticed. Using the geometrical moment's method, only 47% of the non-pecan weevil group could be correctly classified whereas Zernike moments of order 30 could recognize successfully 73% of the same set. The performance of Zernike moments of order 30 is demonstrated in Figure 5.15.

Using the same methodology, four more experiments were conducted using order of 20, 10, 5, and 3 Zernike moments. Their thresholds were found experimentally to be Euclidean distance values of 8.45, 2.42, 1, and 0.8 respectively. These four values satisfied the condition that 80% of the pecan weevils population were recognized at values less than or equal to these thresholds. Figures 5.16 and 5.17 present the results of applying Zernike moments of order 20 and 10 respectively.

The performance of Zernike moments of order 30 is not significantly different than the one of order 20. In fact, at both orders, Zernike moments could achieve the same result in terms of identifying pecan weevils. As far as classifying the insects of non-pecan weevil's group, Zernike moments of order 20 correctly classified 66% of them compared to 73% achieved when Zernike moments of order 30 was used.

Zernike moment is known to be computationally intensive especially at high orders. Therefore, lower orders of moments were evaluated for robustness and processing time. Starting with an order of 5 Zernike moments, the results were significantly improved in terms of recognition rate and processing speed.

Figure 5.18 shows that at threshold of 0.8, 83% of pecan weevils and 100% of other insects were positively classified. It can be noticed from the figure that the distribution pattern of the two testing groups is consistent for each group. That is, it can be seen that the other insects' group is satisfactorily separated from the pecan weevils'

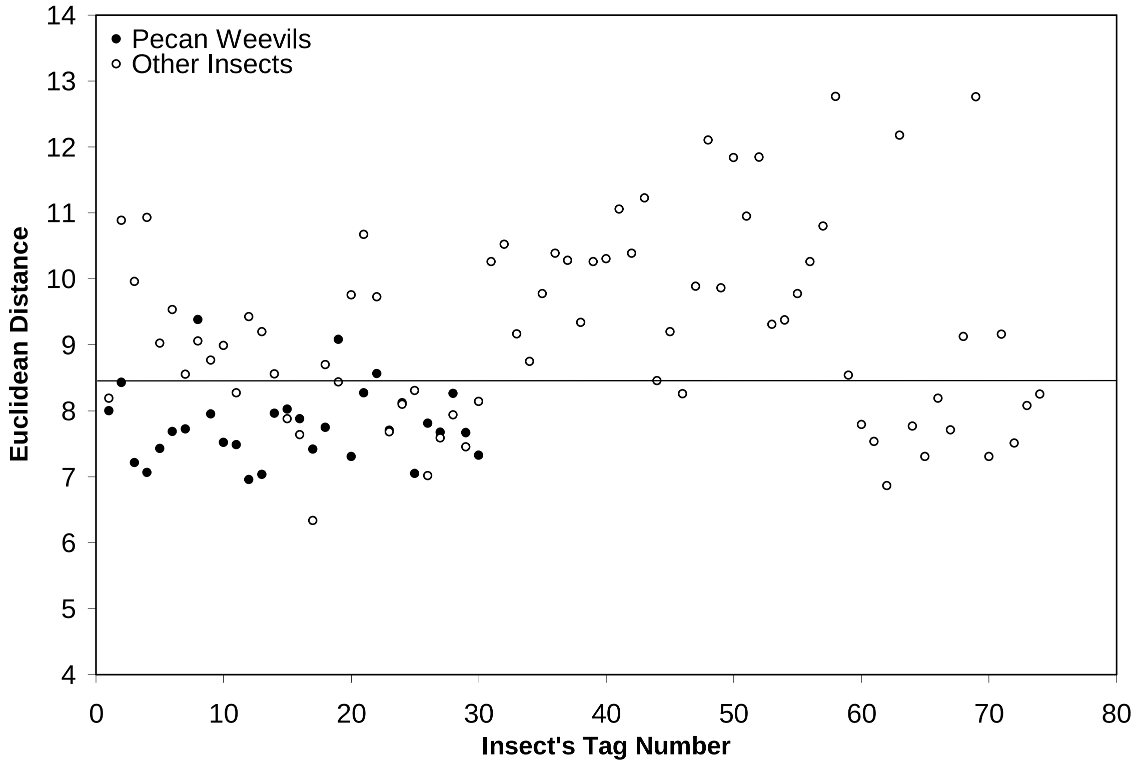



**Figure 5.16 Recognition Results for Pecan Weevils and Other Insects Using the Zernike Moments Method of Order 20**

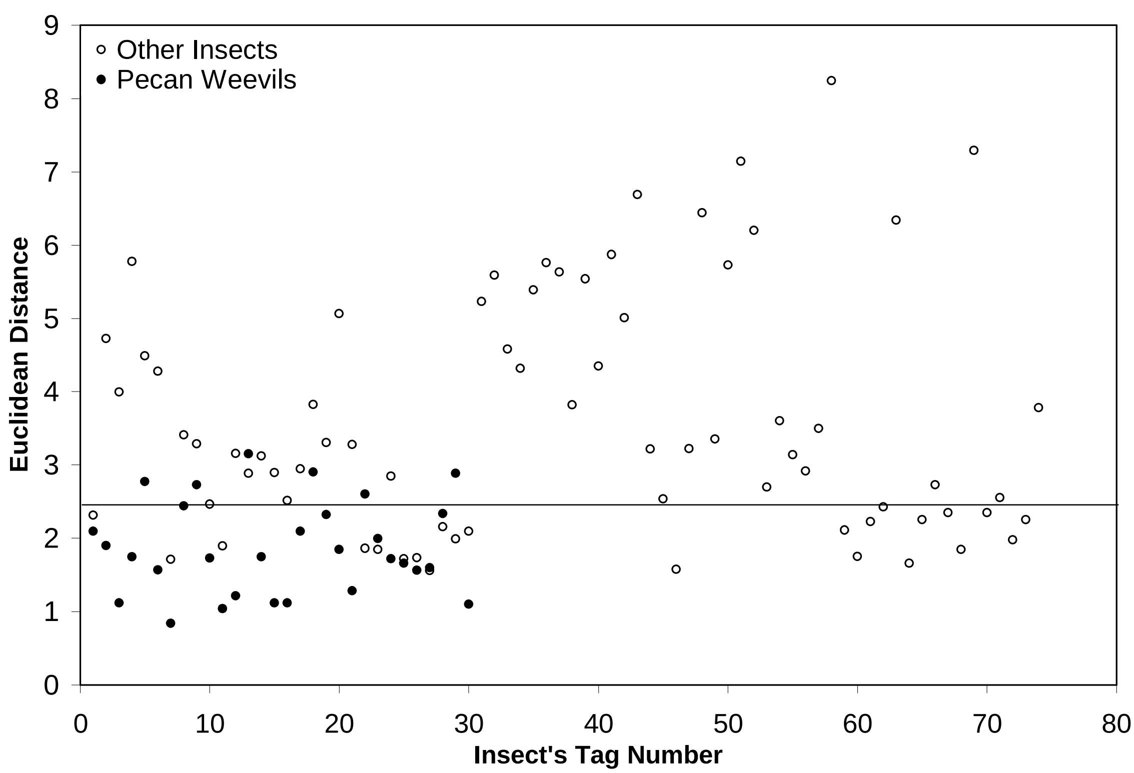



**Figure 5.17 Recognition Results for Pecan Weevils and Other Insects Using the Zernike Moments Method of Order 10**

group. Furthermore, the processing time using this order of Zernike moments was 0.24 sec. on average. This time is 77% less than that required by Zernike moments of order 10.

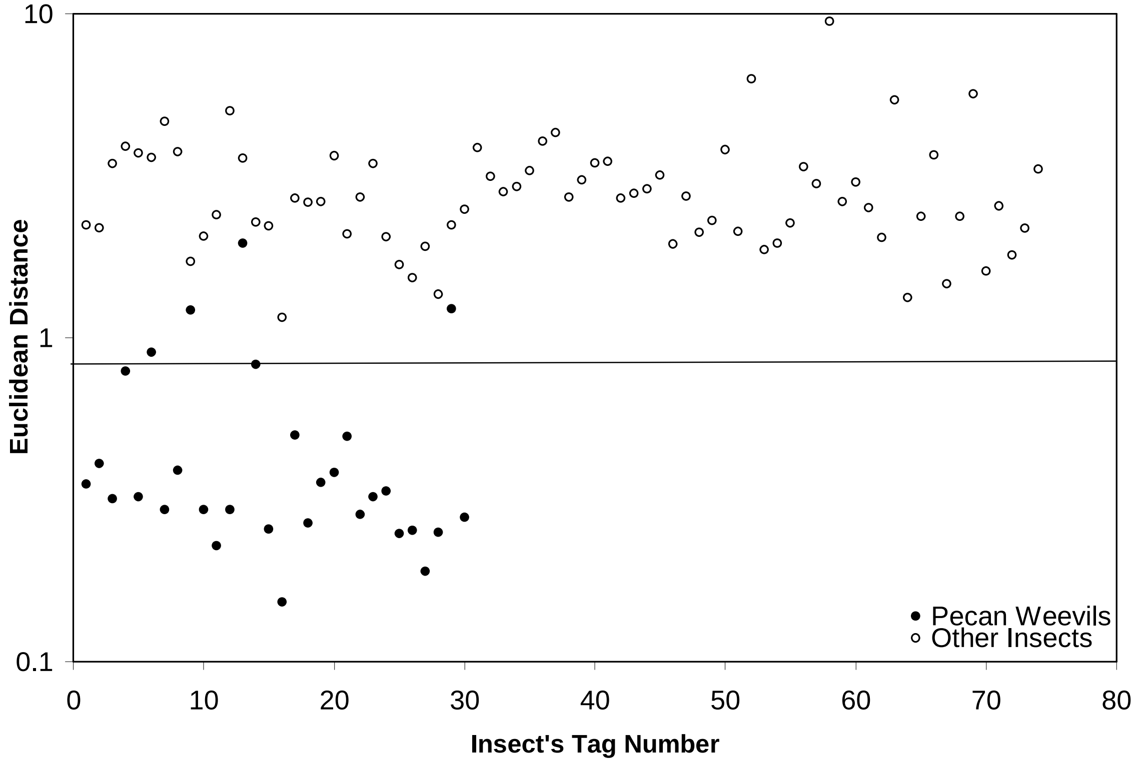



**Figure 5.18 Recognition Results for Pecan Weevils and Other Insects Using the Zernike Moments Method of Order 5**

The last experiment for this method was done using Zernike moments of order three. The results proved that using lower order of Zernike moments for this application would result in better recognition rate. The other advantage of using fewer moments is to reduce the processing time needed for processing the images by reducing the amount of data points. Illustrated in Figure 5.19, the output of this experiment showed that the two testing groups are clearly separated into two different sections. This powerful classification ability of Zernike moments at this order strongly suggests its adoption in the proposed recognition system. In particular, these results are considered to be the best in terms of correct classification rate and speed.

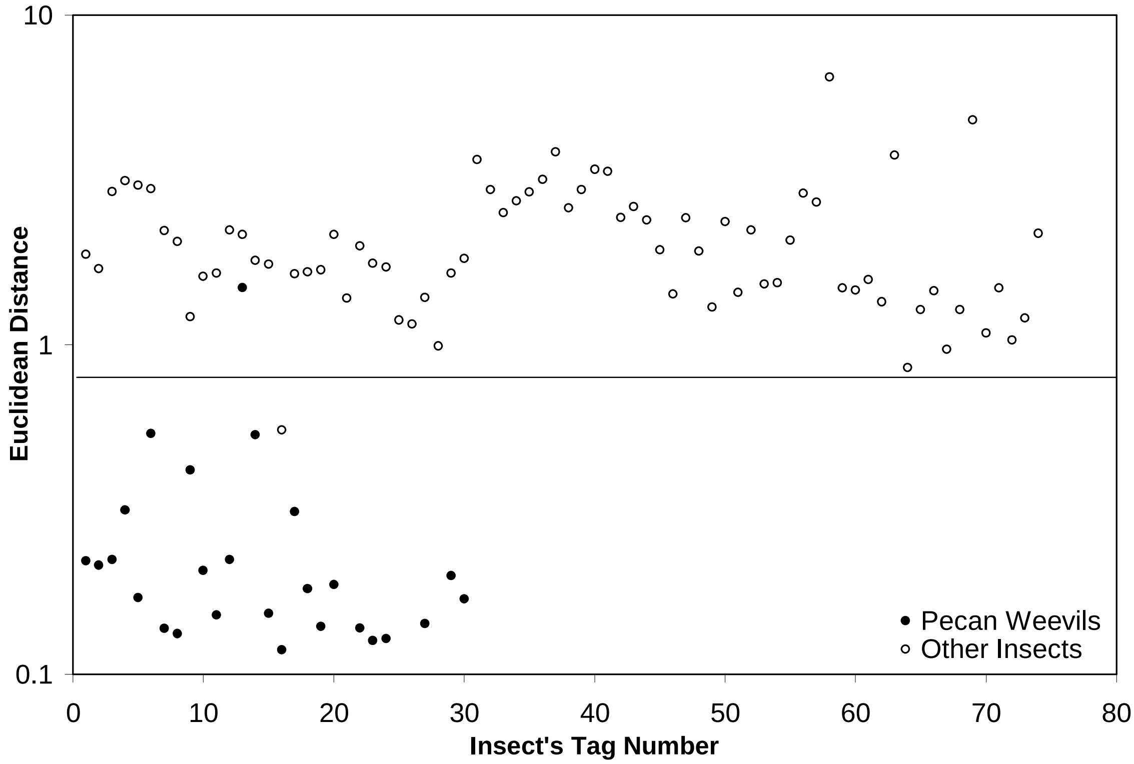



**Figure 5.19 Recognition Results for Pecan Weevils and Other Insects Using the Zernike Moments Method of Order 3**

Table 5.3 summarizes the results of the Zernike moments at the five different orders used in this section. Although the recognition rates for pecan weevils at orders 30 and 20 were 90%, the rates for classifying other insects were not satisfactory at these two orders. The classification rates at Zernike moments of order 10 were low for both pecan weevils (77%) and other insects (70%). At order 5, the positive classification rates were “perfect” (100%) for the pecan weevils’ group and were 83% for the other insects’ group. The optimum results were obtained when performing the classification at order three. There, the recognition rate was 97% for pecan weevils and 99% for other insects. Therefore, Zernike moments at order 3 was implemented in this study because of its excellent and very fast (0.09 sec.) performance in identifying pecan weevils among other insects.

