Chapter from the book Recent Trends for Enhancing the Diversity and Quality of Soybean Products
Downloaded from: http://www.intechopen.com/books/recent-trends-for-enhancing-the-diversity-and-quality-of-soybean-products

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Machine Vision Identification of Plants

George. E. Meyer

University of Nebraska, Department of Biological Systems Engineering, USA

1. Introduction

Weedy and invasive plants cost Americans billions of dollars annually in crop damage and lost earnings. Various Western states have reported annual weed control costs in the hundreds of millions of dollars. Herbicides account for more than 72 per cent of all pesticides used on agricultural crops. $4 billion was spent herbicides in the US in 2006 and 2007 (Grube, et al, 2011). The USDA Economic Research Service reported that adoption of herbicide-tolerant soybeans had grown to 70% from 1996 to 2001, yet significant impacts on farm financial net returns attributable to adoption has yet to be documented. Nebraska is part of regional strategic pest plan published in 2002. During 2001, 97% of the soybean acres in Nebraska were treated with herbicides. One means of improving economic benefit is to develop more efficient management inputs, which may be accomplished with better selection of the kind of pesticide and/or site-specific application of pesticides. Moreover, measuring the impact of various management inputs often depends on manual visual assessment and perhaps this could be automated. One method for estimating impact on crop yield loss includes counting weeds per length of row or determining weed populations by species. In order to improve the weed suppression tactics, accurate mapping and assessment of weed populations within agricultural fields is required. See Figure 1. Weed mapping and taxonomy are major activities and species type found in all regions, which cover much broader ecological areas other than farm fields. These are shown by active websites in Nebraska, Iowa, Pennsylvania, Montana, Nevada, Colorado, and California, as examples. Weed and invasive species mapping also has international implications, (Montserrat, et al, 2003). Efforts of this type support integrated pest management (IPM) programs of both Crops and Risk (CAR) and Risk Avoidance and Mitigation (RAMP) which involve profitability and environmental stewardship and risk management, by providing a tool for timely acquisition of weed information. Research in this area promotes an interdisciplinary, IPM systems approach to weed mapping. There is high labor cost associated with the manual scouting of fields to obtain such maps.

2. Spatial variability of weed populations

Weeds are present in every field and lawn every year. The severity of the weed population is determined by local management practices such as the previous crop in the rotation and the herbicide use. According to a 2002 North Central strategic plan, tillage remained a major tool for controlling perennials, although the dilemma is that tillage contributes to soil erosion. Weed spatial distributions are unique, with monocot infestations more patchy than
Recent trends for enhancing the diversity and quality of soybean products

Dicots (Mortensen et al., 1992 and Johnson et al., 1993, 1995). Monocots differ architecturally from dicots. Most weeds are serious competitors for moisture and soil nutrients. By first classifying the weed as either a monocot or dicot, a herbicide could be selected that most effectively controls that type of plant, resulting in better application efficiencies. Most post-emergent herbicides are selective in controlling one plant type or the other. Wiles and Schweizer (1999, 2002) researched the spatial distribution of weed seed banks using soil samples to map locations of weed seed banks in a given field. Seed banks have been found distributed in a patchy manner. Using the maps as a guide, farmers could treat just the weed patches with minimal amounts of the appropriate chemical. Site-specific weed management could mean a significant reduction in herbicide use, which saves the farmer money and benefits the environment. However, a large number of soil and plant samples are needed to get an accurate map—and that can be costly.

Stubbendick, et al (2003) provided a comprehensive compendium of weedy plants found across the Great Plains of the United States. Color plates were provided of canopy architecture and sometimes close-ups of individual leaves, flowers, and fruit. A hand drawing of canopy architecture was also given. In order to recognize a particular species, one needs to understand the concept of inflorescence and various plant taxonomy terms. There are many existing plant image databases around the United States. However, their suitability as reference images has yet to be determined for machine vision applications. An important application using machine vision is site-specific or spot herbicide application systems to reduce the total amount of chemical applied (Lindquist et al., 1998, 2001 a,b; Medlin, et al, 2000). Therefore, a major need for improved weed IPM and ecological assessment of invasive plant species is the development of a low-cost, but high resolution, machine vision system to determine plant incidence, even when imbedded with other plants, and to identify the species type. Machine vision systems should assist in the creation of plant field maps, leading to valid action thresholds (National Roadmap for IPM 2004).

