A SUPPORT VECTOR MACHINE EMBEDDED WEED IDENTIFICATION SYSTEM

BY

CHUFAN LIN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Agricultural Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Adviser:

Associate Professor Tony E. Grift

ABSTRACT

Over the past decades, the over-reliance on herbicides during corn production has caused severe environmental and biological problems such as pollution in the soil and underground water, and the emergence of the herbicide-resistant weed species. A potential solution to reduce the use of herbicides while maintaining adequate weed control lies in the combined use of chemical and mechanical weeding, in which weeds are controlled adaptively according to their reaction to herbicides. Accurate weed identification is a prerequisite for accomplishing such a control strategy.

A machine vision system for weed identification, which utilized the morphological properties of weed leaves, was developed in this research. The system incorporated a new image segmentation algorithm, termed the 'Pixelwise method' to binarize the color weed images for subsequent image processing and feature extraction procedures. Subsequently, a Support Vector Machine (SVM) based classifier was constructed to distinguish various weed species using seven morphological features.

2,325 indoor images consisting of six weed species were acquired during the first five weeks after emergence of the plants. Among 1,006 test images, the SVM system achieved over 94% accuracy in crop (corn) versus weed discrimination and 95% in grass versus broadleaf weed discrimination. The average classification accuracy for individual weed species was approximately 86%. In addition, the system obtained the best classification result after the second week after plant emergence. In field tests, the SVM classifier based on the indoor image library was able to identify 71.1% of 270 weed plants in the field. With an adaptive median filter to enhance the image quality, the accuracy was raised to 75.9% at the expense of extra image processing time.

Both of the laboratory and field tests showed that the SVM method with reasonable accuracy is feasible for weed identification during their early growth season.

To my family

ACKNOWLEDGMENTS

The fulfillment of this research project would not have been possible without the support of many people. Many thanks to my adviser, Dr. Tony E. Grift, not only for his help in this project, but also for his generous contributions during my stay in the United States. I would like to thank my committee members, Dr. Martin O. Bohn, Dr. Aaron G. Hagar, Dr. Luis F. Rodriguez, and Dr. Lei Tian for their guidance and support. In addition, I would like to thank Heather Lash and Joana Novais for their assistance in setting up the laboratory experiment, and Larry Meyer and David Riecks for their assistance in the field tests. I also thank the C-FAR organization for providing the funds for my graduate study as well as the project. Finally, thanks to my family and Vina, for their continuous encouragements and love during this long process.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Background and Research Motivations

Weed control in corn (*Zea mays* L.) relies heavily on the use of herbicides. From 1999 to 2003, chemical usage surveys (Table 1.1) from several states across the United States show that an average of 94.4% of the corn acres received one or more application of herbicides. The amount of herbicides applied per planted acreage remained at similar levels throughout the survey years (NASS, 2000-2004).

Table 1.1. Herbicide usage in Corn in the United States from 1999-2003.

Calendar Year	Acres Treated		Total Amount Applied (1,000 lbs)	Planted Acreage (million acres)	Application Rate (lbs per acre)	Surveyed States
	Atrazine	Acetochlor				
1999	70%	27%	96,394	68.3	1.41	15
2000	68%	25%	92,371	73.8	1.25	18
2001	75%	26%	102,940	70.7	1.46	19
2002	62%	25%	60,355	51.4	1.17	7
2003	68%	26%	96,531	72.8	1.33	18

The widespread use of herbicides has effectively suppressed weed infestations and promoted agricultural production in the past decades, however, worldwide concerns about the environmental impact of these chemicals have resulted. This research is dedicated to the reduced use of herbicides, by making use of the following properties. First, weeds tend to occur in patches and thus they are less likely to distribute uniformly across the field. Marshall's work (1988) showed that only 20.4% to 72.6% of sampled field areas, depending on the weed species, were infested by weeds. In this context, conventional practice of broadcasting herbicides evenly on an entire field is uneconomical, as areas with no or few weeds would receive the same dose of herbicide as the weed-infested areas.

Secondly, soil-applied herbicides such as atrazine and alachlor, which account for 71.2

percent of corn acres treated on average through 1999 to 2003 (NASS, 2000-2004), pose environmental risks such as soil and groundwater pollution. Increasing public concerns about health risks associated with herbicide residues put great pressure on producers to reduce levels of agrichemicals in streams, rivers and soil.

In addition, the genetic diversity within weed species enables them to cope with pressures imposed by the environment. Thus more and more weed species are found to be herbicide-resistant following repeated treatment with herbicides. For instance, common waterhemp (*Amaranthus rudis*) and tall waterhemp (*A. tuberculatus*) are reported to be the most problematic weed species in Illinois, which is mainly attributed to their high genetic variability, prolific seed production and differential responses to herbicides (Hager et al., 1997).

To address the challenges described above, adaptive weed control with site-specific capacity appears a promising solution. The control method here is defined as one with the ability to perform either mechanical or chemical weeding, dependent upon the weed's response to herbicides and proximity to the weed. In addition, sensors can be implemented such that weeding will not be executed until a weed or patches of weeds are detected, analyzed and classified. In this way, unnecessary herbicide from uniform application can be avoided and in addition, herbicide-resistant weeds can be removed mechanically.

A key requirement for adaptive weed control is accurate weed detection and identification. Many studies have been conducted using machine vision to identify weeds (Dickson et al., 1995; Hemming and Rath, 2001; Meyer et al., 2003; Perez et al., 2000; Woebbecke, 1995). However, most of these researches were conducted under controlled conditions or limited to a specific plant growth stage without considering the change in biological features of weeds as a function of time.

Various forms of classifiers were the primary tools for weed identification, among which the artificial neural network (ANN), Bayesian classifier, decision trees, and discriminant analysis were common in previous researches. In this research, a new method – the support vector machine (SVM) was introduced and tested under various conditions to identify weeds.

The SVM is a novel supervised machine learning method in the field of artificial intelligence (AI), which was developed based on the theory of statistical learning (Vapnik, 1995). In comparison to other popular classifiers such as the ANN, the SVM model is easy to implement. Moreover, through employing a superior structural risk minimization (SRM) principle, SVM achieves better control on the generation error by minimizing an upper bound on the risk function (Gunn et al., 1997). In this way, the common over-fitting problem in ANN models can be prevented (Karimi et al., 2008). Therefore, the SVM method always produces a high accuracy over a wide range of classification cases (Furey et al., 2000; Kim et al., 2002; Tong and Koller, 2002). The hypothesis of this research is that SVM can be implemented with the machine vision technology to identify corn and weed species for the adaptive weed control strategy.

1.2 Outline and Objectives

The overall goal of this research is to develop a methodology to detect and identify inter-row weed plants in the early growth stages of corn. This is to be accomplished with the following activities:

- Develop an automated system for image acquisition;
- Grow weeds in the greenhouse and use an imaging system to construct an image library of the plants in the early growth stages;
- Develop an SVM-embedded, morphology-based machine vision algorithm incorporating a refined image segmentation method for weed identification;
- Conduct laboratory and field experiments to verify the feasibility and efficiency of the proposed method. In addition, through analyzing the identification results, determine the optimal time for weed control in corn production.

1.3 Summary of Results

An SVM based machine vision system, which utilized the geometric characteristics of the plant, was developed. Laboratory and field tests verified that the SVM method with reasonable

accuracy is feasible for weed identification during their early growth stage. A total of 2,325 indoor images consisting of 6 weed species were acquired during the first 5 weeks after emergence of the plants. A new segmentation algorithm – the Pixelwise method – was developed to binarize the raw color images with a correct segmentation rate of 96.3%. Then a SVM classifier was constructed, using seven morphological features extracted from the training set composed of 1,319 images. Among the remaining 1,006 testing images, the SVM system accomplished an overall accuracy of 94.41% in crop (corn) versus weed discrimination, and over 95% in grass versus broadleaf weed discrimination. Among weed species, each weed species achieved a classification rate well above 85%, except for waterhemp, which had an accuracy of 51.6% due to a lack of sample images for the classifier. To investigate the optimal time for inter-row weed identification at the early stage, the corn images were excluded from the test set. The classification accuracy for the remaining 5 weed species as a function of time is shown in Figure 1.1.

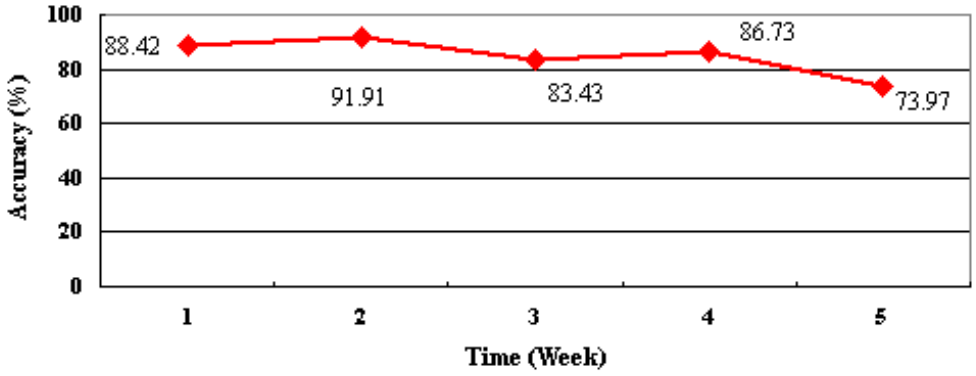


Figure 1.1. Identification accuracy for the undesired weed species at various growth stages during the laboratory experiment.

In the field tests, 107 images including 3 weed species were captured in two separate days during the early growth stage of weeds. Although weed occlusion was not considered in this research, the field images were more complicated due to imperfect segmentation and boundary plants. Therefore, extra image processing procedures were developed to reconstruct

individual plants and eliminate boundary objects before classification. The SVM model, which was based on the indoor weed image library, was able to identify 192 out of the 270 (71.1%) weed plants in all the field images, with a processing time of 1.47s/image. To improve the performance, an adaptive median filter, which reduces blur and distortion in images, was implemented in the Pixelwise algorithm before classification. With the improved images, the SVM system acquired an accuracy of 75.9% at the expense of taking 4.42 second to process each image.

CHAPTER 2

LITERATURE REVIEW

2.1 Weed Interference in Corn

Corn is the major feed grain in the United States, accounting for 94.6 percent of the total feed grain production and use with a planting area of 80 million acres across the country (ERS, 2009). Weeds pose a large threat to corn production, since weeds reduce yields by directly competing against crops for nutrient, moisture, and sunlight. According to a study by Stall (2009), due to weed competition an annual loss of 146 million pounds of fresh market sweet corn and 18.5 million pounds of sweet corn for processing was encountered in the United States from 1975 to 1979. This represents a monetary value of $13,165,000 and $9,155,000, respectively. Thus, weed management is essential in assuring high crop yield and quality.

2.2 Critical Period for Weed Control

Although weed control is crucial in corn production, it is unnecessary and impractical to maintain a weed-free condition during the entire growth season of the crop. This is because if weed control is performed too early when weeds are still small and sparse, late-season weed abundance and competition would result due to their asynchronous emergence characteristic. On the other hand, season-long weed control would be uneconomical in the late crop season when corn grows tall enough to shade and out-compete the weeds. The critical period for weed control (CPWC), defined to be the period in the crop growth cycle during which weeds must be controlled to prevent unacceptable yield losses, is useful for making decisions regarding the need for and timing of weed control (Knezevic et al., 2002).

Numerous researches have been conducted around the world in an attempt to determine CPWC in corn; nevertheless, high variability in the results was found because CPWC is highly subject to weed species and characteristics, weed density, climatic conditions and planting date (Halford et al., 2001; Mahmoodi and Rahimi, 2009; Martin et al., 2001). Those factors were largely attributed to various locations and cultural practices where the studies were carried out.

Hence, the prediction of CPWC should be made site-specifically, taking into account the weeds' generic, environmental and cultural factors.

Related research is also available particularly in Illinois. Williams (2006) used logistic and Gompertz equations to determine the influence of planting date on CPWC in corn. It was found that for corn planted in early May, the CPWC began as early as the V4 and ended at the V8 crop growth stage (CGS) (corresponding to 18 to 31 days after crop emergence (DAE)) to assure yield loss of less than 5%.

2.3 Weed Control Practices

Traditional weed management methods can be categorized into three groups: cultural practices, chemical application and mechanical weeding. Cultural practices include a wide variety of weed management methods such as crop rotation, cover crops, black fallow, planting date, planting density and row spacing optimization (Lyon et al., 1999). However, most of these methods are unable to handle existing or upcoming weeds immediately, or may prohibit continuous cropping (i.e. black fallow) which hampers profitability. As a result, cultural practices are commonly carried out as a pre-treatment prior to chemical or mechanical weeding (Nalewaja, 1999) and are therefore not part of this study.

2.3.1 Chemical Weeding

Since the introduction of synthetic organic chemicals in the late 1940s, U.S. farmers have used herbicides extensively for weed control (Gianessi and Reigner, 2007). The development of selective herbicides and widespread use of glyphosate-resistant crops also contributed to the popularity of chemical weeding in agricultural production. Currently, herbicides are the primary tools to control weeds. It is reported that 87 million ha of cropland receives herbicide treatment and herbicides sprayed for weed control consume up to 60% of the volume and 65% of the expenditures for all pesticides used by U.S. farmers (Donaldson et al, 2002).

The quick adoption of herbicides in the U.S. could be explained by the desire to reduce

weed control costs because labor became deficient and more expensive after World War II. With the help of herbicides, growers in Mississippi were estimated to have saved $10 million per year compared with hiring workers for hand weeding (Gianessi and Reigner, 2007). Besides economic considerations, chemical weeding outperforms other weeding methods in many aspects such as efficiency, and the ability to reduce plant diseases and increase crop yields.

The importance of herbicides is significant in agricultural production. On the other hand, the side effects accompanied with the use of herbicides are not negligible. One major concern about herbicide usage is its potential adverse environmental impacts. Groundwater and surface water pollutions have been reported in many cases during the past decades, and intensive herbicide use was often the major cause (Liu and O'Connell, 2002; Spliid and Koeppen, 1998). Herbicides also pose a potential threat to human health, as there are inevitably pesticide residuals in the crop and water. These concerns have led to legislative directives in several European countries to limit the use of herbicides in agricultural production (Lotz et al., 2002).

2.3.2 Mechanical Weeding

Interest in mechanical weeding has grown steadily over the past two decades, partly because of its reduced impact on the environment. Mechanical weeders range from basic hand tools to sophisticated tractor driven or self-propelled machines. Commonly used mechanisms include duckfoot or "A width" hoes, rotary hoes, rolling cultivators and power take-off (PTO)-driven cultivators (Bowman, 1997). Mechanical weeding is mostly applied in row crops such as sugar beet and corn for inter-row weed control, which is mainly because the spacing among the rows (typically 300 to 700 mm) can prevent the crop plants from being affected by the tools (Mattsson et al., 1990). Many researches have been carried out to assess the efficacy of mechanical weed control methods. Forcella's study (2000) achieved about 50% weed control by purely adopting rotary hoeing in the absence of herbicides. Donald (2007) reported that inter-row mowing systems for controlling both winter annual and summer annual weeds could reduce herbicide inputs by 50%. In addition, cultivation could potentially increase crop yields in that it can reduce tuber greening caused by exposure to sunlight, increase water infiltration

and soil aeration (Bailey et al., 2001).

Although mechanical weeding can help reduce herbicide usage, especially in the scenario of organic farming, it also has potential negative effects both on economy and environment. The disadvantages include potential crop damage, dependence on favorable weather and soil conditions, occasionally high labor requirements, soil erosion and nutrient loss, spread of weed species, and even promoting germination of other weeds while eliminating the existing ones (Belvins et al., 1998; Dallyn, 1971; Hatcher and Melander, 2003). A survey conducted by Napier et al. (2000) showed that only 17% of corn fields were mechanically cultivated in Missouri, which indicated that widespread adoption of mechanical weeding is unlikely.

2.3.3 Adaptive Weeding and Site-Specific Weed Management Strategy

Although herbicides are very effective in controlling weeds, their adverse impacts on environment (pollution) and plant biology (develop of resistance) urge farmers to seek alternative methods of weed management. Meanwhile, mechanical methods are environmental-friendly and efficacious compared to hand weeding, despite decreased weed control consistency, crop yield and economic return if adopted alone (Mount Pleasant et al., 1994). Therefore, the combined use of mechanical and chemical weeding methods is worth investigating, because it has the potential of reducing the over-reliance on herbicides while maintaining satisfactory weed control.

Amador-Ramirez et al. (2001) evaluated weed control and dry bean response to mechanical and chemical treatments. Herbicides were applied exclusively and in combination with rotary hoeing and in-row cultivation. It was found that at low weed densities either mechanical tillage or herbicides spraying were sufficient in suppressing weeds. When weed densities were high however, combined use of herbicide and mechanical weeding methods was required.

Donald et al. (2001) studied the effectiveness of inter-row mowing combined with band-applied herbicide in weed control. In this research, soil residual herbicides such as

atrazine + alachlor were applied shortly before or after the planting of corn. Two or more inter-row mowings were carried out afterwards to control summer annual weeds. The results showed that this weeding method controlled weeds and yielded as well as or better than broadcast application of herbicide at the same rates. Also, the amount of soil-applied herbicides was reduced by 50% as a result of banded treatment in which only 50% of the field area was sprayed.

Other researchers reported that weeds could be controlled by applying lower rates of herbicides in combination with mechanical weeding (Buhler et al., 1995; Mulder and Doll 1993). In fact, the adaptive weeding method, with combined use of mechanical and chemical weeding dependent upon the characteristics of the targeted plant, is an example of the popular integrated pest management (IPM) strategy (Buhler et al., 2000). Both weed control tactics share a common objective, which is to reduce weed density and minimize herbicide input costs without compromising crop yields. A site-specific strategy is an essential approach to realizing this goal, because of the patchy distribution characteristic of weeds in agricultural fields (Cousens and Woolcock, 1997).

Site-specific weed management has been intensively studied and implemented as a herbicide application strategy. Its significance in lowering herbicide use (up to 48%-54%) has been advocated by a number of researchers (Tian et al., 1999; Timmermann et al., 2001). Moreover, site-specific applications also fit well in the framework of Integrated Pest Management (IPM) (Mortensen et al., 1998). This is because the site-specific strategy would only target areas infested by weed patches that would affect crop yield or quality. Once weed patches are located, instead of solely applying herbicides, an adaptive weeding system would activate nozzles or mechanical weeders depending on the weed species and proximity to the crop plants. In this context, the system has the potential of saving chemicals as well as eliminating herbicide-resistant weeds in a mechanical way.

2.4 Weed Identification

An essential part of the adaptive weed control system as discussed previously is the ability to identify weeds in the field in real time. Earlier attempts have focused on simply distinguishing the weeds from the crop and treating everything but the crop as a weed. Due to the emergence of herbicide-resistant species, and the need to decide on treating weeds in a mechanical or chemical way, there is a requirement to identify the weed species in real time. Three types of methods in weed identification can be found in the literature: airborne remote sensing, photo-detector based sensing and machine vision based sensing. Numerous studies have been conducted in the use of these methods; however, few high-accuracy weed identification algorithms or devices have been realized due to the complexity of the field environment, wide variety of species and morphological variation of plants in various growth stages.

2.4.1 Airborne Remote Sensing

Airborne remote sensing (RS) is generally used for locating and identifying weed patches. Sensors mounted on balloons, airplanes, remote control aircraft, and satellites are commonly used for data collection. After the data are processed to create weed maps, decisions can be made regarding where and how much herbicide to apply before the sprayer enters a field.

Conventional color (CC) and color infrared (CIR) photography was the first airborne RS technology utilized to distinguish weeds from agricultural crops (Everitt et al., 1992; Menges et al., 1985). Later, many other tools and techniques were adapted and implemented for RS weed mapping. Everitt et al. (1993) successfully tested the feasibility of using color-infrared photographic, videographic and SPOT satellite images in distinguishing shin oak (*Quercus havardii*) on rangelands. In addition, they concluded that satellite imagery was most useful in mapping large areas of shin oak, while aerial photography and videography were more efficient in small populations of shin oak detection because of better resolution. Lass et al. (1996) used digital images to distinguish yellow starthistle (*Centaurea solstitialis*) and common St.

Johnswort (*Hypericum perforatum*) from other vegetation. Images were obtained from four charge-coupled devices (CCD) with spectral filters mounted in an airplane, with a spatial resolution of 0.5, 1, 2, and 4m respectively. The experiment showed that yellow starthistle and common St. Johnswort were detectable at all those resolutions when their densities were as low as 30% ground cover. Medlin et al. (2000) analyzed the use of multispectral digital images for detecting weed infestation in soybean (*Glycine max*). An aircraft mounted, four-band CCD array camera was utilized for image acquisition, followed by data manipulation using discriminant analysis techniques. It was reported that the proposed remote sensing method had reached at least 75% accuracy in detecting infestations of sicklepod (*Senna obtusifolia*), pitted morningglory (*Ipomoea lacunosa*) and horsenettle (*Solanum carolinense*). Lass et al. (2002) investigated hyperspectral remote sensing with the purpose of detecting spotted knapweed (*Centaurea maculosa*) in Farragut State Park, Idaho. The study used an imaging hyperspectral spectrometer that sampled the reflected solar region of electromagnetic spectrum ranging from 440 to 2543 nm for image recording. The data was classified according to a spectral angle mapper (SAM) algorithm. Field experiments revealed that areas with over 70% spotted knapweed cover were identified; on the other hand, areas with less than 40% spotted knapweed infestation were detected with an overall classification error of 7%.

In airborne remote sensing-based weed detection, differences in spectral reflectance or texture between weeds, crop plants and the background are required. As a result, in most cases, spectral information alone is insufficient for robust crop/weed discrimination due to similarities in weeds and crop reflectance (Zwiggelaar, 1998). In addition, the spatial resolution of the sensor must be high enough, because low spatial resolution will result in spectral mixing, in which plant and soil spectra are combined in the same pixel such that weed patches and crop plants are not discriminable (Brown and Noble, 2005).

2.4.2 Weed Detection Using Photo-Detectors

The physical limitations on infrastructure cost and spectral and spatial resolutions have restricted airborne RS to large-area weed map sensing in most occasions. In contrast, the use of digital cameras or spectral sensors on a ground-based platform to detect weeds could be much more economical, while achieving a higher spatial resolution than airborne or satellite RS. Therefore, ground-based sensing has been extensively studied and photo-detectors were popular tools for early-stage research.

Photo-detectors are non-imaging sensors that distinguish vegetation from a background through calculating the ratios or linear combinations of reflected light in visible and near infrared wavebands. Hooper et al. (1976) designed a photoelectric sensor fitted with a tungsten-halogen light beam to detect plants. The sensor measured the reflected intensities of NIR and visible radiation, which were used for computation of indexes based on spectral band ratios. The ratio for vegetation would be less than that for soil because vegetation absorbed visible red radiation through its chlorophyll. To detect crop plants, a weed-free bed was required because the sensor was unable to discriminate weeds from crop plants. Also, the performance of the sensor was dependent upon sunlight intensity as well as soil reflectance properties. Shropshire et al. (1990) analyzed a "Reflectance Ratio Meter" (RRM) for weed detection. This optical device measured the ratio of NIR red light reflected from a target area, which was converted to a voltage as an indicator of soil or plant. Bargen et al. (1992) developed an optical reflectance sensor that utilized a pair of Red and NIR photo-detectors to detect plants. Reflectance data from five Red (620 nm), NIR (800 or 850 nm) pairs together with the reference reflectance were utilized to compute the normalized difference indices $NDI = (NIR - Red) \big/ (NIR + Red)$, which would provide a strong indication whether plants were present within the field of view of the sensor. Commercial devices such as Detectspray and Weedseeker (NTech Industries, Ukiah, CA) were also available for real-time detection of vegetation patches using the differences in spectral characteristics of plant and background materials. Performance tests on the spray systems built upon these two sensors showed that the

sensor-controlled sprayers reduced herbicides usage from 63% to 85% with the same control effect as conventional system (Hanks and Beck, 1998). However, most of the photo-detector approaches were ineffective in weed/crop and weed/weed discrimination (Felton and McCloy, 1992; Shearer and Jones, 1991). In addition, these devices were sensitive to seedling size, plant density and lighting conditions (Brown and Noble, 2005; Thorp and Tian, 2004). For instance, Wang et al. (2000) tested their spectrometer-based weed detection system under varying field conditions with planting densities of 400, 200, 80, and a single plant per square meter. Their study showed that when the weed density was above 200 plants/m^2, the classifier identified weeds at higher than 70% accuracy. In contrast, the classification rate dropped to below 50% when only a single weed was present against the soil background, with the remaining weeds misclassified as bare soil, which was mainly attributed to the limited spatial resolution of the photo-detectors.