**Table 5.3 Comparison of Recognition Rates Using Zernike Moments at Different Orders**

| Order Value | Threshold Value at 80% Recognition | Processing Time for one Image (sec) | Recognition Rate | |
|---|---|---|---|---|
| | | | Pecan Weevils | Other Insects |
| 30 | 13.8 | 13.28 | 90% | 73% |
| 20 | 8.5 | 4.65 | 90% | 65% |
| 10 | 2.42 | 1.03 | 77 % | 70% |
| 5 | 0.8 | 0.24 | 83 | 100% |
| 3 | 0.8 | 0.09 | 97% | 99% |

Figure 5.20 presents the performance of Zernike moments using the five experiments. The optimum point was found at order 3 of Zernike moments where the highest recognition rate for both pecan weevils and non-pecan weevils could be obtained at the shortest time (0.09 sec.).

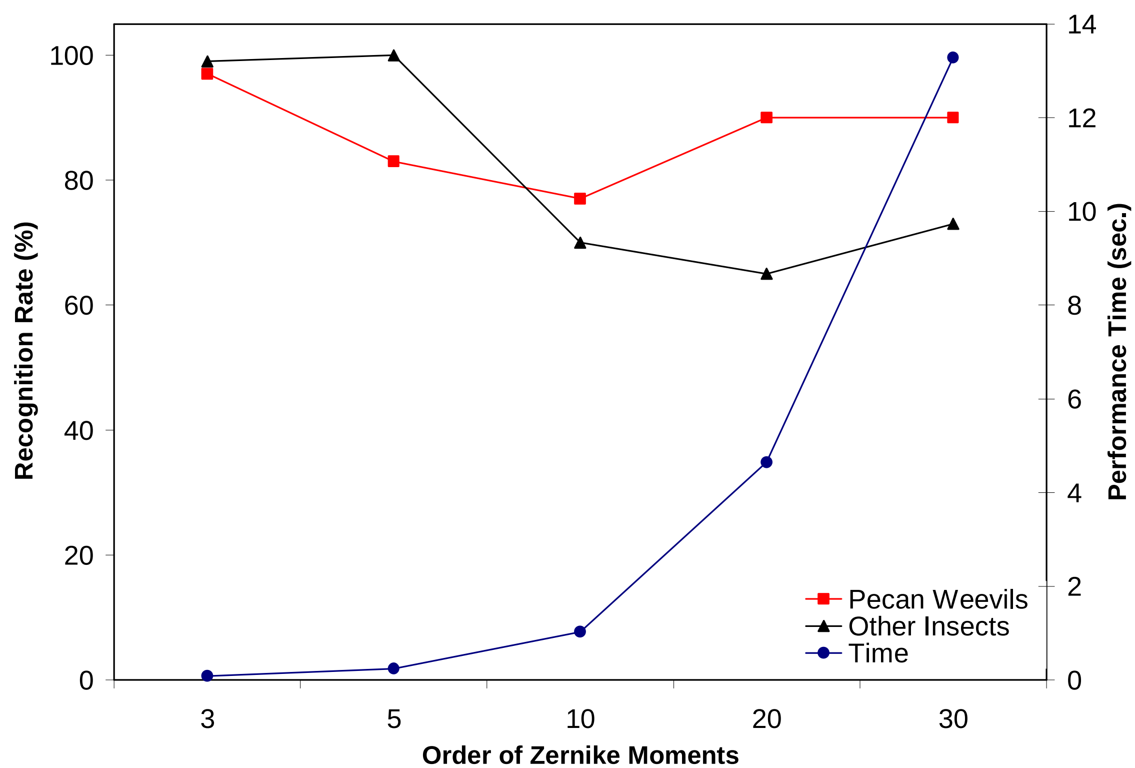



**Figure 5.20 Performance Analysis of Zernike Moments at Different order**

### 5.6 Fourier Descriptors

Fourier descriptors method has been implemented in many recognition and identification applications. With some simple transformation, Fourier descriptors can be modified to account for some variances like rotation, scale, and translation. Fourier Transformation produces the Fourier descriptors to be a unique representation in the frequency domain for a given image. The lower frequency descriptors store the general information of the shape and the higher frequency store smaller details (Sarfraz, 2006). Therefore, lower frequency components of the Fourier descriptors are sufficient for general shape description. The degree of similarity among some insects is very high, especially if they are from the same family. Therefore, higher frequency components of the Fourier descriptors were used in this project.

Figure 5.21 shows a binary image of a pecan weevil and several boundary reconstructed shapes using different Fourier descriptors. It can be noticed that more the descriptors used, the closer the shape would be to the original image. The general shape of pecan weevil was obtained by using the first 30 Fourier descriptors; however, more detailed shaped appeared at higher number of descriptors. However, the performance of recognition algorithm is expected to be slower as large number of Fourier descriptors is used. Therefore, in order to find the optimum number of Fourier descriptors, several experiments were conducted using different number of Fourier descriptors.

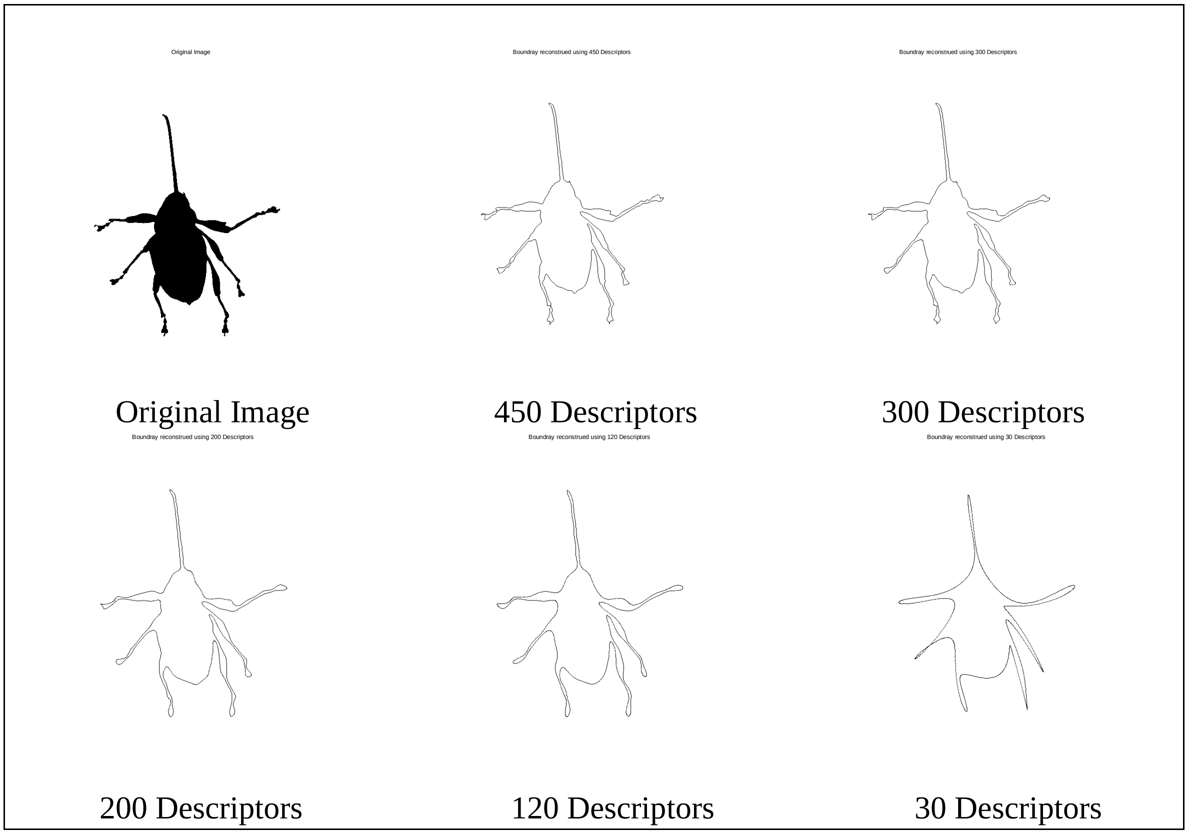



**Figure 5.21 Boundary Reconstructed of Pecan Weevil Insect Using 450, 300, 200, 120, 30 Fourier Descriptors (Out of a Possible 2602 Descriptors)**

The first experiment was performed using 30 descriptors to represent the shape of insects. Four more tests were conducted using 120, 200, 300, and 450 Fourier descriptors to evaluate the performance of each set of descriptors in terms of successful recognition rate. For each test, the threshold of recognition was determined experimentally by using that specific number of descriptors to recognize pecan weevils on the training. The threshold for each experiment was set to be at the Euclidean similarity value where 80% of the training set had a degree of similarity less than or equal to that value.

The result of the first experiment showed a very low recognition rate in classifying the non-pecan weevil's testing group. Although this test could correctly recognize 83% of the pecan weevils, it correctly classified 26% of the non-pecan weevil insects. It can be noticed in Figure 5.22 that the Fourier descriptor method failed to distinguish pecan weevils among other insects when using 30 descriptors. This result suggested that insect's shape need to be represented by more descriptors for better classification rate. This low performance was not unexpected especially with the limited number of descriptors representing the shapes of insects. In fact, 30 descriptors were at maximum representing less than 7% of the insects' shape.  This ratio would have been sufficient if used in representing "regular" shapes which is not the case in this study.

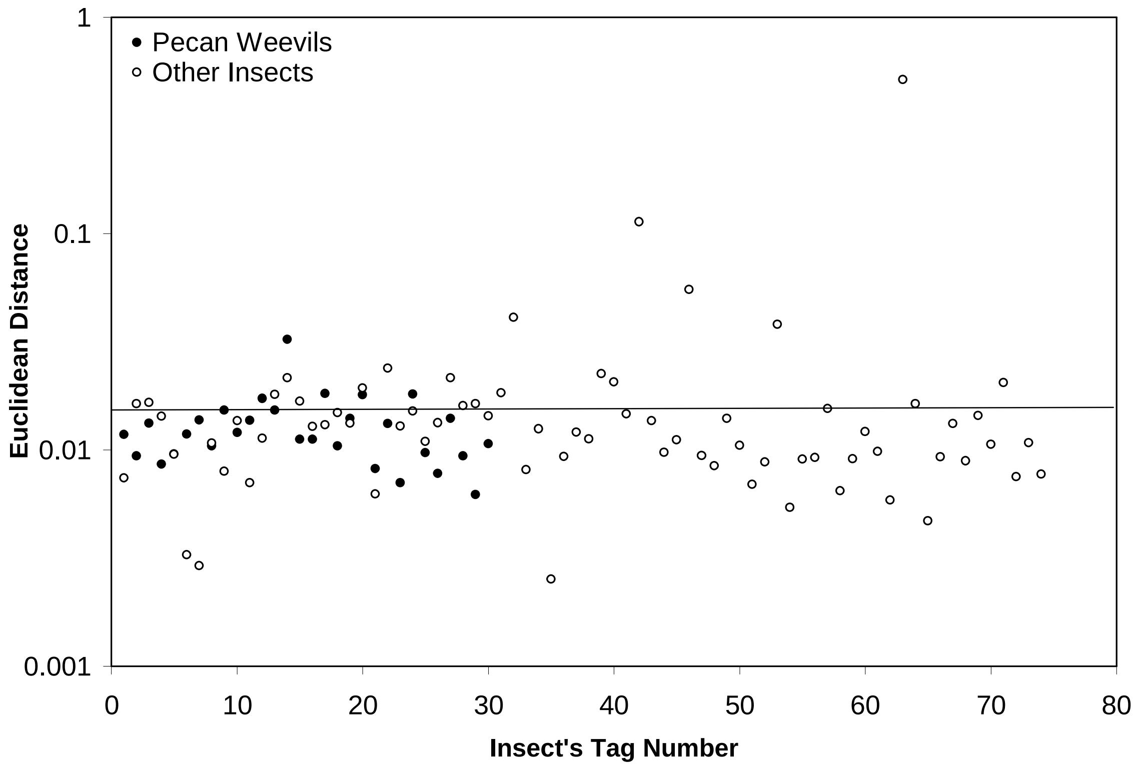



**Figure 5.22 Recognition Results for Pecan Weevils and other Insects Using the Fourier Descriptors Method (30 Descriptors)**

The second test was performed using 120 Fourier descriptors to represent the shapes of insects. Results showed slight improvement in classifying pecan weevils and other insects but they were not significant.  Results also showed that 87% of the pecan weevils and 28% of the non-pecan weevil insects were correctly classified. As the result of recognizing other insects is very low, one can conclude that the number of descriptors used was not sufficient to differentiate the shape of these insects. Figure 5.23 demonstrates the results of this experiment.

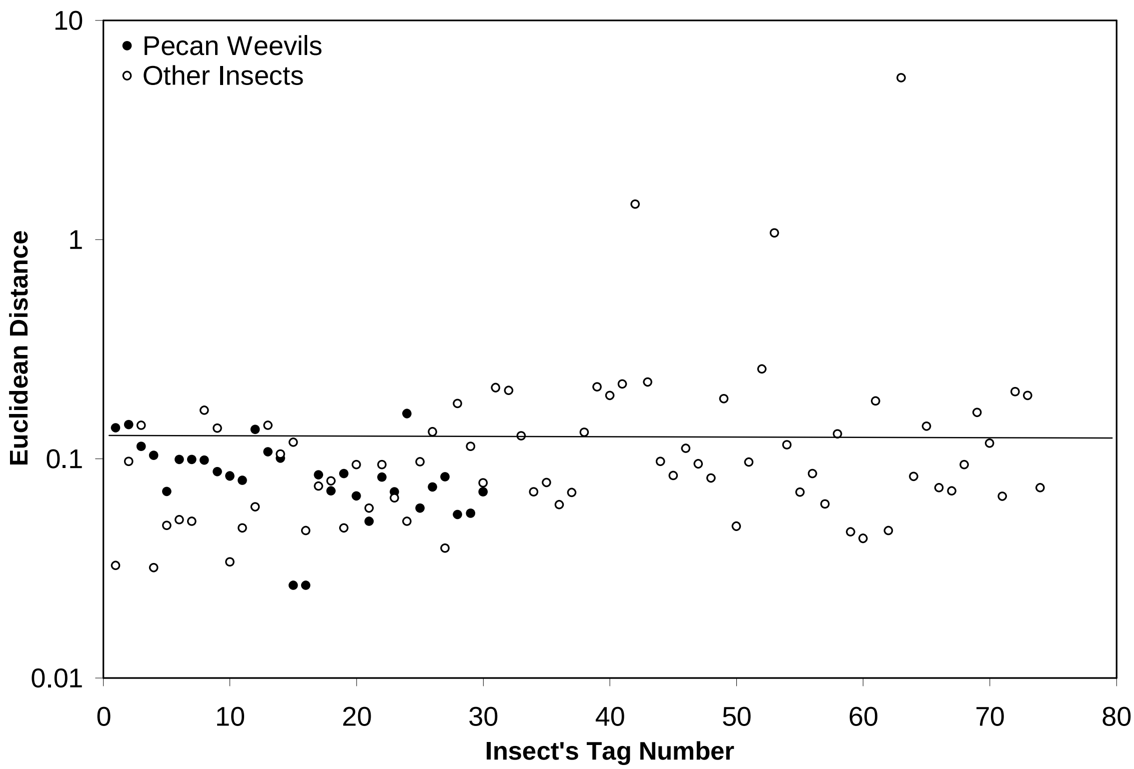



**Figure 5.23 Recognition Results for Pecan Weevils and Other Insects Using the Fourier Descriptors Method (120 Descriptors)**

Three more experiments were carried out using 200, 300, and 400 Fourier descriptors. The performance of the first two experiments was not any better than the previous experiments where 30 and 120 descriptors were used. When using 400 descriptors to represent the insects' shapes, some improvements could be noticed; 35% of the non-pecan weevil insects were correctly classified. Figures 5.24, 5.25, and 5.26 illustrate the results of the three experiments respectively.