3. Machine vision

Field plants, residue, and soil ecosystems are very complex, but, machine vision technology has the potential to systematically unravel and identify plants using optical properties, shape, and texture of leaves (Meyer et al., 1998). Considerable research has been reported using optical or remote sensing sensors to identify crop health by surface reflectance of green plants in agricultural fields (Gausman et al., 1973; Tucker, et al, 1979; Gausman et al., 1981; Thomas, et al, 1988; Storlie et al., 1989, Tarbell and Reid, 1991.; Franz et al., 1991b; and others). Hagger, et al (1983, 1984) reported the first prototype, reflectance-based plant sensor for spraying weeds. Hummel and Stoller (2002) evaluated a later commercial weed sensing system and noted their problems. Tian, et al (1999) developed a simple weed seeker in Illinois. Unfortunately, subsequent optical, non-image, sensor-based weed seekers and spot sprayers have not gained commercial acceptance for various reasons: first, single-element optical sensors can change the size of their field of view based on lens properties and distance to a target. Secondly, sensed reflectance properties may change according to the spatial contents of target components within the field of view Woebbecke, et al (1994); and finally, these sensors therefore may not always distinguish conclusively between crop, weed, or soil residue background. The voltage signal originating from an optical diode or transistor along with the Gaussian lens system used creating the field of view is a weighted average-problem, where the proportions of contributing reflectance and spatial contents are unknown. That problem can be solved only by spatial image analysis.
Image analysis is a mathematical process to extract, characterize, and interpret tonal information from digital or pixel elements of a photographic image. The amount of detail available depends on the resolution and tonal content of the image. The process is iterative, starting with large features followed by more detail, as needed. However, shape or textural feature extraction first requires identification of targets or Regions of Interest (ROI). These regions are then simply classified as green plants or background (soil, rocks, and residue). ROI’s can be also identified with supervised control of the camera or field of view (Woebbecke, et al, 1994, Criner, et al, 1999), using a supervised virtual software window, cropping of selected areas, or unsupervised crisp or fuzzy segmentation procedures. ROI’s are then binarized to distinguish target and background. Binarized images are then used for shape analysis or boundary templates for textural feature analysis. The binary image is combined with tonal intensity images of the targets (Gerhards and Christensen, 2003, Meyer et al., 1999; Kincaid and Schneider, 1983; Jain, 1989; Gonzalez and Woods, 1992; and others). Machine vision offers the best potential to automatically extract, identify, and count target plants, based on color, shape, and textural features (Tillett et al. 2001). However, directing the image analysis process toward the classical botanical taxonomic, plant identification approach has previously required considerable supervised human intervention. A major problem is the presentation of plant features including individual leaves and canopy architecture to a discrimination or classification system. Camargo Neto, et al (2004 a,b; 2005) presented a combination of traditional image processing techniques, fuzzy clustering, pattern recognition, and a fuzzy inference neural network to identify plants, based on leaves. A particular difficult problem was the development of an algorithm to extract individual leaves from complex canopies and soil/residue color images.

If image vegetative/background classification is to be useful for plant species identification, a separated plant region of interest (ROI) must be found to provide important canopy information needed to discriminate at the very least, broadleaf versus grass species (Woebbecke et al., 1995a; Meyer et al., 1998). Four basic steps for a computerized plant species classification system were presented by Camargo Neto (2004). The first step is creating a binary image which accurately separates plant regions from background. The second step is to use the binary template to isolate individual leaves as sub images from the original set of plant pixels (Camargo Neto, et al, 2006a). A third step was to apply a shape feature analysis to each extracted leaf (Camargo Neto, et al, 2006b). The fourth and final step was to classify the plant species botanically using additional leaf venation, textural features acquired during the previous steps (Camargo Neto and Meyer, 2005). Machine vision plant image analysis has been greatly enhanced through the introduction of the automatic color and focusing digital camera (Meyer, et al, 2004). Digital cameras when run in the automatic mode make decisions on “best picture”, and thus are extremely popular as consumer products.