2.4.3 Machine Vision

Machine vision technology has been widely studied and proposed for weed identification. Compared with airborne remote sensing and photo-detectors, machine vision provides sub-centimeter spatial information, as well as spectral and textual information by using high resolution cameras. Image acquisition is accomplished using ground-based camera systems and image processing routines are performed to discriminate weeds against crop and background. In general, three categories of visual characteristics have been used in plant species identification, which include morphology, spectral characteristics, and visual texture.

2.4.3.1 Image Segmentation

The first stage in a typical weed identification procedure is segmentation of vegetation from the soil background. Color based vegetation indices (consisting of Red-Green-Blue or R, G, B components) are frequently considered in this stage because of the fact that vegetation pixels have a strong green component in comparison to background pixels. Woebbecke et al. (1995) developed several color vegetation indices using chromatic coordinates (r, g, b) and

modified the hue component to distinguish living plants from soil and residuals. Experiments showed that the normalized excess green index (*ExG = 2g – r – b*) and modified hue were most efficient in providing a near-binary image outlining the plant region of interest, but modified hue was more computationally expensive. They also found that vegetation indices alone could not discriminate dicotyledon plants from monocotyledon plants consistently. Meyer et al. (1998) defined an excess red vegetative index ($ExR = 1.3R - G$), which was theoretically based on the fact that there are 64% red, compared with 4% blue and 32% green cones in the retina of the human eye, thus the excess red color might facilitate visual perception. However, no further research has proven the validity of this index. Perez et al. (2000) used pixel values in the green and red channel to construct a normalized difference index (*NDI*), which was defined as $NDI = (Green - Red) \big/ (Green + Red)$. Meyer and Neto (2008) proposed an improved vegetation index: Excess Green minus Excess Red (*ExG-ExR*). Segmentation quality tests for both greenhouse and field images of soybean were conducted, in which the accuracy of *ExG-ExR* index was compared to that of *ExG* and *NDI* indices. The contrast experiment results (where quality factor of 1 meant perfect delineation of the region of interest) were shown as follows: for a greenhouse based set, the *ExG-ExR* index had a quality factor of 0.88 ± 0.12 while *ExG* and *NDI* indices had factors of 0.53 ± 0.39 ; for field image sets, both *ExG-ExR* and *ExG* had higher quality factors of 0.88 ± 0.07 while *NDI* had a factor of 0.25 ± 0.08 . In addition, the superiority of the proposed *ExG-ExR* index lied in that sunlight and dry/wet conditions of residuals and soil had little effect on its separation performance. Besides various color vegetation indices derived in the RGB space, other color components or models were introduced. Philipp and Rath (2002) compared six color spaces transformed from the RGB components to optimize the separation of vegetation and background. These models included discriminant analysis (calculation of the probability of each pixel belonging to each group based on a discriminant function), canonical transformation (calculation of the optimal linear combinations of R, G and B to maximize the between-group variance), i1i2i3 (linear transformation of the RGB components by a factor – the covariance matrix of the distribution of the RGB values), HSI (hue, saturation and intensity),

HSV (hue, saturation and value/brightness), and Lab (L: a factor of brightness, a: the content of red or green, and b: the content of yellow or blue). A comparative study regarding the accuracy of segmentation showed that the logarithmic discriminant analysis attained the best classification result with a misclassification rate at about 2% (misclassification of plant pixels: 2.0% and misclassification of soil pixels: 2.1%). However, processing one single image using this color space would take up to 10 min, which was excessively computationally intensive for practical use.

Choosing a color space to process an R-G-B image into a gray-scale sub-binarized image is usually a pre-treatment in a vegetation segmentation procedure. Therefore, an essential step, termed thresholding is needed to binarize the monochrome images. The selection of the thresholding method is often related to the color space or model used to produce the sub-binarized image, because different color spaces result in different image histograms delineating the vegetation and background regions. Lee et al. (1999) utilized a color look-up table (LUT) based on a Bayesian decision rule in the HSI space to segment plant and non-plant regions in an image. However, outdoor performance results for the system showed that 24.2% of tomatoes (*Lycopersicon esculentum*) were misclassified, while only 47.6% of the weeds were recognized. Astrand and Baerveldt (2002) developed a plant perception module to be used on an autonomous mobile robot for mechanical weed control. The color vision system used the normalized green component to obtain a corresponding gray-level image. After that, Otsu's method (Otsu, 1979) in which iterations were used in search of the best parameter to separate classes based on variance, was adopted to find the proper threshold to separate plants from the background in a greenhouse environment. Variable outdoor lighting conditions pose a great challenge in plant segmentation, because direct sunlight causes substantial intensity differences in the form of shadows and highlights within the images (Steward and Tian, 1999). To overcome the influence of unpredictable lighting condition in an outdoor field, Tian and Slaughter (1998) proposed an environmentally adaptive segmentation algorithm (EASA) with automatic look-up table (LUT) generation capacity for vegetation segmentation under natural

illumination. Normalized coordinates were used for image clustering, with the purpose of emphasizing the color of the object. Field experiments indicated that despite a great improvement compared with static segmentation techniques, the EASA algorithm only recognized 45%-66% of all tomato seedlings.

2.4.3.2 Machine Vision Based Weed Identification

After the vegetation is segmented from the background, weeds are to be distinguished from the crop using of various visual plant characteristics. For this purpose, spectral reflectance has been applied. Franz et al. (1991) investigated the use of broadband reflectance to identify soybean, ivyleaf morningglory, velvetleaf and foxtail. Spectral features such as skewness in the red waveband, and the mean and variance in the NIR and blue bands were selected as the optimal set of features for discriminant analysis. Results of the greenhouse experiment showed that when leaf orientation was controlled, the classifier recognized over 93% of the 48 observations. Feyaerts and van Gool (2001) collected multi-spectral images of soybean and five weed species under field conditions. A normalized ratio derived from the NIR and Red wavelength, together with the use of a multi-layer neural network with nonlinear mapping (MLNLM) classifier gave the best identification accuracy, where 80% of the soybean and 91% of the weeds were correctly identified. Vrindts et al. (2002) used a reflectance-measurement based machine vision system to discriminate sugarbeet (*Beta vulgaris*), maize and seven weed species. In laboratory tests, reflectance spectra in the 400 to 2,000 nm wavelength range were recorded under controlled lighting, and a limited number of wavelength band ratios were utilized for classification. With only 3 wavelength ratios, the system achieved over 99% and 98% classification rates for maize/weed and sugarbeet/weed combinations, respectively. Spectra in the 480 to 820 nm range were used for field measurements under natural sunlight, and less than 10% of the crop and weeds were misclassified.

Other than the reflectance properties, a few studies have investigated the use of plant textural information for weed identification. Tang et al. (1999) developed a Gabor wavelet based feature extraction and neural network-based pattern identification system to discriminate

between broadleaf and grass species. The system achieved 100% classification accuracy over 40 sample images. In addition, the algorithm was computationally efficient giving it the potential for real-time application. Burks et al. (2000) utilized the color co-occurrence matrices (CCM) to calculate textural information from soil and five weed species including giant foxtail, crabgrass, velvetleaf, lambsquarters, and ivyleaf morningglory. The system had a classification accuracy of 93% when using 11 texture features in the hue and saturation color spaces, and the computational load was reduced up to one third since the intensity statistic was not included.

An alternative method for weed identification involved the use of morphologic features such as the leaf shapes. Guyer et al. (1986) attempted shape-based machine vision technology to identify plant seedlings. The four features used in their study being complexity, elongatedness, central moment and principle axis moment, were derived from grayscale images of eight plant species. Subsequent studies introduced many more shape features, such as aspect ratio, roundness, circularity, convexity and ratios among length, width, and perimeter dimensions (Guyer et al., 1993; Woebbecke et al., 1995). Recent articles have explored various forms of classifiers or techniques for shape-based weed identification. Cho et al. (2002) evaluated discriminant analysis and artificial neural networks (ANN) in identifying radish from weeds. Among the eight shape features extracted from the plants, aspect ratio, elongation and perimeter to broadness were selected as the significant set for the discriminant analysis model, which showed an identification rate of 92% for the radish and 98% for the weeds. In contrast, the ANN model efficiently identified radish from the weeds with 100% accuracy. Neto et al. (2006) developed an Elliptic Fourier (EF) method in weed identification, based on leaf shape. They found that the EF method combined with principle component analysis and linear discriminant models achieved the best classification results during the third week after the plants' emergence. Here, 77.9% of redroot pigweed (*Amaranthus retroflexus*), 93.8% of sunflower (*Helianthus pumilus*), 89.4% of velvetleaf (*Abutilon theophrasti Medicus*) and 96.5 % of soybean (*Glycine max*) were correctly classified. By combining the leaf information from the second and the third weeks, classification rates for the corresponding plants reduced to

76.4%, 93.6%, 81.6%, 91.5% and 90.9%, respectively. Sogaard (2005) reported a new shape-based machine vision method for classification of 19 weed species. In this work, Active Shape Models (ASM) were generated using a set of points on the boundaries of the leaves for each species, and weed identification was performed by fitting weed images to training sets. Three weed species were tested, each with 100 samples, and the system identified 77% of shepherd's purse (*Capsella bursa-pastoris*), 65% of scentless mayweed (*Tripleurospermum inodorum*) and 93% of charlock (*Sinapis arvensis*). However, the algorithm could only work with weeds within the two-leaf growth stage and required weed images without mutual overlapping.

As mentioned previously, a large proportion of the past plant-shape-based studies that achieved high identification rates were conducted under controlled conditions where the shape of the entire leaf was well displayed. Leaf occlusion, changes of leaf size and shape as a function of growth stage, and leaf orientation relative to the camera still pose the most challenging issues in implementation of the real-time shape-based machine vision technology for weed identification (Thorp and Tian, 2004).

2.5 SVM and Shape-based Machine Vision Weed Identification

Previous researches on the weed identification problem have involved a large number of statistical methods and classifiers. The artificial neural networks (Burks et al., 2005; Tang et al., 2003), bayesian classifier (Marchant and Onyango, 2003; Tian and Slaughter, 1998), decision trees (Goel et al., 2003; Yang et al., 2004), and discriminant analysis (Burks et al., 2000; Cho et al., 2002) are widely studied and implemented with varying success. As an alternative, in this project, a novel approach in Artificial Intelligence – the support vector machines (SVM) is proposed to classify weed species.

SVM is a supervised machine learning method, which was developed in the 1990s based upon the theory of statistical learning (Vapnik, 1995). Essentially, SVM is a binary classifier, which searches for the optimal separating hyperplane that maximizes the margin between two (or more) classes and minimizes the generalization errors. The SVM is well-known for its

robust performance in the presence of sparse and noisy data (Furey et al., 2000), as well as user-friendless (Karimi et al., 2005). Therefore, this technology has been intensively studied for the last few years and applied to a wide range of classification problems, such as text classification (Tong and Koller, 2002), speaker identification (Schmidt and Gish, 1996; Wan and Campbell, 2000), face detection (Kim et al., 2002), tissue classification (Furey et al., 2000), object recognition (Blanz et al., 1996), as well as corn root classification (Zhong et al., 2009). In most of these researches, SVM either substantially outperformed or at least matched the comparable methods. Although SVM is receiving increasing popularity due to its promising classification accuracy, its application in agriculture has not been fully explored (Karimi et al., 2008).

Wu and Wen (2009) investigated the use of SVM in classifying corn seedlings against four weed species at the early growth stage. Ten texture features of the plant were extracted based on the Gray Level Co-occurrence matrix (GLCM) and the histogram distribution from the gray level images. Subsequently, four combinations of these features were selected by the principle component analysis (PCA) as input vectors of the SVM classifier. The SVM classifiers with various feature selections maintained a high classification accuracy ranging from 92.31% to 100%. In comparison, the back-propagation (BP) neural-network model could only achieve 80% accuracy on the same image set. However, this study merely focused on weed-corn discrimination, while the more complicated identification scenarios among weed species were not considered.

Karimi et al. (2005) evaluated the capability of the SVM to analyze hyper-spectral images for identifying weed and nitrogen stresses in corn. Images were collected using an airborne spectrographic imager with 72 wavebands ranging from 408.73 to 947.07 nm during early growth stage. Four weed treatments (no weed control, control of grass, control of broadleaf, and full weed control), as well as three nitrogen application rates (low nitrogen (60 kg N/ha), normal nitrogen (120 kg N/ha), and high nitrogen (250 kg N/ha)) were used separately or in combination, to evaluate the SVM classifier. The results showed that the

support vector machines method obtained 69% accuracy for combined weed and nitrogen application factors: 86% and 81% classification accuracy was achieved when weed and nitrogen treatments were investigated separately. In addition, the SVM model outperformed the competing artificial neural network (ANN) method in each of the three classification categories, where ANN achieved 58.3%, 81.2% and 69.4% accuracy, respectively.

In summary, SVM is an artificial-intelligence method for data mining, which has demonstrated superior classification performance over traditional models in the field of weed-corn classification research. However, more in-depth application of the technology in weed identification among different species is still rare. In addition, most of the previous studies only evaluated the classification models at one or a few random dates, but not on a consistent basis throughout the early crop season. As leaf shapes and other biological features of plants change at different growth stages, identification of weed species against time would be predictably more difficult. Textural and spectral features have also been successfully used for weed classification (Karimi et al., 2005; Wu and Wen, 2009). Whether the third visual characteristic of plant, being morphological features, could improve the SVM classifier to achieve accurate identification performance is worthy of investigation.

CHAPTER 3

METHODS AND MATERIALS

To develop a novel SVM-embedded machine vision system for weed identification, the project was divided into two stages. Firstly, laboratory experiments were conducted from which the identification algorithm was derived and tested using plant samples grown in a greenhouse. Secondly, field trials were conducted, in which outdoor plant images were utilized to assess the feasibility and efficiency of the proposed algorithm. Each stage included procedures such as image acquisition, image processing, and object identification, which will be described in detail in the following sections.

3.1 Laboratory Experiments

Six of the most common weed species in Illinois were selected for this study (Figure 3.1), which included three broadleaf weeds – common lambsquarters (*Chenopodium album*), velvetleaf (*Abutilon theophrasti*) and waterhemp (*Amaranthus rudis*), and three grass weeds – barnyardgrass (*Echinochloa crus-galli*), large crabgrass (*Digitaria sanguinalis*), and corn (*Zea mays*). Since the experiment was conducted in the late spring, this stage of the study was conducted in a greenhouse environment for better control of temperature, humidity, and lighting.

Figure 3.1. Common weed species in Illinois.
Top: common lambsquarters, velvetleaf, and waterhemp;
Bottom: barnyardgrass, large crabgrass, and corn, respectively.

3.1.1 Greenhouse Treatment

All six weeds species were planted in the Turner Hall greenhouse at the University of Illinois, Urbana-Champaign, using Com-Packs Bedding Plant Containers (Hummert International, East City, MO, USA) and a custom 1:1:1 (Soil: peat: torpedo sand) soil mix. Each container consisted of 18 cells containing a single weed species, with a single plant in each cell. Every weed plant was moved to approximately the center of each cell for improved image recording purposes, while excess plants in any cell were eliminated. Room temperature was maintained around 27 degrees Celsius during the growing period, and supplemental lighting, pest control as well as fertilization were utilized to promote vigorous plant growth.

3.1.2 Image Acquisition

Considering the threat of a high humidity environment to electronic equipment in the greenhouse, the images used for this study were collected in Lab 126 of the Agricultural Engineering Science Building, at the University of Illinois. To achieve a high accuracy and efficiency, an automated image acquisition system was devised, which consisted of a digital camera, power supply, computer, control module, and infrastructure module (Figure 3.2).

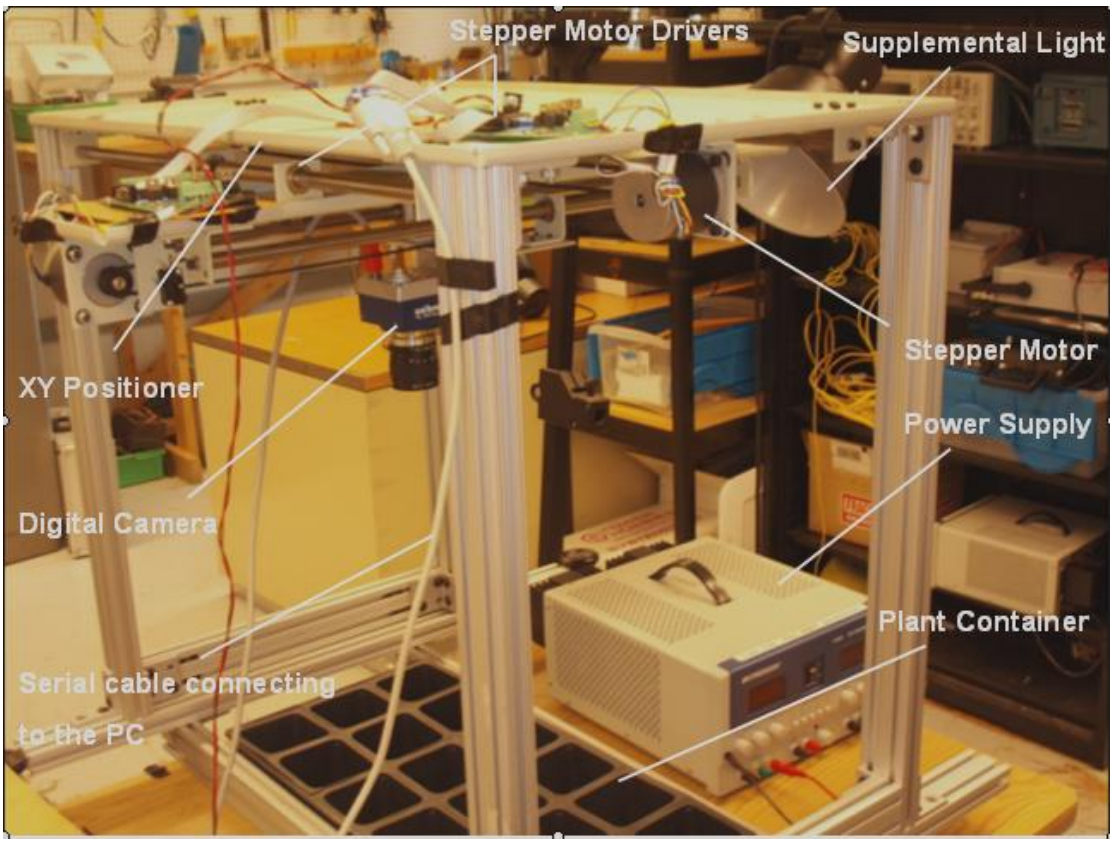


Figure 3.2. Automated image acquisition system devised for imaging plants grown in the greenhouse.

The image acquisition system was based on a dual-axis belt-driven linear positioning table (XY-18, Arrick Robotics, Tyler, TX, USA). The positioning table is commonly used for positioning sensors and performing pick-and-place robotic operations, which was accomplished by activating stepper motors on the X and Y directions, driving a top plate attached to the drive belts. For the purpose of analyzing leaf shape features, overhead images of plants were desired. Therefore, the XY positioner was placed upside-down. To accommodate the weight of the camera module on the positioner, the original motors were replaced by two 12VDC unipolar stepper motors (Jameco Electronics, Belmont, CA, USA), which have a holding torque of 6 kg-cm and a detent torque of 725 g-cm (Figure 3.3a). A Fire-i™ 701c FireWire industrial camera (Unibrain Inc., San Ramon, CA, USA), coupled with a C mount 6mm F1.2 lens (Pentax Co., Golden, CO, USA) constituted the camera module for weed imaging. The camera features

a $1/2"$ interline CCD solid-state image sensor (ICX205AK, Sony Co. LTD, Tokyo, Japan, cell size: $0.465mm \times 0.465mm$), which provides an image size ranging from 320×240 to 1388×1036 pixels. The lens has a focal length of 6 mm, and has a horizontal view angle of 57 degrees. A diaphragm and focal ring allowed manual adjustment of the aperture and focal distance of the lens. The camera module, facing downwards, was mounted on the top plate of the XY positioning table through an anchor plate (Figure 3.3b). Two serial bipolar stepper motor drivers (KTA-5197A, QKits Limited, Kingston, ON, Canada) were used to control the stepper motors on the X and Y axis, through enabling or disabling the motor power output pins on the drivers. The two drivers were linked together using a 2×5 (10-pin) IDC connector cable (MicroController Pros Corp., Reno, NV, USA). The driver boards were connected to a computer (Pavilion a6110n, AMD Athlon™ 64×2 Dual Core Processor 4400+ 2.31 GHz, 2.0 GB of RAM, 320 GB hard drive, Hewlett-Packard, Palo Alto, CA, USA) with a 9-pin straight-through serial cable and an IDC connector cable (Figure 3.3c). A control strategy was established by sending serial signals from the computer to the motor drivers to enable/disable the stepper motors. The camera was controlled through a FireWire cable from a single MATLAB® (The MathWorks, Natick, MA, USA) program. The main purpose of this program was to establish serial communication between the computer and the motor drivers, send commands to move the top plate to designated positions, trigger the camera and save the images (Appendix A1).

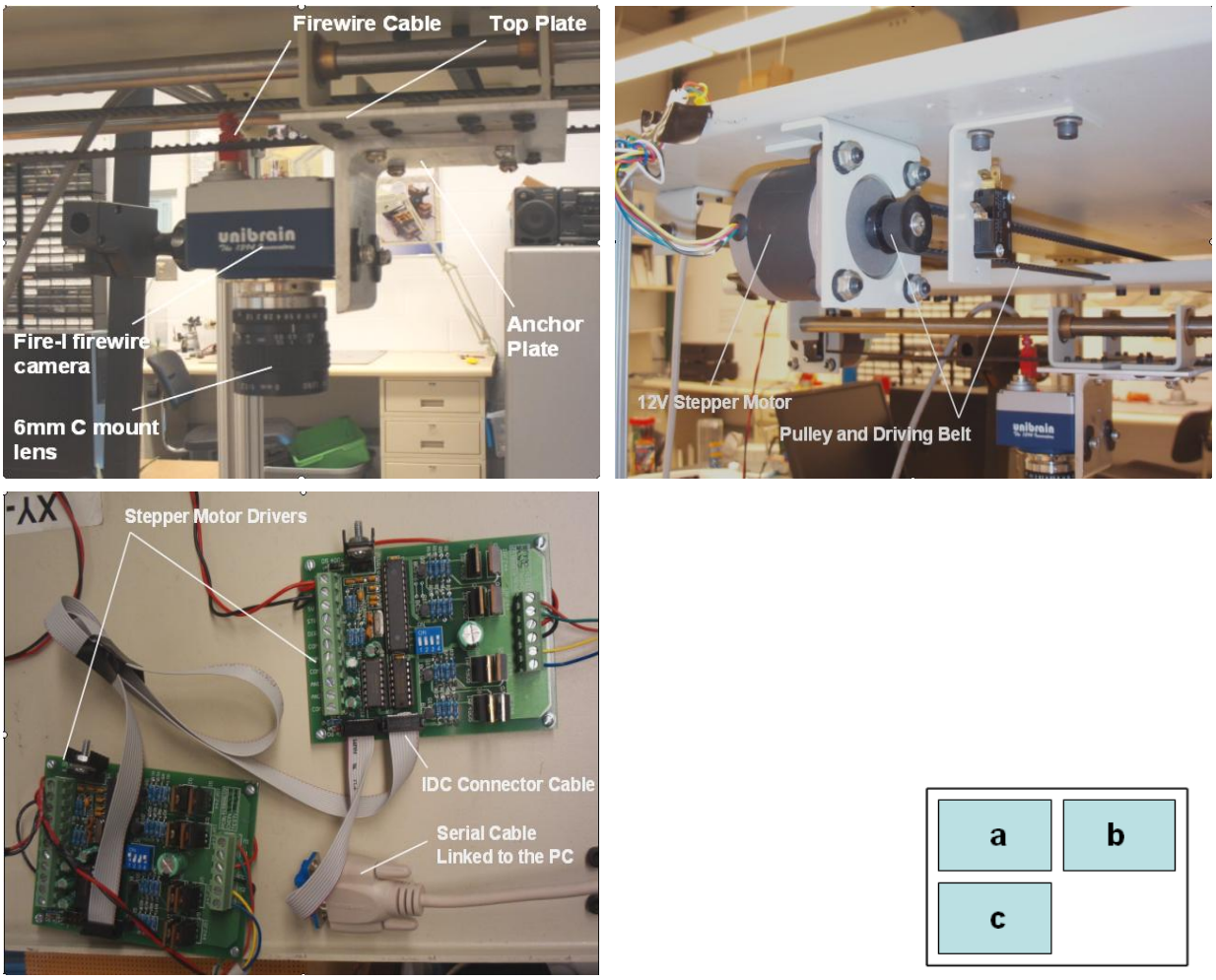


Figure 3.3. Core modules of the automated image acquisition system. From top to bottom:
(a) Stepper motor drives the top plate through the pulley and driving belt;
(b) Camera module mounted on the top plate through the anchor plate;
(c) Linked stepper motor drivers connecting to the PC with a serial cable.