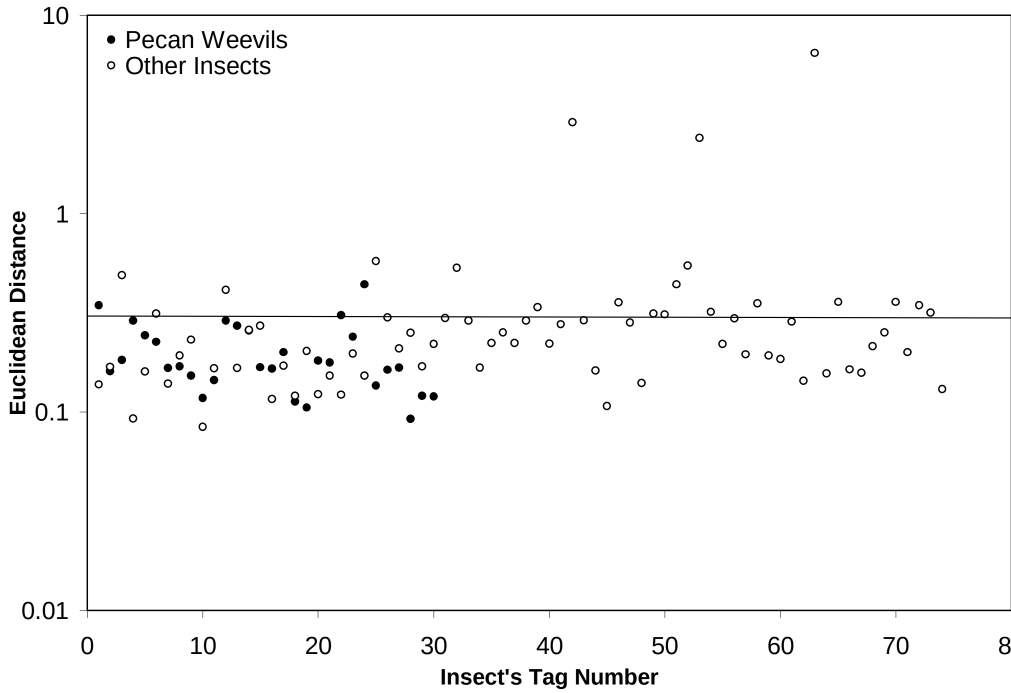



**Figure 5.24 Recognition Results for Pecan Weevils and Other Insects Using the Fourier Descriptors Method (200 Descriptors)**

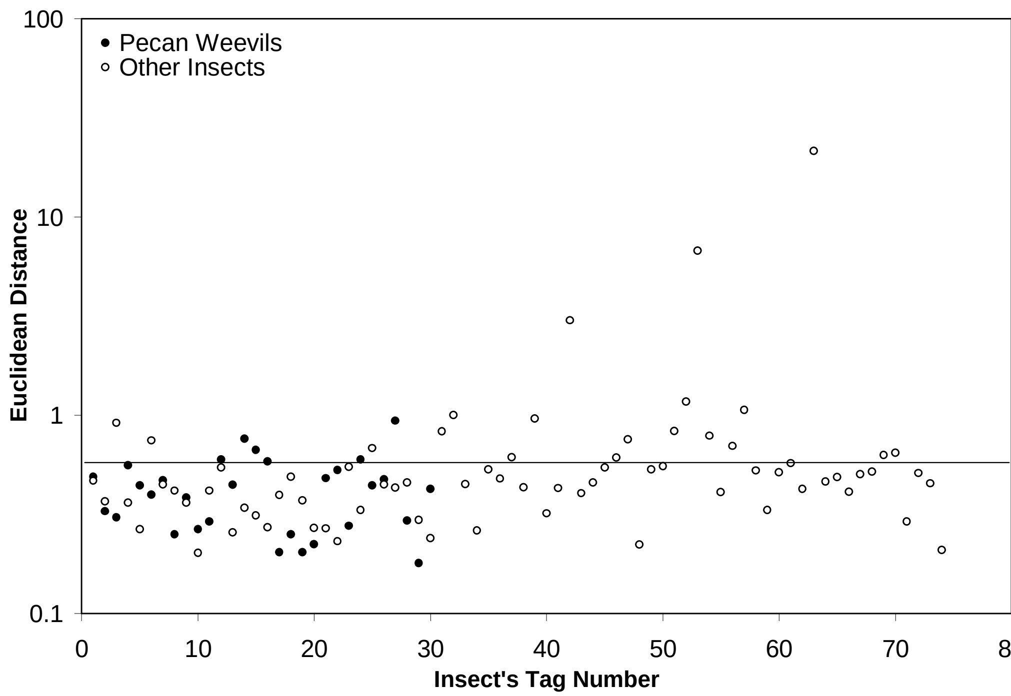



**Figure 5.25 Recognition Results for Pecan Weevils and Other Insects Using the Fourier Descriptors Method (300 Descriptors)**

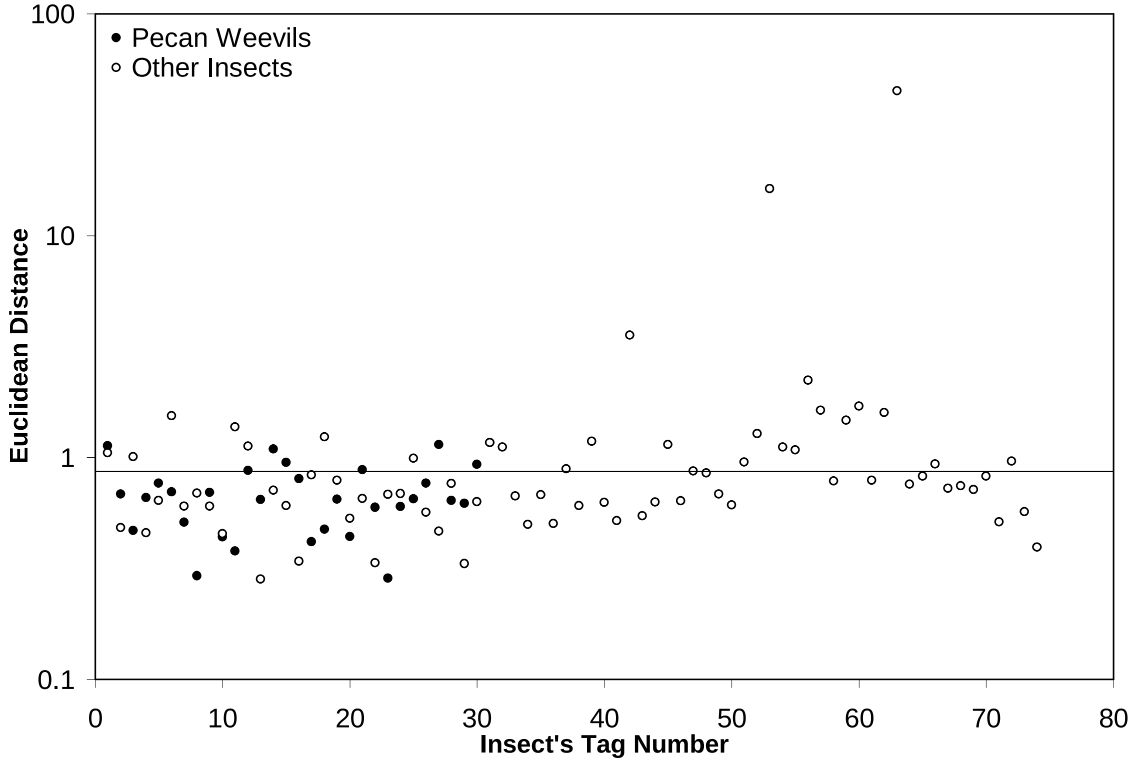



**Figure 5.26 Recognition Results for Pecan Weevils and Other Insects Using the Fourier Descriptors Method (400 Descriptors)**

Finally, an experiment was conducted using 450 Fourier descriptors. This number of descriptors was the maximum possible number of descriptors that could be used. This was because it was the total Fourier descriptors calculated from the boundary of some smaller insects. Using more than 450 descriptors would have required either discarding that sample or adding zeros to the resulting set of Fourier descriptors. In both cases, results would have been negatively affected.

The results of this experiment are presented in Figure 5.27. It can be noticed that 80% of pecan weevils and 41% of the non-pecan weevil insects were correctly classified. Together, these two results were the best recognition rate achieved by the Fourier descriptors method. Although the results of this experiment were the best among others, Fourier descriptor method provided the lowest recognition rate when compared with the other four methods.

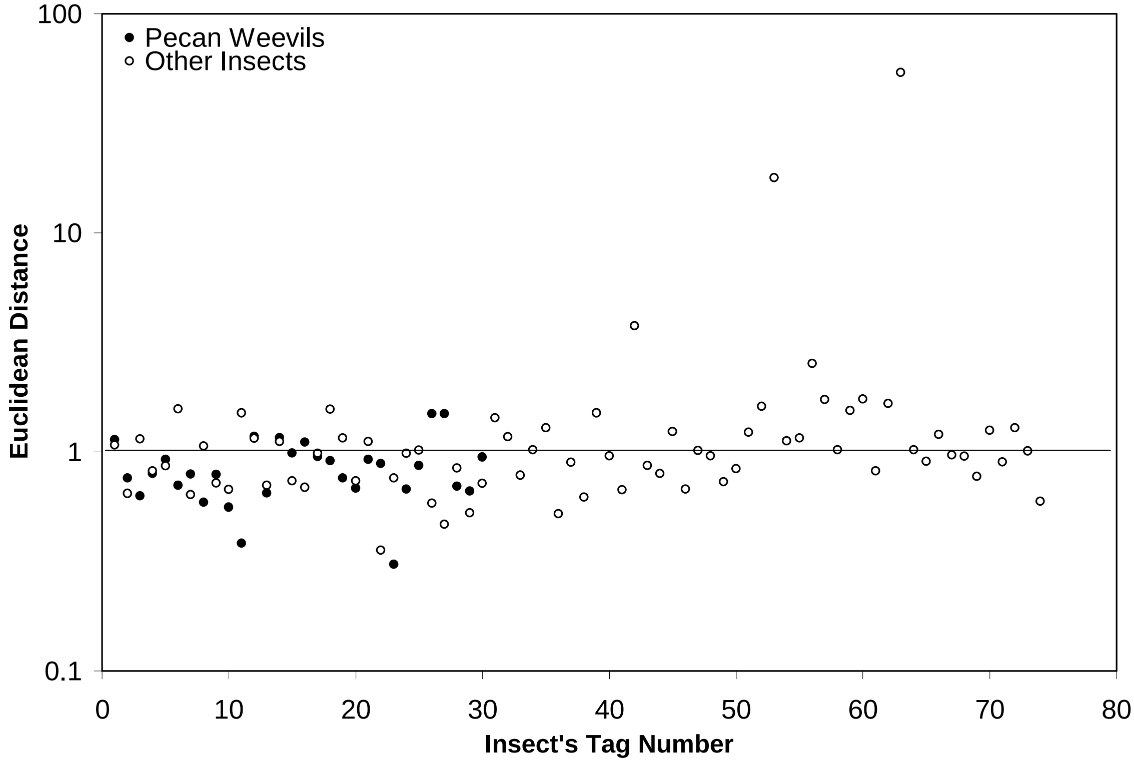



**Figure 5.27 Recognition Results for Pecan Weevils and Other Insects Using the Fourier Descriptors Method (450 Descriptors)**

Table 5.4 summarizes the results of the six experiments conducted in this study using different number of Fourier descriptors. It can be noticed that the successful rate for classifying pecan weevils was always at 80% or higher. On the other hand, the best result of classifying non-pecan weevil insects was 41% when 450 Fourier descriptors were used. Thus, it became clear that the number of Fourier descriptors for any shape depends heavily on its size.

**Table 5.4 Comparison of Recognition Rates Using Different Number of Fourier Descriptors**

| Number of Descriptors Used | Threshold Value at 80% Recognition | Recognition Rate | |
|---|---|---|---|
| | | Pecan Weevils | Other Insects |
| 30 | 0.017 | 83 % | 26% |
| 120 | 0.135 | 87% | 28% |
| 200 | 0.33 | 93% | 20% |
| 300 | 0.59 | 83% | 26% |
| 400 | 0.88 | 80% | 35% |
| 450 | 1.059 | 80% | 41% |

## 5.7 Performance Comparison of Different Methods

Throughout this chapter, the results of using five recognition methods were discussed. These methods varied in their recognition rates and time of processing. Normalized cross-correlation method had shown an excellent classification rate in both testing data sets. The performance of the Region properties method was the closest to the Normalized cross-correlation method. The method of Zernike moments of order 3 was the best in terms of correct classification rates and processing time. String matching and Fourier descriptors were at the fourth and fifth places respectively. Table 5.5 shows the recognition rates of each method when classifying insects of the same testing data sets.

**Table 5.5 Recognition Rates for the Five Methods Used in the Multi-Recognition System**

| Method | Recognition rates | |
|---|---|---|
| | Pecan Weevils | Other Insects |
| Region properties | 90 % | 93% |
| Normalized cross-correlation | 90% | 95% |
| Zernike moments (order 3) | 97% | 99% |
| String matching | 87% | 58% |
| Fourier descriptors (450 Descriptors) | 80% | 41% |

The performance of each implemented method was associated with some margin of errors. These errors can be categorized into two types- Type I and Type II errors. The first one, refers to the case when a pecan weevil is being classified as non- pecan weevil insect (false negative). Type II error, on the other hand, represents the event when a non-pecan weevil insect is being recognized as pecan weevil (false positive). The goal of any recognition system is to minimize the margin of both errors. However, Type I error is more crucial in most applications as it results in rejecting the null hypothesis when it is actually true.