4. Vegetation indices

The use of vegetation indices in remote sensing of crop and weed plants is not new. It represents the first step shown in Figure 2. Studies for crop and weed detection have been performed using different spectral bands and combinations for vegetative indices (Woebbecke et al. 1995b, El-Faki, et al., 2000ab, Marchant et al., 2004; Wang et al., 2001, Lamm et al., 2002; Mao et al., 2003; Yang et al., 2003). Color vegetation indices utilize only the red, green and blue spectral bands. The advantage of using color indices is that they
accentuate a particular color such as plant greenness, which should be intuitive by human comparison. Woebbecke et al. (1995a) was one of the first researchers to test vegetation indices that were derived using color chromatic coordinates and modified hue for distinguishing green plant material in images from bare soil, corn residue, and wheat straw residue. Woebbecke’s indices (without row and column indices of each pixel) included:

\[
\text{Color indices: } (r - g - b, g - b, g - b - r, \text{and } 2 \cdot g - r - b)
\]  
(1)

where: \( r, g, \) and \( b \) are known as the chromatic coordinates (Wyszecki and Stiles, 1982), given as:

\[
r = \frac{R^*}{R^* + G^* + B^*}, \quad g = \frac{G^*}{R^* + G^* + B^*}, \quad \text{and } \quad b = \frac{B^*}{R^* + G^* + B^*}
\]  
(2)

and: \( R^*, G^*, \) and \( B^* \) are normalized RGB values \((0 \text{ to } 1)\), defined as:

\[
R^* = \frac{R}{R_m}, \quad G^* = \frac{G}{G_m}, \quad \text{and } \quad B^* = \frac{B}{B_m}
\]

\( R, G, \) and \( B \) are the actual pixel values obtained from color images, based on each RGB channel or band.

\( R_m, G_m, \) and \( B_m = 255, \) are the maximum tonal value for each primary color.

Woebbecke discovered that the excess green vegetation index \((\text{ExG} = 2 \cdot g - r - b)\) provided an interesting near-binary, tonal image outlining a plant region of interest. Woebbecke’s
excess green (ExG) index has been widely cited in the literature and has been tested in recent studies (Giltelson et al., 2002; Lamm et al., 2002; Mao et al., 2003; and others). ExG plant regions of interest could then be completely binarized using a selected contrast threshold value for each image. Thus, an important condition was the selection of the threshold value. Mao et al. (2003) subsequently tested several indices: ExG, normalized difference index (NDI), and the modified hue for separating plant material from different backgrounds (soil and withered plant residue). In his study, the ExG index was found superior to the other indices tested. A critical step was to select a manual threshold value to binarize the tonal image into a black and white image.

Other color vegetation indices have been reported for separating plants from soil and residue background in color images. For example, the normalized difference vegetation index (NDI) by Perez et al. (2000) uses only the green and red channels and is given as:

\[
\text{NDI} = \frac{G - R}{G + R}
\]  

(2)


Color indices have been suggested to be less sensitive to in lighting variations, and may have the potential to work well for different residues backgrounds (Campbell, 1996). However, a disproportionate amount of redness from various lighting sources may overcast a digital image, making it more difficult to identify green plants with simple RGB indices (Meyer et al, 2004b). For example, image redness may be related to digital camera operation and background illumination, but may also be related to redness from the soil and residue itself. An alternate vegetative index called excess red (ExR = 1.4 · r - g) was proposed by Meyer et al.(1998a), but was not tested until later studies.

Meyer and Camargo Neto (2008) reported on the development of an improved color vegetation index: Excess Green minus Excess Red (ExG-ExR). This index does not require a threshold and compared favorably to the commonly used Excess Green (ExG), and the normalized difference (NDI) indices. The latter two indices used an Otsu threshold value to convert the index near-binary to a full-binary image. The indices were tested with digital color images of single plants grown and taken in a greenhouse and field images of young soybean plants. Vegetative index accuracies were compared to a hand extracted plant regions of interest using a separation quality factor algorithm. A quality factor of one represented a near perfect binary match of the computer extracted plant target compared to the hand extracted plant region. The ExG-ExR index had the highest quality factor of 0.88 ± 0.12 for all three weeks, and soil-residue backgrounds for the greenhouse set. The ExG+Otsu and NDI-Otsu indices had similar quality factors of 0.53 ± 0.39, and 0.54 ± 0.33 for the same set, respectively. Field images of young soybeans against bare soil gave quality factors for both ExG-ExR and ExG+Otsu around 0.88 ± 0.07. The quality factor of NDI+Otsu using the same field images was 0.25 ± 0.08. ExG-ExR has a fixed, built-in plant-background zero threshold, so that it does not need Otsu or any user selected threshold value. The ExG-ExR index worked especially well for fresh wheat straw backgrounds, where it was generally 55
per cent more accurate than the ExG+Otsu and NDI+Otsu indices. Once a binary plant region of interest is identified with a vegetation index, other advanced image processing operations may be applied, such as identification of plant species such as would be needed for strategic weed control.