Besides the core modules described above, other components were added to the system, such as lights for enhancing light intensity in the lab environment, and a 12V DC power supply (1760A, B&K Precision Corp., Yorba Linda, CA, USA) for the stepper motors drivers. Before the image collection procedure, several problems had to be considered:

1. Positioning calibration. In this study, the weed container was placed at a fixed position, and the camera mounted on the top plate was required to access all 18 cells in each container. Thus, the spinning rates of the motors were determined *a priori*, such that the motors would transport the camera to cover each cell in sequence. In addition, due to the

high accuracy of the low-stretch timing belt ($\pm 0.83 mm/m$) and homing procedure performed after imaging each container of plant, the repeatability of the system was high and only one calibration was necessary.

2. Camera calibration. The distance between the camera and the top of the container was approximately 0.37m (19.5") and the zoom factor of the lens was set to 0.4. Thus using an image size of 640×480 pixels, the image could cover an area of $0.21m \times 0.155m$ ($8.4" \times 6.2"$), which was larger than the size of a cell ($0.076m \times 0.076m$). According to the movement pattern and calibration of the positioning system, the camera triggered when centered above each cell. Therefore, the position of a targeted cell in each image was identical (if the positioning calibration was accurate enough), and by cropping each image with the same frame size (281×261 pixels), images containing single plants were obtained.

After the calibration procedures, the automated image acquisition system was established. The layout of the system is shown in Figure 3.4 and the objectives of the system were accomplished using the following substeps.

1. The MatLab® program first initiated the camera channel and serial communication between the computer and the motor drivers, which were powered by the 12VDC power supply.
2. Control commands were sent through the serial port to the motor drivers to move the camera module to the pre-determined position above a cell.
3. Upon arrival of the camera module, the stepper motors were deactivated and held in place. Meanwhile, the camera was triggered through the FireWire cable, resulting in a YCbCr- format 640×480 image. This image was then cropped and converted to an RGB image containing a single cell and plant. [4] Steps [2] and [3] were repeated such that the system handled the remaining cells and at the end, a homing command was sent to direct the camera back to the pre-set starting point. At this point, the imaging of a container was completed.

The planting date was April 2nd, 2009 for all weed species, and daily image collections started from April 7th, 2009 as soon as any weed plant emerged from the soil. The greenhouse experiment lasted approximately 5 weeks, corresponding to the VE to V7 growth stages of corn. A total of 2,325 plant images were collected during this period.

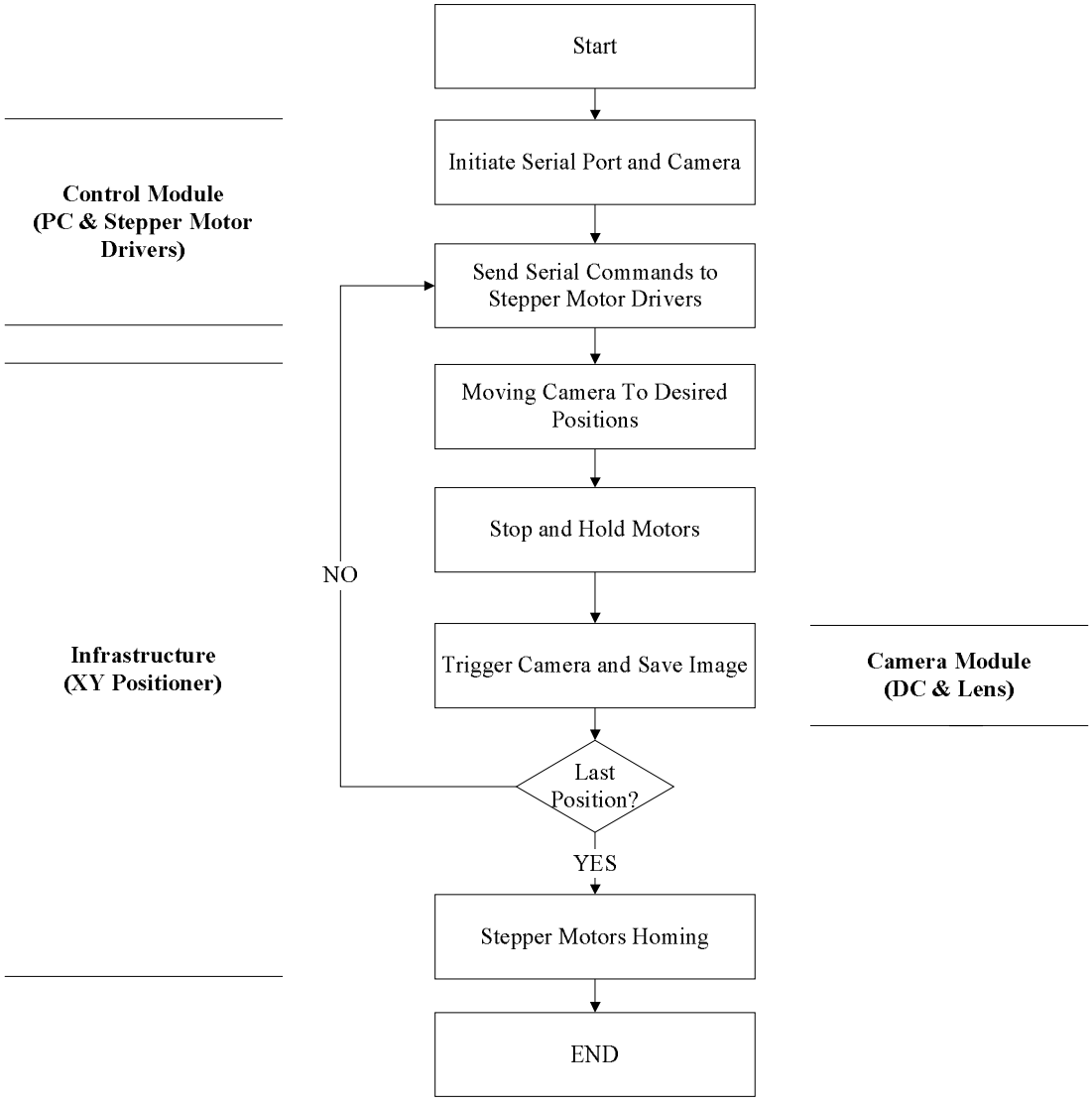


Figure 3.4. Flow chart of the automated image acquisition system.

3.1.3 Image segmentation

After the image acquisition procedure, the next step was image segmentation. This step is used to produce binary images from the acquired RGB images with the goal of separating the plant-related pixels from background-related pixels. This is a critical step in morphology-based weed identification systems, because high-quality images provide important leaf shape information for the feature extraction and plant classification procedures (Meyer and Neto, 2008). Numerous studies were conducted in the attempt of delineating the plant region from non-plant background using various color spaces. Normalized Excess Green and Modified Hue were frequently reported (Jafari et al., 2006; Meyer et al., 1998; Tang et al., 2003; Woebbecke et al., 1995) as the superior methods due to their low sensitivity to background noise and lighting conditions. In this section, the fundamentals and limitations of these two methods will be discussed, and a new Pixelwise plant segmentation method is proposed and compared to the Normalized Excess Green and Modified Hue methods.

3.1.3.1 Normalized Excess Green (NExG)

The derivation of NExG index used the RGB color space. However, as non-normalized RGB coordinates were sensitive to illumination (Woebbecke et al., 1995), a better way to define NExG was through the use of chromatic coordinates (or chromaticities):

$$NExG = 2 \times g - r - b \tag{3-1}$$

Where *r*, *g*, and *b* are the chromaticities obtained from the transformations:

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B}, \quad b = \frac{B}{R+G+B} \tag{3-2}$$

Where *R*, *G*, and *B* are the non-normalized red, green, and blue channel pixel intensities, respectively, with the constraint: if $R+G+B=0$, then $NExG=0$.

After the chromaticity conversion, the original RGB image was converted to a grey-scale image with the plant region being highlighted. To binarize the grey-scale image, a threshold value, which maximized the variance between the plant group and non-plant group pixels, was chosen using OTSU's method (Gonzalez et al., 2004). An alternative thresholding method

involved visual examination of the NExG histogram, where the “valley” position would be selected manually as the threshold value for plant segmentation.

A program was written in MatLab® to implement the NExG method with both automatic and manual thresholding (Appendix A2). Computer trials with the weed images taken during the laboratory image acquisition procedure showed the following: In general, NExG was capable of handling shadows or dark parts on the weed leaves, while high-intensity spots remained problematic. This was because the plant pixels on those reflecting spots had almost identical R-G-B values ($R \approx G \approx B$ for white color), which would be grouped to the background pixels after thresholding ($2g - r - b \approx 0$). In addition, NExG could be erroneous in processing images with a low green plant pixel rate and a high background pixel rate, because in that case, the histogram of NExG no longer had the characteristic “valley” which separates the plant from the background. This made it difficult to choose a threshold value in either thresholding method (Figure 3.5). To summarize, NExG with automatic OTSU was efficient in outlining the approximate shapes of the plant leaves at the expense of losing some details such as small stems. In comparison, manual histogram thresholding achieved better segmentation results with most of the weed images, but it is highly undesirable for real-time weed identification.

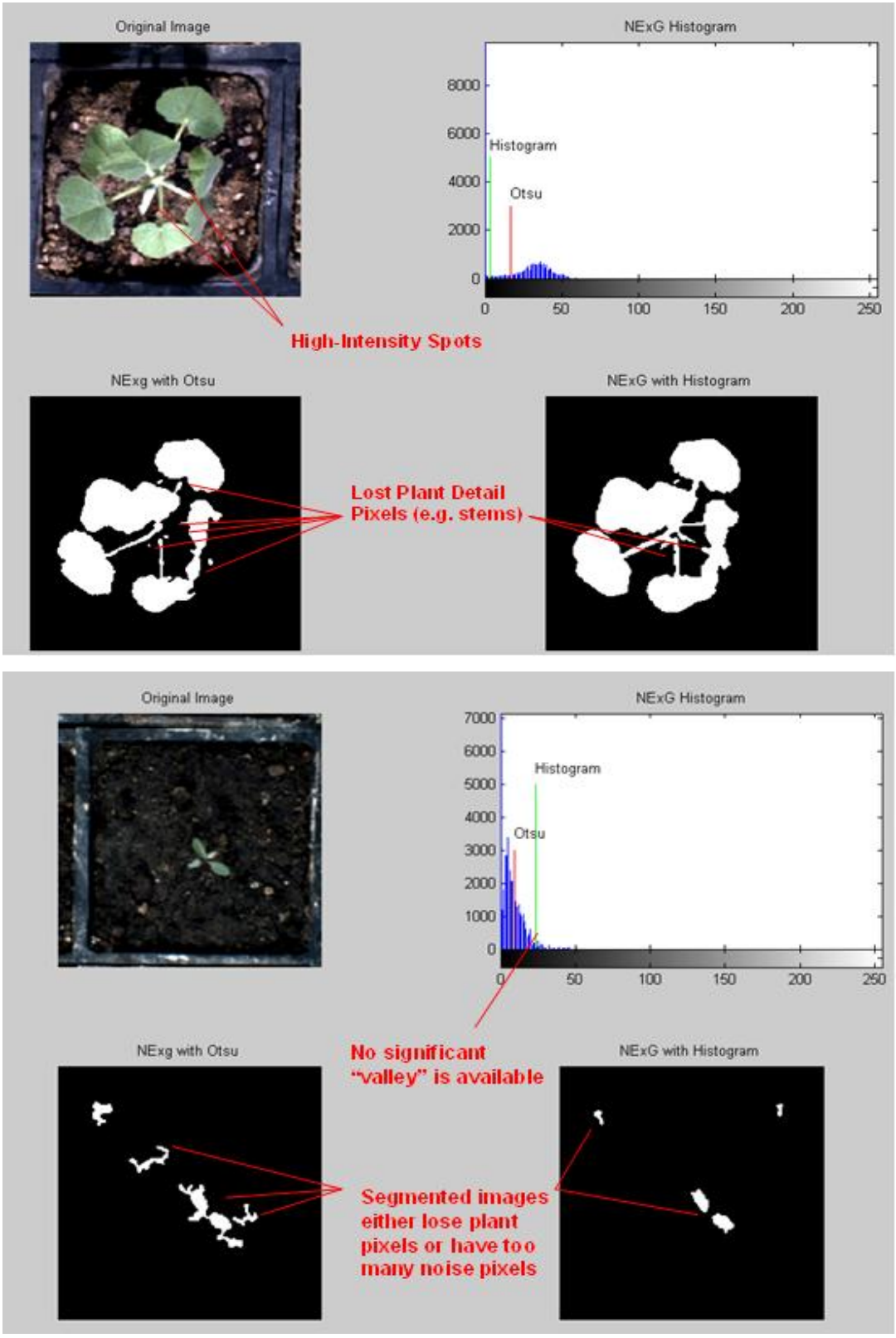


Figure 3.5. Comparison of NExG algorithm with automatic (Otsu)/manual (Histogram) thresholding methods in processing images of weed at various growth stages.

3.1.3.2 Modified Hue

HSI (hue, saturation, intensity) is a color system that describes color as points in a cylinder (Figure 3.6). This color system is commonly considered more appropriate to describe color than the RGB system as far as the way human perceive and interpret color sensations is concerned (Gonzalez et al., 2004). Of the three components that constitute the HSI space, hue is regarded as a key component because it represents the dominant wavelength in mixed light waves (Tang et al., 2000) and it is not subject to highlights and shadows (Cheng et al., 2001). In general, hue is defined to be an angle between a reference color line and the selected color point ranging from 0 to 360 degrees, which can be converted from the RGB coordinates as:

$$H = \theta \qquad \text{if} \quad B \le G$$

Or

$$H = 360 - \theta \qquad \text{if} \quad B > G \tag{3-3}$$

Where

$$\theta = \cos^{-1}\left(\frac{2R - G - B}{2[(R-G)^2 + (R-B)(G-B)]^{1/2}}\right) \tag{3-4}$$

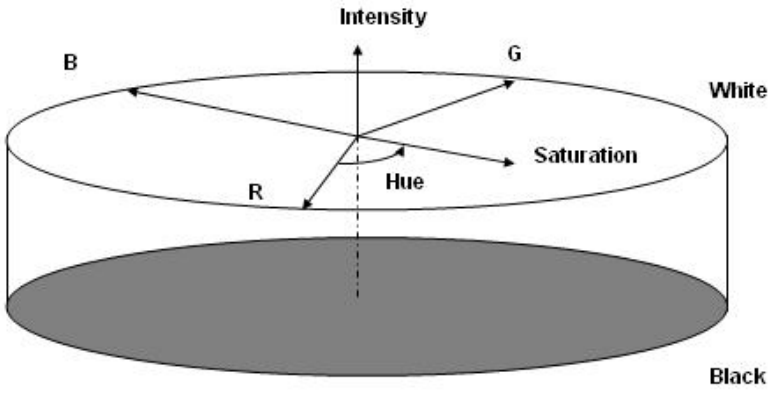


Figure 3.6. The HSI color cylinder represented by the three components, which includes Hue, Saturation, and Intensity.

The Hue value can be further normalized to the range of [0, 1] through dividing H by 360^{o}. Nevertheless, non-removable singularities exist along the intensity axis (R = G = B) of the HSI color cylinder, which makes the RGB-to-HSI transformation sensitive to subtle changes in input values and causes discontinuities in the representation of colors (Cheng et al., 2001). Thus in the modified hue algorithm the hue value was forced to 0 at the singularity, where $Saturation = 1 - \frac{3 \times \min(R,G,B)}{R+G+B} = 0$.

A MatLab® program was written encoding the Modified Hue algorithm (Appendix A3). To simplify the thresholding method, a versatile range in the histogram of hue channel specifying the "plant greenness" was to be determined. 50 sample images were randomly selected from the weed image collection and the distribution of hue component for each sample image was inspected. With hue and saturation being normalized to 0-255 and hue being 0 at the singularity, an empirical hue range at [65, 120] for plant segmentation was concluded through histogram segmentation based on manual selection of thresholds. Computer trials were conducted using the same set of weed images as used in the NExG test. The segmentation results from adopting the empirical range of hue, together with that from manually adjusting the thresholds showed that modified hue with empirical thresholding worked only partially satisfactorily. In most cases, one or both threshold values had to be manually adjusted. In addition, high intensity leaf and background pixels as well as dark leaf pixels caused segmentation errors (Figure 3.7). This is because bright background pixels, although 'white' in RGB space, can contain considerable greenness, which make it possible for such pixels to fall in the selected 'plant greenness' range. On the other hand, shaded leaf pixels, whose hue values might be out of the range due to poor intensity of light reflection, would be erroneously classified as the background.

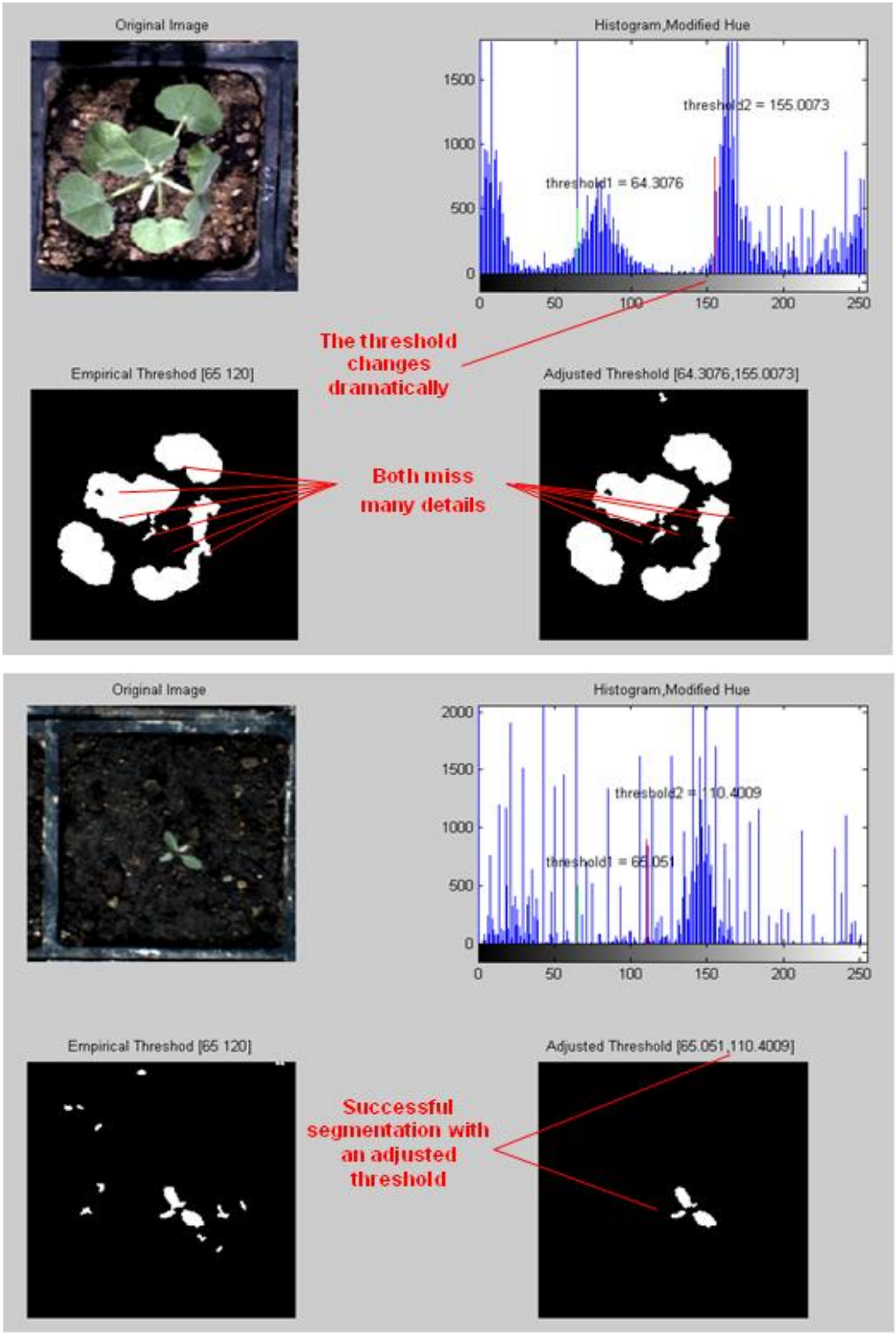


Figure 3.7. Examples of using the Modified Hue method with empirical thresholding and manually-adjusted thresholding to segment weed images at various growth stages.

3.1.3.3 Pixelwise Segmentation Method

Neither the NExG nor the Modified hue algorithm with global thresholding was able to perform high quality image segmentation consistently. This is because global thresholding, where the complete image is used to determine the threshold value, is prone to fail when the background illumination is uneven (Gonzalez et al., 2004). Another essential problem lies in the fact that both the NExG and the Modified hue algorithm reduce a three-dimensional matrix (RGB) to a one-dimensional vector. This results in loss of spatial and color information, which can be crucial in the plant segmentation procedure. For instance, background pixels being highlighted were frequently misclassified as plant related pixels, because both of the algorithms relied on 'greenness' as the only measure. These segmentation errors are difficult to remedy even with manual thresholding, due to the similar chromatic property of the noise and the plant pixels. To alleviate this problem, the interrelationship among the R-G-B channels of each pixel belonging to the plant and non-plant region was investigated. The aim was to find possible complementary classification criteria to delineate weeds from the soil and residual background.

44 sample images at four imaging dates were chosen from the weed database for segmentation analysis. Among those images, 10 images were selected from date 04/11, 10 from date 05/01, 12 from date 04/15, and 12 from date 04/23. The first two groups included five weed species, and the remaining groups included six weed species, depending on availability of weed images, which accounted for various germination times among the weed species. The sample images were first segmented using the NExG index with OTSU's automatic thresholding method. Then manual adjustment on the threshold was applied based on the histogram until the optimal segmentation result (by visual judgment) was achieved (Appendix A4). At this point, plant regions were delineated with all '1's and background with '0's in the binarized images, so were the corresponding regions in the original color images. The means and ratios of the R, G, and B components were calculated for pixels within the plant and non-plant regions as shown in Table 3.1.

Table 3.1. Comparison of the means and ratios of the RGB channels between plant pixels and background pixels.

(a) Statistical summary

Type	Date	R	G	B	R/B	G/B	G/R
Plant	04/11	116.13	150.49	116.13	1.02	1.32	1.30
	04/15	116.66	150.10	116.66	1.03	1.34	1.30
	04/23	104.04	136.52	104.04	1.13	1.47	1.31
	05/01	104.97	135.95	104.97	1.15	1.50	1.30
Non-Plant	04/11	43.97	43.10	43.51	1.00	0.99	1.00
	04/15	31.15	29.75	31.42	1.01	0.97	0.96
	04/23	21.98	21.16	22.46	0.97	0.94	0.98
	05/01	18.63	17.63	19.18	1.00	0.93	0.95

(b) Mean of the above statistics

Type	R	G	B	R/B	G/R	G/R
Plant	110.45	143.27	110.44	1.08	1.41	1.30
Non-Plant	28.93	27.91	31.64	1.00	0.96	0.97

In generating the Pixelwise segmentation method based on interrelationship of the RGB components, a few major concerns should be pointed out:

1. It is near impossible to achieve a 100% segmentation rate especially without manual thresholding, due to uneven illumination (even in the laboratory environment) and color variation among weed species or even among the same species at various growth stages. Thus the objective of developing this new algorithm was to achieve automatic performance while maintaining an accurate segmentation rate compared to other methods such as NExG and Modified Hue.
2. Table 3.1 showed steady distribution patterns of the RGB components of the pixels within the plant and non-plant regions such as the R, G and B components, as well as the G/B, G/R ratios of plant pixels that had greater mean value than that of the non-plant pixel. However, extreme conditions should also be considered because the statistics given on the table were mean values derived from pixels under varying lighting conditions. For instance, bright background pixels would have high R-G-B values and dark plant pixels have low values, which indicated that proper adjustment and offsets on

the limits of the means and ratios for the algorithm are necessary.