Figure 5.28 illustrates the two types of errors for each individual method. The Normalized cross-correlation and Region properties methods have recorded 10% of Type I error whereas String Matching and Fourier descriptors methods produced 17% and 20% Type I error respectively. The figure also shows that Zernike moments method has the lowest Type I (3%) and Type II errors (1%) whereas these are the highest (20% and 59%) for Fourier descriptors method.

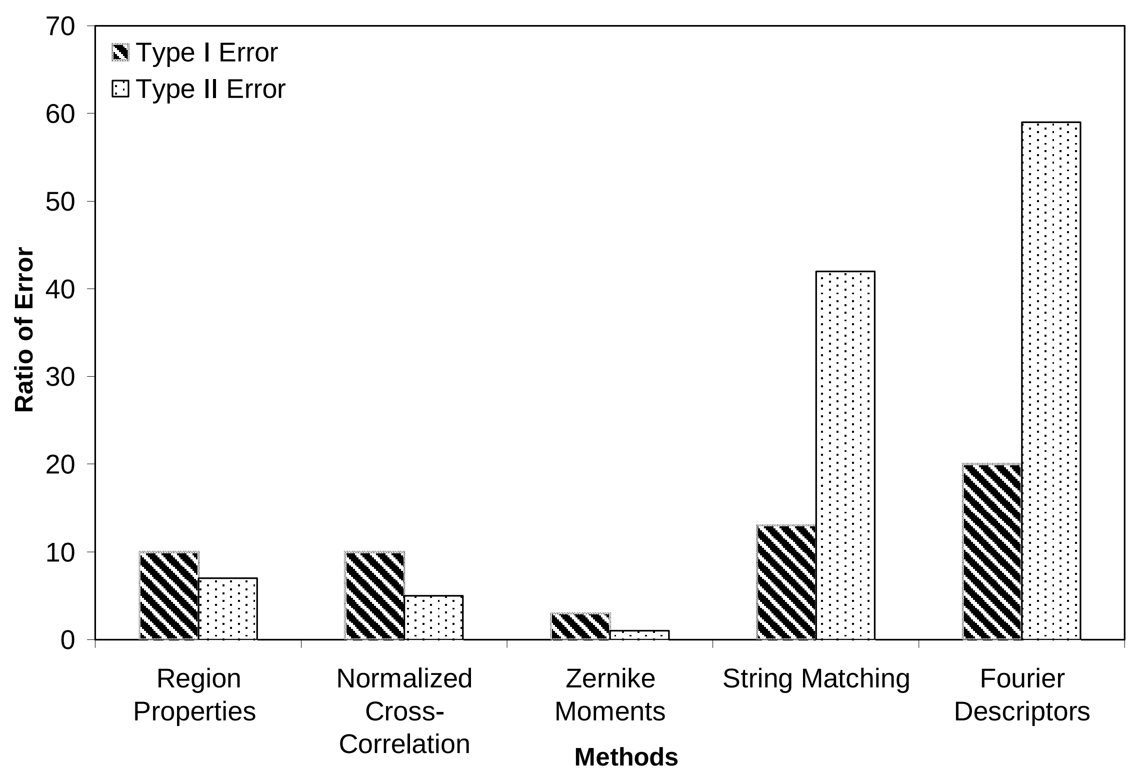



**Figure 5.28 Type I and Type II Errors for the Five Methods**

To enhance the performance of the recognition system, these errors have to be minimized. One way of doing so is to use a combination of more than one method. Table 5.6 demonstrates several possibilities of combining more than one method and different recognition criteria. The first combination included three methods- Region properties, Normalized cross-correlation, and Zernike moments. In this case, a positive match was considered when an input image is correctly classified by two of the three methods. The results showed that 100% of both pecan weevils and the other insects were correctly classified. These findings showed that using the three methods have enhanced the recognition rate and eliminated all errors in this case.

In the second case, four methods were combined and used for testing their ability in classifying insects from the two data sets. Here, correct classification is considered when three out of the four methods positively recognized an input image. The correct classification rate for pecan weevils was 100% where it was 89 % for the other insects' data set. The low correct classification rate for non-pecan weevil group was due to the stricter condition for confirming a positive match.

The rates of recognition in the third case were very promising when using all five methods where positive match was considered when two of them correctly classified a given image. The results showed that 100% of both the pecan weevils and of the non-pecan weevil insects were correctly classified. The pecan weevils were classified at the same recognition rate (100%) when input image was correctly classified by three methods out of the five. Evidently, in this case study, all methods were used and Type I error was 0%. The more stricter criterion of recognition (three out of five) caused the recognition rate for non-pecan weevil insects to be 91%; less than the one of the third case (100%) where the recognition criterion was two positive matches out of five.

**Table 5.6 Different Combinations of Methods and Recognition Criteria**

| Case No. | Methods Involved | Recognition Criteria Image Recognized by | Recognition Rates | |
|---|---|---|---|---|
| | | | Pecan weevils | Other Insects |
| 1 | RP, NCC, and ZM | Two Methods out of the Three | 100% | 100% |
| 2 | RP, NCC , ZM, and SM | Three Methods out of the Four | 100% | 89% |
| 3 | RP, NCC , ZM, SM, and FD | Two Methods out of the Five | 100% | 100% |
| 4 | RP, NCC , ZM, SM, and FD | Three Methods out of the Five | 100% | 91% |
| 5 | RP, NCC , ZM, SM, and FD | Four Methods out of the Five | 93% | 73% |
| 6 | RP, NCC , ZM, SM, and FD | Five Methods out of the Five | 53% | 22% |
| 7 | All methods except NCC | Three Methods out of the Four | 93% | 74% |
| 8 | RP and ZM | Two out of two | 87% | 92% |
| 9 | RP and ZM | One out of two | 100% | 100% |

The five recognition methods were used in the fifth and sixth cases where their recognition criteria were four out of five and five out of five respectively. The output of these experiments showed significant reduction of the recognition rates for the non-pecan weevil insects (73%). These results were not unexpected as the recognition criterion became stricter. In particular, 53% of the pecan weevils and 22% of the non-pecan weevil insects were successfully classified, when a positive match was required from all five methods.

Normalized cross-correlation method is a very reliable recognition method but is known to be time intensive. Therefore, the seventh case involved evaluating the recognition rates of all methods except Normalized cross-correlation (four methods). Positive match here was considered when three out of these four methods correctly classified an input image. The results showed that 93% of the pecan weevils and 74% of the non-pecan weevils were correctly classified. Even though the recognition rate for pecan weevils was promising, the performance of this case in classifying non-pecan weevil insects was associated with 26% Type II error. This degree of inaccuracy may discourage adopting this combination.

The last experiment was performed using only Region properties and Zernike moments as they produced the best classification results in the shortest processing time. The positive match here is confirmed when both of these methods correctly classify an input image. The result showed that 87% of the pecan weevils and 92% of the other insects were correctly classified. These rates reached 100% correct classification if the recognition criterion was to affirm positive match when an input image is recognized by any of the two methods. It is worth noting here that the combination of these two methods has resulted in highest classification rates.

Figure 5.29 shows the 7 insects that failed to be classified using the recognition criterion of case #4. Among these insects, 6 were in fact weevils, with very similar shape and size as pecan weevils. These 6 weevils are responsible for about 86% of the total misclassification incidents.

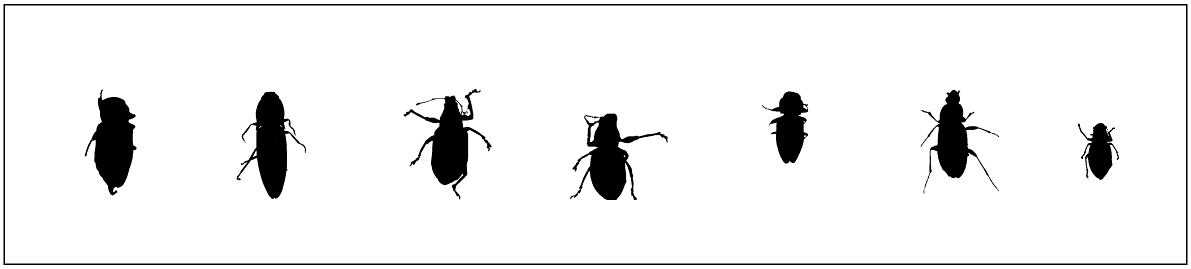


**Figure 5.29 Images of the Seven Misclassified Insects**

## 5.8 Processing Time

Processing time is a major factor in evaluating the performance of any recognition system. It depends on several aspects including the type of image (color, gray, or binary, etc), size of the images, the number of samples used on the template, the capacity of the platform, and the limitation of the software used in processing. The processing time for an input image was recorded for each individual method. Figure 5.30 presents the performance time, on average, to classify one image for each of the five methods. It also presents the corresponding Type I and Type II errors. It can be noticed that Normalized cross-correlation method requires 22 sec. to correlate a given image with the template of 205 images of pecan weevils (size of images were 67×57 pixels).

Compared with the other four methods, Normalized cross-correlation method was the most time intensive method whereas Zernike moments method was the fastest in classifying images. Within 0.09 sec., on average, this method could measure the set of moments of a given image and evaluate its similarity with the ones of each pecan weevil in the training data set. The other times were 0.35 sec., 2.5 sec., and 0.5 sec. for Region properties, String matching, and Fourier descriptors methods respectively. Furthermore, it can be noticed that Zernike moments has the least processing time as well as the lowest Type I and II errors.

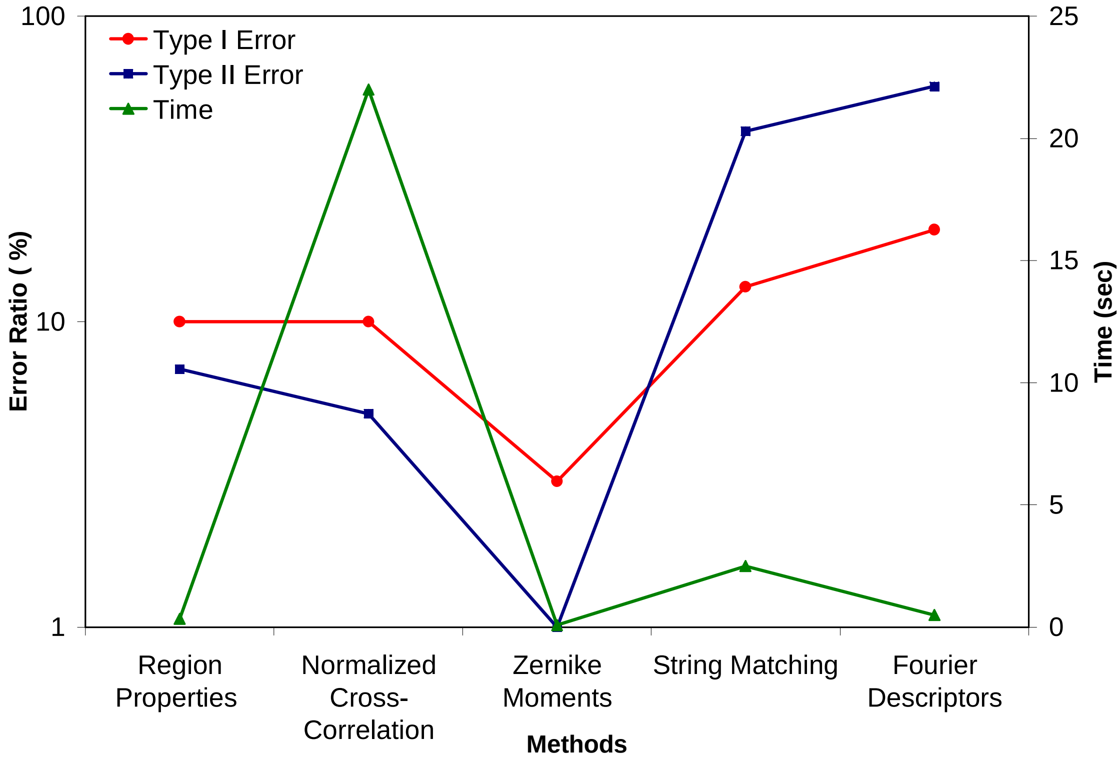



**Figure 5.30 Processing Time and Type I and II Errors for the Five Recognition Methods**

## 5.9 Conclusions

Overall, it can be concluded that a combination of more than one method is **essential** for a robust recognition system since no single method yielded the desired detection rates. Zernike moments at order 3 was found to have the highest recognition rates for pecan weevils and other insects. This method also yielded the lowest Type I and II errors. Also, this method required the least processing time. Region properties method showed similar advantages to the Zernike moments; thus, 100% successful recognition rate for pecan weevils was achieved using a combination of Zernike moments and Region properties methods.

Fourier descriptors method using 450 descriptors was found to be the least successful of these methods and yielded the highest Type I and II errors. Moreover, region-based methods were found to represent insects' shapes better than boundary-based method. As a result, the recognition rates of the region-based methods were higher than the boundary-based method.

## CHAPTER VI

## OVERALL ALGORITHM

The overall algorithm for this work is primarily based on the levels of accuracy and time intensiveness desirable from such a system. A careful analysis of the literature had shown that a combination of methods is **essential** to account for the variations found among the objects of interest. Moreover, using multiple methods also overcomes the inherent limitations of a given method. Therefore, based on their performance, five methods were tested to serve in this recognition system; namely, Normalized-cross-correlation, Region property, String matching, Zernike moments (order 3), and Fourier descriptors (450 descriptors).

### 6.1 Initial Algorithm

These selected methods, when used in a combination, were found to be very effective in classifying pecan weevils, since they deal with different aspects of shape representation. For example, the correlation-based template matching procedure computes the similarity of two images; however, it is not (without spatial transformations) invariant to rotation, scale, or translation. On the other hand, the moment invariants method depends only on the values of the moments regardless of the description of the boundaries. Similarly, Region properties method describes specific properties of a given object and is independent of the boundary pattern. Fourier descriptors method depends on the contour of the object's shape regardless of the interior part of that shape. Unlike the cross-correlation method which deals with patterns quantitatively and does not take into account structural relationships present in a pattern's shape, String matching utilizes such structural relationships to achieve pattern recognition. Therefore, implementing these five methods may result in the methods complementing the overall performance of the algorithm.

Therefore, as a first approximation, the initial algorithm for identifying pecan weevils among other insects incorporated all five methods explained in this chapter. The order in which these methods were implemented was based on their performance in identifying pecan weevils. The first step in this algorithm was Zernike moments of order 3. The recognition rates for pecan weevils and other insects were found to be 97% and 99% respectively. This method of moment invariants outperformed the geometrical moment invariants which had performed poorly in classifying non-pecan weevil insects (successful rate was 41%). At this order of moments, the processing time for classifying one sample was found to be 0.09 sec., on average.