Near-Infrared (NIR) along with color bands have been used in vegetative indices for satellite remote sensing applications. However, NIR is less human intuitive, since the human eye is not particularly sensitive to the NIR spectrum which begins with red light. The human eye is only able to discern color (retinal sensors called cones). The eye also contains rods which are essentially receptive to small amounts of blue light that may exist after sundown. NIR is also not readily available with an RGB color digital camera. NIR usually requires a special monochromatic camera with a silicon-based sensor that can detect light up to one micron in wavelength with an NIR band pass filter. Hunt, et al (2011) has experimented with extracting near infrared out of RGB digital cameras. They developed a low-cost, color and color-infrared (CIR) digital camera that detects bands in the NIR, green, and blue. The issue still remains as to how does one verify the accuracy of infrared-image-based vegetative index without comparison to vegetation observed in a corresponding color visual image? So, the verification process of existence of plant material either returns to color images or some other non-optical method.

Two additional problems tend to exist with previous research regarding vegetative indices (a) the disclosure of the manual or automatic threshold used during the near-binary to binary conversion step, and (b) generally, the lack of reporting of vegetation index accuracy. Gebhardt, et al (2003) suggested that it was not necessary to classify vegetation on a pixel basis with digital imaging. However, if there are too many plant pixels mixed up with background pixels, accuracy may be reduced. Hague, et al (2006) suggested a manual comparison of vegetative areas from high resolution photographs. To date, very few vegetative index studies have reported validation accuracy of detecting plant material in independent images from other sources. This problem becomes particularly apparent, when these indices are applied to the collection of photographic plant databases currently available.

Plant classification might be expanded to hyper spectral imaging (Okamoto, et al. 2007). Wavelet along with discriminant analyses were used to identify spectral patterns of pixel samples for a 75–80 percent classification rate of five young plant species. Typically, hyper spectral cameras are expensive.

In summary, color image classification systems utilize the red (R), green (G), and blue (B) tonal intensity components. Color is a special form of spectral reflectance, which can be derived from spectral measurements (Wyszecki and Stiles, 1982; Murch, 1984; Jain, 1989; Gonzalez and Woods, 1992; Perry and Geisler, 2002). Perceived (human) color is based on the (RGB) primary colors. Woebbecke et al. (1995) discovered that the excess green index (2·G-R-B) could provide excellent near-binary segmentation of weed canopies over bare soil for canopy shape feature analysis. El-Faki et al. (2000b) studied different RGB indices, as potential weed detection classifiers, but none possibly as good as excess green. The best correct segmentation rates (CCR) found were around 62%, while some misclassification rates were less than 3%. Meyer et al. (1999, 2004) proposed an excess red index (1.3·R-G), based on physiological, rod-cone proportions of red and green. This index also provides near-binary silhouettes of plants under natural lighting conditions. Marchant, et al, (2004) proposed additional procedures for dealing with machine vision and natural lighting. The
utilization spectral wave bands and color components have been used arithmetically and called vegetation indices. The index Meyer and Camargo Neto (2008) is an advanced color vegetation index.

5. Computerized single leaf extraction

Only a few methods of unsupervised leaf extraction from canopy images have been reported in the literature. Franz et al. (1991b) reported the use of curvature functions and the Fourier-Mellin correlation to identify completely visible and partially occluded sets of leaves. Leaf statistical features of mean, variance, skewness, kurtosis were computed, using spectral wavebands of red, green, blue, and near infrared. These features were used to discriminate leaf types of unifoliolate soybean, ivy, morning glory cotyledons, velvetleaf cotyledons, foxtail, first leaf of ivy, morning glory, and the first leaf of velvet leaf. Franz et al. (1995) further developed an algorithm to extract boundaries of occluded leaves using an edge detection technique to link the end points of leaf edge segments. User intervention was required at various steps of the algorithm. The fractions of individual leaves obtained were reported to be 0.91, 0.87, 0.95, and 0.71 for velvetleaf, soybean, ivy leaf morning glory, and foxtail, respectively.