3. Introduction of additional judging criteria would aid accurate segmentation but in return require more handling time, which is not desirable for real-time operation.

Refined judging criteria were proposed based on statistics from Table 3.1 as well as empirical modifications on the limits of the means and ratios of the RGB components, which were obtained from repeated computer trials with the weed database using different limits. The final algorithm was defined as:

$$(G > 1.02 \times R) \,\&\, \&(G > 1.02 \times B) \,\&\, \&(R > 0.6 \times B) \,\&\, \&(G > 35) \qquad (3\text{-}5)$$

As mentioned, the proposed segmentation method was based on sampling of images obtained under similar illumination conditions such as during laboratory or shaded imaging in the field. Therefore, to apply formula 5 to a different scene, the parameters are to be determined using the sampling method described above on a few images taken in the new scene. Despite of this, an essential improvement of the Pixelwise method lied in that all of the three channels in the RGB image were used, thus potential loss of color information due to dimension reduction could be avoided. Nevertheless, only a raw binary image resulted for any color image that had undergone the Pixelwise segmentation procedure. Small salt and pepper noise and other random noise remained because of the existence and scatter of high intensity background pixels. Therefore, multiple image filtering was implemented to minimize segmentation errors, among which the median filter was used to handle salt and pepper noise, and a size filter based on pixel area was utilized to remove aggregated random noise. The image segmentation was finalized with the completion of the filtering procedure. The complete segmentation algorithm is shown in Figure 3.8.

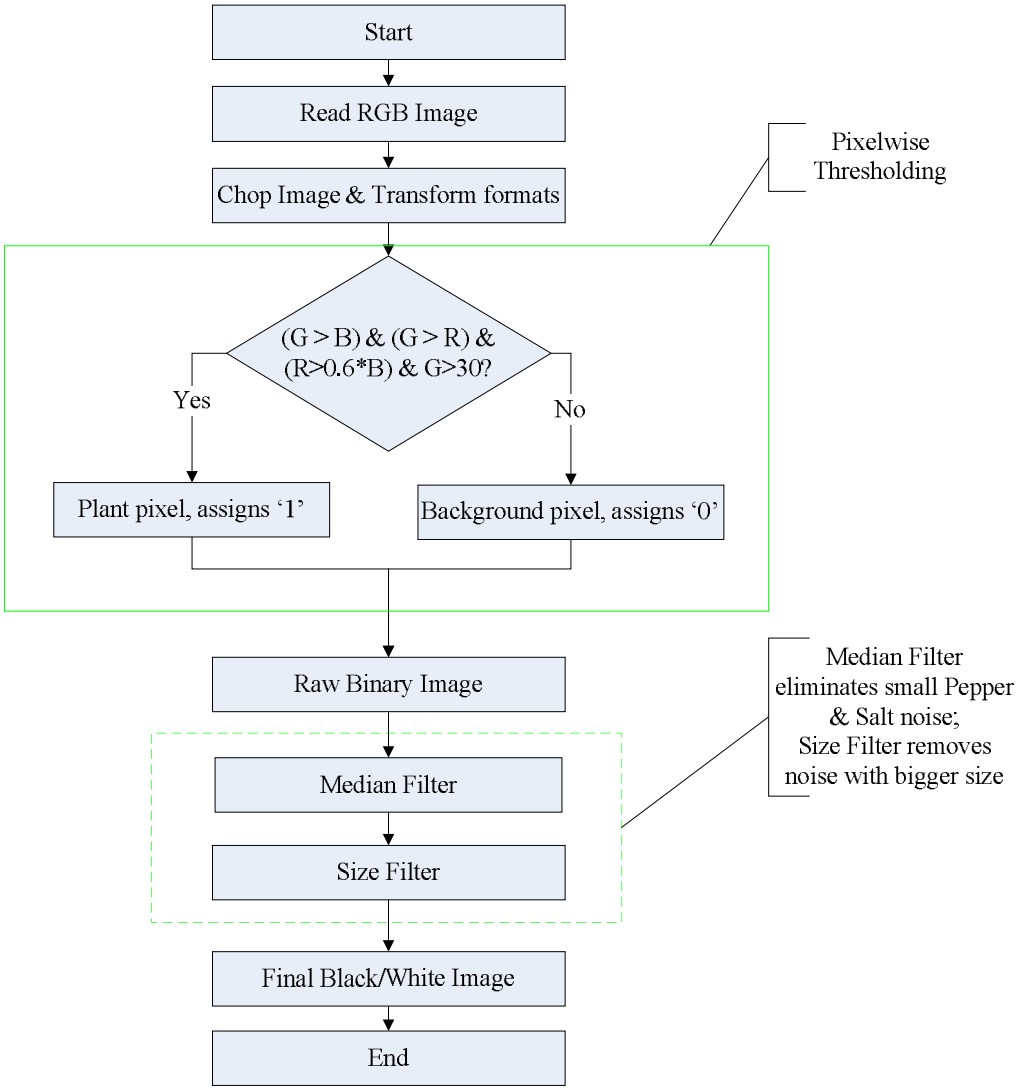


Figure 3.8. Flow chart of the complete image segmentation procedure.

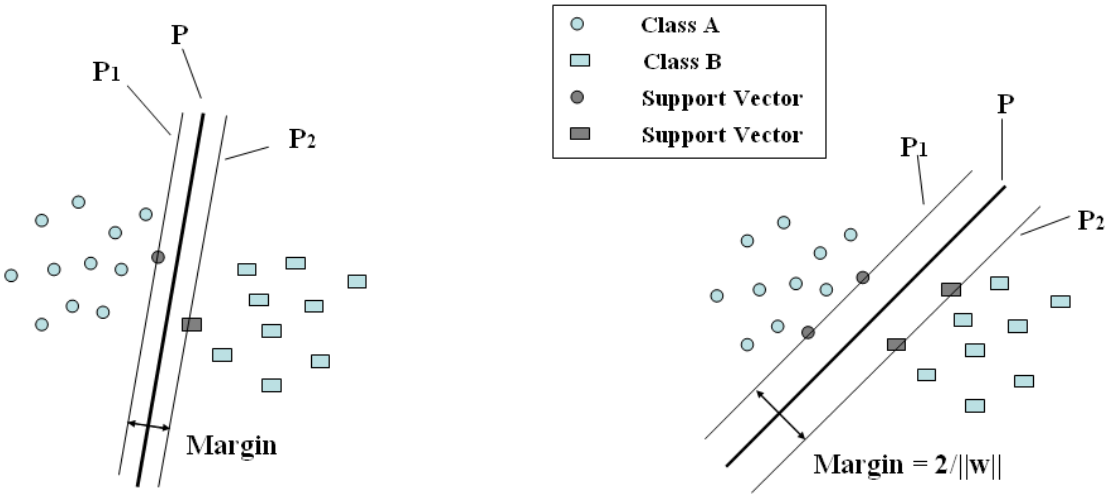
3.1.4 Weed Identification

With the completion of vegetation segmentation, binary images resulted where the plant region was displayed in white and the background in black. To implement the Support Vector Machine (SVM) method, morphological features were extracted from the images as the input vector of the system. In addition, internal parameters of the SVM model needed to be determined prior to the weed classification procedure.

3.1.4.1 Fundamentals of the SVM Algorithm

Linear Support Vector Machines

The working of the Support Vector Machine can be summarized as the search for an optimal separating plane that maximizes the margin and minimizes the generalization errors among classes. Figure 3.9 illustrates the principle of SVM through a simple linear binary classification problem.



(a) A hyperplane with small margin. (b) Optimal hyperplane with maximal margin.

Figure 3.9. Schematic of a linear hyperplane between two classes.

Given is a set of training data $\{(x_i, y_i), i = 1, 2, \ldots, n\}$, where $y_i \in \{+1, -1\}$ denotes the class labels and $x_i \in R^d$ are the d-dimensional input vectors. To separate the two classes, any hyperplane in this example can be expressed as the set of points x satisfying:

$$w \cdot x - b = 0 \qquad (3\text{-}6)$$

Where w is a normal vector to the hyperplane; $|b|/\|w\|$ represents the distance from the origin to the hyperplane along the normal w, and $\|w\|$ is the Euclidean norm of w (Burges, 1998).

It is obviously possible to have an infinite number of hyperplanes able to separate the data groups (see plane P in Figure 3.9 a and b); however, geometrically, the ideal separating

plane should have the largest distance from both of the classes so that minor disturbances of the input data would not affect the accuracy of the system.

Define the parallel supporting planes as $w \cdot x - b = 1$ and $w \cdot x - b = -1$ for class +1 and -1, respectively, such that all points belonging to one class are on or on one side of the plane (see P_1, P_2 in Figure 3.9). In this context, the perpendicular distances from the origin to the supporting planes will be $|1-b| / \|w\|$ and $|-1-b| / \|w\|$. Thus the hyperplane margin being $2 / \|w\|$, can be calculated via subtraction of the two. Here, the optimal classification problem becomes maximizing the margin $2 / \|w\|$, in other words, minimization of the norm $\|w\|$ with the constraint that no data points fall in the margin. This is equivalent to a quadratic programming optimization (QP) problem with $\|w\|$ substituted by $\frac{1}{2}\|w\|^2$ without changing the solution (Burges, 1998):

Minimize: $\frac{1}{2}\|w\|^2$ (in w, b) (3-7)

Subject to: $w \cdot x - b \geq 1$ for $x_i \in$ class +1 and (3-8)

$w \cdot x - b \leq -1$ for $x_i \in$ class - 1. (3-9)

With a Lagrangian formulation of the problem, the optimization problem can be transformed into the Lagrangian dual form, which is only a function of the support vectors (Burges, 1998):

Maximize: $\sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j} a_i a_j y_i y_j x_i \cdot x_j$ (in a_i) (3-10)

Subject to: $\sum_{i=1}^{n} a_i y_i = 0$ and $a_i \geq 0$ for any $i = 1, 2, \ldots, n$ (3-11)

Where a_i are the support vectors (also called Lagrangian multipliers), the training data that lie on the supporting planes. The desired Lagrangian multipliers a_i^* can be obtained by solving equation (10) with constraint (11). Hence the optimal separating hyperplane is given by (Gunn,

1998):

$$w^* = \sum_{i=1}^{n} a_i^* y_i x_i \quad \text{and} \quad b^* = -\frac{1}{2}[w^* \cdot (x_p + x_q)] \tag{3-12}$$

Where x_p and x_q are any support vectors from each class satisfying:

$$a_p, a_q > 0, \quad y_p = -1, y_q = 1 \tag{3-13}$$

The corresponding classifier can be defined as:

$$f(x) = \text{sgn}(w^* \cdot x - b) \tag{3-14}$$

For more general cases where input data overlap each other, it is virtually impossible to find a hyperplane that can linearly separate two data sets at 100% accuracy. Instead, SVM will trade off part of the accuracy and seek for an optimal balance between maximizing the margin and minimizing the classification errors. This is achieved by introducing an upper bound on the number of training errors through the use of the "positive slack variables" (Cortes and Vapnik, 1995) in the constraints.

Non-Linear Support Vector Machines

For training samples that are not linearly separable, the input vectors are mapped into a higher dimensional space (referred to as "feature space"), where the optimal separating hyperplane can be constructed. Such non-linear transformation incorporates the use of a "kernel function", $K(x_i, x_j)$ to resolve the computational complexity on the feature space (Boser et al., 1992). In this context, the optimization problem in equation (10) becomes:

Maximize: $$\sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j} a_i a_j y_i y_j K(x_i, x_j) \tag{3-15}$$

Subject to: $$\sum_{i=1}^{n} a_i y_i = 0 \quad \text{and} \quad 0 \le a_i \le C \tag{3-16}$$

Where $K(x_i, x_j)$ replaces the dot product $x_i \cdot x_j$, and C is a user-chosen parameter which reflects the noise in the data and determines the tolerance to misclassification errors (Gunn,

1998). The classifier corresponding to the optimal separating hyperplane in the feature space becomes:

$$f(x) = \text{sgn}(\sum_{i \in SVs} a_i y_i K(x_i, x) + b) \quad (3\text{-}17)$$

where

$$w^* \cdot x = \sum_{i=1}^{n} a_i^* y_i K(x_i, x) \quad (3\text{-}18)$$

$$b^* = -\frac{1}{2} \sum_{i=1}^{n} a_i y_i [K(x_i, x_p) + K(x_i, x_p)] \quad (3\text{-}19)$$

Previous discussions are based on the binary classification schemes; however, the concept of SVM can also be extended to multi-category classification, where the number of classes is larger than 2. The primary idea is to reduce the single multi-category problem into multiple binary schemes, which is commonly accomplished by introducing a decomposition and reconstruction procedure. In general, there are two decomposing methods: [1] "1-versus-rest", where one class has the label +1 and the remaining patterns are labeled -1; [2] "1-versus-1", in which $M = N(N-1)/2$ sets of binary machines are to be constructed, where N is the number of classes (Hsu and Lin, 2002). These decomposing treatments are followed by a parallel reconstruction phase with different decision strategies for classification: for "1-versus-rest" case, a "winner-take-all" strategy will be applied, where the class of the instance is decided by the classifier with the highest output function; while in the "1-versus-1" case, a "max-wins voting" strategy determines the instance classification by counting the most votes from each of the binary classifiers (Angulo et al., 2000).

3.1.4.2. Feature extraction

Several shape features describing the geometric properties of the weed canopy were extracted from the binary images obtained from the image segmentation procedure. Table 3.2 lists the calculation of six shape features used for weed identification for a given binary image shown in Figure 3.10.

Table 3.2. Description of various shape features of weed.

Feature	Description
Area ($pixel^2$)	Total number of pixels in each segmented region
Perimeter (*pixel*)	Number of the boundary pixels in each segmented region
Width (*pixel*)	Horizontal dimension of the segmented region
Height (*pixel*)	Vertical dimension of the segmented region
Major Axis Length (*pixel*)	Length of the major axis of the ellipse that has the same normalized second central moments as the segmented region
Minor Axis Length (*pixel*)	Length of the minor axis of the ellipse that has the same normalized second central moments as the segmented region

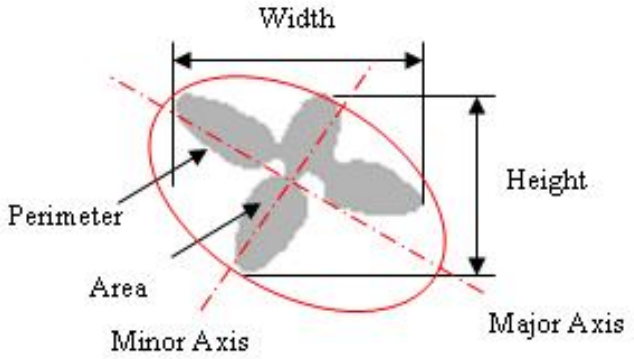


Figure 3.10. Geometric definition of the weed canopy.

While the objective of this project is to identify weed species in the early growth stage, the features listed above may not be appropriate for direct use as the input of a classifier. This is because these features tend to change along with time and the growth stage of the targeted weed plant. Therefore, several geometric parameters (Cho et al., 2002; Lee et al., 1999; Tian et al., 1997) aiming at better describing the plant pattern instead of outward dimensions, were used as follows:

1. Ratio of Area to Length (ATL): ratio of the segmented area to the major axis length.

$$\text{ATL} = \frac{\text{Area}}{\text{Major Axis Length}} \quad (pixel) \tag{3-20}$$

2. Compactness (CMP): ratio of the segmented area to the perimeter squared.

$$\text{CMP} = 16 \times \frac{\text{Area}}{\text{Perimeter}^2} \tag{3-21}$$

3. Elongation (ELG): difference of the best-fit ellipse axis lengths divided by the sum of the of the axis lengths.

$$\text{ELG} = \frac{\text{Major Axis Length - Minor Axis Length}}{\text{Major Axis Length + Minor Axis Length}} \tag{3-22}$$

4. Aspect (ASP): ratio of the major axis length to the minor axis length.

$$\text{ASP} = \frac{\text{Major Axis Length}}{\text{Minor Axis Length}} \tag{3-23}$$

5. Logarithm of the Ratio of Height to Width (LHW): common logarithm of the ratio of the vertical to horizontal dimensions of the segmented area.

$$\text{LHW} = \log_{10} \frac{\text{Height}}{\text{Width}} \tag{3-24}$$

6. Ratio of Perimeter to broadness (PTB): measurement of a convex region.

$$\text{PTB} = \frac{\text{Perimeter}}{2 \times (\text{Height + Width})} \tag{3-25}$$

7. Ratio of Length to Perimeter (LTP): measurement of the 2-D distribution pattern of the boundary of the segmented region.

$$\text{LTP} = \frac{\text{Major Axis Length}}{\text{Perimeter}} \tag{3-26}$$

3.1.4.3. Feature Selection

Feature selection is the technique with which a subset of relevant features will be selected for building robust learning models. It is an effective way of reducing the training sample size and the operational time, although at the risk of affecting the generalizing rate of the classifier. Thus with a real-time facilitation purpose, it is necessary to explore the interrelationship between the classification accuracy and the efficiency of the system.

The minimum-Redundancy-Maximum-Relevance method (mRMR) is a unique feature selection technique for machine learning (Peng et al., 2005). To construct an optimal feature subset that describes the statistical property of a target classification variable, the mRMR

method selects features that are mutually dissimilar to each other, while highly correlated to the classification variable. Such a scheme was used in this project to determine the most critical features. In the implementation of the mRMR method, the continuous feature data should be discretized first to achieve accurate results (Peng et al., 2005). Table 3.3 lists the selection results according to various threshold values for data discretization, from which the ATL, PTB, and ELG were selected as the best feature combination.

Table 3.3. Feature selection results using mRMR with various thresholds.

Priority	K=0	K=0.5	K=1
1	ATL	ATL	ATL
2	PTB	PTB	PTB
3	ELG	LHW	ELG
4	ASP	ELG	LHW
5	LHW	ASP	ASP

Note: K is the threshold chosen to discretize the data, i.e., mean +/- K*STD.

3.1.4.4. Implementation of the SVM Method

To evaluate the performance of the SVM method, it is common practice to separate the complete data set into two parts. A training set is used to develop a predictive model after which the unknown testing set is used to assess the validity of the model and the performance of the SVM classifier.

The first step in implementing the SVM method is to choose the kernel function (Table 3.4). For a particular multi-class categorization problem, the Radial Basis Function (RBF) kernel is an ideal first choice in most cases. This is because the RBF kernel is capable of handling nonlinear SVM problems, is less prone to numerical difficulties, and easy to implement (Keerthi and Lin, 2003).

The next step is to construct the characteristic model based on the training samples. Two parameters have to be identified and optimized such that the classifier can predict unknown data precisely. Of this parameter pair (C, γ), C is the regularization parameter that defines the error bound, while γ is the characteristic parameter of the kernel function that specifies the Gaussian model structure. A common method to accomplish this is through the use of cross validation. In

this project, a five-fold cross-validation (CV) method (Chang and Lin, 2007) was adopted, in which the training set was divided into five subsets, and each subset (the validation set) was tested using the classifier trained on the remaining four subsets. The CV accuracy was then the average of prediction accuracy on the validation set. In this context, each instance of the whole training set was predicted once and accordingly the best (C, γ) pair could be determined by comparing the cross-validation accuracy.

Finally, to determine the relationship between the number of features used against accuracy and operational time of the SVM classifier, two distinct models based on the three selected features and the complete seven features were constructed with the same training set. The two models were then used to classify the testing set consisting of 1,006 images. The complete identification procedure, including cross-validation, model-construction and label-prediction, was carried out in MatLab®, incorporating the use of the LIBSVM MatLab® Toolbox (Chang and Lin, 2001).

Table 3.4. Some common kernels for nonlinear SVM.

Kernel	Function
Linear	$K(x, y) = x \cdot y$
Sigmoid	$K(x, y) = \tanh(\gamma x \cdot y + c)$
Polynomial	$K(x, y) = (\gamma x \cdot y + c)^d, \gamma > 0$
Radial Basis Function (RBF)	$K(x, y) = \exp(-\gamma(x - y)^2), \gamma > 0$

Note: γ, c, and d are kernel parameters.

3.2 Field Testing

Since the proposed identification algorithm is expected to be implemented on an adaptive weed suppression device for real-time weed control, field tests were conducted. In addition, with the aim of achieving inter-row weeding, weed species spreading between the cornrows were the focus of this project.

3.2.1 Field Experiment Setup and Data Acquisition

Five weed varieties, including three broadleaf weed species – lambsquarters, velvetleaf and waterhemp, and two grass species – barnyardgrass and large crabgrass were manually planted at the University of Illinois Agricultural Engineering Research Farm on the 1st July, 2009. Each species was planted in a plot, measuring $0.76m \times 6.35m$ ($30" \times 250"$), which was separated by corn rows (Figure 3.11).

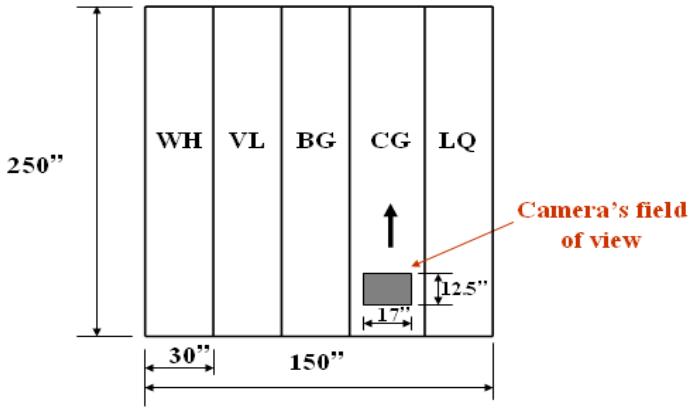


Figure 3.11. Field test layout: VL – velvetleaf, LQ – lambsquarters, WH – waterhemp, BG – barnyardgrass, CG – crabgrass.

A Fire-i™ industrial camera (model 701c, Unibrain Inc., San Ramon, CA) featuring a C mount 6mm F1.2 lens (Pentax Co., Golden, CO, USA) was employed to acquire field images. The camera was mounted at a height of 1.87 m (73.5") on a custom-made camera holder, which was attached to a utility tractor (model 1024D, New Holland North America Inc., New Holland, PA) via a supporting frame. A portable computer (model Studio™ 15, Intel Core™ 2 Duo Processor, 2.96 GB of RAM, Dell Inc, Round Rock, TX) was used to control the camera through a MatLab® program (Appendix A5). A Fire-repeater (model 400™ 1394a, Unibrain Inc., San Ramon, CA), powered by a 12v battery (model BP 12-12, B&B Battery USA Inc., Commerce, CA), was used to link the camera with a 1394a interface and the computer with a 1394b interface. The computer, battery, and Fire-repeater were placed on a custom holder plate, whose extended connecting piece was attached to the tractor (Figure 3.12 a – d).

Figure 3.12. Facilities used for image acquisition during the field test.
(a) Image acquisition platform;
(b) Laptop, Battery and Fire-repeater;
(c) Custom holder plate and its connecting piece;
(d) Custom camera holder and supporting frame;
(e) Custom shading plate.

For image acquisition, the camera was used to capture an area of $0.425m \times 0.315m$ ($17inch \times 12.5inch$) at a resolution of 800×600 pixels, with a 0.3 m zoom setting on the lens. Thus the spatial resolution for the imaging system was approximately $5.3mm \times 5.3mm$ ($0.2" \times 0.2"$) per pixel. The aperture was fixed at F7 and exposure was set at 40 to reduce light absorption. However, under strong sunlight, these settings still resulted in poor image quality due to the automatic white balance function of the camera. A simple solution was to create an artificial shadow over the imaging area by adding a shading plate made of photography exclusive ripstop nylon fabric (Hancock Fabrics, Champaign, IL), (Figure 3.12 e). In this way,

excessive sunlight was blocked, leaving a homogeneous illumination condition over the imaging area (Figure 3.13). Images were taken while the tractor was moving at a low speed of around 0.7 km/h (0.4 miles/hour). This travel speed, in combination with an image capturing rate of 1,200 frames per second minimized speed blur to an acceptable level.

(a) Over-illuminated waterhemp image (b) Shaded waterhemp image

Figure 3.13. Contrasting images of waterhemp taken without and with the shading plate.