The Region properties method was selected as the second step in this recognition system. The excellent performance of this method and its simplicity motivated this decision. The results of this method had shown that 90% of the pecan weevils and 93% of the other insects' group were correctly classified. Moreover, this method can be used to correct for any rotation of the pecan weevil's appearance on the image prior to the correlation process. This can be done simply by determining the rotation angle of a given image. Furthermore, the algorithm of this method is simple and fast. It required an average of 0.35 second to process one sample.

Normalized Cross-Correlation method was chosen to be the third step in the recognition system for its promising results. Using this method, the successful recognition rates were 90% for pecan weevils and 95% for other insects. However, correlation process is time intensive; but this issue was resolved by resizing the images to be 65 × 57 pixels. By doing so, the time required for identifying one insect was 22 sec., on average. This time involved correlating that image with a template of 205 pecan weevil images. As the purpose of this algorithm was to identify pecan weevils and help in their monitoring in a field area, the time intensiveness of this process was not as critical as in inspection applications like the scanning of products in an assembly line. Therefore, such process time was acceptable. Furthermore, the accurate identification of pecan weevils was deemed far more important than the actual process time required to detect them. Moreover, this study evaluated some other alternatives that employed all methods except the Normalized cross-correlation. Therefore, when the criterion for a positive match is confirmation from three out of the four methods, the recognition rate was 74% and 93% for pecan weevils and other insects respectively. In this case, the maximum processing time would be 3.44 sec.

String matching method was implemented as the fourth step in the recognition process. The results of this method showed that 87% of the pecan weevils and 58% of the non-pecan weevil insects were correctly classified. The processing time for classifying one sample was 2.5 sec., on average.

Fourier descriptors method was used as the fifth step of this recognition system. Using maximum possible descriptors (450 descriptors), the correct classification rates were 80% for pecan weevils and 41% for non-pecan weevil insects. It required less than 0.5 sec. to process and classify one image.

Figure 6.1 illustrates the performance of each step (method) in the recognition system. The arrows in Figure 6.1 refer to pecan weevils that were classified correctly at that particular step whereas diamond-edged lines indicate the pecan weevils not recognized positively. As can be noticed from the figure, each pecan weevil was correctly classified by at least three methods. As shown in Figure 6.1, the sample image was correctly classified by at least three steps to be regarded as an overall correct classification.

For example, Figure 6.1 shows that pecan weevil #2 was positively recognized by the Region properties, Zernike moments, String matching, and Fourier descriptors methods; whereas Normalized cross-correlation method failed to classify it correctly. Thus, this case study has shown that an algorithm consisting of these five methods can yield a reasonable success rate in varying data sets.

The results of applying the five recognition steps on non-pecan weevil insects are presented in Figure 6.2. Here, arrows represent the correctly classified insect at that specific step and the diamond-edged lines represent a misclassified insect. This diagram emphasized the importance of implementing these five methods. It can be noticed that each method has contributed to the recognition process. For example, Fourier descriptors correctly classified insect #1, which could have been misclassified if it had failed this test.

String matching method, on the other hand, classified positively insect #22. Without using this method, this insect would have been misclassified as a pecan weevil and would have reduced the system efficiency. The first three methods played the main role of classifying insects. Region properties method was successful in recognizing all the insects, except 5 of them whereas Normalized Cross-correlation misclassified only 4 out of 74 insects.  Zernike moments misclassified only one pecan weevil and one non-pecan weevils insect.

Figure 6.3 presents the **initial** algorithm of the recognition system in this study. The sequence starts by loading a new image of insect which will directly be processed by the Zernike moments of order three and its six moments would be calculated. The similarity degree between these moments and the moments of pecan weevils will be measured. If this degree is greater than or equal to a threshold value of 0.8, the input image will be classified as pecan weevil. A value of 1 will be assigned to the counter (S=1) and the algorithm will do the next step. In case, an insect does not match any pecan weevil of the training set, the algorithm will keep S=0 and move to the next step.

The input image then would be analyzed by the Region properties method in the second stage.  After measuring the three properties of that insect (area, major and minor axis), their similarity to each pecan weevil of the training set will be evaluated. If the degree of similarity is greater than or equal to the threshold of recognition (1.0), this insect will be recognized as pecan weevil and the counter will add 1 to its value.  In case that insect does not match any pecan weevil of the training set, the algorithm will keep the value of S unchanged and move to the next step. Thus, the value of the counter S would either be 0, 1, or 2 at the second step.

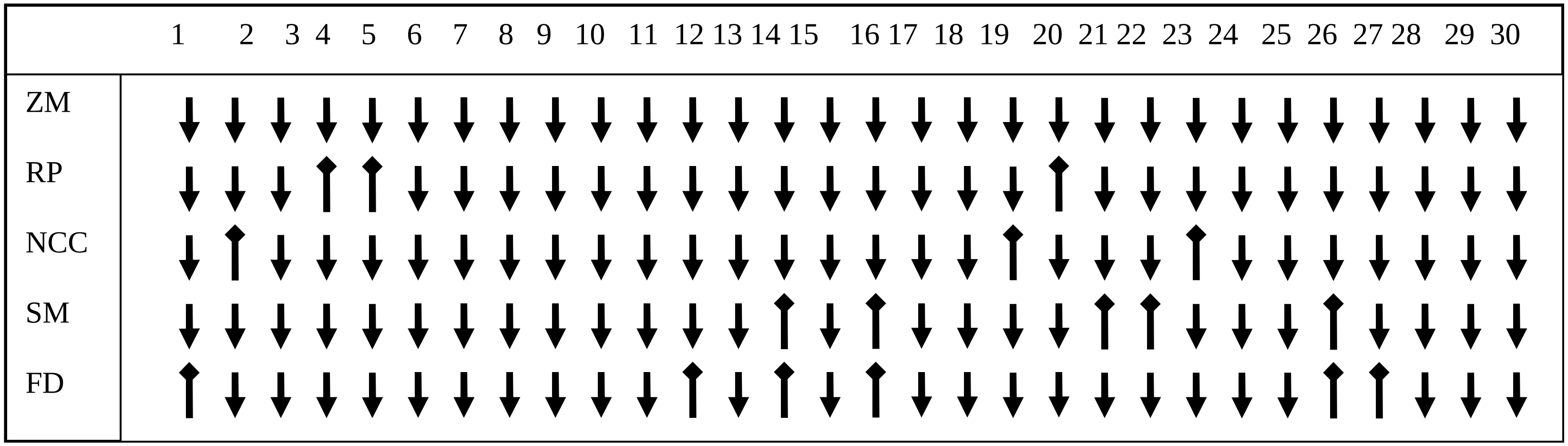



**Figure 6.1 Recognition Results of Pecan Weevils Testing Set Using the Five Methods**

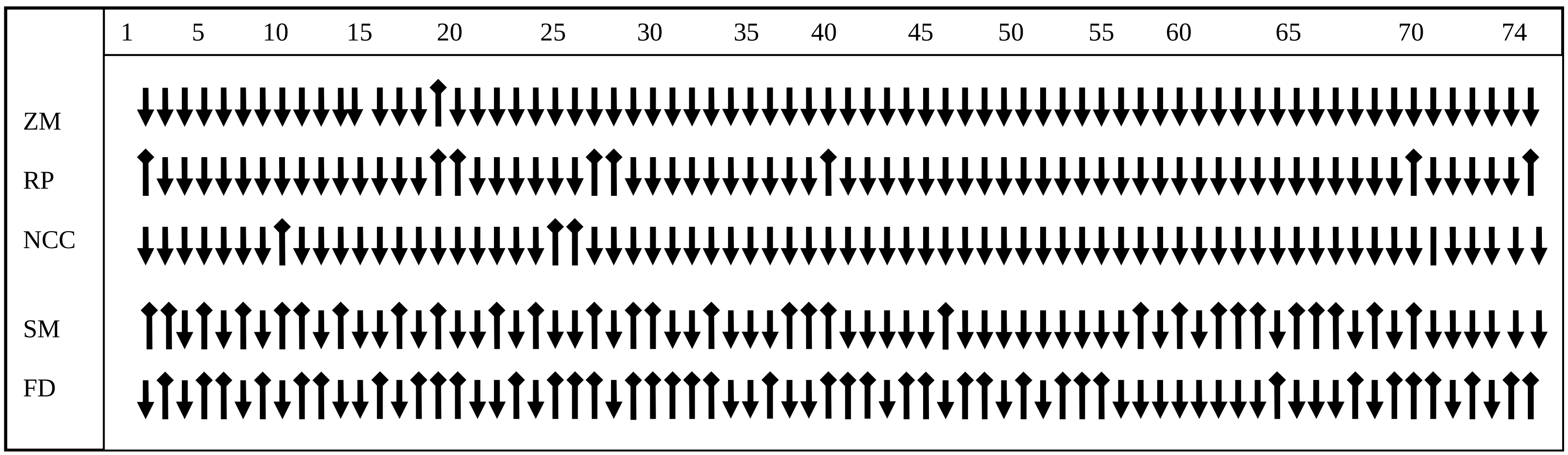



**Figure 6.2 Recognition Results of Non-Pecan Weevil Insects Testing Set Using the Five Methods**

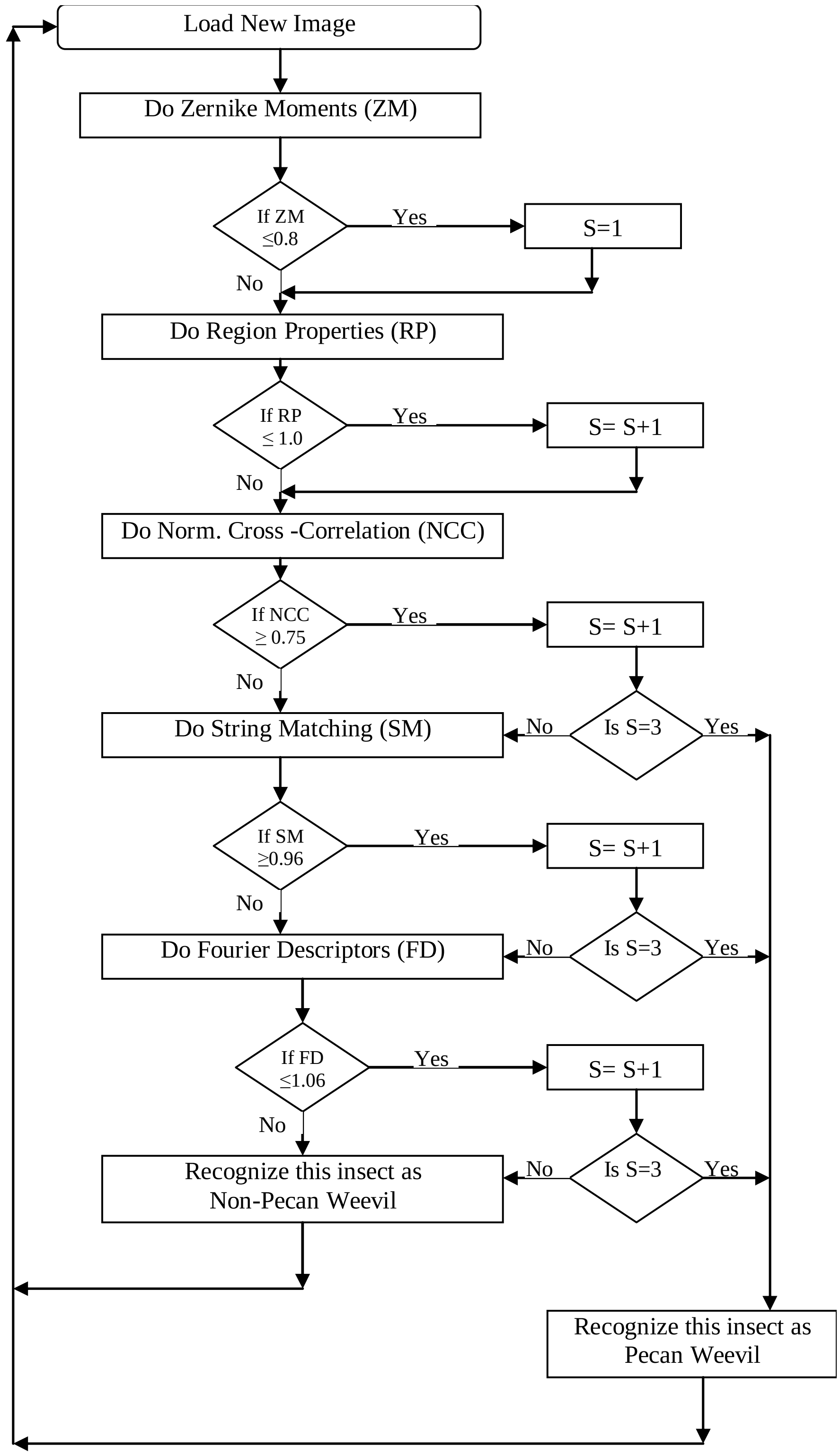



**Figure 6.3 Flow Diagram of the Initial Algorithm for Identifying Pecan Weevils**

In the third step, Normalized cross-correlation method will be used. If the correlation value between the input image and any pecan weevil image is greater than or equal to 0.75, this image will be recognized as pecan weevil and the correlation process will stop. The counter value will increase by one (S=S+1). At this stage, the counter S can have possible values of 3, 2, 1, or zero. For the case when S equals to 2 or 1, the algorithm will go to the fourth step. On the other hand, if S equals to 3, which means input image was recognized by all three previous methods, the algorithm will recognize this image as pecan weevil ending the recognition process of the input image. The program would then be ready for the next image. Moreover, if the S value is 0, which indicates that the input image was not positively classified by any of the three methods, the algorithm will classify this insect as non-pecan weevil insect ending the recognition process and would be ready for a new image.

The String matching method at the fourth step will process the input image only if the counter value is either S=1 or S=2. If the similarity measure of the string of this image and any other string of the training set is greater than or equal to 0.96, this image will be regarded as pecan weevil insect. In this case, the counter value will be either S=2 or S=3. In the first case, the algorithm will go to the fifth step whereas in the second case the insect will be confirmed as pecan weevil. However, if this insect did not match any pecan weevil of the training set, the counter value will remain as either S=1 or S=2. The input image will then go through the fifth method at S=2 even though it was not recognized at this level. However, when S=1, the algorithm will classify the input image as non-pecan weevil ending the recognition process.

At the last step, Fourier descriptors method will process the image if the counter value is S=2. This method will calculate the Fourier descriptors (450 descriptors) of the input image and measure the similarity between this set of descriptors and those of the training data set. If the similarity measure is greater than or equal to the threshold of 1.059, this image will be classified as pecan weevil. In this case, the counter will add one to its value (S=3), and hence the image will be confirmed as a pecan weevil insect. Otherwise the input image will be regarded as non-pecan weevil insect. In both cases, the recognition process for that image will be complete and the system would be ready for a new input image.