To clarify this issue, occluded or partial fractions of leaves are probably not that useful for species identification. However, all canopies will exhibit whole individual leaves at the canopy apex, which can be seen in overhead photographs. Some leaves may standout by
themselves (non-concealed) against the soil-residue background. Others will have vegetation from occluded leaves around them, which we will call concealed leaves. The latter would represent a difficult image processing problem, not easily solved by traditional algorithms such as edge detection, erosion, dilation, and such. Deformable templates using active contours were used by Manh, et al. (2001) to locate boundaries of green foxtail leaves. Manh's process attempted to combine color separation and shape feature analysis into a single operation. The procedure began with identification of a leaf tip, and followed by shape analysis across the rest of the green material. However, a manually selected energy level or color was needed. Segmentation accuracy for a single species of foxtail leaves was reported to be 84%. No other species were studied. Individual, whole, and fragments of leaves were isolated using the Gustafson-Kessel fuzzy clustering method over bare soil, corn stalks, and wheat straw color images (Hindman, 2001, Meyer et al., 2004b, Gustafson and Kessel, 1979). Zadeh intensification of the fuzzy cluster membership functions resulted in definitive green canopy areas, but not individual leaves. However, Camargo Neto, et al (2006) used the Gustafsen-Kessel fuzzy leaf cluster fragmentation method on green canopy regions of interest. He also developed a reassembling method of the green cluster fragments resulting in individual leaves using a genetic algorithm (Holland, 1975).

6. Shape feature analysis

If the process of image vegetative/background classification is to be useful, the separated plant region of interest (ROI) must provide important canopy or leaf shape feature or property information to at least discriminate between broadleaf and grass species (Woebbecke et al., 1995b; Meyer et al., 1998a; Meyer et al., 1998b). Supervised leaf and single plant canopy shape feature analysis has been studied the most. Petry and Kuhbauch (1989) found shape parameters using five canonical indices found distinctly different for several weed species. Guyer et al. (1986, 1993) used image shape feature analysis on individual leaves to distinguish between weed species and corn. Guyer et al. (1993) using only leaf and canopy shapes, reported a 69% correct identification rate for 40 weeds and agricultural crop species. Guyer found that no single shape feature alone could distinguish corn from all other species. Franz et al. (1991 a,b) identified plants based on individual leaf shape at two growth stages using the Fourier-Mellin correlation. Woebbecke et al. (1995a, b) used basic image shape feature analysis to discriminate between broadleaf and grassy plant canopies. Woebbecke found that broadleaf and grass shape features best appeared to a vision system at early stages of growth or within a specific window of time, from 1-4 weeks after emergence. Downey, et al (2004) described a field canopy shape identification system which used a binary canopy erosion technique to discriminate between grasses and broadleaf plants. Yonekawa et al. (1996) presented a set of classical shape features for a leaf taxonomy database. Chi, et al (2002) fitted Bezier curves to different leaf boundary shapes. Mclellan and Endler (1998) compared several morphometric methods for describing complex shapes. They found that approximately 20 harmonics of the elliptic Fourier method accurately depicted shapes of Acer saccharinum, Acer saccharum, and Acer palmatum leaves. A leaf shape image retrieval systems was also reported by Wang, et al (2003). Du et al (2005, 2006, 2007) proposed the Douglas-Peucker approximation algorithm for leaf shapes and the shape representation was used to form the sequence of invariant attributes.
A modified dynamic programming (MDP) algorithm for shape matching was proposed for the plant leaf recognition. Oide and Ninomiya (2000) used the Elliptic Fourier (EF) method to classify soybean varieties, using a normalized leaf shape. The EF method using a chain-coded, closed contour, invariant to scale, translation, and rotation was first introduced by Kuhl and Giardina (1982). EF has been used in recent studies to describe the shape of objects. Innes and Bates (1999) used an Elliptical Fourier descriptor to demonstrate an association between genotype and morphology of shells. Chen et al. (2000) used Elliptic Fourier descriptors to describing shape changes in the human mandible for male and female at different ages. Most methods previously investigated ignore leaf edge serration. Leaf serration or edgeness is an important morphologic feature used for identifying plant species. For example, the curvature functions developed by Franz et al. (1991b) were found generally inadequate where leaflet serration was quite pronounced. Camargo Neto, et al, 2006b applied the Elliptic Fourier shape feature analysis to extracted leaves of velvet leaf Abutilon theophrasti, pig weed Amaranthus retroflexus L., sunflower Helianthus annus, and soybean Glycine max. A velvet leaf example is shown in Figure 3.