The experimental field was maintained weed-free except for the planted species. However, during the field experiment, few lambsquarters survived due to the influence of the growth season, management and extreme weather conditions. In addition, at the date of imaging, it was discovered that smooth crabgrass was planted in the field whereas large crabgrass was planted in the greenhouse. Hence, this experiment only considers the remaining species, being barnyardgrass, velvetleaf and waterhemp. Two sets of images were recorded on two days in the fourth and fifth weeks after seeding, which corresponded to the V4 and V5 growth stage of corn. The first set of images (including 60 samples) was taken in the afternoon of July 26th, 2009 under cloudy conditions, and the other set (consisting of 47 images) was taken at noon of July 29th, 2009 under sunny conditions. The raw images were in the YCbCr format, which were converted to 24-bit RGB images and saved for future analysis.

3.2.2 Weed Identification

Weed identification in the field followed similar steps compared to the laboratory experiment, those being image segmentation, feature extraction and weed classification. Modifications in the algorithm were made to accommodate real-time application and images containing multiple weeds.

In the laboratory experiments, the complete image set was divided into two parts. The training set was used to construct the SVM model in use of the known labels (weed species) and the corresponding morphological features as input vectors: The test set was used to classify the samples where predictions of the unknown labels were made based on the SVM model and features input from the test set. However, in the field experiment, all weed images were considered as the unknown test set. Label prediction was based on the model using the complete greenhouse image set excluding corn images. Subsequently, cross validation was performed on the greenhouse image set, and the resulting optimal parameters were utilized to construct the SVM model. This model was consequently preloaded to the system before any test image input.

For image segmentation, new thresholds obtained from sampling of 30 field images were used to implement the Pixelwise method for all images. A significant difference between the laboratory and field segmentation procedures is that for the greenhouse image only a single plant was present and every segmented component other than the background was considered part of the plant. However, in a field scenario, every field image may contain multiple as well as partially recorded plants crossing the image boundary. In addition, a single plant might be divided into several parts due to segmentation errors. Therefore, to reunite the separated parts of a plant and various plants in an image, the centroid for each segmented part was calculated. Then the adjacent parts were joined together or labeled with diverse numbers depending on the distances of their corresponding centroids among each other. Partially recorded plants were discarded from each image (Figure 3.14), because the shape-based machine vision system relies on feature information extracted from the complete weed plant or plant leaf.

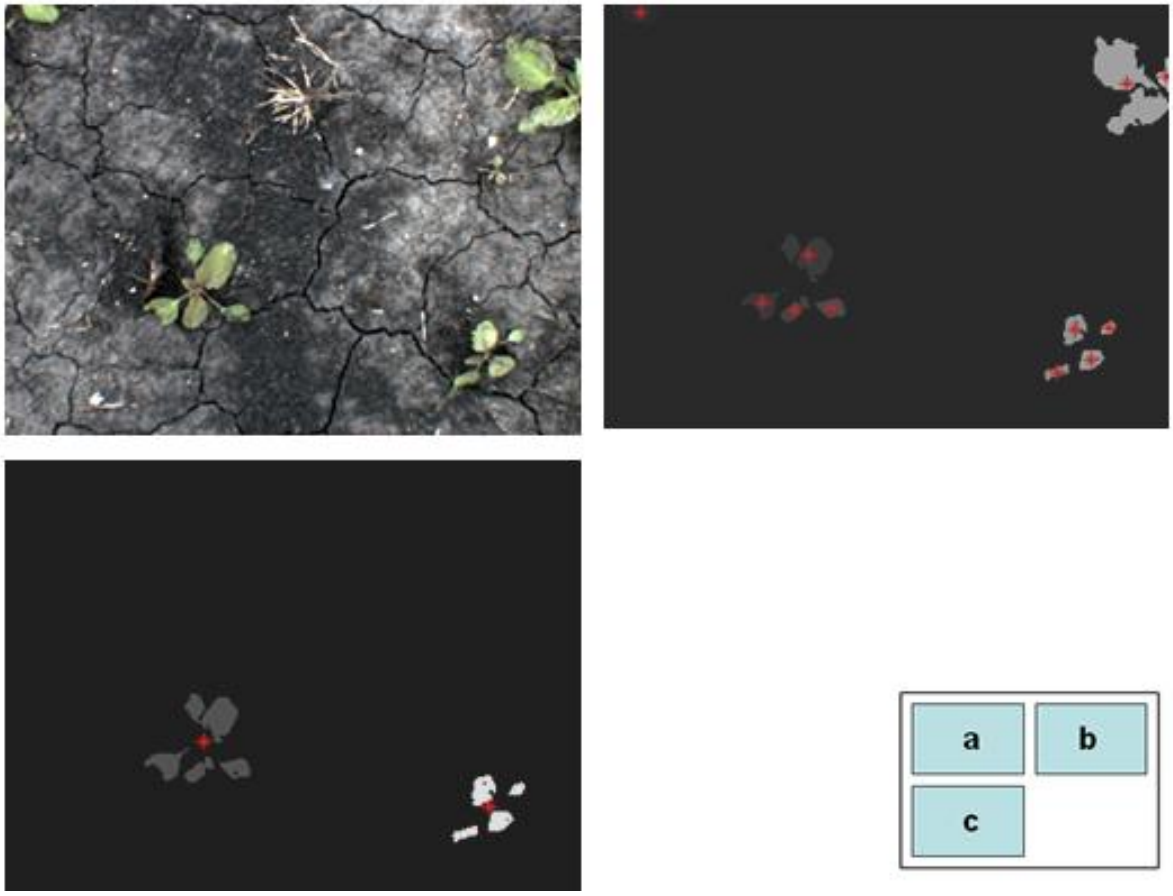


Figure 3.14. The detailed image segmentation procedure for field images.
(a) Original RGB image.
(b) Segmented image with centroids of each separated part marked.
(c) Reunion of parts and separation of plants.

An adaptive median filter (Gonzalez et al., 2004), which was able to choose the filtering window size automatically according to the characteristics of the image, was an optional additional component to the Pixelwise method. The advantage of the adaptive median filter is its ability to eliminate salt and pepper noise while preserving the sharpness and detail of the image. A disadvantage aspect lies in its high computational cost, which is undesirable in a real-time scenario.

As soon as the binary image was acquired, each plant labeled with various numbers passed through the feature segmentation procedure, where the seven features, being the Ratio of Area to Length (ATL), Compactness (CMP), Elongation (ELG), Aspect (ASP), Logarithm of

the Ratio of Height to Width (LHW), Ratio of Perimeter to broadness (PTB), and Ratio of Length to Perimeter (LTP) were computed. Finally, the preloaded model was used to predict the class/classes of each weed plant in the input image based on the calculated features. The weed identification process was completed automatically through a MatLab® program (Appendix A7). A flowchart is shown in Figure 3.15.

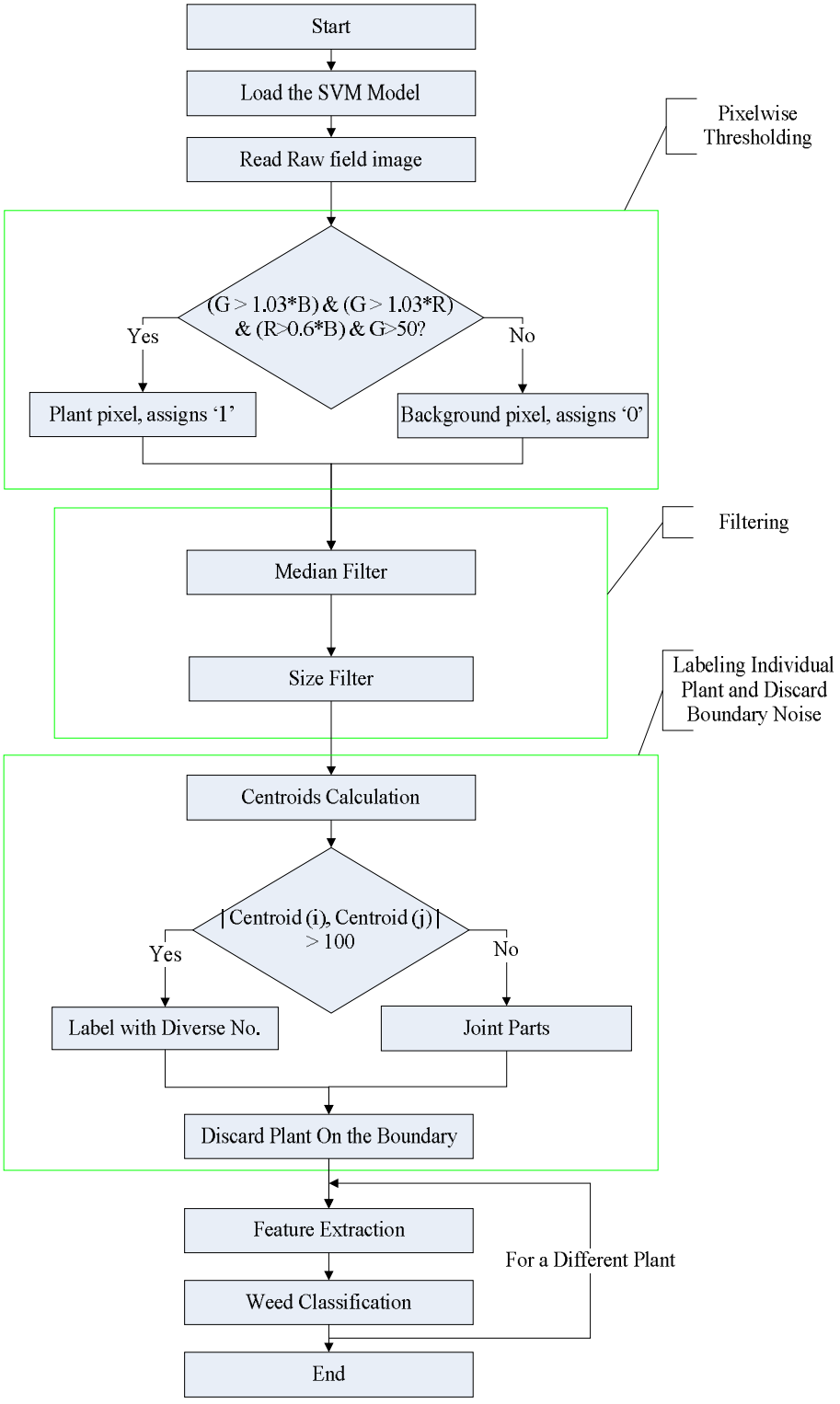


Figure 3.15. Flow Chart of the weed identification algorithm for field images.

CHAPTER 4

RESULTS

4.1 Results of the Laboratorial Image Segmentation

To evaluate the proposed Pixelwise segmentation method, 101 laboratory images were chosen randomly from growth dates 04/08, 04/16 and 04/30 (covering various growth stages of the weeds), and manually segmented until optimal segementation results were achieved. In the process, three automatic segmentation methods including the NExG with OTSU's method, Modified Hue with fixed color range [65, 120] and the Pixelwise method with empirical limits based on sampling were implemented. Each segmented image was compared pixel-by-pixel with the corresponding hand-segmented image. Two variables were determined in this experiment, where the correct segmentation rate (CSR) was defined as the ratio of the number of plant pixels segmented in agreement with hand segmentation and the total number of plant pixels obtained from hand segmentation. The incorrect segmentation rate (ISR) was defined as the ratio of the sum of plant pixels misclassified as background and background pixels misclassified as plant, relative to the total number of plant pixels obtained from hand segmentation (Figure 4.1).

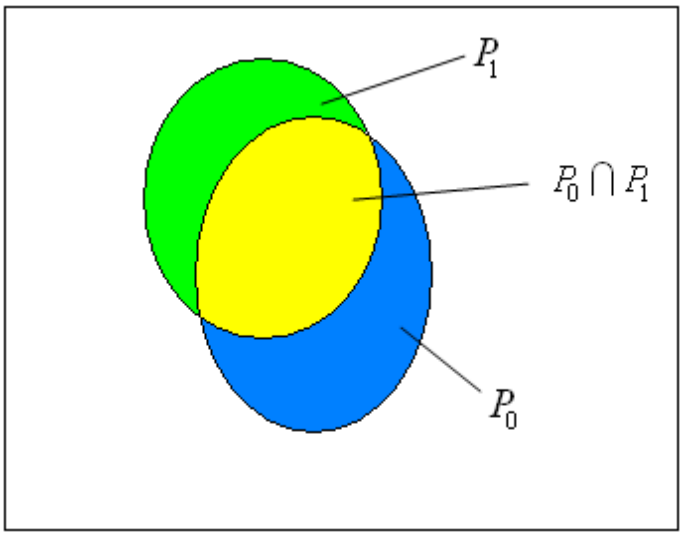


Figure 4.1. Hand-segmented plant image (P_0) VS. Image segmented using other automatic Methods (P_1)

Test results are shown in Table 4.1 as well as Figure 4.2, where CSR and ISR were computed as:

$$\text{CSR}=\frac{P_0 \bigcap P_1}{P_0} \quad , \quad \text{ISR}=\frac{P_0 \bigcup P_1 - P_0 \bigcap P_1}{P_0} \tag{4-1}$$

Where

P_0 = number of pixels hand-segmented as plant in the image.

P_1 = number of pixels segmented as plant in the image segmented as plant in the image using NExG, Modified Hue or Pixelwise methods..

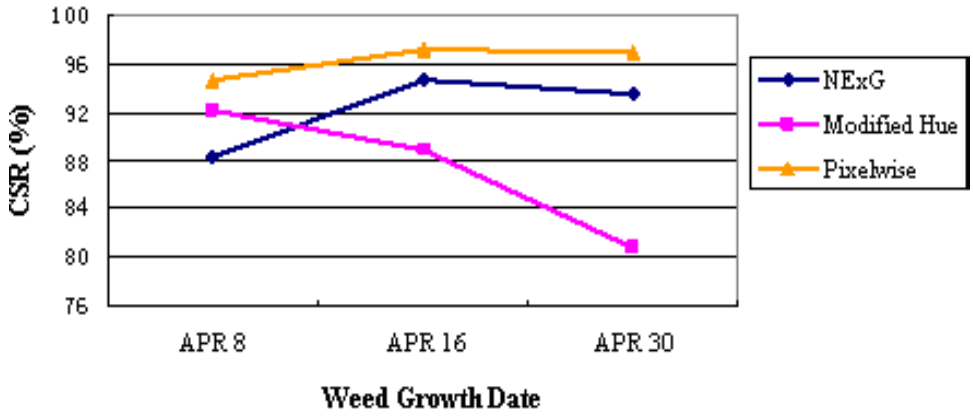
Table 4.1. Mean values of the CSR and ISR for various segmentation methods.

(a) Segmentation performance across various growth stages of weeds

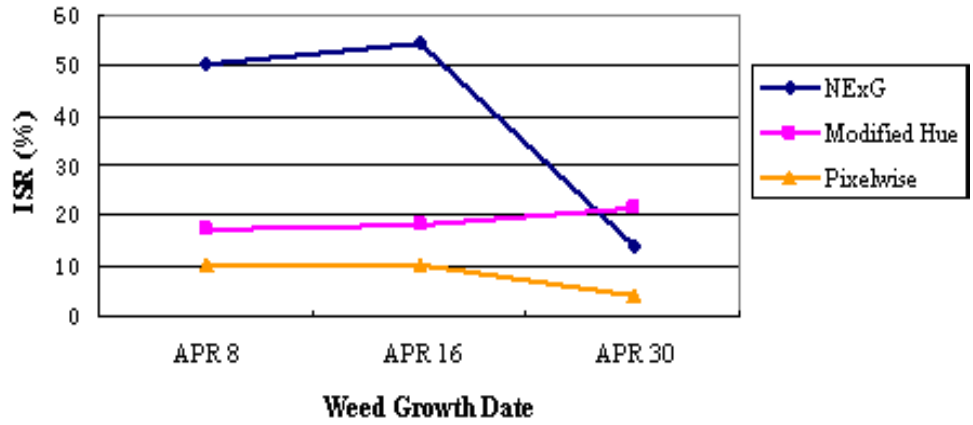
Sample Images	Methods	CSR (%)	ISR (%)
04/08 (17 samples)	NExG	88.3	50.1
	Modified Hue	92.2	17.2
	Pixelwise	94.6	10.5
04/16 (49 samples)	NExG	94.6	54.4
	Modified Hue	88.9	18.1
	Pixelwise	97.2	10.1
04/30 (35 samples)	NExG	93.5	14.2
	Modified Hue	80.8	21.5
	Pixelwise	96.9	4.4

(b) Overall comparison among the three segmentation methods

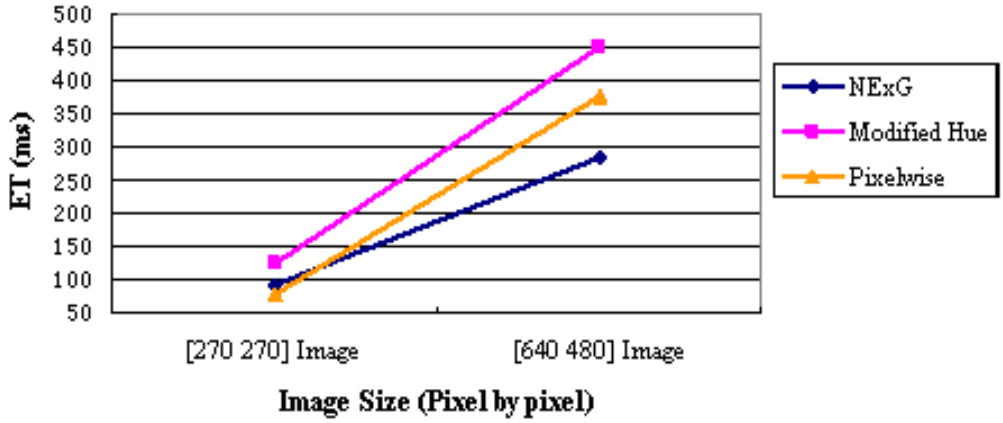
Sample Images	Methods	CSR (%)	ISR (%)	Elapsed Time (ms per image)
Total (101 samples)	NExG	92.1	39.6	187.8
	Modified Hue	87.3	18.9	287.6
	Pixelwise	96.3	8.4	226.6



(a) Comparison of CSR values across various plant growth stages



(b) Comparison of ISR values across various plant growth stages



(c) Comparison of program elapsed times for images with various sizes

Figure 4.2. Segmentation performance of the NExG, Modified Hue and Pixelwise methods

Based on the experiment results listed in the tables as well as the corresponding charts, the following conclusions were drawn:

1. Both of the CSR and ISR values are important parameters denoting the efficiency of the segmentation methods. Of these two indicators, CSR quantifies the percentage of plant pixels that are segmented correctly by the automatic segmentation methods: ISR represents the rate of the sum of plant pixels that were classified as background and background pixels classified as plant relative to the 'real' plant pixels. Thus an ideal segmentation method should achieve a high CSR value and maintain a low ISR level.
2. The Pixelwise segmentation method outperformed the NExG and Modified Hue methods in both of the categories, achieving the highest CSR values ranging from 94.6% to 96.9%, and lowest ISR values ranging from 4.4% to 10.5%. In addition, the NExG and Modified Hue methods were not consistent in plant segmentation across various growth stages of the weeds (notice the abrupt rise and decline on the CSR or ISR curves for these two methods in Figure 4.2 (a) and (b)). However, the proposed Pixelwise algorithm attained a consistent performance in producing high CSR and low ISR levels throughout the experiment.
3. The processing time was another important consideration due to the ultimate goal for real-time weed identification. Two image formats were used in the test: cropped size (281×261) for small weeds confined within a cell and full size (640×480) for larger plants with leaves stretching out of the cell. It is clear from Figure 4.2 (c) that the elapsed time for all of the three methods remained at a low level in processing cropped images, while they increased dramatically for full-size images. Of the three, the NExG method accomplished the best time-efficiency in the test and the Modified Hue attained the worst. Nevertheless, there was less than 0.1 second (0.0998 s) difference between the mean values of processing time among these segmentation methods.

In summary, the Pixelwise method achieved significantly higher accuracy in weed segmentation compared to the NExG and Modified Hue methods. On the other hand, although

not as fast as the NExG method, the Pixelwise algorithm still performed efficient segmentation with an average handling time of 226.6 ms/image, which was only 0.0387 s slower than the NExG algorithm.

4.2 Results of the Laboratory Weed Identification

1,319 binary images (56.7%) were randomly selected from the weed images database to form the training set, and the remaining 1,006 images (43.3%) were kept for evaluation. Two SVM models were derived using the three selected features (ATL, PTB, and ELG) and the complete seven features (ATL, PTB, ELG, CMP, LHW, LTP, and ASP). The kernel parameter γ and error bound C for the SVM models were determined through five fold Cross Validation (CV), where the CV search ranges were set at $C \in (0.5, 8{,}192)$ and $\gamma \in (0.00098, 2)$. The returned optimal parameter pairs were (8,192, 0.002) with a maximum CV accuracy of 68.84% for the three-feature model, and (8,192, 0.001) with maximum CV accuracy of 82.48% for the seven-feature model, respectively.

However, the parameter pair (8,192, 0.002) was found later to have achieved a higher accuracy than (8,192, 0.001) for the seven-feature model in identifying the unknown weed spices in the testing set. This is because during CV, the subdivision of the training set is completely random; subsequently the parameter pair that has the highest accuracy for the "one-versus-rest" classification according to such random data arrangement will be considered the best. On the other hand, because CV may be affected by factors such as coupling among features and skewness of the data set (i.e. the number of classes far exceeds the number of features or *vice versa*), it is possible that the returned parameters would slightly deviate from the optimal values. In this context, the "mistakenly-selected parameter" is theoretically explainable due to the different constitution of data between the CV and the prediction procedures as well as the nature of the classification problem itself. Therefore, the optimal parameters for the both of the models were determined to be (8,192, 0.002), and the classification accuracies, training time and predicting time for the three-feature and seven-feature models against the testing image set were listed in Table 4.2.

Table 4.2. Overall classification results for the three-feature and seven-feature models.

Model Basis	Parameter Pair	Accuracy (%)	Training Time (s)	Predicting Time (s)
3-features	(8192, 0.002)	70.48	3.281	0.109
7-features	(8192, 0.002)	86.58	2.781	0.125

It is clear that the SVM model using the complete seven features achieved higher classification accuracy (as much as 16%) over the one using the selected three features. On the other hand, the three-feature model has a slight advantage (0.016s) over the seven feature model in prediction speed, while it is slower than the complete model in model construction by 0.5s. Considering the three factors that determine the performance of a classifier, the seven-feature SVM model performs equal or better than the three-feature model in every category. Thus the seven-feature model was used for model construction, and the following discussion will only focus on this complete model.

To explore the experiment results in more detail, the parameter pair (8192, 0.002) that acquires a classification rate of 86.58% was used to construct the SVM model with the complete seven features. Two variables were defined to quantify the error terms for each interested category. For a given weed species A, the calculation of the error terms are as follows:

$$\text{Commission Error (CE)} = \frac{\#\text{ of Other Weed Species Misclassified as Weed A}}{\#\text{ of Weed A}} \qquad (4\text{-}2)$$

$$\text{Omission Error (OE)} = \frac{\#\text{ of Weed A being Classified as Other Species}}{\#\text{ of Weed A}} \qquad (4\text{-}3)$$

The classification results are shown in Table 4.3:

Table 4.3. Classification results split using the seven-feature model with parameter (8192, 0.002)

(a) Classification rate between crop and weeds (all five species).

Weed Species	Accuracy	CE	OE
Crop (Corn)	94.41%	5.59%	5.59%
Undesired Weed	99.07%	0.93%	0.93%

Table 4.3. cont.

(b) Classification rate for Grass and Broadleaf weed.

Weed Species	Accuracy	CE	OE
Grass	95.87%	4.13%	4.59%
Broadleaf	96.49%	3.51%	3.16%

Notice: Grass includes: corn, barnyardgrass, and crabgrass;

Broadleaf includes: lambsquarters, velvetleaf, and waterhemp.

(c) Classification rate for each weed species.

Weed Species	Accuracy	CE	OE
Corn	94.41%	5.59%	5.59%
Barnyardgrass	86.36%	14.77%	13.64%
Crabgrass	90.24%	10.24%	10.73%
Lambsquarters	85.96%	14.47%	17.45%
Velvetleaf	88.56%	11.81%	29.52%
Waterhemp	51.56%	48.44%	7.81%

(d) Classification rate for each weed species versus time (on a weekly scale).