The maximum processing time for one image through the five methods is 25.44 sec., one average. However, the system may require shorter time because an input image may not need to be matched with all pecan weevil images in the template, if it positively matches any one of them. Furthermore, the algorithm was designed to end the recognition process when the input image fails the first three steps (methods). In this case, the processing time was found to be 22.44 sec.

The above algorithm was successfully implemented in this recognition system and it yielded promising results for the data sets investigated.

## 6.2 Revised Algorithm

Although, the algorithm discussed above showed reasonable classification rates, an alternative algorithm was also tested. This alternative algorithm implemented only two methods, namely, Zernike Moments and Region properties. The recognition criterion was chosen to be a positive match from either of the two methods. The results were found to be superior to the older algorithm, for the data set tested. In addition to better performance, the revised algorithm required very small processing times (0.44 sec. as compared to 25 sec. in the old algorithm, on average). Moreover, the target monitoring system of this study is a wireless network wherein the data transfer rates are an important factor. Therefore, using fewer methods in the algorithm is highly recommended.

Based on the above findings and a careful analysis of the system requirements, it is concluded that the application of the two methods, namely, Zernike Moments and Region properties would yield the desired success rates for identifying pecan weevil, in field applications.

Figure 6.4 illustrates the revised algorithm implemented in this study. It shows the two methods mentioned above applied in a sequential order. It can be seen that a positive match form either of these two methods was used as the selection criterion.

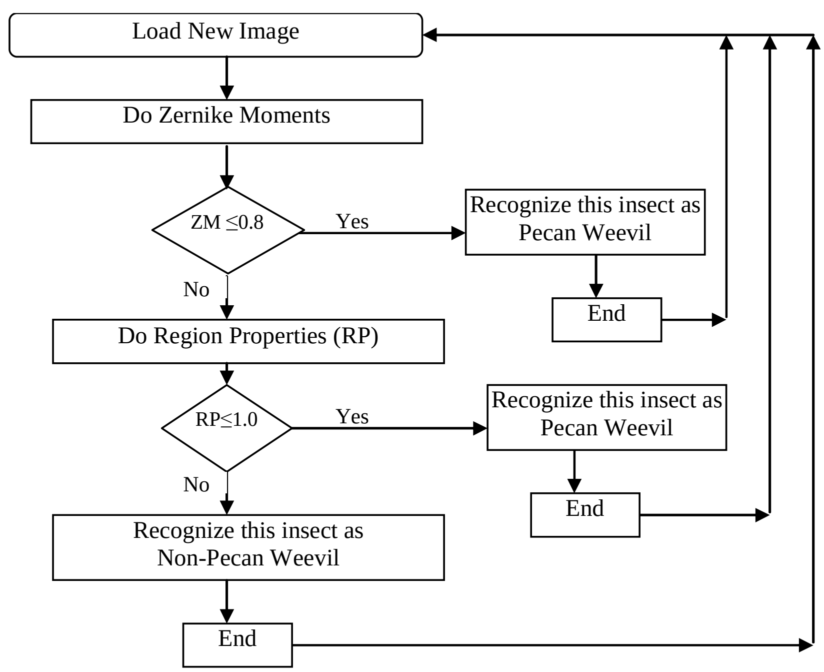



**Figure 6.4 Flow Diagram of the Revised Algorithm for Identifying Pecan Weevils**

As discussed above, this algorithm yielded the best results when compared to other combinations of methods. Therefore, this algorithm is expected to be implemented in a wireless monitoring system for field applications.

CHAPTER VII

# CONCLUSIONS AND RECOMMENDATIONS

## 7.1 Conclusions

The aim of this work was to develop the software part of a wireless network imaging system that can automatically identify pecan weevils in the field. This study resulted in developing a new recognition algorithm which uses image processing techniques based on template matching to identify pecan weevils among other insects that are naturally present in the pecan habitat. This system is anticipated to replace the manual insects' monitoring techniques currently in use, and is expected to be a valuable device for pest management, in general.

Five recognition methods, namely, Zernike moments, Region properties, Normalized cross-correlation, String matching, and Fourier descriptors methods were used in this recognition system. It was found that no single method was sufficiently robust to yield the desired recognition rate, especially in varying data sets. It was also found that region-based shape representation methods were better suited in recognizing insects than boundary-based methods.

The most favorable recognition rate was found to be when an input image was required to be positively matched by either one of the two methods, namely, Zernike moments and Region properties. This evaluation criterion resulted in accurate and reliable classification of pecan weevil images. Specifically, this criterion ensured 100% correct identification rate for both pecan weevils and other insects. Therefore, it is concluded that a reliable and accurate recognition system for identifying pecan weevils has been obtained by implementing the new algorithm.

### 7.2 Recommendation and Further Study

This study focused on distinguishing pecan weevils among other insects that are naturally present in the pecan habitat. Evaluating the system for a wider variety of insects would add to the reliability of this algorithm. Further, it was noticed that the majority of the misclassification rate originated from misclassifying some weevils as pecan weevils. Therefore, the inclusion of a template of such weevils (as a training data set) would greatly enhance the capability of the algorithm to distinguish between different types of weevils.

REFERENCES

Abu-Mostafa, Y. and D. Psaltis 1984. Recognitive Aspects of Moment Invariants. *IEEE TRANS. Pattern Anal. Mach. Intellig.* . PAMI-6 (6): 698-706.

Adair, H. 1932. Black Pit of the Pecan and Some Insects Causing It. *USDA Circ.* 234: 1-14.

Aguirre-Uribe, A. and L. Alberto 1979. *Biology of the Immature Stage of the Pecan Weevil Curculio Caryae (Horn) and Ovipositions Habits of the Adult Weevil*. College Station: Evans Library Annex.

Arbuckle, T., et al. 2001. Biodiversity Informatics in Action: Identification and Monitoring of Bee Species Using Abis. In *15th Int. Symp. Informatics for Environmental Protection*, 425–430.

Baker, W.L. 1972. Eastern Forest Insects. *USDA Forest Serv. Misc. Publ* . 1175: 642.

Belkasim, S., et al. 1991. Pattern Recognition with Moment Invariants: A Comparative Study and New Results Pattern Recognition. *Pattern Recognition*. 24 (12): 1117-1138.

Bilsing, S. 1940. Notes on Some Common Pecan Insects. In *Proc.Texas Pecan Growers Assoc.*, 49-50.

Bissell, T. 1935. How to Fight the Pecan Weevil. In *Proc. TX. Pecan Growers Association* 33-34.

Boddy, L., et al. 2000. Identification of 72 Phytoplankton Species by Radial Basis Function Neural Network Analysis of Flow Cytometric Data. *Marine Ecology Progress Series*. 195: 47-59.

Boethel, D.J., et al. 1974. Sampling Pecan Weevil Nut Infestations: Efects of Tree, Height and Sector. *Environ Entomol.* 3208-3210.

Borror, D.J., et al. 1976. *An Introduction to the Study of Insects. Holt, Reinhart and Winston*. 4th. Ed. New York.

Brison, F. 1975. *Pecan Culture*. Austin, Texas: Capitol Printing.

Brooks, F. 1910. Snout Beetles That Injure Nuts. *W.V. Agric. Exp. Stan. Bull*. 128:162-5.

Brunelli, R. and T. Poggio 1993. Face Recognition: Features Versus Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 15 (10): 1042 – 1052.

Calcote, V. R. 1975. Pecan Weevil: Feeding and Initial Oviposition as Related to Nut Development. *J. Econ. Entomology*. 68: 4-6.

Chau, K. 1949. Study of the Feeding Habits and Investigation Soil Fumigation as Means of Control for Pecan Weevil. M.S. Thesis. Oklahoma State University.

Chellappa, R. and R. Bagdazian 1984. Fourier Coding of Image Boundaries. *IEEE Transaction on pattern analysis and machine intelligenc*. 6 (1): 102-105.

Choi, M. S. and W.Y. Kim 2002. A Novel Two Stage Template Matching Method for Rotation and Illumination Invariance. *Pattern Recognition* 35: 119–129.

Collins, J. K., et al. 1996. Mating Behavior and Peak Mating Activity of the Pecan Weevil Curculio Caryae (Horn). *Southwestern Entomologist* 21 (4): 479-481.

Craighead, F. 1950. Classification of the Nut Curculios (Formerly Balaninus) of Boreal America. *Entomology Am.* 17: 129 –53.

Criswell, J., et al. 1975. Longevity, Puncturing of Nut, and Ovipositional Activities by the Pecan Weevil on Three Cultivars of Pecans. *Economical Entomology* 68: 173-7

Daly, H. V., et al. 1998. *Introduction to Insect Biology and Diversity*. Second edition. Oxford: Oxford University Press.

Darrell, T., et al. 1996. Task- Specific Gesture Analysis in Real-Time Using Interpolated Views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 18 (12): 236-242.

Di Ruberto, C. and A. Morgera 2005. Image Analysis and Processing. In *Proceedings Lecture Notes in Computer Science*. 212-219.

Dietrich, C. and C. Pooley 1994. Automated Identification of Leafhoppers (Homoptera: Cicadellidae: Draeculacephala Ball. *Annals of the Entomological Society of America*. 87 (4): 412-423.

Do, M., et al. 1999. A Test of a Pattern Recognition System for Identification of Spiders. *Bull. Entomology Res*. 89: 217-224.

Dudani, S., et al. 1977. Aircraft Identification by Moment Invariants. *Institute of Electrical and Electronics Engineers (IEEE) Trans. Computers*. 26: 39-45.

Eikenbary, R. and H. Raney 1973. Intra Tree Dispersal of the Pecan Weevil. *Environ. Entomology*. 2: 927-930.

El-Khaly, F. and M. Sid-Ahmed 1990. Machine Recognition of Optically Captured Machine Printed Arabic Text. *Pattern Recognition*. 23 (11): 1207-1214.

Ellis, H. 2007. Pecan Weevil. http://www.bugwood.org/factsheets/99-007.html

Embleton, K. V., et al. 2003. Automated Counting of Phytoplankton by Pattern Recognition: A Comparison with a Manual Counting Method. *Journal of Plankton Research*. 25 (6): 669-681.

Ezer, N., et al. 1994. A Comparative Study of Moment Invariants and Fourier Descriptors in Planar Shape Recognition. *Institute of Electrical and Electronics Engineers (IEEE) Trans.* 1: 242-245.

Farag, A., et al. 2004. Detection and Recognition of Lung Abnormalities Using Deformable Templates. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)* 738-741.

Farmery, M.J. 1981. Optical Studies of Insect Flight at Low Altitude. Ph.D. Thesis York, U.K.: University of York.

Flusser, J. and T. Suk 1994. Affine Moment Invariants - a New Tool for Character-Recognition. *Pattern Recognition Letters* 15 (4): 433-436.

France, I., et al. 2000. A New Approach to Automated Pollen Analysis. *Quatern. Sci. Rev*. 19: 537–546.

Gdalyahu, Y. and D. Weinshall 1999. Flexible Syntactic Matching of Curves and Its Application to Automatic Hierarchical Classification of Silhouettes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 21 (12): 1312-1328.

Gibson, L.P. 1969. Monograph of the Genus Curculio in the New World (Coleoptera: Curculionidae). Part 1. United States and Canada. *Miscellaneous Publications of the Entomological Society of America*. 6: 241-285.

Gill, J. 1924. Important Pecan Insects and Their Control. *U.S. Dept. Agr. Farmers Bull.* 1364: 48.

Gonzalez, R. and R. Woods 2001. *Digital Image Processing*. 2nd Edition. New Jersey: Prentice Hall.

Gonzalez, R., et al. 2004. *Digital Image Processing Using Matlab*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Goshtasby, A. 1985. Template Matching in Rotated Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 7: 338-344.

Goshtasby, A., et al. 1984. A Two-Stage Cross Correlation Approach to Template Matching. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. 6 (3): 374-378.

Gossard, H. 1905. Insects of the Pecan *Fla. Bull*. 79: 308.

Granlund, G.H. 1972. Fourier Processing for Hand Print Character Recognition. *IEEE Transactions on Computers C*. 21 (2): 195-201.

Hamilton, J. 1890. Balaninus (Pecan Weevil) – Its Food Habits. *Can. Entomology* 22: 1-8.

Harp, S. 1970. Biology and Control of Pecan Weevil, Curculio Caryae (Horn), in Texas. Ph.D. Dissertation. Texas A&M University.

Harp, S. and H. Van Cleave 1976. Biology of Pecan Weevil. *Southwestern Entomology*. 1: 21-30.

Harris, M. 1975. Pecan Weevil Distribution in Some Texas Soils. *Environmental Entomology* 48: 49-53.

Harris, M. 1976. Pecan Weevil Adult Emergence, Onset of Oviposition and Larval Emergence from the Nut as Affected by the Phenology of Pecan *Journal of Economic Entomology* 69 (2): 167-170.

Harris, M. 1976. Pecan Weevil Infestation of Pecans of Various Sizes and Infestations. *Environment Entomology* 52: 48-50.

Harris, M. 1979. Pecan Weevil Distribution on Pecan across the Pecan Belt. *South. Coop. Set. Bull.-Tex. Agrie. Exp. Stn* 238: 12.

Harris, M. and N. Mcglohon 1972. Pecan Insects and Diseases and Their Control. *Ibid. Bull*. 644: 1-20.

Harris, M. and D. Ring 1980. Pecan Weevil (Coleoptera, Curculionidae) Emergence Related to Soil-Moisture. *Journal Of Economic Entomology* 73 (2): 339-343.

Harris, M., et al. 1981. Longevity of Post-Emergent Adult Pecan Weevil in the Laboratory and Field *Environmental Entomology* 10 (2): 201-205.

Hatfield, L., et al. 1982. Mating-Behavior of the Pecan Weevil, Curculio-Caryae (Horn) (Coleoptera, Curculionidae). *Journal of The Georgia Entomological Society*. 17 (2): 275-279.

Hendricks, D. 1990. Electric Systems to Automatically Detect Insect Pests That Respond to Species-Specific Chemical Sex Pheromone Baits. *Institute of Electrical and*

*Electronics Engineers*. 1: 287-293.

Hendricks, I., et al. 2005. Image Analysis Techniques: A Tool for the Identification of Bivalve Larvae. *Journal of Sea Research*. 54: 151-162.

Herrera, E. 2000. Designing a Pecan Orchard Guide H-607. http://cahe.nmsu.edu/pubs/_h/h-607.html

Hill, D. S. 1983. *Agricultural Insect Pests of the Tropics and Their Control*. Second Edition. Cambridge University Press.

Hinrichs, H. A. 1965. Pecan Investigations in Oklahoma. *Annual Report Nut Growers Assn*. 56: 44-51.

Hinricks, H. and H. Thompson 1955. Insecticide Tests for Pecan Weevil Control. *Oklahoma Agricultural Experimental Station Bull.* 450: 1-12.