Hearn (2009) used a database of 2,420 leaves from 151 plant species for a plant leaf shape analysis. Using metrics derived during Fourier and Procrustes analyses, it was found that a minimum of ten leaves for each species, 100 margin points, and ten Fourier harmonics were required to develop any accuracy using the leaf shape of a species. His results indicated a success rate of 72% correct species identification for all 151 species used. This may mean that more than leaf shape is needed for classification.

7. Textural feature analysis

Color and/or leaf shape features alone may not be sufficient to consistently distinguish between young weed and crop plant species. Textural features may supply some additional botanical information, such as leaf venation, leaf pubescence, but also leaf disease and insect damage. The color or tonal detail for texture was first described by quantification of co-occurrence of tonal pairs or contrast also known as spatial tonal frequency (Haralick, 1978 and 1979). Wavelet analysis and energy have been recently suggested as a frequency based textural analysis for segmenting weeds imbedded in canopies (Chang and Kuo, 1993, Strickland and Hahn, 1997, Tang, et al, 2003). Shearer and Holmes (1990) used color co-occurrence matrix method to identify the textural features of isolated plants. Shearer and Jones (1991) proposed a texture-alone plant detection system based upon hue-saturation-intensity (HSI) images. Oka and Hinata (1989) used side view images of rice to distinguish between old and new Japanese rice cultivars. Zhang and Chaisattapagon (1995) tested a combination color, shape, and texture approach for detecting weeds in wheat fields and found that leaf-surface coarseness indices defined by Fourier spectra may be effective in differentiating wheat from broad-leaf weeds. Meyer et al. (1999) showed that combined color, shape, and textural statistical discriminate analysis system could separate grasses from broadleaf canopies against bare soil backgrounds. Major problems for obtaining botanical textural detail involve image resolution, leaf orientation or rotation, shadows, bidirectional reflectance of leaf surfaces, and background lighting. Uneven lighting for example, could obscure venation - mesophyll leaf detail. Diffuse lighting could provide more even illumination than direct-beam lighting. Fu and Chi (2006) presented an algorithm for extracting leaf vein details from detached leaves under artificial light. Park, et al (2008) described a prototype system for classifying plants based on leaf venation features. Their
Recent Trends for Enhancing the Diversity and Quality of Soybean Products

Fig. 3. Elliptic Fourier shape approximations for velvetleaf (*Abutilon theophrasti Medicus*), a) original leaf image, b) 1st EF harmonic, c) 1st + 2nd + 3rd + 4th EF, d) 1st + … + 8th EF harmonics, e) 1st + … + 16th EF harmonics, and f) 1st + … + 30th EF harmonics.

method detected the differences between tree and parallel venations in leaves, and thus could be considered as an enhancement to the classification tool set.

De Oliveira Plotze (2009) combined computer vision techniques and plant taxonomy protocols, these methods are capable of identifying plant species. The biometric measurements are concentrated in leaf internal forms, specifically in the venation system. The methodology was tested with eleven species of passion fruit of the genus *Passiflora*. The features extracted from the leaves were then applied to a neural network system to develop a classification of species. The results were very accurate in correctly differentiating among species with 97% of success. Zheng and Wang (2009, 2010) presented the results of mathematical morphology used on images of single leaf samples. Mathematical morphology provides four fundamental operations of dilation, erosion, opening, and closing in image processing. Their goal was to extract only leaf veins using hue and intensity information. Camargo Neto and Meyer (2005) classified the plant species botanically sing additional leaf venation textural features acquired during the previous steps. One thing is clear, lack of care in the photography of a leaf may affect image textural properties and classification.
8. Plant species classification

Most studies in the last 20-years have addressed the classification of only two crop-weed classes or general cases of broad leaf versus grasses and in other cases, crop row versus between crop row (Tang, et al, 2003). However, to precisely classify a plant species that may be imbedded within other different species of plants in an image is a botanically challenging exercise.

Agarval, et al (2006) described an ongoing project to digitize information about plant specimens that would become available to field botanists and crop managers. They indicated that the first step required acquisition of digital images and possibly plant architectural models, along with an effective retrieval method and mobile computing mechanisms for accessing this information. At that time they had indicated progress in developing a digital archive of the collection of various plant specimens at the Smithsonian Institution.