Weed		Week 1	Week 2	Week 3	Week 4	Week 5	Sum
Corn	Test Size	56	61	26			143
	Accuracy	87.5%	100%	96.15%			94.41%
Barnyard grass	Test Size	19	19	21	22	7	88
	Accuracy	68.42%	94.74%	100%	90.91%	57.14%	86.36%
Crabgrass	Test Size	39	39	46	56	25	205
	Accuracy	97.44%	92.31%	93.48%	89.29%	72.0%	90.24%
Lambsquarters	Test Size	43	49	50	60	33	235
	Accuracy	83.72%	91.84%	70%	91.67%	93.94%	85.96%
Velvetleaf	Test Size	53	55	37	58	68	271
	Accuracy	94.34%	92.73%	91.89%	91.38%	76.47%	88.56%
Waterhemp	Test Size	10	11	15	15	13	64
	Accuracy	80.0%	81.82%	53.33%	33.33%	23.08%	51.56%

(e) Summary of the classification rate for undesired weed species

All 5 weeds	Week 1	Week 2	Week 3	Week 4	Week 5	Sum
Testing Size	164	173	169	211	146	863
Accuracy	88.42%	91.91%	83.43%	86.73%	73.97%	85.28%

Note: Corn is excluded.

Table 4.3 (a) shows the performance of the SVM classifier in discriminating a crop (corn) plant from all other desired weeds (5 species). The system has demonstrated high classification ability: 138 out of 143 (94.41%) corn images were correctly classified, and 855 out of 863 (99.07%) non-crop images were classified as weeds. Combined with the prediction accuracy of corn against time as shown in Table 4.3 (d), it can be concluded that the first week after the emergence of corn plants is the most error-prone time for corn-weed discrimination, as most of the misclassification cases of corn in the experiment took place during this period.

Table 4.3 (b) illustrates the classification results of the SVM classifier for a binary case, in which three weed species were grouped as the “Grass” category, and the remaining three as the “Broadleaf” class. The system again showed good performance, since merely 18 Grasses were misclassified as Broadleaf, and 20 Broadleafs were misclassified as Grasses among the testing set containing 1,006 weed images.

Table 4.3 (c) compares the performance of the SVM method for each weed species considered in this project. Among the six species, the SVM classifier achieved the best prediction in corn with an accuracy of 94.41%, a commission error of 5.59% and omission error of 5.59%. Theoretically, this is largely due to the highest growth rate and largest canopy area of corn, which make it the easiest to identify based on morphology. The lowest classification rate is in waterhemp with 51.56% prediction accuracy and up to 48.44% commission and 7.81% omission errors. The main reason for this result is the limited number of weed samples: Only five out of 18 waterhemp plants survived during the laboratory imaging process, which lasted for more than five weeks. Therefore, compared to other species, there were insufficient waterhemp prototypes to construct the SVM model, which led to low classification accuracy. Secondly, the error terms for most of the species except corn were quite high, ranging from 7.81% to 48.44%. However, in comparison to Table 4.3 (b), it seems reasonable to conclude that most of the misclassification occurred in the “Within Category”. For example, lambsquarters possesses a higher probability of being classified as one of the other three broadleaf species, than as one of the grass species. This phenomenon may caused by the similar

morphological characteristics within the weed categories.

Table 4.3 (d) shows the classification accuracy per species as a function of time. From the table it is evident that the classification rates for corn, lambsquarters and velvetleaf remain relatively constant throughout the experiment. For crabgrass and barnyardgrass, the accuracies stay at high levels in the early stages until week five, where the rate dropped significantly because of the reduction of available weed images for both model construction and prediction. Finally, waterhemp, which lacked image quantity, achieved the lowest and decreasing accuracies as shown in Figure 4.3.

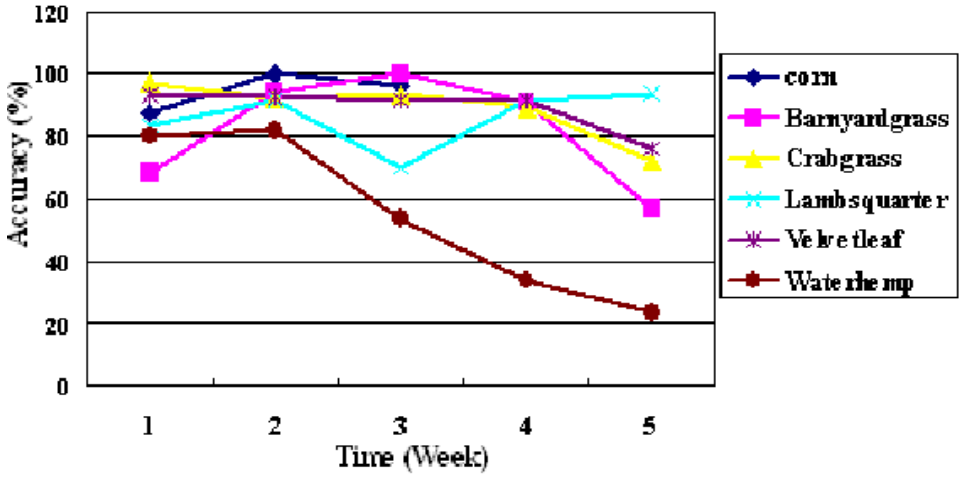


Figure 4.3. Classification accuracy for individual weed species at various growth stages after emergence.

Since the main purpose of this research is to apply the SVM method for inter-row weed identification, weed species spreading between the corn rows are a major concern. Hence, the accuracy variation versus time for the undesired weed species is listed in Table 4.3 (e) and illustrated in Figure 4.4. The system demonstrates even classification performance over the 5 weeks experiment period, which validates the feasibility and stability of the SVM method. The drop in accuracy during the final week was undoubtedly caused by a reduced sampling size.

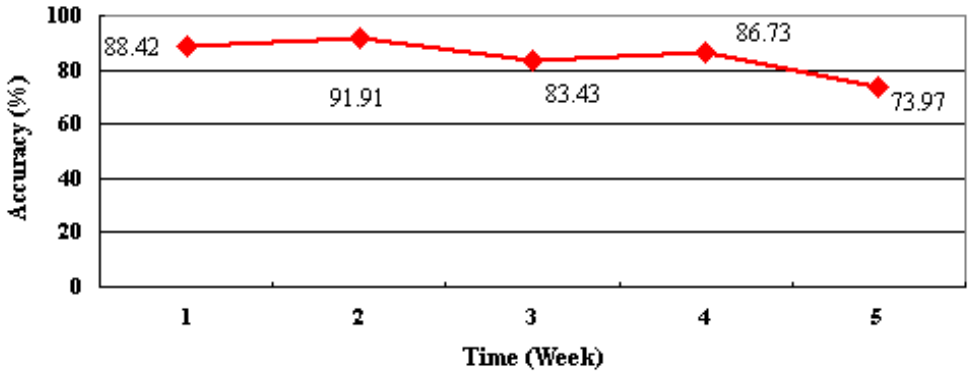


Figure 4.4. Classification accuracy for the collective undesired weed species at various growth stages after emergence.

4.3 Results of the Field Test

As the focus of this project was on inter-row weed identification, 306 corn images were removed from the weed image library. To maintain generality, the remaining images consisting of five species were used to construct the SVM model, although only three weed species were present in the field. In use of the complete seven features, the cross validation had the highest CV accuracy of 85.34%, with the selected optimal parameters (4096, 0.0078) for the model. 166 weed plants were included in the 60 images captured on July 26th, and 104 were included in the 47 images taken on July 29th. The number of each weed species on either image set varied, which was decided upon availability. The identification results are shown in Table 4.4.

Table 4.4. Results for weed identification field test based on SVM.

(a) Test results for the SVM method without the adaptive median filter.

Date	Species	# of images	# of weed plants	Accuracy (%)	Average Time (s/Image)
07/26	Barnyardgrass	14	22	31.8%	1.09
	Waterhemp	17	38	57.9%	2.95
	Velvetleaf	29	106	85.8%	1.19
07/29	Barnyardgrass	15	22	50.0%	0.83
	Waterhemp	11	21	47.6%	1.74
	Velvetleaf	21	61	80.3%	1.03
	Total	107	270	71.1%	1.47

(b) Test results for the SVM method with the adaptive median filter.

Date	Species	# of images	# of weed plants	Accuracy (%)	Average Time (s/Image)
07/26	Barnyardgrass	14	22	31.8%	4.10
	Waterhemp	17	38	62.4%	5.76
	Velvetleaf	29	106	88.7%	4.20
07/29	Barnyardgrass	15	22	59.1%	3.76
	Waterhemp	11	21	61.9%	4.71
	Velvetleaf	21	61	85.3%	4.01
	Total	107	270	75.9%	4.42

Both of the SVM models have achieved a lower accuracy compared to the laboratory experiment, which was largely due to various interfering factors during field tests. The morphological difference between the weed plants grown in the field and greenhouse has a significant influence on the classification results, which were obtained using the SVM model based on the greenhouse image library. This is a common challenge for shaped-based machine vision systems due to inherent variability present in the biological realm. In addition, natural factors such as wind, rain, contamination, and time of day may change the original morphology of the plants. Other systematic factors such as the angles at which the images are taken, and segmentation flaws affecting the quality of the images, will also jeopardize successful identification.

In spite of these challenges, the proposed SVM method yielded a reasonably accurate

and consistent identification performance (see Figure 4.5) over the 107 sampling images and 270 weed plants over two imaging days. The original SVM model obtained an average classification accuracy of 71.1%, with a low processing time of 1.47 s. This was improved by adding the adaptive median filter, which raised the accuracy to 75.9%, by trading off to processing time, which was as much as 2.95s solely for the filtering procedure.

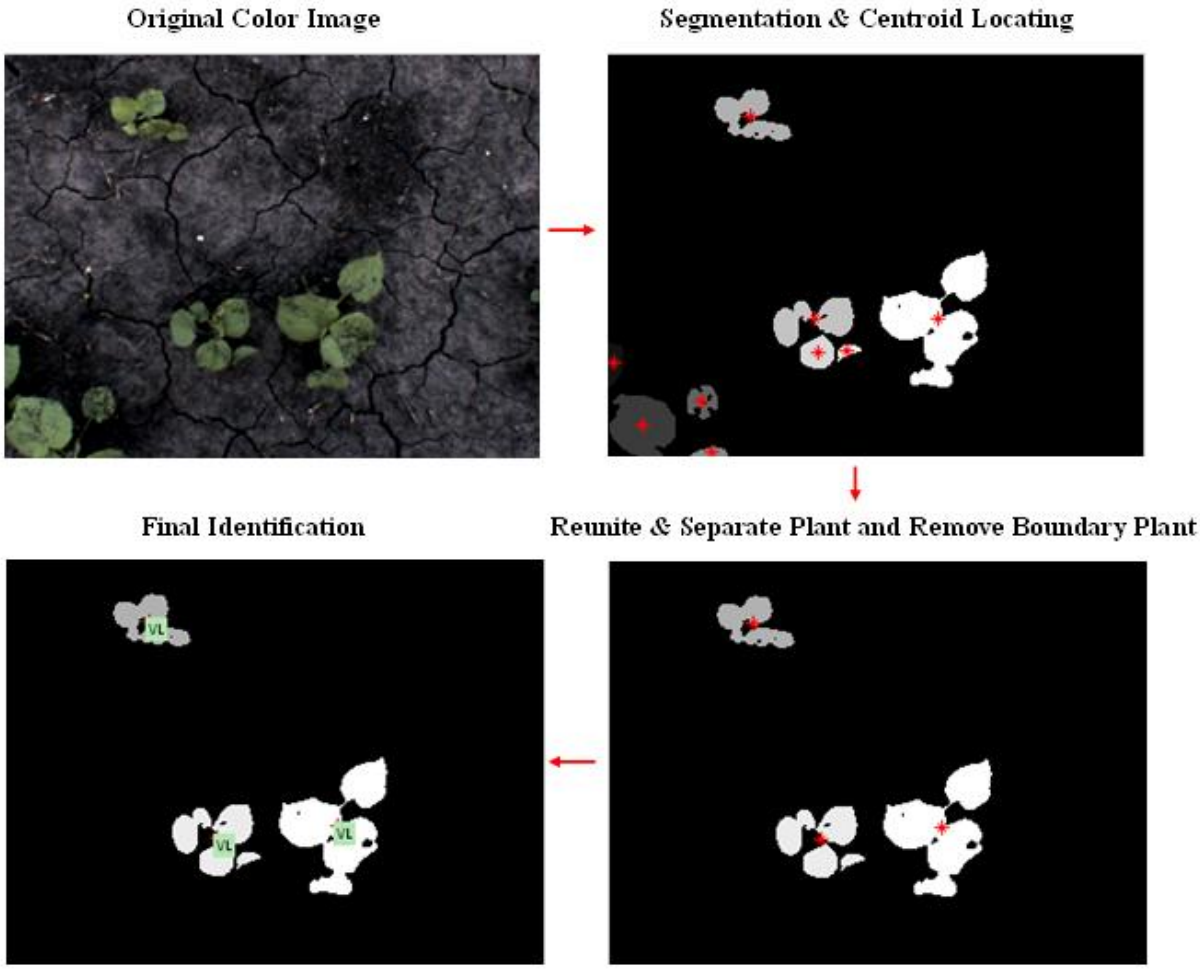


Figure 4.5. An example of successful weed identification procedure by the SVM method.

CHAPTER 5

Discussion and Recommendation for Future Research

An automated Pixelwise image segmentation algorithm and shape-based SVM classifier were developed to identify weeds in their early growth stage. The Pixelwise method is a modified Normalized Excess Green method, where refined thresholding criteria on individual RGB component are imposed based on advanced sampling to binarize color images. Seven features including the Ratio of Area to Length (ATL), Compactness (CMP), Elongation (ELG), Aspect (ASP), Logarithm of the Ratio of Height to Width (LHW), Ratio of Perimeter to Broadness (PTB), and Ratio of Length to Perimeter (LTP) were calculated in use of the geometrical parameters extracted from the binary images. The SVM model was subsequently constructed based on these morphological features as well as kernel parameters acquired from the cross validation procedure.

Six weed species were planted in the greenhouse, whose images were recorded on a daily basis using an automated image acquisition system throughout their early growth stages. 101 sample images were randomly selected among three imaging days to evaluate the performance of the segmentation algorithm. Comparative results showed that the Pixelwise method achieved a correct segmentation rate (CSR) of 96.3% and an incorrect segmentation rate (ISR) as low as 8.4%, which were both superior to the classic normalized excess green (NExG) and modified hue methods (Table 5.1).

Table 5.1. Comparison of segmentation performance among three segmentation methods

Sample Images	Methods	CSR (%)	ISR (%)	Elapsed Time (ms per image)
Total (101 samples)	NExG	92.1	39.6	187.8
	Modified Hue	87.3	18.9	287.6
	Pixelwise	96.3	8.4	226.6

To investigate the classification ability of the SVM method, laboratory and field experiments were conducted. In the laboratory experiments, 2,325 indoor weed images were acquired, among which 1,319 were used as the training set, while the remaining 1,006 images were treated as the test set for classification. The results showed that the SVM algorithm was highly effective in crop-weed, and grass-broadleaf weed classification, both with accuracy over 94% and errors below 6%. For each individual weed species, the accuracy ranged from 51.6% (waterhemp) to 94.4% (corn), which was largely dependent upon availability of plant images to construct the SVM model. In addition, the contrast between the high classification error rate for each weed species and the low percentage for each weed category (i.e. grass and broadleaf) indicated that most of the classification error occurred within-category rather than among-category. To determine the best time to identify weeds for the requirement of inter-row weeding, the relationship of the classification accuracy for the undesired weeds against time was investigated (see Figure 5.1). The SVM method achieved a consistently high classification performance during the entire early growth season of weeds, and weeds in the second week after emergence were classified with the highest accuracy of 91.9%.

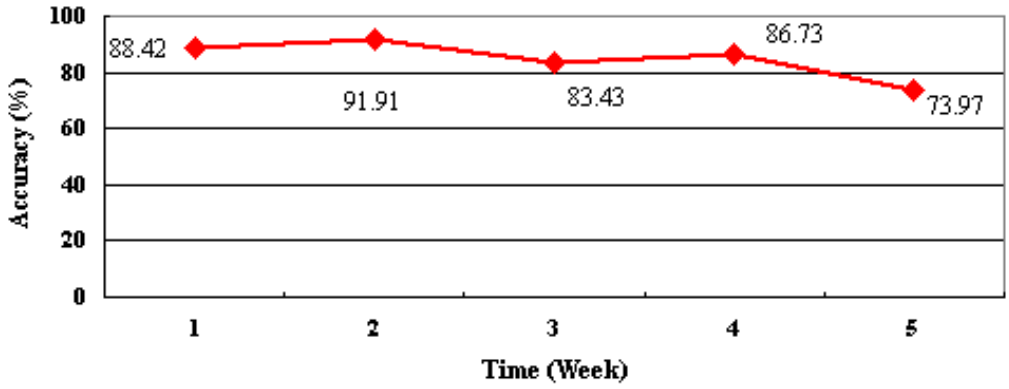
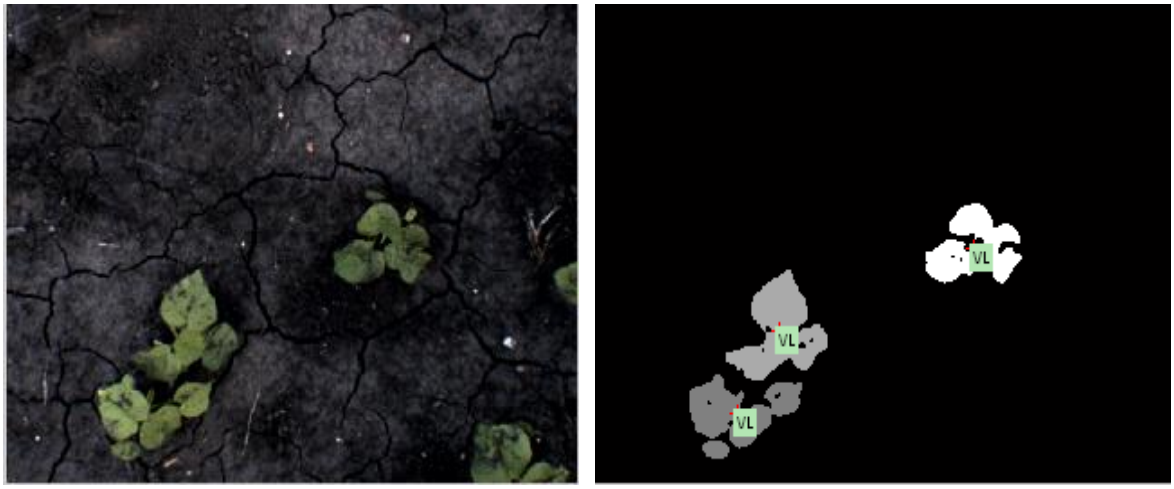


Figure 5.1. Classification accuracy of the SVM method for the undesired weed species at various growth stages after emergence.

During the field trials, 107 images of three weed species were collected using a digital camera mounted on a utility tractor. These images were processed offline by firstly applying the

Pixelwise segmentation method and subsequent exposure to the SVM model based on the entire greenhouse weed image library (except corn). Because more interfering factors were involved in the field scenario, two additional measures were applied: first, plant reunion was applied to reconnect parts of plant images that were disconnected in the imaging process. Secondly, boundary plants in each image were removed (Figure 5.2). Classification results revealed that the SVM classifier system was able to reach a classification rate of 71.1% at a processing time of 1.47s/image. With an optional adaptive median filter used to improve image quality, the accuracy could be raised to 75.9% at a cost of increasing the processing time to 4.42s/image.



(a) Original Image (b) Classifying individual weed plant

Figure 5.2. Field weed identification using the SVM method.

In summary, the proposed method in this research was effective in identifying various weed species. The classification accuracy was reasonable, considering the task was executed during the early growth season. Overall, the SVM algorithm has shown great potential in agricultural applications, especially in assisting in real-time weed identification systems. For future research, several recommendations are made upon this research:

- To construct a more solid SVM model, larger numbers of weed images are needed. An important reason for the low classification accuracy in waterhemp was that only five plants survived during the laboratory image acquisition procedure. As the natural characteristics of biology, the variation in leaf shape among individual weed plants (even within the same

species) is expected. Therefore, the larger the weed image library, the easier the SVM can find the support vectors and locate the hyperplanes to separate each species.

- For indoor image acquisition, the distance between the camera and the weed plant should be increased. Some weed species (such as corn) in this research grew rapidly within a short period, such that their leaves extended out of the imaging area of the camera. To assure a clear view of weeds when they are small and keeping the whole plant in the field of view when they grow, an appropriate distance needs to be determined or a platform with adjustable height or a zoom lens would be desirable.
- An effective and efficient image enhancement method needs to be implemented in the image processing procedure. Even though the Pixelwise algorithm performed accurate segmentation in this research, there existed segmentation errors in the form of “plant holes”, noise background pixels and leaf boundary distortions. The comparative field experiments with or without the adaptive median filter showed that images with higher quality increased the classification accuracy, albeit at the expense of processing time. While the real-time identification practice requires both accuracy and efficiency, measures such as erosion, dilation or other forms of filters that might improved image quality with a short processing time should be considered.
- A study needs to be conducted using an SVM model built on a field image library to identify weeds in the field. The shapes and development rate of the greenhouse plants compared to and field plants can vary, which attributed to the differences in cultivating environment such as soil, lighting, moisture, and cell depth. Hence, the SVM model constructed based on the laboratory weed images may not be able to reflect the real features of the field plants, and lower classification accuracy for the field test should be expected.
- A study should be conducted combining the use of other features such as spectral and textural features to identify weeds. An important disadvantage of the proposed algorithm is its inability to handle the occlusion problem, which is a common difficulty for all shape-based machine vision systems. Incorporating additional visual characteristics would

enable the system to detect among-species occlusion, rather than considering the overlapped leaves as a belonging to a single plant. On the other hand, by introducing more features to construct the SVM model, it is likely that the classification rate would be increased. This is because there would be more limitations to define the hyperplanes so results that are more accurate should be expected.

REFERENCES

Amador-Ramirez, M. D., R. G. Wilson, and A. R. Martin. 2001. Weed control and dry bean (*Phaseolus vulgaris*) response to in-row cultivation, rotary hoeing, and herbicides. *Weed Technology* 15(3): 429-436.

Angulo, C., X. Parra, and A. Catala. 2000. K-SVCR. A support vector machine for multi-class classification. In *Proceedings of the 11th European Conference on Machine Learning*, 31-38. Mantaras, R. and E. Plaza, eds. Springer-Verlag, Berlin Heidelberg.

Astrand, B. and A. Baerveldt. 2002. An agricultural mobile robot with vision-based perception for mechanical weed control. *Autonomous Robots* 13(1): 21-35.

Bailey, W. A., H. P. Wilson, and T. E. Hines. 2001. Influence of cultivation and herbicide programs on weed control and net returns in potato (*Solanum tuberosum*). *Weed Technology* 15(4): 654-659.

Bargen, K. V., G. E Meyer, D. A. Mortensen, S. J. Merritt, and D. M. Woebbecke. 1992. Red-near infrared reflectance sensor system for detecting plants. *Optics in Agriculture and Forestry* 1836: 231-238.

Belvins, R. L., R. Lai, J. W. Doran, G. W. Langdale, and W. W. Frye. 1998. Chapter 4: Conservation tillage for erosion control and soil quality. In *Advances in Soil and Water Conservation*, 51-68. F. J. Perce and W. W. Frye, eds. Ann Arbor, MI: Sleeping Bear Press.

Blanz, V., B. Scholkopf, H. Bulthoff, C. Burges, V. Vapnik, and T. Vetter. 1996. Comparison of view-based object recognition algorithms using realistic 3D models. In Artificial Neural Networks-ICANN'96, 251-256. Malsburg, C. von der, W. von Seelen, J. C. Vorbruggen, and B. Senhoff eds. Springer-Verlag, Berlin Heidelberg.

Boser, B. E., I. M. Guyon, and V. V. Vapnik. 1992. A training algorithm for optimal margin classifiers. In *Proceeding of the 5th Annual workshop on Computer learning theory*, 144-152. Pittsburgh, PA: ACM Press.

Bowman, G. 1997. *Steel in the field: A farmer's guide to weed management tools*. Burlington, VT: Sustainable Agriculture Publications, University of Vermont.

Brown, R. B., and S. D. Noble. 2005. Site-specific weed management: sensing requirements – what do we need to see? *Weed Science* 53(2): 252-258.

Buhler, D. D., J. D. Doll, R. T. Proost, and M. R. Visocky. 1995. Integrating mechanical weeding with reduced herbicide use in conservation tillage corn production system. *Agron J* 87: 507-512.

Buhler, D. D., M. Liebman, and J. J. Obrycki. 2000. Theoretical and practical challenges to an IPM approach to weed management. *Weed Science* 48(3): 274-280.

Burges, C. 1998. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery* 2(2): 121-167.

Burks, T. F., S. A. Shearer, and F. A. Payne. 2000. Classification of weed species using color texture features and discriminant analysis. *Transactions of the ASAE* 43(2): 441-448.

Burks, T. F., S. A. Shearer, J. R. Heath, and K. D. Donohue. 2005. Evaluation of neural-network classifiers for weed species discrimination. *Biosystems Engineering* 91(3): 293-304.

Chang, C. C. and C. J. Lin. 2001. LIBSVM: a library for support vector machines. Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Cheng, H. D., X. H. Jiang, Y. Sun, and Jingli Wang. 2001. Color image segmentation: advances and prospects. *Pattern Identification* 34(12): 2259-2281.

Cho, S. I., D. S. Lee, and J. Y. Jeong. 2002. Weed-plant discrimination by machine vision and artificial neural network. *Biosystems Eng* 82(3): 275-280.

Cisneros, J. J., B. H. Zandstra. 2008. Flame weeding effects on several weed species. *Weed Technology* 22(2): 290-295.

Cortes, C., and V. N. Vapnik. 1995. Support vector networks. *Machine Learning* 20(3): 273-297.

Cousens, R. D. and J. L. Woolcock. 1997. Spatial dynamics of weeds: and overview. In *Proceedings of the Brighton Crop Protection Conference – Weeds,* 613-618. Farnham, UK: British Crop Protection Council.

Dallyn, S. L. 1971. Weed control methods in potatoes. *American Journal of PotatoResearch* 48(4): 116-128.

Dickson, M. A., W. C. Bausch, and M. S. Howarth. 1995. Classification of a broadleaf weed, a grassy weed, and corn using image processing techniques. *Proceedings of SPIE* 2345: 297-305.

Donald, W. W., N. R. Kitchen, and K. A. Sudduth. 2001. Between-row mowing + banded herbicide to control annual weeds and reduce herbicide use in no-till soybean (*Glycine max*) and corn (*Zea mays*). *Weed Technology* 15(3): 576-584.

Donald, W. W. 2007. Between-row mowing systems control summer annual weeds in no-till grain sorghum. *Weed Technology* 21(2): 511-517.

Donaldson, D., T. Kiely, and A. Grube. 2002. Pesticides industry sales and usage, 1998 and 1999 market estimates. Washington, DC: U.S. Environmental Protection Agency, Office of Pesticide Programs.

ERS. 2007. Net farm income, up over 50 percent in 2007, sets new record. Farm Income and Costs: 2007 Farm Sector Income Estimates. Washington, DC: Agricultural Statistics Board and USDA. Available at: http://www.ers.usda.gov/Briefing/FarmIncome/2007incomeaccounts.htm

ERS. 2009. Corn: background. Briefing Rooms. Washington, DC: Agricultural Statistics Board and USDA. Available at: http://www.ers.usda.gov/Briefing/Corn/background.htm

Everitt, J. H., M. A. Alaniz, D. E. Escobar, and M. R. Davis. 1992. Using remote sensing to distinguish common (*Isocoma coronopifolia*) and Drummond goldenweed (*Isocoma drumondii*). *Weed Sci* 40: 621-628

Everitt, J. H., D. E. Escobar, R. Villarreal, M. A. Alaniz, and M. R. Davis. 1993. Canopy light reflectance and remote sensing of shin oak (*Quercus havardii*) and associated vegetation. *Weed Science* 41(2): 291-297.

Felton, W. L. and K. R. McCloy. 1992. Spot spraying. *Agric. Eng* 73: 9-12.

Feyaerts, F. and L. van Gool. 2001. Multi-spectral vision system for weed detection. *Pattern Identification Letters* 22(6): 667-674.

Forcella F. 2000. Rotary hoeing substitutes for two-thirds rate of soil-applied herbicide. *Weed Technology* 14(2): 298-303.

Franz, E., M. R. Gebhardt, and K. B. Unklesbay. 1991. Shape description of completely visible and partially occluded leaves for identifying plants in digital images. *Transactions of ASAE* 34(2): 673-681.

Furey, T. S., N. Cristianini, N. Duffy, D. W. Bednarski, M. Schummer, and D. Haussler. 2000. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics* 16(10): 906-914.

Gunn, S. 1998. Support vector machines for classification and regression. University of Southampton. Available at: http://www.svms.org/tutorials/Gunn1998.pdf

Gunn, S., M. Brown, and K. M. Bossley. 1997. Network performance assessment for neurofuzzy data modeling. In *Intelligent Data Analysis*, 313-323. X. Liu, P. Cohen, and M. Berthold, eds. Springer-Verlag, Berlin Heidelberg.

Gianessi, L. P. and N. P. Reigner. 2007. The value of herbicides in U.S. crop production. *Weed Technology* 21(2): 559-566.

Goel, P. K., S. O. Prasher., R. M., Patel, J. A. Landry, R. B. Bonnell, and A. A Viau. 2003. Classification of hyperspectral data by decision trees and artificial neural networks to identify weed stress and nitrogen status of corn. *Computers and Electronics in Agriculture* 39(2): 67-93.

Gonzalez, R. C., R. E. Woods, and S. L. Eddins. 2004. *Digital image processing using MATLAB®*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Guyer, D. E., G. E. Miles, L. D. Gaultney, and M. M. Schreiber. 1993. Application of machine vision to shape analysis in leaf and plant identification. *Transactions of the ASAE* 36(1): 163-171.

Guyer, D. E., G. E. Miles, M. M. Schreiber, O. R. Mitchel, and V. C. Vanderbilt. 1986. Machine vision and image processing for plant identification. *Trans. Am. Soc. Agric. Eng* 29(6): 1500-1507.

Hager, A. G., L. M. Wax, F. W. Simmons, and E. W. Stoller. 1997. Waterhemp management in agronomic crops. Champaign, IL: University of Illinois Bull X855.

Halford, C., A. S. Hamill, J. Zhang, and C. Doucet. 2001. Critical period of weed control in no-till soybean and corn. *Weed Technology* 15(4): 737-744.

Hanks, J. E. and J. L. Beck. 1998. Sensor-controlled hooded sprayer for row crops. *Weed Technology* 12(2): 308-314.

Hatcher, P. E. and B. Melander. 2003. Combining physical, cultural and biological methods: prospects for integrated non-chemical weed management strategies. *Weed Research* 43(5): 303-322.

Hemming, J. and T. Rath. 2001. Computer-vision-based weed identification under field conditions using controlled lighting. *J. agric. Engng Res* 78(3): 233-243.

Hooper, A. W., G. O. Harries and B. Ambler. 1976. A photoelectric sensor for distinguishing between plant material and soil. *Journal of Agricultural Engineering Research* 21: 145-155.

Hsu, C. W. and C. J. Lin. 2002. A comparison of methods for multiclass support vector machines. *IEEE Transactions on neural networks* 13(2): 415-425.

Jafari, A., S. S. Mohtasebi, H. E. Jahromi and M. Omid. 2006. Color segmentation scheme for classifying weeds from sugar beet using machine vision. Iranian Journal of Information Science & Technology 4(1): 1-12.

Karimi, Y., S. O. Prasher, A. Madani, and S. Kim. 2008. Application of support vector machine technology for the estimation of crop biophysical parameters using aerial hyperspectral observations. *Canadian Biosystems Engineering* 50: 713-720.

Karimi, Y., S. O. Prasher, R. M. Patel, and S. H. Kim. 2005. Application of support vector machine technology for weed and nitrogen stress detection in corn. *Computers and Electronics in Agriculture* 51(1): 99-109.

Keerthi, S. S. and C. J. Lin. 2003. Asymptotic behaviors of support vector machines with Gaussian kernel. *Neural Computation* 15(7): 1667-1689.

Kim, K. I., K. Junq, and H. J. Kim. 2002. Face recognition using kernel principal component analysis. *IEEE Signal Processing Letters* 9(2): 40-42.

Knezevic, S. Z., S. P. Evans, E. E. Blankenship, R. C. Van Acker, and J. L. Lindquist. 2002. Critical period for weed control: the concept and data analysis. *Weed Science* 50(6): 773-786.

Lass, L. W., D. C. Thill, B. Shafii, and T. S. Prather. 2002. Detecting spotted knapweed (*Centaurea maculosa*) with hyperspectral remote sensing technology. *Weed Technology*. 16(2):426-432.

Lass, L. W., H. W. Carson, and R. H. Callihan. 1996. Detection of yellow starthistle (*Centaurea solstitialis*) and common St. Johnswort (*Hypercum perforatum*) with multispectral digital imagery. *Weed Technology* 10(3): 466-474.

Lee, W. S., D. C. Slaughter, and D. K. Giles. 1999. Robotics weed control system for tomatoes. *Precision Agriculture* 1(1): 95-113.

Liu, F. H. and N. V. O'Connell. 2002. Off-site movement of surface-applied simazine from a citrus orchard as affected by irrigation incorporation. *Weed Science* 50(5): 672-676.

Lotz, L. A. P., R. Y. van der Weide, G. H. Horeman, and L. T. A. Joosten. 2002. Weed management and policies: from prevention and precision technology to certification of individual farms. In *Proceedings of the 12th EWRS Symposium*, 2-3. Wageningen, Papendal, The Netherlands.

Lyon, D. J., A. R. Martin, and R. N. Klein. 1999. Cultural practices to improve weed control in winter wheat. NebGuide: University of Nebraska – Lincoln. Available at: http://www.ianrpubs.unl.edu

Mahmoodi, S. and A. Rahimi. 2009. Estimation of critical period for weed control in corn in Iran. *Proceedings of world academy of science, engineering and technology* 37: 67-72.

Marchant, J. A. and C. M. Onyango. 2003. Comparison of a Bayesian classifier with a multilayer feed-forward neural network using the example of plant/weed/soil discrimination. *Computers and Electronics in Agriculture* 39(1): 3-22.

Marshall, E. J. P. 1988. Field-scale estimates of grass weed populations in arable land. *Weed Research* 28(3): 191-198.

Martin, S. G., R. C. Van Acker, and L. F. Friesen. 2001. Critical period of weed control in spring canola. *Weed Sci* 49(3): 326-333.

Mattsson, B., C. Nylander and J. Ascard. 1990. Comparison of seven inter-row weeders. V*eroff Bundesanst Agrarbiol Linz/Donau* 20: 91-107.

Medlin, C. R., D. R. Shaw, P. D. Gerard, and F. E. LaMastus. 2000. Using remote sensing to detect weed infestations in Glycine max. *Weed Science* 48(3): 393-398.

Menges, R. M., P. R. Nixon, and A. J. Richardson. 1985. Light reflectance and remote sensing of weeds in agronomic and horticultural crops. *Weed Sci* 33(4): 569-581.

Meyer, G. E. and J. C. Neto. 2008. Verification of color vegetation indices for automated crop imaging applications. *Computer and electronics in agriculture* 63(2): 282-293.

Meyer, G. E., J. C. Neto, D. D. Jones, and T. W. Hindman. 2003. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and Electronics in Agriculture* 42(3): 161-180.

Meyer, G. E., T. W. Hindrman, and K. Lakshmi. 1998. Machine Vision detection parameters for plant species identification. In *Proceedings of SPIE*, 327-335. Meyer, G. E. and J. A. Deshazer, eds. Bellingham, WA.

Mount Pleasant, J., R. F. Burt, and J. C. Frisch. 1994. Integrating mechanical and chemical weed management in corn (*Zea mays*). *Weed Technol* 8(2): 217-223.

Mortensen, D. A. and J. A. Dieleman, and G. A. Johnson. 1998. Weed spatial variation and weed management. In *Proceedings of Precision Weed Management in Crops and Pasture*, 14-19. Medd, R. W. and J. E. Pratley, eds. Wagga Wagga, Australia.

Mulder, T. A. and J. D. Doll. 1993. Integrating reduced herbicide use with mechanical weeding in corn (*Zea mays*). *Weed Technol* 7(2): 382-389.

Nalewaja, J. D. 1999. Cultural practices for weed resistance management. *Weed Technology* 13(3): 643-646.

Napier, T. L., M. Tucker, and S. McCarter. 2000. Adoption of conservation production systems in three Midwest watersheds. J. *Soil Water Conserv* 55(2): 123-134.

NASS. 2000 – 2004. Agricultural Chemical Usage – Vegetables. Washington, DC: Agricultural Statistics Board and USDA. http://usda.mannlib.cornell.edu/MannUsda/viewDocumentInfo.do?documentID=1561

Neto, J. C., G. E. Meyer, D. D. Jones, and A. K. Samal. 2006. Plant species identification using Elliptic Fourier leaf shape analysis. *Computer and Electronics in Agriculture* 50(2): 121-134.

Otsu, N. 1979. A threshold selection method from gray-level histograms. *IEEE Trans. On Systems Man and Cybernetics* 9(1): 63-66.

Peng, H. C., C. Ding, and F. H. Long. Minimum redundancy maximum relevance feature selection. *IEEE Intelligent Systems* 20(6): 70-71.

Peng, H. C., F. H. Long, and C. Ding. 2005. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(8): 1226-1238.

Perez, A. J., F. Lopez, J. V. Benlloch, and S. Christensen. Colour and shape analysis techniques for weed detection in cereal fields. *Computers and Electronics in Agriculture* 25: 197-212.

Philipp, I. and T. Rath. 2002. Improving plant discrimination in image processing by use of different colour space transformations. *Computer and Electronics in Agriculture* 35: 1-15.

Schmidt, M. and H. Gish. 1996. Speaker identification via support vector classifiers. In *Processing of 1996 IEEE International Conference on Acoustics, Speech, and Signal*, 105-108. Atlanta, GA, USA.

Scholkopf, B., and Smola, A. 2002. *Learning with kernels: support vector machines, regularization, optimization and beyond*. Cambridge, MA: MIT Press,

Shearer, S. A. and P. T. Jones. 1991. Selective application of post-emergence herbicides using photo-electrics. *Transactions of the ASAE* 34(4): 1661-1666.

Shropshire, G. J., K. Von Bargen, and D. A. Mortensen. 1990. Optical reflectance sensor for detecting plants. *Proceedings of SPIE (The International Society for Optical Engineering): Optics in Agriculture* 1379: 222-235.

Sogaard, H. T. 2005. Weed classification by active shape models. *Biosystems Engineering* 91(3): 271-281.

Spliid, N. H. and B. Koeppen. 1998. Occurrence of pesticides in Danish shallow ground water. *Chemosphere* 37(7): 1307-1316.

Stall, W. M. 2009. Weed Control in Sweet Corn. HS 197: Gainesville, University of Florida. Available at: http://edis.ifas.ufl.edu/pdffiles/WG/WG03800.pdf

Swanton, C. J., K. J. Mahoney, K. Chandler, and R. H. Gulden. 2008. Integrated weed management: knowledge-based weed management systems. *Weed Science* 56: 168-172.

Steward, B. L. and L. F. Tian. 1999. Machine-vision weed density estimation for real-time, outdoor lighting conditions. *Transactions of the ASAE* 42(6): 1897-1909.

Tang, L., L. Tian, and B. L. Steward. 2003. Classification of broadleaf and grass weeds using gabor wavelets and an artificial neural network. *Transactions of the ASAE* 46(4): 1247-1254.

Tang, L., L. F. Tian, and B. L. Steward. 2000. Color image segmentation with genetic algorithm for in-field weed sensing. *Transactions of the American Society of Agricultural Engineers* 43(4): 1019-1027.

Tang, L., L. F. Tian, B. L. Steward and J. F. Reid. 1999. Texture-based weed classification using gabor wavelets and neural network for real-time selective herbicide applications. ASAE Paper No. 99-3036. St. Joseph, Mich.: ASAE.

Tian, L., D. C. Slaughter, and R. F. Norris. 1997. Outdoor field machine vision identification of tomato seedlings for automated weed control. *Transactions of the ASAE* 40(6): 1761-1768.

Tian, L. F., J. F. Reid, and J. W. Hummerl. 1999. Development of a precision sprayer for site-specific weed management. *Trans. Am. Soc. Agric. Eng* 42: 893-900.

Tian, L. F. and D. C. Slaughter. 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture* 21(3): 153-168.

Thorp, K. R., and L. F. Tian. 2004. A review on remote sensing of weeds in agriculture. *Precision Agriculture* 5(5): 477-508.

Timmermann, C., R. Gerhards, P. Krohmann, M. Sokefeld, and W. Kuhbauch. 2001. The economical and ecological impact of the site-specific weed control. In *Third European Conference on Precision Agriculture*, 563-568.G. Grenier and S. Blackmore, eds. Montpelier, France: Agro Montpellier.

Tong, S. and D. Koller. 2002. Support vector machine active learning with applications to text classification. *The Journal of Machine Learning Research* 2: 45-66.

Vapnik, V. N. 1995. *The Nature of statistical learning theory*. Springer Verlag, New York.

Vrindts, E., J. De Baerdemeaeker, and H. Ramon. 2002. Weed detection using canopy reflection. *Precision Agriculture* 3(1): 63-80.

Wan, V. and W. M. Campbell. 2000. Support vector machines for speaker verification and identification. In *Proceedings of the 2000 IEEE signal Processing Society Workshop*, 775-784. Sydney, NSW, AU.

Wang, N., N. Zhang, D. Peterson, and F. Dowell. 2000. Design of an optical weed sensor using plant spectral characteristics. *Biological quality and precision agriculture II, Proc. SPIE* 4023: 63-72.

Williams, M. M. 2006. Planting date influences critical period of weed control in sweet corn. *Weed Science* 54(5): 928-933.

Woebbecke, D. M., G. E. Meyer, K. Von Bargen, and D. A. Mortensen. 1995. Shape features for identifying young weeds using image analysis. *Trans. Am. Soc. Agric. Eng* 38(1): 271-281.

Wu, L. and Y. Wen. 2009. Weed/corn seedling recognition by support vector machine using texture features. *African Journal of Agricultural Research* 4(9): 840-846.

Yang, C. C., S. O. Prasher, and P. K. Goel. 2004. Differentiation of crop and weed by decision-tree analysis of multi-spectral data. *Transactions of the ASAE* 47(3): 873-879.

Zhong, D. X., J. Novais, T. E. Grift, M. Bohn, and J. Han. 2009. Maize root complexity analysis using a support vector machine method. *Computers and Electronics in Agriculture* 69(1): 46-50.

Zwiggelarr R. 1998. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. *Crop Protection*. 17(3):189-206.

APPENDIX

A1. MATLAB® Program for Laboratory Image Acquisition

AcquireImage.m

```
% AcquireImage records indoor plant images using the automated image
% acquisition system based on the XY positioning device. Plants are
% recorded according to inputs of the record date and name of the plant.
% Author : Chufan Lin
% Date   : Feb 21st, 2009

clc
close all
clear all

% reset imaging and instrument devices
imaqreset
instrreset

% ===== Setting camera properties =====
vid = videoinput('dcam',1,'Y422_640x480');
triggerconfig(vid,'manual');
set(vid,'FramesPerTrigger',1);
set(vid,'TriggerRepeat',inf);
start(vid);

% ===== Setting up Serial Communication =====
s = serial('COM4');
set(s,'Baudrate',9600);
set(s,'Databits',8);
set(s,'Parity','none');
set(s,'StopBits',1);
set(s,'Terminator','CR');

% adjust speed of the motors
fopen(s)
fprintf(s,'@00 rate 150');
fprintf(s,'@01 rate 150');
```

```
% ===== Timer Setup =====
t = timer('StartDelay',2,'TimerFcn'...
    ,'disp([''Taking Image of Pot No. '' num2str(PotNum)])');
% @00 long shaft motor, horizontal shaft, Col;
% @01 short shaft motor, vertical shaft, Row;

% ===== Image Acquisition =====
Date = input('Type in recording date, eg. 0216: ','s');
Type = input('Type in plant species, eg. waterhemp: ','s');

% Make Folder for recording
folder = ['C:\Documents and Settings\lin33\Desktop\Research\MatLab®
\Final\image ' num2str(Date)];
pass = exist(num2str(folder),'file');
if pass == 0
   mkdir(['image ' num2str(Date)]);
end
dir = ['C:\Documents and Settings\lin33\Desktop\Research\MatLab®
\Final\image ' num2str(Date) '\'];

while Type ~= '0'
    Col = 1;
    Row = 1;
    PotNum = 1;
    % Distance from one pot to another
    MoveStep = 690;
    while Col <= 3
        while Row <= 6
          start(t);
          wait(t);
          trigger(vid);
          image = getdata(vid,1);
          image = ycbcr2rgb(image);
          imshow(image,[])
          % Chopping the image
          figure,imshow(image(110:370,190:450,1:3))
          title(['figure' num2str(PotNum)])
         % Save the image
          imwrite(image, [num2str(dir) num2str(Type) ...
                  num2str(PotNum) '.bmp'], 'bmp');
```

```
          Row    = Row + 1;
          PotNum = PotNum + 1;
          pause(2)
            if Row < 7
                fprintf(s,['@00 rmov ' num2str(MoveStep)])
               pause(5)
               disp('done moving to next row')
            end
        end
        MoveStep = MoveStep*(-1);
        Col = Col + 1;
        Row = 1;
        if Col < 4
            fprintf(s,['@01 rmov ' num2str(abs(MoveStep))])
            pause(12)
            disp('done moving to next column')
        end
    end

% ====== Moving Home ======
    pause(2)
    clc
    disp('Finish image aquisition, moving home')
    fprintf(s,'@00 rmov -3450')
    fprintf(s,'@01 rmov -1380')

    close all
    pause(25)
    disp('Procedure done')
    Type = input('Type in next weed species, type ''0'' to quit: ','s');
end

clc
disp('Work Done')
stop(vid);
fclose(s);

% ======== End ========
```

A2. MATLAB® Implementation of the Normalized Excess Green Method

NExG.m

```
% NExG.m transfers a color image into gray-scale using the
% normalized excess green index. Then threshold values are selected using
% automatic OTSU's method or manual pick-up value based on observation of
% the NExG histogram. The step is accomplished by the callback function
% 'getthreshold.m'.
% Author: Chufan Lin
% Date  : Feb 27th, 2009

clc
clear all
close all
Index = 1;
next  = 1;

while Index ~= 0
   % Open a dialog and select an image file
[FileName,FilePath,Index] = uigetfile('*.bmp', 'Open Imagefile ')
   if Index == 0
      disp('Procedure Done')
      break;
   end
   f  = imread([num2str(FilePath) FileName]);
   [a,b,c] = size(f);
   f  = f(100:a-100,180:(b-180),1:c);   % for 6mm lens
   figure, imshow(f,[]), title('Original Image')
   f1 = f;

% Normalized Excess Green Method
R = im2double(f(:,:,1));
G = im2double(f(:,:,2));
B = im2double(f(:,:,3));
[m,n] = size(B);
Den = R+G+B;
% to avoid R = G = B = 0, such that denominator equals to zero
for i = 1:m
   for j = 1:n
      if Den(i,j) == 0
```

```
            Den(i,j) = 0.001;
        end
    end
end

% obtain chromatic coordinates
r = R./Den;
g = G./Den;
b = B./Den;

% normalized excess green representation
ex_g = 2.0*g - r - b;
ex_g = im2uint8(ex_g);

% use a 3 by 3 median filter for eliminating ‘Pepper’ noise
for i = 1:3
    ex_g = medfilt2(ex_g,[3 3]);
end

figure,imshow(ex_g,[]),title('NExG grayscale')
% call function ‘getthreshold’ to figure out both of the threshold values
 [Otsu, Hist] = getthreshold(ex_g);

% figure generated by OTSU's Method
ExG2  = im2bw(ex_g,Otsu/255);
for i = 1:3
     ExG2 = medfilt2(ExG2,[3 3]);
end
figure,imshow(ExG2,[]), title('Exg with Otsu')

% figure generated based on the observation of the histogram
ExG3  = ex_g;
for i = 1:m
    for j = 1:n
        if ExG3(i,j)  <= Hist
            ExG3(i,j)   = 0;
        else ExG3(i,j) = 1;
        end
    end
end
for i = 1:3
```

```
    ExG3 = medfilt2(ExG3,[3 3]);
end
figure,imshow(ExG3,[]), title('ExG with Histogram')
end
```

Getthreshold.m

```
function [Otsu, Hist] = getthreshold(image)
% Getthreshold.m implements different thresholding methods,
% including OTSU and Histogram methods
% Author: Chufan Lin
% Date  : Feb 27th, 2009

%  OTSU's Method: pick a value that maximizes the between-class
%  variance
Otsu = graythresh(image)*255;

% Based on Histogram: threshold value equals to the first valley
% value of the histogram
figure, imhist(image),title('adjusted Histogram')
[x,y] = ginput(1);
Hist = x;
```