Hu, M. 1962. Visual Pattern Recognition by Moment Invariants. *IRE Trans. Inf. Theory* 8: 179-187.

Hupkens, T.M. and J. De Clippeleir 1995. Noise and Intensity Invariant Moments. *Pattern Recognition Letters* 16 (4): 371-376.

Hwang, S.K. and W.Y. Kim 2006. A Novel Approach to the Fast Computation of Zernike Moments. *Pattern Recognition* 39 (11): 2065-2076.

Kadir, T., et al. 2004. An Affine Invariant Salient Region Detector. In *European Conference on Computer Vision (ECCV-04)*, 228-241.

Kauppinen, H., et al. 1995. An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2d Shape Classification. *IEEE Transactions Pattern Analysis and Machine Intelligence*. 17 (2): 201 – 207.

Kern, F. 1949. The Pecan Weevil Larva, Curculio Caryae (Horn); Studies on Its Morphology and Biology. M.S. Thesis. Stillwater: Oklahoma State University.

Khotanzad, A. and Y.H. Hong 1990. Invariant Image Recognition by Zernike Moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 12 (5): 489-497.

Kim, S., et al. 2003. Choquet Integral-Based Aggregation of Image Template Matching Algoritnms. In *22nd International Conference of North American Fuzzy Information Processing Society*, 143-148.

Kim, W.Y. and Y.S. Kim 2000. A Region-Based Shape Descriptor Using Zernike Moments. *Signal Processing: Image Communication*. 16 (1-2): 95-102.

Kohandani, A., et al. 2006 *A Fast Algorithm for Template Matching* 398-409.

Springer Berlin.

Langston, J. 1930. Some Experiments with the Pecan Weevil. *The Quart. Bull. State Plant Board of Miss.* 9 10.

Larios, N., et al. 2007. Automated Insect Identification through Concatenated Histograms of Local Appearance Features. In *Institute of Electrical and Electronics Engineers, Application of Computer Vision*, 26: IEEE.

Lee, J.D. and J.Y.  Lee 1998. A New Algorithm for Two-Dimensional Object Inspection Using String Matching. *Mathl. Comput. Modelling* 27 (1): 101-116.

Leiby, R. 1925. Insect Enemies of Pecans in North Carolina. *N.C. Dept. Agric. Bull.* . 67.

Lewis, J. 1995. Fast Normalized Cross-Correlation. *Vision Interface*. 120-123.

Liao, S. and M.  Pawlak 1996. On Image Analysis by Moments. *IEEE Trans. on Pattern Analysis and Machine Intelligence*. 18 (3): 254-266.

Lin, T.W. and Y.F. Chou 2003. A Comparative Study of Zernike Moments. In *Proceedings of the  IEEE/WIC International Conference on Web Intelligence*, 516-519.

Maitra, S. 1979 Moment Invariants. In *Proceedings of the IEEE.* 697 – 699.

Mayo, M. and T.  Watson 2007. Automatic Species Identification of Live Moths. *Artificial Intelligence* 20 (2): 195-202.

Mikolajczyk, K. and C. Schmid 2004. Scale and Affine Invariant Interest Point Detectors. *IJCV*. 60 (1): 63-86.

Mizell, R. F. 2003. Traps for Monitoring Plum Curculio and Pecan Weevils. http://ufinsect.ifas.ufl.edu/weevil-trapping.htm

Mody, N., et al. 1976. Volatile Components of Pecan Leaves and Nuts, Carya Illinoensis Koch. . *J.Agric. Food Chem.* 24 (1): 175-177.

Morgan, A., et al. 1998. Evaluation of Artificial Neural Networks for Fungal Identification, Employing Morphometric Data from Spores of Pestalotiopsis Species. *Mycol. Res.* 102 (8): 975-984.

Moznette, G., et al. 1931. Insects of the Pecan and How to Combat Them *USDA Farmers Bull.* . 16: 54-59.

Mukundan, R. and K.R.  Ramakrishnan 1993. Attitude Estimation Using Moment Invariants. *Pattern Recognition Letters* 14 (3): 199-205.

Mukundan, R. and K.R. Ramakrishnan 1996. An Iterative Solution for Object Pose Parameters Using Image Moments. *Pattern Recognition Letters* 17 (12): 1279-1284.

Mulder, P. 2004. Pecan Weevil Season in Full Swing.
http://entoplp.okstate.edu/Pddl/2004/PDIA3-15.pdf

Neel, W., et al. 1975. Pecan Weevil: Late Season Emergence in Arkansas. *Ibid*. 9-20.

Osburn, M., et al. 1963. Controlling Insects and Diseases of Pecan. *USDA Agriculture Handbook* 204: 5-8.

Padilla-Vivanco, A., et al. 2007. Comparative Analysis of Pattern Reconstruction Using Orthogonal Moments. *Optical Engineering* 46 (1): 17002-17017.

Palaniappan, R., et al. 1999 Noise Tolerant Moments for Neural Network Classification. *IEEE International Joint Conference on Neural Networks*. 4: 2802-2807.

Park, J.S. and T.Y. Kim 2004. Shape Image Retrieval Using Invariant Features. In *Lecture Notes In Computer Science 3332:* 146-153.

Paugam, A., et al. 2006. The Limits of Morphometric Features for the Identification of Black-Lip Pearl Oyster (Pinctada Margaritifera) Larvae. *Journal of Shellfish Research*. 25 (3): 959-967.

Pavlidis, T. 1979. Filling Algorithms for Raster Graphics. *Computer Graphics and Image Processing*. 10: 126-141.

Pavlidis, T. and F. Ali. 1979. A Hierarchical Syntactic Shape Analyzer. *IEEE Transactions on Pattern Recognition and Machine Intelligence*. 1 (1): 2-9.

Pech-Pacheco, J. and J. Alvarez-Borrego 1998. Optical-Digital System Applied to the Identification of Five Phytoplankton Species. *Mar Biol* 132: 357–365.

Persoon, E. and K. Fu 1986. Shape Discrimination Using Fourier Descriptors. In *IEEE Transactions on Pattern Analysis and Machine Intelligence. 8:* 388-397.

Polles, S. and J. Payne 1974. Pecan Weevil, Hickory Shuckworm and Yellow Pecan Aphids. In *S. E. Pecan Growers Association* 79-94

Price, W.S.1939. Observation of Pecan Weevil. In *Tex. Pecan Growers Association*, 30.

Prokopy, R. , et al. 2003. Odor-Baited Trap Trees: A New Approach to Monitoring Plum Curculio (Coleoptera: Curculionidae). *Journal of economic entomology*. 96 (3): 826-834.

Raney, H. 1969. Flight Habits of the Pecan Weevil, Curculio Caryae (Horn) (Coleoptera: Curculionidae). M.S. Thesis. Stillwater: Oklahoma State University.

Raney, H. and R. Eikenbary 1968. Investigations on Flight Habits of the Pecan Weevil Curculio Caryae (Coleoptera: Curculionidae). *Canada Entomology*. 100: 1091-1095.

Raney, H., et al. 1970. Population Density of Pecan Weevil under Stuart Pecan Trees. *Journal of Economical Entomology* 63: 697-700.

Ree, B., et al. 2000. Controling the Pecan Weevil.
http://insects.tamu.edu/extension/bulletins/l-5362.html

Reeves, A., et al. 1988 Three-Dimensional Shape Analysis Using Moments and Fourier Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 10 (6): 937-943.

Reynolds, D. and Riley J. 2002. Remote-Sensing, Telemetric and Computer-Based Technologies for Investigating Insect Movement: A Survey of Existing and Potential Techniques. *Computers and Electronics in Agriculture*. 35 (2): 271-307.

Ridgway, C., et al. 2001. Imaging for the High-Speed Detection of Pest Insects and Other Contaminants in Cereal Grain in Transit. *American Society of Agricultural and Biological Engineers*. Paper 01-3056: 7.

Riley, J.R. 1993. *Flying Insects in the Field.* 1-15. London: Chapman and Hall.

Ring, D. R. 1978. Biology of Pecan Weevil Emphasizing the Period from Oviposition to Larval Emergence. M.S. Thesis. Texas A&M University.

Rui, Y., et al. 1999. Modified Fourier Descriptors for Shape Representation – a Practical Approach. *Journal of visual communication and image representation*. 10 (4): 39-47.

Sarfraz, M. 2006. Object Recognition Using Fourier Descriptors: Some Experiments and Observations. In *Proceedings of the International Conference on Computer Graphics, Image Processing and Visualization*, 1706 – 1708.

Schaefer, G.W. and G.A. Bent 1984. An Infra-Red Remote Sensing System for the Active Detection and Automatic Determination of Insect Flight Trajectories (Iradit). *Bull. Entomol. Res.* 74 261–278.

Schouest, L.P. and T.A. Miller 1994. Automated Pheromone Traps Show Male Pink Bollworm (Lepidoptera: Gelechiidae) Mating Response Is Dependent on Weather Conditions. *J. Econ. Entomol*. 87: 965–974.

Schroder, S., et al. 2002. *The New Key to Bees: Automated Identification by Image Analysis of Wings.* Kevan. Ministry of Environment: Brasilia.

Shen, D.G. and H.S. Ip 1999 Discriminative Wavelet Shape Descriptors for Recognition of 2-D Patterns. *Pattern Recognition*. 32 (2): 151-165.

Sluzek, A. 1988 Using Moment Invariants to Recognize and Locate Partially Occluded 2d Objects. *Pattern Recognition Letters* 7(4): 253-257.

Sluzek, A. 1995 Identification and Inspection of 2-D Objects Using New Moment-Based Shape Descriptors. *Pattern Recognition Letters* 16 (7): 687-697.

Sluzek, A. 2005. On Moment-Based Local Operators for Detecting Image Patterns Shape Descriptors. *Image and Vision Computing* 23 (3): 287-298.

Storring, M. and T.B. Moeslund 1997. An Introduction to Template Matching Adapted from Fixation and Tracking Using Active Cameras with Foveated Wide-Angle Lenses. M.S. Thesis.

Swingle , H. 1934. The Pecan Weevil. *Alabama Agricultural Experimental Station Leaflet*. 4: 4.

Sze, T.W. and Y.H. Yang 1981. A Simple Contour Matching Algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 3(6): 676-678.

Teague, M.R. 1980. Image-Analysis Via the General-Theory of Moments. *Journal of the Optical Society of America* 70 (8): 920-930.

Tedders, W. and M. Osburn 1970. Notes on Biology and Control of the Pecan Weevil. In *S. E. Pecan Growers Association 63:* 35.

Tedders, W., et al. 1969. Effects of Color and Trunk-Wrap on Pecan Weevil Catch in Pyramidal Traps. *Journal of entomological science* 31 (4): 414-419.

Teh, C.H. 1988. On Image Analysis by the Methods of Moments. *IEEE Trans. On Pattern Analysis and Machine Intelligence*. 10 (4): 496-530.

Tsirikolias, K. and B.G. Mertzios 1993. Statistical Pattern-Recognition Using Efficient 2-Dimensional Moments with Applications to Character-Recognition. *Pattern Recognition*. 26 (6): 877-882

United States Department of Agriculture, 2004.
www.ers.usda.gov\briefing\fruitandtreenuts\fruitnutpdf\pecansfts304.pdf

University of Georgia 2005. Pecan-Carya Illinoensis.
http://www.uga.edu/fruit/pecan.html

Van Cleave, H. and S. Harp 1971. The Pecan Weevil: Present Status and Future Prospects. *SE Pecan Growers Assoc.* 11: 64-99.

Van Cleave, H. and C. Shaykewick 1970. The Influence of Soil Water Metric Potential on the Strength Properties of Unsaturated Soil. In *Soil Science Society Am.* 21-30.

Wallace, A. 1988. A Comparison of Approaches to High-Level Image Interpretation. *Pattern Recognition* 21 (3): 241-259.

Wang, Lz. and G. Healey 1998. Using Zernike Moments for the Illumination and Geometry Invariant Classification of Multispectral Texture. *IEEE Transactions On Image Processing* 7(2): 196-203.

Wang, X., et al. 2005. The Correlation Template Matching Algorithm Based Td Filter and Eso Filter. In *International Conference on Machine Learning and Cybernetics*, 5361-5365.

Watson, A., et al. 2003. Automated Identification of Live Moths (Macrolepidoptera) Using Digital Automated Identification System (Daisy). *System Biodiv.* 1: 287–300.

Weeks, P., et al. 1997. Automating the Identification of Insects: A New Solution to an Old Problem. *Bull. Entomology. Res.* 87: 203–211.

Weeks, P., et al. 1999. Species-Identification of Wasps Using Principal Component Associative Memories. *Image and Vision Computing*. 17 (12): 861-866.

White, R. and L. Winokur 2003. Quantitative Description and Discrimination of Butterfly Wing Patterns Using Moment Invariant Analysis. *Bulletin of Entomological Research* 93 (4): 361-374.

Wilkins, M.F., et al. 1999. Identification of Phytoplankton from Flow Cytometry Data by Using Radial Basis Function Neural Networks. *Appl. Environ. Microbiol.* 65: 4404–4410.

Wolfson, H. 1990. On Curve Matching *IEEE Transaction on Pattern Analysis and Machine Intelligence*. 12 (5): 483-489.

Wong, W.H., et al. 1995. Generation of Moment Invariants and Their Uses for Character Recognition. *Pattern Recognition Letters*. 16: 115-123.

Wu, W.Y. 2001. Two-Dimensional Object Recognition through String Matching. *Imaging Science Journal*. 49 (4): 213-221.

Wu, W.Y. and M.J. Wang 1999. Two-Dimensional Object Recognition through Two-Stage String Matching. *IEEE Transaction on Image Processing*. 8 (7): 978-981.

Yu, D., et al. 1992. Identification of Ichneumonid Wasps Using Image Analysis of Wings. *Systematic Entomology*. 17 (4): 389-395.

Zayas, I. and P. Flinn 1998. Detection of Insects in Bulk Wheat Samples with Machine Vision. *American Society of Agricultural and Biological Engineers*. 41 (3): 883-888.

Zhang, D.S. and G.J. Lu 2002. A Comparative Study on Shape Retrieval Using Fourier

Descriptors with Different Shape Signatures. In *International Conference on Multimedia and Distance Education*, 646-651.

Zhang, D.S. and G.J. Lu 2002. Shape-Based Image Retrieval Using Generic Fourier Descriptors. *Signal Processing: Image Communication*. 17 (10): 825-848.

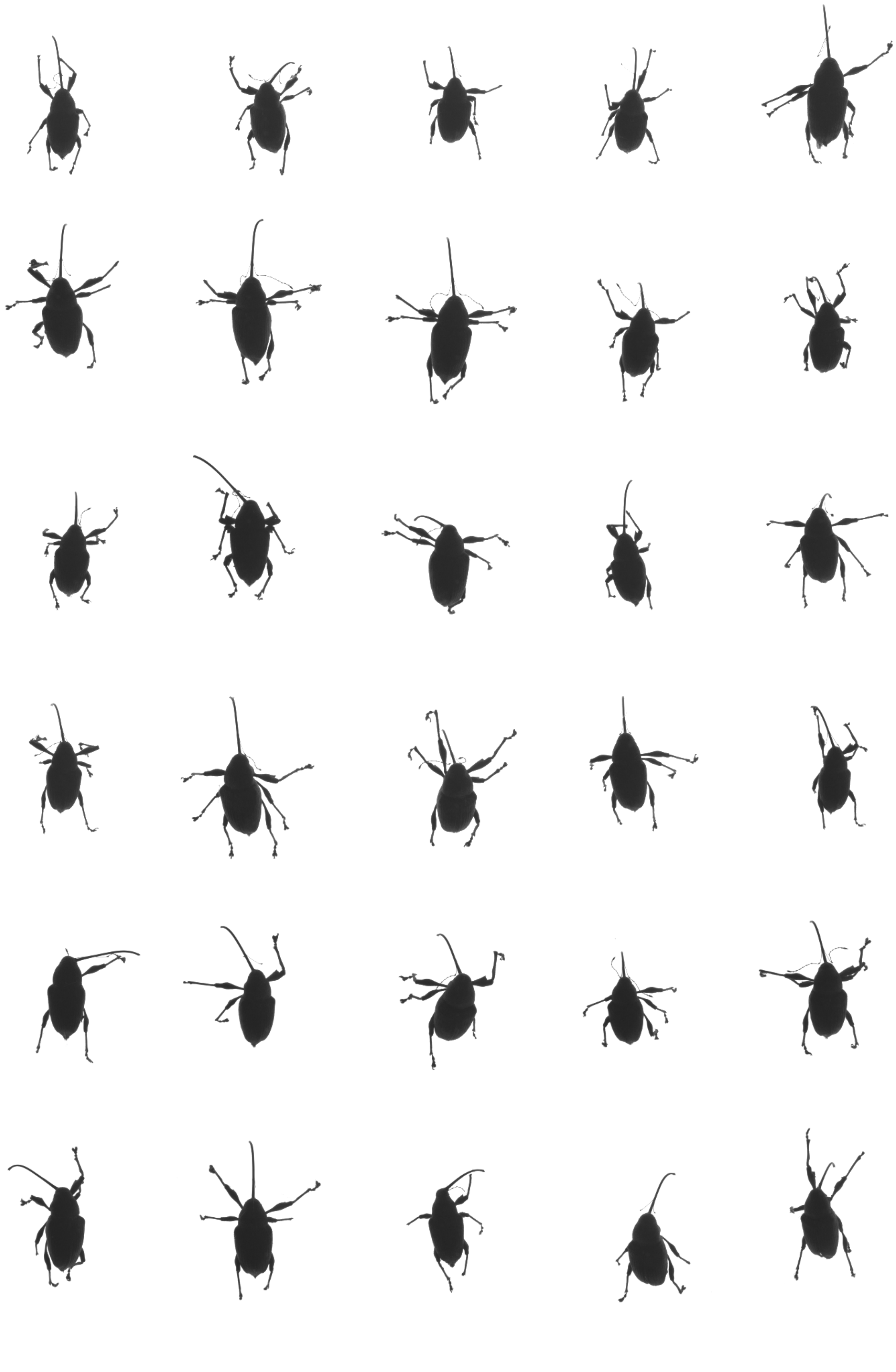
Zhang, D.S. and G.J. Lu 2003. A Comparative Study of Curvature Scale Space and Fourier Descriptors for Shape-Based Image Retrieval. *Journal of Visual Communication & Image Representation*. 14: 41-60.

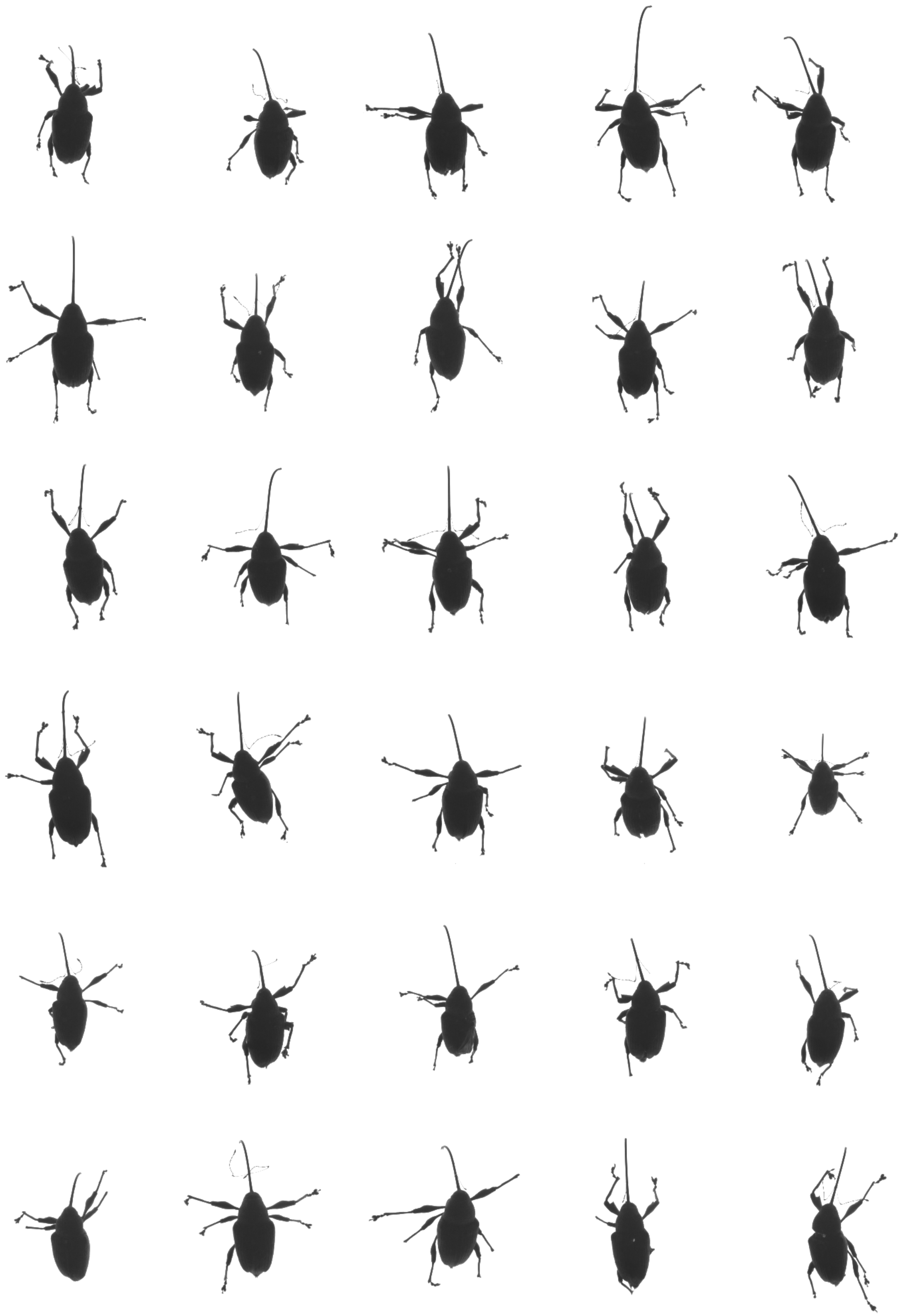
Zhang, D.S. and G.J. Lu 2004. Review of Shape Representation and Description Techniques. *Pattern Recognition*. 37 (1): 1-19.

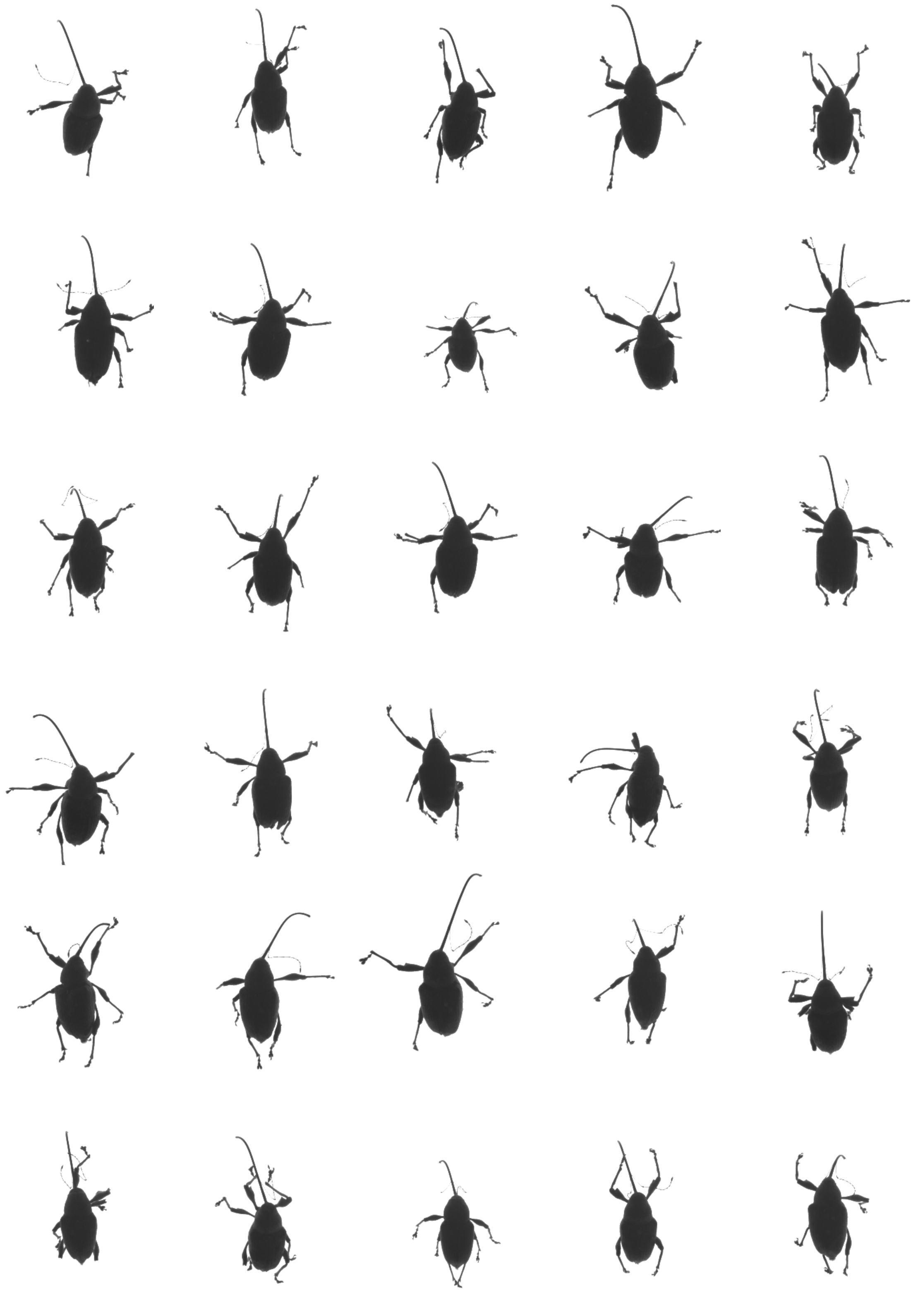
Zhenjiang, M. 2000. Zernike Moment-Based Image Shape Analysis and Its Application. *Pattern Recognition Letters*. 21 (2): 169-177.

# APPENDIX A

The Training Data Set of 205 Pecan Weevils







VITA

Saleh M. Ashaghathra

Candidate for the Degree of

Doctor of Philosophy

Thesis: IDENTIFICATION OF PECAN WEEVILS THROUGH IMAGE PROCESSING

Major Field: Biosystems and Agricultural Engineering

Biographical:

Personal Data: Born in Riyadh, Saudi Arabia on July 31$^{st}$, 1973, the son of Mofleh Ashaghathra and Norah Al-Suliman. Married to Aisha Ashughithry on January 9$^{th}$, 2000. Daughter, Nora born in Al-aflaj, Saudi Arabia on October 25$^{th}$, 2001; Daughter, Jana born in Oklahoma City, Oklahoma, USA on November 17$^{th}$ 2002; Son, Mofleh born in Stillwater, Oklahoma, USA, on August 26$^{th}$ 2007.

Education: Bachelor of Agricultural Engineering from the college of Agriculture, King Saud University, Riyadh, Saudi Arabia, 1996. Received Master of Science in Biosystems and Agricultural Engineering at Oklahoma State University, Stillwater, Oklahoma, in December 2003. Completed the requirements for the Doctor of Philosophy in Biosystems and Agricultural Engineering at Oklahoma State University, Stillwater, Oklahoma in May, 2008.

Experience: Worked as a manager in the machinery operations and maintenance division at the Agricultural Research Station, College of Agriculture, King Saud University, Riyadh, Saudi Arabia, from 1996 to 1999. Taught in the Department of Agricultural Engineering between 1999 and 2000.

Professional Membership: American Society of Biological and Agricultural Engineers

Name: Saleh Ashaghathra Date of Degree: May, 2008

Institution: Oklahoma State University Location: Stillwater, Oklahoma

IDENTIFICATION OF PECAN WEEVILS THROUGH IMAGE PROCESSING

Pages in Study: 135 Candidate for the Degree of Doctor of Philosophy

Major Field: Biosystems and Agricultural Engineering

Scope and Method of Study: The scope of this study is to develop a recognition system that can serve in a wireless imaging network for monitoring pecan weevils. The recognition methods used in this study are based on template matching. Five recognition methods were implemented in this study; namely, Normalized cross-correlation, Fourier descriptors, Zernike moments, String matching, and Regional properties. The training set consisted of 205 pecan weevils and the testing set included 30 randomly selected pecan weevils and 74 other insects which typically exist in pecan habitat.

Findings and Conclusions: It is found that Region-based methods are better in representing and recognizing biological objects such as insects. Moreover, different recognition rates are obtained at different order of Zernike moments. The optimum result among the tested orders of Zernike moments is found to be at order 3. The results also show that using different number of Fourier descriptors may not significantly increase the recognition rate of this method. The most robust and reliable recognition rate is achieved when the two recognition methods, namely, Zernike moments and Region properties are used in a combination. The results indicate that a positive match from either of these two independent tests would yield reliable results; therefore, 100% recognition could be achieved by adopting the proposed algorithm. In addition, the processing time for such recognition is 0.44 sec., on average.

ADVISER'S APPROVAL: Dr. Paul Weckler