Analytical tools are improving for classifying plant species. The artificial neural network (ANN) has been proposed for many classification activities. Plotze and Bruno (2009) have also proposed a plant taxonomy system. Yang et al., (2000, 2002, 2003) used RGB pixel intensities as inputs for a fuzzy artificial neural network (ANN) for distinguishing weeds from corn, with success rates as high as 66% for corn and 85% for weeds. To encompass the uncertainty of image classification processes, fuzzy set theory (FST) has been proposed for plant classification (Gottimukkala et al., 1999). FST provides a possibilitistic alternative (different, but in many cases complementary) to the probabilistic or statistical approaches. FST embraces virtually all (except one) of the definitions, precepts, and axioms that define classical sets that supports common mathematics, (Ross, 2004). It uses variables in the form of membership functions with degrees of support for fuzziness, incorporating uncertainty (Zadeh, 1965; Mamdani, 1976; Li and Yen, 1995). Pal, et al 1981, 1994, Bezdek, 1973, 1993) summarized the use of a FST neural network for pattern recognition, generating membership functions, performing fuzzy logic (FL) operations, and then deriving inference rule sets. Jang (1993) invented the artificial neural fuzzy inference system (ANFIS) for training membership functions and rule sets that could be used for classification (Figure 4).

Fuzzy logic machine vision classification systems are intended to imitate human perception or vision and to handle uncertainty. In the weed discrimination example, expert human perception or scouting validation is required for ground truthing. Bhutani and Battou (1995) and Tizhoosh (1998, 2000) provide computational overviews and various examples of fuzzy logic applied to image processing. Incorporating unsupervised fuzzy logic clustering and image analysis into site-specific technologies has tremendous potential (Kuhl and Giardina, 1982, Gath and Geva, 1989, De and Chatterji, 1998, Babuska, 1998, Manthalkar, at al, 2003, Meyer, et al. 2004). The very nature of site-specific data collection, image analysis, decision-making, etc., is characterized by uncertainty, ambiguity, and vagueness, which may be over overcome with these techniques.

or fuzzy ANN (ANFIS) model is that they can be designed to mimic signal errors and random noise data too well, especially with an inadequate size of the training data set. Fuzzy inference systems can also incorporate the “I do not know” result.

Fuzzy clustering refers to unsupervised partitioning of data into subclasses for pattern recognition (Ross, 2004). Babuska (1998) presented six different clustering techniques that might be used to organize tonal image data with their limitations. These included the fuzzy c-means, the Gustafson-Kessel, fuzzy maximum likelihood, fuzzy c-varieties, fuzzy c-elliptotypes, and possibilistic clustering that might be used on tonal images. Moghaddamzadeh et al. (1998) described a fuzzy nearest-neighbor, clustering method for segmenting color images. Townsend (2000) discussed methods for making comparisons of fuzzy ecological pattern recognition methods. Beichel, et al. (1999) discussed the use of an unsupervised Gath-Geva clustering method for Landsat thematic mapper (TM) images. Classification accuracy reached a maximum value of 86 % with five clusters. Individual, whole, and fragments of leaves were isolated using the Gustafson-Kessel fuzzy clustering method over bare soil, corn stalks, and wheat straw color images (Meyer et al., 2004b). Zadeh intensification of the membership functions resulted in definitive green canopy areas.

Fig. 4. Advanced Species Classifier Method– Fuzzy Logic- Neural Network using Image metrics and others.

A machine vision system with unsupervised image analysis and mapping of features was presented by Camargo Neto (2006a) and Camargo Neto, et al. (2006b). A classification system was trained using statistical discriminant analysis which was tested using individual test leaves and clusters from several plants. As many as 75 percent of exposed whole leaves were extracted, and can be further species identified at 75% or better. When such a system is improved and validated with scientific-based methods, it could dramatically assist understanding crop-weed relationships, growth, competition, and control. A machine vision system certainly should be able to identify and distinguish weed species that are 7 - 21 days old, a time when post emergence herbicides are most effective.
9. Linking machine vision with weed management systems

Predicting crop yield loss due to weed competition is one critical component of dynamic decision making for integrated weed management. Moreover, spatial variation in weed occurrence must be accounted for to accurately predict crop yield loss (Lindquist et al. 1998, 2001 a,b). The fuzzy logic machine vision classification system will be extremely useful where weeds are distinguished from crop plants and precisely mapped within a farm field. Shape feature analysis also provides a means for determining the relative surface area of weed plants relative to crop plants. Kropff and Spitters (1991) argued that the competitive strength of a species is determined by its share in leaf area at the moment when interspecific competition begins. Kropff et al. (1995) presented an equation that expresses yield loss (YL) as a function of weed and crop LAI. This approach has recently been expanded to relate yield loss to weed and crop relative volume (Conley et al. 2003) and could easily be used to relate yield loss to weed and crop relative surface area obtained from our image analysis. This kind of detail requires close-in imaging within a few meters with current high pixel rate digital cameras.