A3. MATLAB® Implementation of the Modified Hue Method

ModifiedHue.m

```
% ModifiedHue.m converts the weed images from RGB into HSI space, among
% which Hue chanel is modified and used for plant segmentation. The
% thresholding method is based on visual judgment and empirical value over
% a couple of sample images.
% Author: Chufan Lin
% Date  : Feb 28th, 2009

clc
clear all
close all
scrsz = get(0,'ScreenSize');
Index = 1;
next  = 1;

while Index ~= 0
 % Open a dialog and select an image file
[FileName,FilePath,Index] = uigetfile('*.bmp', 'Open Imagefile ')     if
    Index == 0
        disp('Procedure Done')
        break;
    end
    f = imread([num2str(FilePath) FileName]);
    [a,b,c] = size(f);
    f = f(100:a-100,180:(b-180),1:c);     % for 6mm lens
    figure('Position',[0 -50 scrsz(3) scrsz(4)]),
   subplot(2,2,1),imshow(f,[]), title('Original Image')

f = im2double(f);
R = f(:,:,1);
G = f(:,:,2);
B = f(:,:,3);
Den  = R+G+B;
[m,n]=size(R);

% Convert to the hue component
num1 = 0.5 *((R-G)+(R-B));
Den1 = sqrt((R-G).^2+(R-B).*(G-B));
```

```
theta1 = acos(num1./(Den1+eps));
H1 = theta1;
H1(B>G) = 2*pi-H1(B>G);
H1 = H1/(2*pi);

% Convert to the saturation component
num3 = min(min(R,G),B);
den3 = R+G+B;
den3(den3 == 0) = eps;
S1 = 1 - 3.*num3./den3;
H1(S1==0)=0;
H1 = im2uint8(H1);
H2 = H1;
% empirical threshold values originated from sample images
threshold_low  = 70;
threshold_high = 120;
for i = 1:m
    for j = 1:n
        if H2(i,j)<= threshold_high && H2(i,j)>=threshold_low
             H2(i,j)=1;
        else H2(i,j)=0;
        end
    end
end

for i = 1:5
    H2 = medfilt2(H2,[3 3]);
end
subplot(2,2,3),imshow(H2,[]),title('B/W Modified Hue')

subplot(2,2,2),imhist(H1),title('Histogram,Modified Hue')
satisfy = 1;
% change the color range if the segmentation isn't satisfactory
while satisfy == 1
         H3 = H1;
    satisfy = input('press 1 to change the range, press 0 to go to next image :
              ');
    if satisfy == 0
       break
    else
    subplot(2,2,2),imhist(H1),title('Histogram,Modified Hue')
```

```
            [x,y] = ginput(2);
    threshold_low  = min(x(:));
    threshold_high = max(x(:));
        value_low  = 1:1000;
       value_high = 1:1500;
       hold on
        plot(threshold_low,value_low,'g',threshold_high,value_high,'r')
text(threshold_low-20,1300,['threshold1 = ' num2str(threshold_low)]);
text(threshold_high-20,1800,['threshold2 = ' num2str(threshold_high)]);
       hold off

    for i = 1:m
     for j = 1:n
        if  H3(i,j)<= threshold_high && H3(i,j)>=threshold_low
             H3(i,j)=1;
        else H3(i,j)=0;
        end
     end
    end

     for i = 1:5
        H3 = medfilt2(H3,[3 3]);
    end
    subplot(2,2,4),imshow(H3,[]),title('B/W Modified Hue')
    end
end

     % terminate the program or proceed into next image
    next = input('next image? press Enter: ','s');
    if next == '0'
       break
    else
       close all
       pause(0.2)
       continue
    end

end
```

A4. Explore the Interrelationship among the RGB Channels

RGBrelation.m

```
% RGBrelation.m explore the interrelationship of the RGB components
% through manual-thresholding using Normalized Excess Green Method.
% In applying the NExG Method, first turn the RGB info into
% chromaticity coordinates, then using manually selected threshold value
% for segmentation between the plants and soil. After Best segmentation
% result is achieved, calculate the mean and ratios using the R, G, B value
% of the soil and plant pixels respectively. In the end, the acquired values
% will be saved in an Excel file % for analysis.
% Author: Chufan Lin
% Date  : April 3rd,2009

clc
clear all
close all

Index = 1;
scrsz = get(0,'ScreenSize');
filepath = 'C:\Documents and Settings\lin33\Desktop\lately used files\';
filename = 'RGBrelation.xls';
filedir   = [filepath filename];

% initiate communication between MatLab®  and excel
channelactivity = 0;
channel = ddeinit('excel',filedir);
if channel == 0
    error(['Please Open File: ' num2str(filename)]);
else
    channelactivity = 1;
end

% when the communication channel is active
while channelactivity ~= 0
    if channelactivity == 0
        disp('===== Procedure Done =====')
        pause(0.2)
        break
```

```
    end
    ddepoke(channel,'r1c1:r1c1', 'IamgeID')
    ddepoke(channel,'r1c2:r1c2', 'R_Soil' );
    ddepoke(channel,'r1c3:r1c3', 'G_Soil' );
    ddepoke(channel,'r1c4:r1c4', 'B_Soil' );
    ddepoke(channel,'r1c5:r1c5', 'R_Plant' );
    ddepoke(channel,'r1c6:r1c6', 'G_Plant' );
    ddepoke(channel,'r1c7:r1c7', 'B_Plant' );
    ddepoke(channel,'r1c8:r1c8', 'R_B_Soilratio' );
    ddepoke(channel,'r1c9:r1c9', 'G_B_Soilratio' );
    ddepoke(channel,'r1c10:r1c10', 'G_R_Soilratio' );
    ddepoke(channel,'r1c11:r1c11', 'R_B_Plantratio' );
    ddepoke(channel,'r1c12:r1c12', 'G_B_Plantratio' );
    ddepoke(channel,'r1c13:r1c13', 'G_R_Plantratio' );

    row = 2;
   while Index ~= 0
   % Open a dialog and select an image file
[FileName,FilePath,Index] = uigetfile('*.bmp', 'Open Imagefile ');
    if Index == 0
        channelactivity = 0;
        break;
    end
    f = imread([num2str(FilePath) FileName]);
    figure('Position',[1 1 scrsz(3) scrsz(4)]), subplot(2,2,1)
    imshow(f,[]), title('Original Image')

    % chop the image to remove the edges of the container
        [x,y] = ginput(4);
        hor_min = round(min(y));
        hor_max = round(max(y));
        ver_min = round(min(x));
        ver_max = round(max(x));
        if hor_min < 1
            hor_min = 1;
        end
        if hor_max > a
            hor_max = a;
        end
        if ver_min < 1
            ver_min = 1;
```

```
        end
        if ver_max > b
            ver_max = b;
        end
    f = f(hor_min:hor_max,ver_min:ver_max,:);
    [a,b,c] = size(f);
    hold off
    subplot(2,2,1),imshow(f,[]), title('Original Image')

% Normalized Excess Green Method
    R = im2double(f(:,:,1));
    G = im2double(f(:,:,2));
    B = im2double(f(:,:,3));
    [m,n] = size(B);
    Den = R+G+B;

% to avoid R = G = B = 0, such that denominator equals to zero
for i = 1:m
    for j = 1:n
        if Den(i,j) == 0
            Den(i,j) = 0.001;
        end
    end
end

% obtain chromatic coordinates
r = R./Den;
g = G./Den;
b = B./Den;

% normalized excess green representation
ex_g  = 2.0*g - r - b;
ex_g  = im2uint8(ex_g);
for i = 1:3
    ex_g = medfilt2(ex_g,[3 3]);
end
subplot(2,2,2), imhist(ex_g),title('NExG histogram')

% Initial Thresholding using OTSU's Method
Otsu = graythresh(ex_g)*255;
ExG2 = im2bw(ex_g,Otsu/255);
```

```
hold on
value = 1:3000;
plot(Otsu,value,'r');
text(Otsu,3500,['Otsu = ' num2str(Otsu)]);
hold off
for  i = 1:3
  ExG2 = medfilt2(ExG2,[3 3]);
end
subplot(2,2,3),imshow(ExG2,[]), title('Exg with Otsu')

% Thresholding using Histogram
Hist =1;
while Hist ~= 0
    ExG3 = ex_g;
    Hist = input('Enter threshold, press 0 to process next image : ');
    if Hist == 0
       break
    else
    for i = 1:m
        for j = 1:n
            if  ExG3(i,j) <= Hist
                 ExG3(i,j) = 0;
            else ExG3(i,j) = 1;
            end
        end
     end
     for i = 1:3
         ExG3 = medfilt2(ExG3,[3 3]);
     end
     [ExG3_Num, num] = bwlabel(ExG3,8);
     % remove remaining noise, in the view of object size
     for k = 1:num
         if numel(find(ExG3_Num == k)) < 50
              ExG3_Num(ExG3_Num == k) = 0;
         end
     end
subplot(2,2,4),imshow(ExG3_Num,[]), title('NExG with Histogram')
     end
end

% discriminate plant pixels and non-plant pixels; then calculate the mean
```

```
% and different ratios among R,G,B components
PlantIndex = find(ExG3_Num ~= 0);
SoilIndex  = find(ExG3_Num == 0);
R1 = im2uint8(R);
G1 = im2uint8(G);
B1 = im2uint8(B);

MeanPlantRed   = mean2(R1(PlantIndex));
MeanPlantGreen = mean2(G1(PlantIndex));
MeanPlantBlue  = mean2(B1(PlantIndex));
MeanSoilRed    = mean2(R1(SoilIndex));
MeanSoilGreen  = mean2(G1(SoilIndex));
MeanSoilBlue   = mean2(B1(SoilIndex));

R_B_PlantRatio = MeanPlantRed/MeanPlantBlue;
G_B_PlantRatio = MeanPlantGreen/MeanPlantBlue;
G_R_PlantRatio = MeanPlantGreen/MeanPlantRed;
R_B_SoilRatio  = MeanSoilRed/MeanSoilBlue;
G_B_SoilRatio  = MeanSoilGreen/MeanSoilBlue;
G_R_SoilRatio  = MeanSoilGreen/MeanSoilRed;

ddepoke(channel,['r' num2str(row) 'c1:r' num2str(row) 'c1'],
FileName)
ddepoke(channel,['r' num2str(row) 'c2:r' num2str(row) 'c2'],
MeanSoilRed);
ddepoke(channel,['r' num2str(row) 'c3:r' num2str(row) 'c3'],
MeanSoilGreen);
ddepoke(channel,['r' num2str(row) 'c4:r' num2str(row) 'c4'],
MeanSoilBlue);
ddepoke(channel,['r' num2str(row) 'c5:r' num2str(row) 'c5'],
MeanPlantRed);
ddepoke(channel,['r' num2str(row) 'c6:r' num2str(row) 'c6'],
MeanPlantGreen);
ddepoke(channel,['r' num2str(row) 'c7:r' num2str(row) 'c7'],
MeanPlantRed);
ddepoke(channel,['r' num2str(row) 'c8:r' num2str(row) 'c8'],
R_B_SoilRatio);
ddepoke(channel,['r' num2str(row) 'c9:r' num2str(row) 'c9'],
G_B_SoilRatio);
ddepoke(channel,['r' num2str(row) 'c10:r' num2str(row) 'c10'],
G_R_SoilRatio);
```

```
ddepoke(channel,['r' num2str(row) 'c11:r' num2str(row) 'c11'],
R_B_PlantRatio);
ddepoke(channel,['r' num2str(row) 'c12:r' num2str(row) 'c12'],
G_B_PlantRatio);
ddepoke(channel,['r' num2str(row) 'c13:r' num2str(row) 'c13'],
G_R_PlantRatio);

    next = input('next image? press Enter: ');
    if next == 0
        channelactivity = 0;
        break
    else
        close all
        disp('================================')
        row = row + 1;
        pause(0.2)
        continue
    end
    end
end
close all
```

A5. MATLAB® Implementation of the Pixelwise Segmentation method

PixelwiseSeg.m

```
% PixelwiseSeg.m used the multiple criteria based on the R,G,B
% interrelationship for image segmentation. The parameters for
% those criteria originates from sampling in use of the program in A4.
% Author: Chufan Lin
% Date  : April 20th,2009

clc
clear all
close all

Index = 1;
scrsz = get(0,'ScreenSize');
while Index ~= 0
% Open a dialog and select an image file
[FileName,FilePath,Index] = uigetfile('*.bmp', 'Open Imagefile ');
    if Index == 0
        disp('Procedure Done')
        break;
    end
    f = imread([num2str(FilePath) FileName]);
    % chop the image to remove the edges of the container
    [a,b,c] = size(f);
    f = f(120:a-120,200:(b-200),1:c);
    figure('Position',[1 1 scrsz(3) scrsz(4)]), subplot(2,2,1)
    imshow(f,[]), title('Original Image')

    f = double(f);
    R = f(:,:,1);
    G = f(:,:,2);
    B = f(:,:,3);
    [m,n]  = size(B);
    Image  = zeros(m,n);
    Image1 = zeros(m,n);

for i = 1:m
```

```
    for j = 1:n
        if   G(i,j) > 1.0*B(i,j) && G(i,j) > 1.0*R(i,j) && R(i,j) >0.6*B(i,j)
             && G(i,j)>30
             Image(i,j) = 1;
        else Image(i,j) = 0;
   end
end
subplot(2,2,2)
imshow(Image,[]),title('Pixelwise threholding ')

% median filter remove small salt & pepper noise
 for i = 1:3
     Image = medfilt2(Image,[3 3]);
 end
 [Image_Num, num] = bwlabel(Image,8);

 % size filter remove remaining noise, in the view of object size
 for k = 1:num
    if numel(find(Image_Num == k)) < 350
        Image_Num(Image_Num == k)   = 0;
    else
        Image_Num(Image_Num == k)   = 255;
    end
 end
 subplot(2,2,3)
 imshow(Image_Num,[]),title('final segmentation')

   next = input('next image? press Enter: ');
   if next == 0
      channelactivity = 0;
      break
   else
      close all
      disp('==================================')
      pause(0.2)
      continue
   end
end
close all
```

A6. MATLAB® Program for Acquiring Field Images

FieldImgAcq.m

```
% FieldImgAcq waits for "Enter" to take an image and save it during the
% field experiment.
% Author: Chufan Lin
% Date  : June 20th, 2009.
clc
close all
clear all

% Disconnect and Delete imaging and instrument objects
imaqreset
% Initiate camera & set camera properties
vid = videoinput('dcam',1,'Y422_800x600');
triggerconfig(vid,'manual');
set(vid,'FramesPerTrigger',1);
set(vid,'TriggerRepeat',inf);
camera = getselectedsource(vid);
set(camera,'shutterMode','Manual');
set(camera,'AutoExposure',40);
set(camera,'Shutter',1200);

preview(vid)
start(vid);
figure

% Image Aquisition
Date = input('Type in recording date, eg. 0216: ','s');
% Make Folder
folder = ['C:\Documents and Settings\lin33\Desktop\Experiment\image '
num2str(Date)];
mkdir(num2str(folder));
dir = ['C:\Documents and Settings\lin33\Desktop\Experiment\image '
num2str(Date) '\'];

Type = 1;
t    = timer('StartDelay',1,'TimerFcn'...
```

```
        ,'disp([''Taking Image of Pot No. '' num2str(Position)])');

while Type ~= '0'
      Type = input('Type in name of the plant species: ','s');
    if Type == '0'
       disp(' ===== Procedure Done ===== ')
       pause(0.8)
       clc
       break
     end
     mkdir([num2str(dir) 'image ' num2str(Type)]);
     dir2 = ([dir 'image ' num2str(Type) '\']);
     Position = 1;
 while  Position ~= 0
       Position = input(['Type in the Position number of the current weed
                  species -- ' num2str(Type) ': ' ]);
       if Position == 0
          disp([' ===== Image Acquisition Done for ' num2str(Type) ' =====
                '])
          pause(0.8)
          clc
          break
       end

          start(t)
          wait(t)
          trigger(vid);
          image = getdata(vid,1);
          image = ycbcr2rgb(image);
          imshow(image,[])
          title(['image triggered at position: ' num2str(Position)
                 ' for plant: ' num2str(Type)])
          imwrite(image, [num2str(dir2) 'Position '
                  num2str(Position)'.bmp'], 'bmp');
  end
end

closepreview
stop(vid);
close all
% ======== End ========
```

A7. MATLAB® program for Weed Identification using the SVM Method

FieldWeedIdentify.m

```
% FieldWeedIdentify.m uses the all of the weed images obtained from
% laboratorial experiment to construct the SVM model. And such model is
% used to classify the weed image captured in the field.
% Those weed image include:
% 2, Barnyardgrass
% 3, Crabgrass
% 4, Lambsquarters
% 5, Velvetleaf
% 6, Waterhemp
% notice: 1, corn is not include
% Author: Chufan Lin
% Date  : Oct 17th, 2009

close all
clear all
clc

nocorn = xlsread('C:\ClassifiedImages\Fieldimages\wocorn.xls');
label1 = nocorn(:,16);
inst1  = nocorn(:,9:15);

% [bestcv1,bestc1,bestg1] = CVselect(label1,inst1);
% returned CVs and best (c,g):
% for nocorn data: best (c,g) = (4096, 0.0078), CV = 85.34%;
% Accuracy = 91.53%;
load model1.mat

time  = [];
scrsz = get(0,'ScreenSize');
plant = 1;
date = 1;
while date ~= 0;
    date = input('type in the date to be processed: ');
    if date == 0
        disp('procedure done, exiting...')
```

```
    close all
    break
end

dir_old = ['C:\ClassifiedImages\Fieldimages\image 0' num2str(date)
           '\'];
dir_new = ['C:\ClassifiedImages\Fieldimages\BWimages\image 0'
          num2str(date) '\'];

if exist(num2str(dir_new),'file') == 0
    mkdir(num2str(dir_new));        % make a new directory
end
while plant ~= '0'
      plant = input('type in weed name: ','s');
      if plant == '0'
         disp(['*** processing for date 0' num2str(date) ' is
               done ***'])
      close all
      break
      end
    switch plant
       case 'corn'
           label = 1;
       case 'barnyardgrass'
           label = 2;
       case 'crabgrass'
           label = 3;
       case 'lambsquarters'
           label = 4;
       case 'velvetleaf'
           label = 5;
       case 'waterhemp'
           label = 6;
       otherwise
           disp('unknown method, please type in the weed species
                 again..')
           weed = input('Type in plant name: ','s');
    end
    folder = ['image ' num2str(plant) '\'];
    weedpredict = zeros(100,1);
    figure('Position',[0 -50 scrsz(3) scrsz(4)])
```

```
for pos = 1:30
    if exist([ dir_old folder 'Position ' num2str(pos)
             '.bmp'],'file')==0
       disp(['image ' plant num2str(pos) 'does not exit'])
       continue
    else

f = imread([num2str(dir_old) num2str(folder) 'Position '
           num2str(pos) '.bmp']);
[a,b,c] = size(f);
 subplot(2,2,1),imshow(f,[]),
 title(['Original Image for Position ' num2str(pos)])

f1 =f;
t1 = cputime;
% Automatic Pixelwise Segmentation
f  = double(f1);
R  = f(:,:,1);
G  = f(:,:,2);
B  = f(:,:,3);
[m,n] = size(B);
Img1  = zeros(m,n);

 for i = 1:m
     for j = 1:n
         if G(i,j)>1.03*R(i,j) && G(i,j)>1.03*B(i,j) && R(i,j)
            >0.60*B(i,j)&& G(i,j) >50
                Img1(i,j) = 1;
         else   Img1(i,j) = 0;
         end
     end
 end
t1 = cputime - t1;
t2 = cputime;
for i = 1:3
    Img1 = medfilt2(Img1,[3 3]);
end
[Img1, num1] = bwlabel(Img1,8);
for k = 1:num1
    if numel(find(Img1 == k)) < 240
       Img1(Img1 == k) = 0;
```

```
            end
        end
%       Fil_Img1=Img1;
        Fil_Img1 = adpmedian(Img1,7);
        t2 = cputime - t2;
        subplot(2,2,2)
        imshow(Fil_Img1,[]),title('adaptive filtering ')
        hold on
        t3 = cputime;
         s = regionprops(Fil_Img1,'centroid');
%          group plants
      for k = 1:numel(s)    %isnan
          i = 1;
          while i <= numel(s) && i ~= k
                dist = sqrt((s(k).Centroid(1)-s(i).Centroid(1))^2
                       +(s(k).Centroid(2)-s(i).Centroid(2))^2);
                if dist < 100
% if the distance between two centroid is less than a certain value
% consider the two objects belong to one
                   Fil_Img1(Fil_Img1==i)=k;
                end
                i = i+1;
          end
          plot(s(k).Centroid(1),s(k).Centroid(2),'r*');
     end
     hold off
     s = regionprops(Fil_Img1,'centroid');
     for k = 1:numel(s)
         if ~isnan(s(k).Centroid)
         % if s(k).Centroid is not a NaN
            [r,c] = find(Fil_Img1 ==k);
         % that's f_BW == k exist where k ~= 0
             r_min = min(r);
             r_max = max(r);
             c_min = min(c);
             c_max = max(c);
           if r_min == 1 || r_max == a || c_min == 1 || c_max == b
              Fil_Img1(Fil_Img1 == k) = 0;
              continue
           else
```

```
 width  = c_max - c_min;
 height = r_max - r_min;
 f_sep  = Fil_Img1(r_min:r_max, c_min:c_max);
 ind    = find(f_sep ~= k);
 % remove other pixvels in that territory
 f_sep(ind) = 0;
 area = regionprops(f_sep,'Area');
 perimeter = regionprops(f_sep,'Perimeter');
 majoraxis = regionprops(f_sep,'MajorAxisLength');
 minoraxis = regionprops(f_sep,'MinorAxisLength');
% feature calculation
 par(1) = area(k).Area;
 par(2) = perimeter(k).Perimeter;
 par(3) = majoraxis(k).MajorAxisLength;
 par(4) = minoraxis(k).MinorAxisLength;
 par(5) = height*(0.000326/0.000528);
 par(6) = width*(0.000326/0.000528);
 % ratio of area to length(ATL)
 par(7) = par(1)/par(3);
 % compactness (CMP)
 par(8) = 16*par(1)/(par(2))^2;
 % elongation (ELG)
 par(9) = (par(3)-par(4))/(par(3)+par(4));
 % logarithm height to width (LHW)
par(10)= log10(par(5)/par(6));
% ratio of perimeter to broadness(PTB)
par(11)= par(2)/(2*(par(5)+par(6)));
% ratio of length to perimeter(LTP)
par(12)= par(3)/par(2);
% ratio of major axis to minor axis length (ASP)
par(13)= par(3)/par(4);

inst = [par(7), par(8), par(9), par(10), par(11), par(12),
       par(13)];
disp('unscaled result...')
[predict,accuracy,d] = svmpredict(label, inst, model);
switch predict
 case 1
    name = 'corn';
 case 2
    name = 'barnyardgrass';
```

```
            case 3
                name = 'crabgrass';
            case 4
                name = 'lambsquarters';
            case 5
                name = 'velvetleaf';
            case 6
                name = 'waterhemp';
           end
           weedspecies(k) = predict;
           disp(['weed is ' num2str(plant) ';  prediction is ' name]);
           end
       end
end

      subplot(2,2,3)
      imshow(Fil_Img1,[]),title('after grouping ')
      hold on
      s = regionprops(Fil_Img1,'centroid');
      for k = 1:numel(s)
          plot(s(k).Centroid(1),s(k).Centroid(2),'r*');
      end
      hold off
      subplot(2,2,4)
      imshow(Fil_Img1,[]),title('weed identification ')
      hold on
      s = regionprops(Fil_Img1,'centroid');
      for k = 1:numel(s)
          switch weedspecies(k)
          case 1
               name = 'corn';
          case 2
               name = 'BG';
          case 3
               name = 'CG';
          case 4
               name = 'LQ';
          case 5
               name = 'VL';
          case 6
               name = 'WH';
```

```
           end
plot(s(k).Centroid(1),s(k).Centroid(2),'r*');
text(s(k).Centroid(1),s(k).Centroid(2), num2str(name),
'Fontsize',6,'BackgroundColor',[.7 .9 .7],'HorizontalAlignment',
'left','VerticalAlignment','top')
     end
     hold off

     satisfy = input('if satisfied, type in 0, otherwise 1 to quit: ');
     if satisfy == 1
        break
     else
        clc
        pause(0.1)
        continue
      end
 end
         t3 = cputime - t3;
         t = [t1,t2,t3];
         time = [time;t];
       end
   end
end
```