Holst, et al (2006) reviewed the progress of weed population modeling and of course the use is similar: strategic decision making for weed management. Freckleton and Stephens (2009) discussed the use if dynamic plant models for weed management. They concluded that there exist a discrepancy in the field of weed population modeling; many of the problems faced by weed ecologists require detailed quantitative predictions, but few modelers are attempting to provide such predictions. FST has also been used for modeling biological systems. Ambue! et al. (1994) used FL to develop a crop yield simulator for assessing spatial field variability for accuracy and optimizing pestice application rates. Weed plant growth and plant population models that also describe the canopy architecture would be very helpful for weed classification.

10. Conclusions

The literature is rich in selected or component ideas for machine vision, plant species identification. Now is the time to put together a complete robust system that essentially mimics the human taxonomic, plant identification keying method. If one returns to Stubbendick, et al (2003), one can verify that the human classification process requires metrics on leaves, stems, flowers, inflorescence, and a picture of the plant. Leaf shape and venation images alone may not close the classification process. Shape analysis for image processing is very well-understood and computer algorithms are readily available. The leaf angle in the plane of the canopy is of interest (the first elliptic Fourier harmonic), and that is a critical angle for rotationally invariant leaf texture or venation analysis. Additional studies regarding leaf orientation relative to the camera lens might help to reduce classification errors. Modern digital cameras are capable of acquiring large amounts of image-pixel data. Future studies need to determine minimal digital image resolutions needed to maintain the highest species discrimination performance.

Fuzzy logic, cluster algorithms and cluster reassembly routines work well for extracting convex leaf shapes from plant canopy images. However, for more botanically diverse leaf shapes, such as species with complex leaves, lobed margins (indent), trifoliolates, etc., new fitness criteria need to be developed to accommodate these leaf shapes. Undoubtedly, integration of specific shape and textural feature analyses as fitness criteria may be a key to improvement of this process. New leaf extraction/species classification algorithm can
become especially useful, if acceptance criteria can be designed to accommodate more a
tensive leaf taxonomy digital library (shape and texture of single and compound leaves).
Work has been extended on utilizing digital canopy architecture metrics in three dimensions
which is important plant taxonomy.
Species classification and mapping has been tested using a neural-fuzzy inference model,
which can be improved with inclusion of additional training information, including: stage of
growth, expected canopy architecture, distance from a designated crop row, crop row
spacing and direction.
Studies and discussion should be conducted to determine if older photographic plant image
data bases can be used as references for new unknown digital plant images. Considerable
field testing and validation are always needed for plant identification studies using machine
vision.

11. References

Agarwal, G., Belhumeur, P., Feiner, S., Jacobs, D., Kress, W.J., Ramamoorthi, R., Bourg, N.A.,
Dixit, N., Ling, H., Mahajan, D., Russell, R., Shirdhonkar, S., Sunkavalli, K. and
1597-1599.
Center, Cornell University, Ithaca.
Pattern Recognition Letters. 16:901-909.
Canopy Images using Gustafson-Kessel Clustering and a Genetic Algorithm. Computers
and Electronics in Agriculture (Elsevier) 51:65-85.
Elliptic Fourier Analysis. Computers and Electronics in Agriculture (Elsevier),
50:121-134.
Mapping Weed Species in Minimum Tillage Systems. Unpublished PhD Dissertation,
University of Nebraska, Lincoln, NE.
computer extracted individual leaves. In: Chen, Y.R., Meyer, G.E., Tu S. (Eds.),
Optical Sensors and Sensing Systems for Natural Resources and Food Safety and Quality,


Li, H.X., and V.C. Yen, 1995. Fuzzy sets and fuzzy decision-making. CRC.
Recent Trends for Enhancing the Diversity and Quality of Soybean Products


This book presents new aspects and technologies for the applicability of soybean and soybean products in industry (human food, livestock feed, oil and biodiesel production, textile, medicine) as well as for future uses of some soybean sub-products. The contributions are organized in two sections considering soybean in aspects of food, nutrition and health and modern processing technologies. Each of the sections covers a wide range of topics. The authors are from many countries all over the world and this clearly shows that the soybean research and applications are of global significance.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: