sustainability

MDPI

Article

Detection of Corn and Weed Species by the Combination of Spectral, Shape and Textural Features

Fenfang Lin [1], Dongyan Zhang [2,*], Yanbo Huang [3], Xiu Wang [4] and Xinfu Chen [5]

1 School of Geography and Remote Sensing, Nanjing University of Information Science & Technology, Nanjing 210044, China; 002173@nuist.edu.cn
2 Anhui Engineering Laboratory of Agro-Ecological Big Data, Anhui University, Hefei 230601, China
3 USDA-ARS Crop Production Systems Research Unit, Stoneville, MS 38776, USA; yanbo.huang@ars.usda.gov
4 Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China; wangx@nercita.org.cn
5 Fujian Vocational College of Agriculture, Fuzhou 350007, China; cxinfu2003@126.com
* Correspondence: zhangdy@ahu.edu.cn; Tel.: +86-159-1050-8416

Received: 12 June 2017; Accepted: 26 July 2017; Published: 4 August 2017

Abstract: Accurate detection of weeds in farmland can help reduce pesticide use and protect the agricultural environment. To develop intelligent equipment for weed detection, this study used an imaging spectrometer system, which supports micro-scale plant feature analysis by acquiring high-resolution hyper spectral images of corn and a number of weed species in the laboratory. For the analysis, the object-oriented classification system with segmentation and decision tree algorithms was utilized on the hyper spectral images to extract shape and texture features of eight species of plant leaves, and then, the spectral identification characteristics of different species were determined through sensitive waveband selection and using vegetation indices calculated from the sensitive band data of the images. On the basis of the comparison and analysis of the combined characteristics of spectra, shape, and texture, it was determined that the spectral characteristics of the ratio vegetation index of R677/R710 and the normalized difference vegetation index, shape features of shape index, area, and length, as well as the texture feature of the entropy index could be used to build a discrimination model for corn and weed species. Results of the model evaluation showed that the Global Accuracy and the Kappa coefficient of the model were both over 95%. In addition, spectral and shape features can be regarded as the preferred characteristics to develop a device of weed identification from the view of accessibility to crop/weeds discriminant features, according to different roles of various features in classifying plants. Therefore, the results of this study provide valuable information for the portable device development of intelligent weed detection.

Keywords: hyper spectral imaging; object-oriented; decision tree; corn; weed

1. Introduction

Farmland weeds are harmful to the growth of crops. They can cause developmental disorders of the plant, such as short seedling, thin stem, and yellow plant leaves. Further, they can cause the reduction of grain yield up to 10% per year [1]. To develop sustainable agriculture to ensure food security, weed detection for variable herbicide application has been an important research topic in precision agriculture.

In recent years, machine vision and remote sensing techniques have been used for weed identification. The machine vision technique captures the intensity of the reflected light from the target plant using a digital camera. Through image processing, plant location, color, shape, texture, and other features are extracted to identify the target plants. This method can focus on the whole plant or part of the plant and is able to differentiate the weed from other objects in real-time, accurately and automatically [2–4]. However, the factors such as illumination condition, leaf shape, wind speed, and direction consistently limit the extraction of the color, shape, texture, and other features of the plants from the images, so rapid image processing and accurate weed identification are two major challenges [5–8]. Compared with machine vision, remote sensing methods can distinguish crops and weeds by reflectance data, especially hyper spectral remote sensing. The hyper spectral sensor with nano-scale spectral resolution can detect the subtle differences of plants using certain feature wavelengths or electromagnetic radiation in the visible and near-infrared (NIR) region. Up to date, there are many studies on identification of weeds from crops using the sensitive spectral bands with encouraging results. However, the identification accuracy is low in cases when the spectral difference between the crop and the weed is not obvious, or the reflection of leaves is affected by factors of water content, plant disease, and growth stage [9–13]. Therefore, to more effectively discriminate weeds from crops, the combination of multiple features, such as the combination of shape and textural, shape and spectral, and spectral and textural features, should be considered.

The hyper spectral imaging technique with high spectral and spatial resolutions, can be used to extract the spectral, spatial and structural features synchronously. The combination of the extracted features is good for improved identification of weeds from the crop in a complex environment. Studies have been done using spectral and shape features simultaneously for classification of plants based on the analysis of hyper spectral images. Li et al. [14] utilized the method of combination of shape-based analysis and spectral angle match to identify the weeds in a watermelon field, which obtained a good identification accuracy but the shape and spectral features were used separately and the texture features were not included. Tits et al. [15] pointed out that when the shape features were considered, the differences between classes could be enhanced for the recognition of targets. Somers et al. [16] used hyper spectral mixture analysis with spectral and shape features to detect weeds in an orchard field. In addition, textural features are also important in studying spatial information of different targets. Alchanatis et al. [17] considered that spectral and textural features could detect weeds in a cotton field, but the textural features were only used for weed identification after the soil and plants were separated by spectral features.

To promote the development of weed identification technology for designing an intelligent weed identification system and accurately spraying pesticides, the research objectives of this study are:

- To characterize shape, textural and spectral features to differentiate between corn and a number of weeds;
- To compare and validate the identification models of corn and weeds and determine the optimal feature combination to develop an intelligent weed identification system.

2. Materials and Methods

2.1. Imaging System

In this study, a field imaging spectrometer system (FISS) was used to collect the imaging spectrometer data [18]. The system consists of an enclosed optomechanical subsystem, a computer, and an electronic system. The optomechanical component includes a scanning mirror, a charge-coupled device (CCD) camera (Model INFINITY3-1), a dispersing unit with a "prism-grating-prism" (PGP), and an objective lens. The computer system includes hardware and software for storing imaging data, operating the FISS and processing basic data. The electronic system includes the power supply and the motor control circuit. The FISS has 344 spectral channels in the range of 446 nm to 912 nm. The sampling interval of the spectrum is about 1.4 nm and the spectral resolution is 4–7 nm. The signal-to-noise ratio for 60% bands is greater than 500. The spatial resolution of FISS is related to both the parallel and perpendicular directions of the entrance slit. The spatial resolution of the direction parallel to the entrance slit depends on the size of the imaging unit in the CCD camera focal plane and the focal length of the objective lens, and the other one depends on the width of the entrance slit and the focal length of the objective lens. The space resolution of the FISS is better than 2 mm.

2.2. Data Collection

Sample collection was carried out at the National Demonstration Base for Precision Agriculture (40°10′31.6″ N, 116°26′44.4″ E), Beijing, China. Corn and a number of commonly known weeds were chosen as the research targets. Corn was sampled at the 5-leaf stage, an important time for weed control. The two categories of weeds were monocotyledonous and dicotyledonous plants. Monocotyledonous weeds included foxtail (*Setaria viridis* L., SETVI) and goose grass (*Eleusine indica* L., ELEIN), and dicotyledonous weeds were round leaf pharbitis (*Pharbitis purpurea* L. *Voigt*, PHAPU), lobed leaf pharbitis (*Pharbitis nil* L. *Choisy*, PHANI), redroot amaranth (*Amaranthus retroflexus* L., AMARE), purslane (*Portulaca oleracea* L., POROL) and lambs quarters (*Chenopodium album* L., CHEAL). These seven weeds are widely distributed in North China, and are major weeds that impact corn yield.

Whole plants were collected in the field and sent to the laboratory as quickly as possible to minimize analytical errors resulting from the wilting of the leaf blade. In the laboratory, fully expanded leaves were clipped and wiped clean for measurement with FISS. Lighting was achieved by two 500 W halogen tungsten filament lamps. Images of white reference were simultaneously acquired with the leaf samples to minimize illumination changes among different measurements and converting the digital number to relative reflectance in every image. A sample image is shown in Figure 1. We collected 10 sheets of imaging spectrometer data when the spectrometer was 1.0 m above the samples.

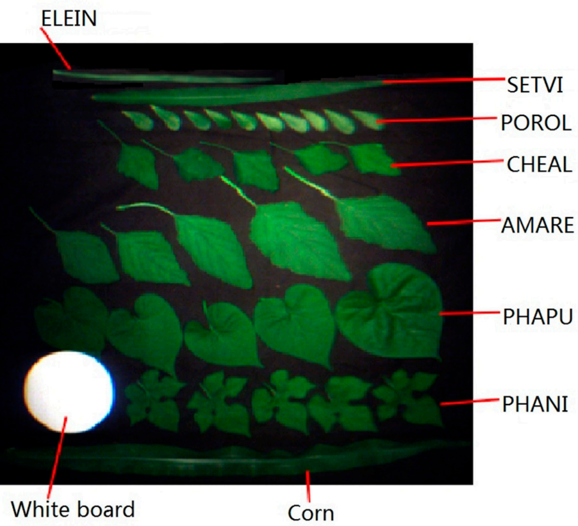

Figure 1. Leaf image for corn and seven weed species. This is one of images acquired by FISS. FISS: Field imaging spectrometer system, SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

2.3. The Process for Crop/Weeds Identification

To extract the image features for identifying the crop and weeds, a procedure of data processing was formulated in sequence as data pre-processing, sample objectivation, feature extraction, identification modeling, and model validation (Figure 2). Data processing included Minimum Noise Fraction (MNF) and data normalization by the imaging spectrometer data of the white reference. The effectiveness of such processing is in removing noises resulting from the instrument and illumination. The imaging data has 344 bands containing both useful and irrelevant information for the corn/weeds identification. Thus, segmented principal component analysis (SPCA) was adopted to extract sensitive bands and calculate the vegetation index for reducing the data dimensionality. Segmentation was then executed on each image in order to obtain leaf object samples and extract spectral, shape, and texture features of the object by the eCognition software (Trimble Navigation

Ltd., Broomfield, CO, USA). Finally, a decision tree model for leaf object identification of plants was constructed based on the above optimal features by the C 5.0 algorithm.

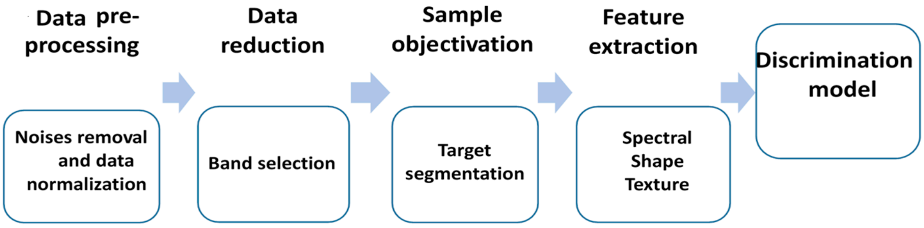

Figure 2. Flow of data processing. The identification process is based on leaf objects except for two parts of image preprocessing and data reduction.

2.4. Analysis Method

2.4.1. Data Pre-Processing

Data pre-processing consists of two steps: data noise removal with MNF and filter, and data normalization [14]. Every band of the imaging spectrometer data was checked so as to remove noises using MNF. After MNF, the eigenvalues, which were greater than the slope of the MNF eigenvalue curve, were selected to the adaptive filtering process and transformed back to the original spectral space with the noises removed. To eliminate the impacts of variations in spectral values due to illumination changes among areas of view and spectrometer, we performed data normalization on imaging spectral datasets such that the white reference board was used to calibrate the digital number to the relative reflectance.

2.4.2. Data Reduction

SPCA was applied to maintain the most vital and useful information in the imaging spectral datasets of corn/weeds [19]. The complete data set was first partitioned into several highly correlated subgroups by a correlation matrix and then PCA was conducted separately on each subgroup of data. Thirdly, the contribution was calculated of the sum of the square of the correlation coefficient between each band and each important principle component within a subgroup. Finally, the optimal band was selected based on the maximum of the contributions in each subgroup. The method is an effective way to retain desired features for corn/weeds identification, compared with conventional PCA. In the study, five wavelengths were selected, at 516 nm, 677 nm, 710 nm, 749 nm, and 843 nm.

2.4.3. Sample Objectivation

Before classifying plant species, plant object samples were obtained by segmentation in the eCognition Developer 8.7 software (Trimble Navigation Ltd., Broomfield, CO, USA). Combination of multiresolution segmentation and spectral difference segmentation was studied to subtract the background of plants and make a complete leaf of each species an object for sample selection. Images for segmentation were the single-band near infrared image at the wavelength of 910 nm and the NDVI (Normalized Difference Vegetation Index) image calculated from the bands of 710 nm and 762 nm. The NDVI image layer weight was set to 2, while the near-infrared image to 1. Scale parameter was set to 0.2 in the multiresolution segmentation while that of spectral difference segmentation was 0.19. This was conducted by the rule set in the software.

2.4.4. Spectral, Shape, and Texture Feature Extraction of Plant Objects

We measured a number of features from the raw corn/weeds images to be later used for classification purposes. A total of 11 features were considered to be most representative of the classes of interest, which correspond to spectral, shape, and textural characteristics of corn and weeds. They are described below.

(1) Spectral Features

It was clear that it was useful to use certain spectral indices or ratios to distinguish between weeds and crop based on selected wavebands. In this study, four vegetation indices were selected and calculated as spectral properties of plant objects. These indexes were Ratio Vegetation Index (RVI), Red Index (RI), the Ratio of the band 677 nm and 710 nm (M1), and the normalized difference index between the band 749 nm and 710 nm (M2), separately (Table 1).

Table 1. Formulas, names, and citations of vegetation indices used to calculate spectral features.

Vegetation Indices	Formulas	Reference Indices	Reference Formulas
RVI	R843/R677	RVI	NIR/R [20]
RI	(R710 − R516)/(R710 + R516)	RI	(R − G)/(R + G) [21]
M1	R677/R710	R680	R680/R710 [22]
M2	(R749 − R710)/(R749 + R710)	NDI	(R750 − R705)/(R750 + R705) [23]

Note: R, G, and NIR indicate the red, green, and near-infrared band, respectively; NDI is the normalized difference index, R680 is the reflectance at 680 nm, R516 is the reflectance of 516 nm, R677 is the reflectance of 677 nm.

(2) Shape Features

Shape features for object identification typically require invariance, which indicates that the values of the shape features will not be changed after the object is translated, rotated, rescaled, or repositioned [13]. Therefore, four kinds of features were determined in the study, including area, shape index, length/width, and length. The details of these features are shown in the Reference Book of eCognition Developer 8.7 [24].

The area describes the entire size of the object, which is the number of pixels multiplied by the size of the pixel.

The shape index describes the smoothness of an image object border. The smoother the border of an image object is, the lower the shape index. It is calculated from the border length feature of the image object divided by four times the square root of its area.

$$\alpha = \mathrm{e}/4\sqrt{A} \tag{1}$$

where α is the shape index; e is the perimeter; A is the area.

Length/width is the length-to-width ratio of an image object. The covariance matrix of the coordinates of the boundary pixels is the foundation for calculating length/width. The form of the covariance matrix is

$$\mathrm{S} = \begin{vmatrix} Var(X) & Cov(XY) \\ Cov(XY) & Var(Y) \end{vmatrix} \tag{2}$$

where X is the vector composed of the x coordinates of the boundary pixel; Y is the vector composed of the y coordinated of the boundary pixel; both X and Y use the same order. $Var(X)$ and $Var(Y)$ are the variances of X and Y, and $Cov\ (XY)$ is the covariance of X and Y.

After the covariance matrix is obtained, length/width can be represented by Equation (3).

$$\gamma = eig_1(S)/eig_2(S) \tag{3}$$

where $eig_1(S)$ is the maximum eigenvalue of the matrix S; $eig_2(S)$ is the minimum eigenvalue of the matrix S.

The length of a 2D image object is calculated using the length-to-width ratio.

$$\mathrm{l} = \sqrt{A\gamma} \tag{4}$$

where l is the length; γ is the length-to-width ratio; A is the area of the object.

(3) Texture Features

The texture is an important image feature that shows the regularity of the gray or color distribution for the image. In the macro scale, we can see different crop and weed cluster textures when the leaf shows different texture traits in the micro level. For example, the vein texture of the monocotyledonous plants is parallel or curved; that of the dicotyledonous plants is reticular. Therefore, the texture information of the plant leaf is useful to identify corn/weeds.

Gray level co-occurrence matrix (GLCM) is a statistical technique for texture analysis, which was presented by Haralick et al. [25]. The eCognition software generates the gray level co-occurrence matrix (GLCM) algorithm, which is a tabulation of how often different combinations of pixel gray levels occur in a scene. A different co-occurrence matrix exists for each spatial relationship. Calculation of textures is dependent upon the direction of the analysis and the distance. The gray co-occurrence can be specified in a matrix of relative frequencies P (i, j, s, β) with which two neighboring resolution cells separated by distances occur on the image, one with gray tone i and the other with gray tone y. The direction β includes four directions, Direction 0°, Direction 45°, Direction 90°, Direction 135°. Such matrices of spatial gray tone dependence frequencies are symmetric and a function of the angular relationship between the neighboring resolution cells as well as a function of the distance between them. Texture features are based on the GLCM matrix and algorithm of a particular feature, such as GLCM homogeneity being locally homogeneous in the image, GLCM contrast being the opposite of the homogeneity and a measure of the amount of local variation in the image, GLCM dissimilarity similar to contrast but increasing linearly, GLCM entropy representing equal distribution of the elements of GLCM.

It will be time-consuming and generate redundant information if the GLCM algorithm is performed on the imaging data of each waveband. Through testing, it was determined that three textural features of GLCM contrast, entropy, and homogeneity, based on three sensitive wavelengths of 516 nm, 677 nm, and 843 nm as well as the 90° direction, would be used (Table 2).

Table 2. Formulas and names of selected textural features.

Texture Features	Wavelengths	Formulas	Angle
Contrast	843 nm	$\sum_{i,j=0}^{N-1} P_{i,j}(i-j)^2$	90°
Entropy	677 nm	$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right)$	
Homogeneity	516 nm	$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$	

Note: i is the row number; j is the column number; $P_{i,j}$ is the normalized value in the cell i, j; N is the number of rows or columns.

2.4.5. Classifier and Validation

Identification of corn and weed species was done with a C 5.0 algorithm based on plant leaf objects. C 5.0 is an extension of the C 4.5 algorithm, and it is more efficient and uses less memory [26]. C 5.0 constructs the classification trees from discrete values based on the "information gain" calculated by the entropy. The C 5.0 model can split samples on the basis of the biggest information gain field. The sample subset that is obtained from the former split is split afterwards. The process continues until the sample subset cannot be split and is usually according to another field. Finally, on examining the lowest level split, those sample subsets that do not have a remarkable contribution to the model are rejected [27]. The decision tree method automatically discovers classification rules by using machine learning techniques. It uses the "information gain ratio" to determine the splits at each internal node of the decision tree [28]. In the study decision tree classifier construction options included 10 trials of boosting and global pruning.

All plant object samples were randomly split into two parts, 70% of which was training dataset and the other was testing dataset. The training set was useful to build the discrimination model, whereas the testing set was used for validating the classification. The classification quality was quantitatively assessed through the Global accuracy, Kappa coefficients, user's accuracy, and producer's accuracy, were all extracted from the confusion matrix.

3. Results and Discussion

3.1. Spectral Signatures of Corn and Weed Species

The spectral response curves for corn and seven weed species had similar change trends and no obvious difference in absorption bands as measured by FISS (Figure 3). The spectral curves of monocotyledonous plants were close together and separate from dicotyledonous plants, resulting from the morphological characteristics and internal structure of leaves for these two groups of plants. In the visible range (446–760 nm), the spectral curves of the corn and other weeds basically overlapped, except the ones of POROL and AMARE. However, in the near infrared range (760–912 nm), the difference was visible among corn and weeds, except that between corn and ELEIN. Therefore, it is difficult to distinguish corn and all weeds by the spectrum in the whole band range. Also, it is necessary to do dimensionality reduction in order to quantitatively analyze spectral features of corn and various weeds.

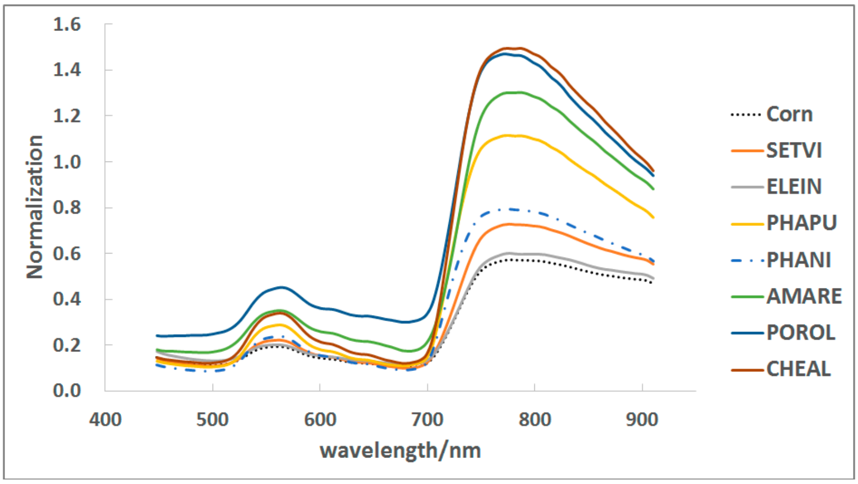

Figure 3. Spectral response curves of corn and seven weed species. Each curve is the mean of normalized values in all leaf images of each plant after data preprocessing. SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

In the study, the SPCA algorithm was used to select the most appropriate wavelengths for corn and seven weed species. The five selected wavelengths were 516 nm, 677 nm, 710 nm, 749 nm, and 843 nm, which is consistent with the previous research [14,29–31]. These wavelengths are closely related to the cytochrome composition and internal structure of leaves. The 516 nm waveband locates in the blue-green spectral region (400–550 nm), which is related to the anthocyanin content of the leaf, while 677 nm, 710 nm, and 749 nm wavelengths are located in the red light region, and are related to the chlorophyll content of the leaf. The 843 nm waveband is in the near infrared range, which is sensitive to the cell organization structure (the cell wall, intercellular space etc.). The normalized reflectance means of corn and weeds for each selected wavelength and vegetation index highlighted that there were similar changes for corn/weeds at two wavelengths of 516 nm and 677 nm, which showed no significant difference between corn and other weeds except POROL and AMARE (Table 3). However, the mean difference was significant between corn and PHAPU, or AMARE, or POROL, or CHEAL at the sensitive bands of 749 nm and 843 nm. The variation was more obvious at the 710 nm wavelength between corn and all dicotyledonous weeds, or between POROL and other weeds. In a word, it was still not ideal for completely differentiating between the corn and the seven weeds or between weeds only according to these selected wavebands.

Four vegetation indices, RVI, RI, M1, and M2, were calculated to improve the identification accuracy of the corn and weed species (Table 3). Compared to selected wavelengths, vegetation indices can increase the differences in the normalized reflectance means between corn and weed species or, to some extent, between weeds.

Table 3. Normalized reflectance means for corn and seven weed species for each selected wavelength and vegetation index used to improve plant leaf object identification (n = 155).

Wavelengths and Vegetation Indices	Corn	SETVI	ELEIN	PHAPU	PHANI	AMARE	POROL	CHEAL
R516	0.12cde	0.14bd	0.12cde	0.11cde	0.10c	0.18b	0.24a	0.14be
R677	0.10cde	0.13bd	0.10cde	0.11cde	0.09cde	0.17b	0.28a	0.12be
R710	0.17cd	0.23bd	0.16cd	0.29b	0.25b	0.31b	0.44a	0.33b
R749	0.55be	0.75bcd	0.44e	0.99ad	0.73bd	1.11ac	1.26a	1.30a
R843	0.55bc	0.75bd	0.47c	0.93ade	0.68be	1.07ad	1.17a	1.21a
RVI	5.88e	5.96de	4.65f	8.72b	7.42c	6.36d	4.32f	9.93a
RI	0.20de	0.23d	0.14e	0.42a	0.43a	0.27c	0.29b	0.40a
M1	0.55b	0.55b	0.66a	0.37c	0.38c	0.53b	0.63a	0.37c
M2	0.51bde	0.52bde	0.47c	0.54ae	0.49bc	0.55ad	0.49bc	0.58a

Note: Values within each row followed by the same letter are not significantly different through the Least Significant Difference (0.05) test. SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters. R516: the normalization value at 516 nm, others are similar, RVI: R843/R677, RI: (R710 − R516)/(R710 + R516), M1: R677/R710, M2: (R749 − R710)/(R749 + R710).

3.2. Shape Signatures of Corn and Weed Species

Shape features are important information for detection of crop and weed species based on a plant object's shape. In the study, four features of area, shape index, length, and length/width were selected and calculated for discriminating between corn and weeds, which describe the spatial attributes of plants. Area and length are simple parameters, while shape index and length/width are the recombination of length, width, perimeter, and area, the values of which are related to these parameters (Table 4). It was found that the differences between corn and various weeds were obvious using these shape features. The mean differences between corn and each dicotyledonous weed reached significance at the 5% level for these four features and the ones between corn and each monocotyledonous weed were significant in area and length but not in shape index and length/width. Corn is a gramineous plant, leaves of which are longer and wider than varieties of weeds, especially in the late period of corn growth. The leaf length and width of monocotyledonous weeds including SETVI and ELEIN are much smaller than the ones of corn, but their shape index and length/width are comparable to the ones of corn. Shape signatures of the dicotyledonous weeds differ from the monocotyledonous weeds, leaf length, and width of which are similar. The mean differences of the four shape indices between them were significant at the 5% level. However, the mean differences were not the same between varieties of monocotyledonous weeds or between dicotyledonous weed species, which resulted from the same types with similar shape features. Thus, it can be seen that shape features, especially area and length, were readily used to differentiate between corn and many kinds of weeds but not between weeds. Furthermore, leaf images were acquired on the premise of full expansion of leaves in the laboratory, from which these shape features were easily obtained. However, these features of the whole leaves are not easily calculated when data are collected in the field due to the influence of leaf blades covering each other and illumination changes, which reduce the effectiveness of shape features and limits their further application.

Table 4. Comparison of means of each shape feature for corn/weeds leaf objects (n = 155).

Shape Features	Corn	SETVI	ELEIN	PHAPU	PHANI	AMARE	POROL	CHEAL
Area	8876a	2957bd	909c	3012b	1948d	2646bd	180e	706c
Shape index	2.29c	2.39ac	2.84a	1.24b	1.57d	1.24b	1.32e	1.23b
Length	338a	224c	150d	70b	58g	75b	24e	38f
Length/width	9.17a	11.41a	14.55a	1.19d	1.16d	1.53b	2.09c	1.42bd

Note: Values within each row followed by the same letter are not significantly different using the Least Significant Difference (0.05) test. SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

3.3. Textural Signatures of Corn and Weed Species

The texture features of an object reflect the properties of the object itself, which is useful for differing from other objects. The monocotyledonous and dicotyledonous plants provide different patterns of veins in a leaf, which can be described by textural parameters. The study selected and calculated three textural features including entropy, contrast, and homogeneity. Table 5 illustrates the comparison results of the mean values for three textural features. For entropy, PHAPU had higher values than other weeds, and the differences of the mean value between corn and each variety of monocotyledonous weeds were significant. However, the differences between weed species were not the same, such as the significant difference between SETVI and ELEIN or between AMARE and POROL, and no difference between PHAPU and PHANI. For contrast, CHEAL had the highest value, while corn had the smallest value. The multiple comparison results showed that CHEAL was significantly different from other weeds except PHANI, and corn significantly differed from various dicotyledonous weeds. For homogeneity, corn had the highest mean value, while POROL and CHEAL had smaller values. This indicated that the textures of POROL and CHEAL are rougher than that of corn. From the view of multiple comparisons, corn was significantly different from various dicotyledonous weeds, whilst the differences between weed species were not the same. Differences between corn and each monocotyledonous weed existed in the textural feature of entropy, and differences between corn and each dicotyledonous weed occurred for contrast and homogeneity, but differences among weed species were not consistent. The vein pattern of the corn is parallel, which is different from the reticulate texture in dicotyledonous plants. In addition, texture is closely related to the shape of leaves. This demonstrates that texture is a useful characteristic for classifying plant species.

Table 5. Mean values for three textural features of entropy, contrast, and homogeneity for corn and seven weed species (n = 155).

Texture Features	Corn	SETVI	ELEIN	PHAPU	PHANI	AMARE	POROL	CHEAL
Entropy	7.67a	7.45c	6.92b	7.87a	7.75a	7.61ac	5.84d	7.06b
Contrast	34.33e	51.83ce	126.67bce	192.01bd	287.20ab	86.66cd	219.61b	445.73a
Homogeneity	0.24a	0.24ab	0.18bd	0.14cd	0.11cf	0.16d	0.07e	0.09ef

Note: Values within each row followed by the same letter are not significantly different using the Least Significant Difference (0.05) test. SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

3.4. Discrimination Model of Corn and Weed Species by Combining Spectra, Shape, and Texture Features

The identification model for corn and weed species was constructed based on the above stated spectral, shape, and textural features (Figure 4). As shown in the decision tree workflow, the model utilized spectral, shapen and textural signatures, which were respectively M1, M2, shape index, area, length, and entropy. The set of rules was defined by these attributes and their respective thresholds in the decision tree. The structure of the relationship between attributes and thresholds was implemented automatically by C 5.0. Firstly, M1 and the shape index divided corn/weeds into the two categories of monocotyledonous and dicotyledonous plants. The spectral feature played an important role in this process. Secondly, each type was subdivided. Two rules with length and entropy characteristics were generated for distinguishing various monocotyledonous plants, whilst four features of M2, area, length, and entropy were used for detecting varieties of dicotyledonous weeds. The rules and classifying characteristics for identifying corn and weed species were simple in the model, which meets the design requirements for a portable and intelligent crop/weeds detecting system. The tree structure enabled an easy understanding of the form of decisions used to classify corn and weed species. Thus, the combination of spectral, shape, and texture attributes presented a potentially useful way to solve the problem of corn/weeds detection. In addition, among the attributes extracted by the eCognition platform, some attributes were not used, such as RVI, RI, length/width, contrast, and homogeneity. We believe that this occurred due to the similarity of the characteristics between corn and weed species or between weed species. It is also of note that a new decision tree workflow for plant identification should be made with the distribution of features, due to the crops and the weeds having different spectral, shape, and textural features at different stages of growth.

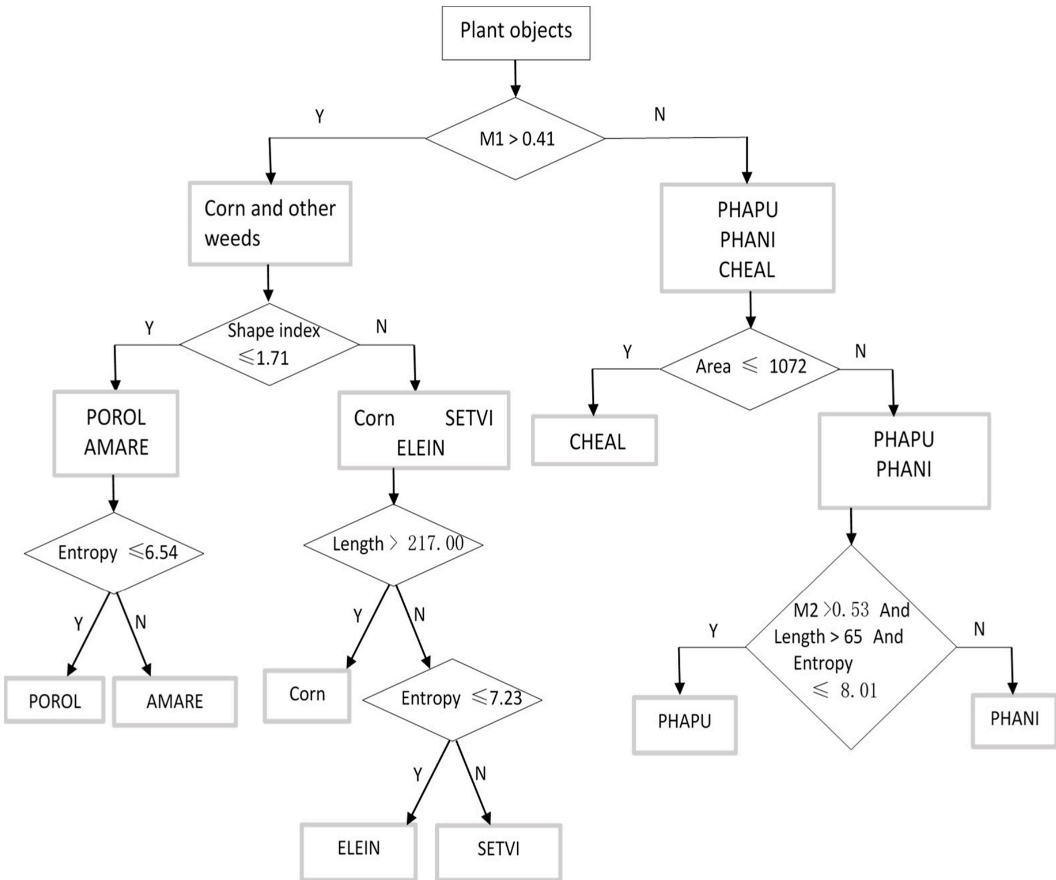

Figure 4. Decision tree workflow for discriminating corn and weed species by spectral, shape, and texture features. The decision rules are from the training database in all spectrometer images (n = 100). SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters, M1: the ratio of the band 677 nm and 710 nm, M2: the normalized difference index between the band 749 nm and 710 nm, Y: yes, N: no.

Assessment of the discriminating quality of the model further demonstrated the reliability and accuracy of the decision tree model with test samples alone (Table 6). Both Global Accuracy (GA) and Kappa (k) coefficients of the model were respectively higher than 95%. The value found for the Kappa coefficient indicates conformity of the classification with the reference data. Producer's and user's accuracies are ways of representing individual category accuracies instead of just the overall classification accuracy. Producer's accuracy presents the accuracy of classification, which is the fraction of correctly classified pixels or objects with regard to all pixels or objects of that ground truth class. User's accuracy presents the reliability of classes in the classified image, which is the fraction of correctly classified pixels or objects with regard to all pixels or objects classified as this class in the classified image. Corn, SETVI, PHAPU, AMARE, POROL, and CHEAL can be identified. However, only 83% of ELEIN and 88% of PHANI objects in the classified image actually represented ELEIN or PHANI on the ground. For SETVI and PHAPU, ground truth objects also appeared as in the classified image and were 80% and 86%, respectively. Leaves of SETVI and ELEIN are very narrow and easily roll-folded when isolated from the whole plant, which resulted in difficult data acquisition. The number of features for detecting PHAPU and PHANI was more than that of other plants and involved all three categories signatures, which affected the distinguishing accuracy.

Table 6. Identification accuracy for corn and weed species by spectral, shape, and textural features.

Species	Global Accuracy (%)	Kappa Coefficient	User's Accuracy (%)	Producer's Accuracy (%)
Corn	96	0.96	100	100
SETVI			100	80
ELEIN			83	100
PHAPU			100	86
PHANI			88	100
AMARE			100	100
POROL			100	100
CHEAL			100	100

Note: SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

3.5. Comparison of the Identification Models for Corn and Weeds Species with Different Feature Combinations

Each attribute has a different impact on the identification models for corn and weeds species (Table 7). Various accuracies of the model with these combinations were lower than that with the combination of spectra, shape, and texture features, and the number of features used was more. The model with three categories of features utilized six characteristics including M1, M2, shape index, area, length, and entropy, while the ones with two categories of features used seven or eight features. Furthermore, the decision tree models with combinations of two categories of features were more complicated than that with the combination of spectra, shape, and texture features. The model with three categories of features did not need boosting and only contained a classifier, while the ones with two categories of features consisted of 10 boosting trials. However, it was found that each variety of features played different roles in the identification models from Table 7. The Global accuracy and Kappa of the models with shape features involving all reached more than 90%, while the one of integrating spectra and texture features was more than 85%. As expected, shape is the principle characteristic for corn/weeds identification. In fact, it is not easy to obtain shape parameters of the whole leaf in the field due to the problems of leaf shielding, illumination difference, and morphological changes during plant growth. This reduces the effectiveness of shape features for discriminating between crops and weeds. The results demonstrated that the integration of spectral features can increase the accuracy of the models, especially the combination of spectral and shape characteristics. This indicated that the spectral attribute can provide important support for improving the overall identification accuracy and the identification accuracy of an individual plant. Meanwhile, the extraction of spectral features is easier than that of shape features. In addition, it was found that the models with textural feature participation can reduce the identification error rate of the individual plant although the Global accuracy was not increased. This indicated that texture can be considered to be used to obtain a good performance.

Table 7. Identification accuracy of corn and weed species.

Species	Global Accuracy (%)			Kappa Coefficient			User's Accuracy (%)			Producer's Accuracy (%)		
	Spectral + Shape	Spectral + Texture	Shape + Texture	Spectral + Shape	Spectral + Texture	Shape + Texture	Spectral + Shape	Spectral + Texture	Shape + Texture	Spectral + Shape	Spectral + Texture	Shape + Texture
Corn	95	90	93	0.94	0.87	0.92	100	100	100	100	83	100
SETVI							100	60	100	100	60	80
ELEIN							100	100	83	100	80	100
PHAPU							86	88	100	86	100	71
PHANI							83	100	86	71	86	86
AMARE							100	78	80	100	88	100
POROL							100	91	100	100	100	100
CHEAL							88	100	100	100	100	100

Note: SETVI: foxtail, ELEIN: goosegrass, PHAPU: round leaf pharbitis, PHANI: lobed leaf pharbitis, AMARE: redroot amaranth, POROL: purslane, CHEAL: lambs quarters.

4. Conclusions

The viability of integrating spectral, shape, and texture characteristics to identify corn and seven weed species from the imaging spectrometer data was investigated. The identification approach combines the object-oriented algorithm and the decision tree C 5.0. The former can generate plant leaf objects by segmentation and is used to extract the attributes considered to be most representative of the classes of interest, while the latter is a classifier to identify eight plant species. We selected characteristics among three categories of spectral, spatial, and textural features. Classification results showed that the decision tree model combining these three types of characteristics containing simpler decision rules, had higher levels of exactitude and used less attributes than the ones with combination of two kinds of features. In addition, we observed the importance of various features in classifying plants. It was found that spectral and shape features can be regarded as the preferred characteristic to develop the means of weed identification from the view of accessibility to crop/weeds discriminant features. Additionally, the identification error rate of some weeds was able to be reduced when the texture features were considered. Our preliminary results can provide valuable support for the portable device development of intelligent weed identification. However, many factors may affect the features of extraction of crops and weeds in the natural environment, such as leaf blocking, leaf angle inclination, and solar altitude. As to whether this method could be applied in the field requires further investigation.

Acknowledgments: The study was funded by National Natural Science Foundation of China (Grant No. 41301505, 41301471), Special Fund for Agro-scientific Research in the Public Interest, Ministry of Agriculture (Grant No. 201303031) and Specialized Research Fund for the Doctoral Program of Higher Education of China (20123228120004).

Author Contributions: Fenfang Lin and Dongyan Zhang conceived and designed the experiments; Dongyan Zhang performed the experiments; Xiu Wang and Xinfu Chen analyzed the data; Fenfang Lin wrote the paper; Yanbo Huang revised the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Qiang, S. *Weed Science*, 2nd ed.; China Agriculture Press: Beijing, China, 2009.
2. Cope, J.S.; Corney, D.; Clark, J.Y.; Remagnino, P.; Wilkin, P. Plant species identification using digital morphometrics: A review. *Expert Syst. Appl.* **2012**, *39*, 7562–7573. [CrossRef]
3. Granitto, P.M.; Navone, H.D.; Verdes, P.F.; Ceccatto, H.A. Weed seeds identification by machine vision. *Comput. Electron. Agric.* **2002**, *33*, 91–103. [CrossRef]
4. Taghadomi-Saberi, S.; Hemmat, A. Improving field management by machine vision—A review. *Agric. Eng. Int. CIGR J.* **2015**, *17*, 92–111.
5. Borregaard, T.; Nielsen, H.; Nørgaard, L.; Have, H. Crop-weed discrimination by line imaging spectroscopy. *J. Agric. Eng. Res.* **2000**, *75*, 389–400. [CrossRef]
6. Guerrero, J.M.; Pajares, G.; Montalvo, M.; Romeo, J.; Guijarro, M. Support vector machines for crop/weeds identification in maize fields. *Expert Syst. Appl.* **2012**, *39*, 11149–11155. [CrossRef]
7. Lamb, D.W.; Brown, R.B. Remote-sensing and mapping of weeds in crops. *J. Agric. Eng. Res.* **2001**, *78*, 117–125. [CrossRef]
8. Meyer, G.E. Machine vision detection parameters for plant species identification. *Proc. SPIE* **1999**, *3543*. [CrossRef]
9. Mao, W.; Hu, X.; Zhang, X. Weed detection based on the optimized segmentation line of crop and weed. In *Computer and Computing Technologies in Agriculture, Volume ii, In Proceedings of the First ifip tc 12 International Conference on Computer and Computing Technologies in Agriculture (CCTA 2007), Wuyishan, China, 18–20 August 2007*; Li, D., Ed.; Springer: Boston, MA, USA, 2008; pp. 959–967.
10. Meyer, G.E. Design of an optical weed sensor using plant spectral characteristics. *Trans. ASAE* **2001**, *44*, 409–419.

11. Moshou, D.; Kateris, D.; Pantazi, X.E.; Gravalos, I. Crop and weed species recognition based on hyperspectral sensing and active learning. In *Precision Agriculture '13*; Stafford, J.V., Ed.; Wageningen Academic Publishers: Wageningen, The Netherlands, 2013; pp. 555–561.
12. Piron, A.; Leemans, V.; Kleynen, O.; Lebeau, F.; Destain, M.F. Selection of the most efficient wavelength bands for discriminating weeds from crop. *Comput. Electron. Agric.* **2008**, *62*, 141–148. [CrossRef]
13. Symonds, P.; Paap, A.; Alameh, K.; Rowe, J.; Miller, C. A real-time plant discrimination system utilising discrete reflectance spectroscopy. *Comput. Electron. Agric.* **2015**, *117*, 57–69. [CrossRef]
14. Li, Y.; Zhang, L.; Wei, Y.; Huang, C. Weed identification using imaging spectrometer data. *J. Remote Sens.* **2013**, *78*, 87–89.
15. Tits, L.; Keersmaecker, W.D.; Somers, B.; Asner, G.P.; Farifteh, J.; Coppin, P. Hyperspectral shape-based unmixing to improve intra- and interclass variability for forest and agro-ecosystem monitoring. *ISPRS J. Photogramm. Remote Sens.* **2012**, *74*, 163–174. [CrossRef]
16. Somers, B.; Delalieux, S.; Verstraeten, W.W.; Verbesselt, J.; Lhermitte, S.; Coppin, P.; Entekhabi, D.; Kerekes, J.P.; Miller, E.L.; Reising, S.C. Magnitude- and shape-related feature integration in hyperspectral mixture analysis to monitor weeds in citrus orchards. *Geosci. Remote Sens. IEEE Trans.* **2009**, *47*, 3630–3642. [CrossRef]
17. Alchanatis, V.; Ridel, L.; Hetzroni, A.; Yaroslavsky, L. Weed detection in multi-spectral images of cotton fields. *Comput. Electron. Agric.* **2005**, *47*, 243–260. [CrossRef]
18. Tong, Q.; Xue, Y.; Wang, J.; Zhang, L.; Fang, J.; Yang, Y.; Liu, X.; Hongxing, Q.I.; Zheng, L. Development and application of the field imaging spectrometer system. *J. Remote Sens.* **2010**, *14*, 409–422.
19. Jia, X.; Richards, J.A. Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 538–542.
20. Pearson, R.L.; Miller, L.D. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. In Proceedings of the Eighth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, USA, 2–6 October 1972; pp. 1357–1381.
21. Escadafal, R.; Huete, A. Improvement in remote sensing of low vegetation cover in arid regions by correcting vegetation indices for soil "noise". *C. R. Acad. Sci.* **1991**, *312*, 1385–1391.
22. Maccioni, A.; Agati, G.; Mazzinghi, P. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *J. Photochem. Photobiol. B Biol.* **2001**, *61*, 52. [CrossRef]
23. Gitelson, A.; Merzlyak, M.N. Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *J. Photochem. Photobiol. B Biol.* **1994**, *22*, 247–252. [CrossRef]
24. GmbH, T.G. *Ecognition Developer 8.7: Reference Book*; Trimble: Munich, Germany, 2011.
25. Haralick, R.M. Statistical and structural approaches to texture. *Proc. IEEE* **1979**, *67*, 786–804. [CrossRef]
26. Quinlan, J.R. *C 4.5:Programs for Machine Learning*; Morgan Kaufmann: San Francisco, CA, USA, 1992.
27. Patil, N.; Lathi, R.; Chitre, V. Comparison of C 5.0 & Cart Classification Algorithms Using Pruning Technique. *Int. J. Eng. Res. Technol.* **2012**, *1*, 1–5.
28. Boulle, M. Khiops: A statistical discretization method of continuous attributes. *Mach. Learn.* **2004**, *55*, 53–69. [CrossRef]
29. Liu, B.; Fang, J.Y.; Liu, X.; Zhang, L.F.; Zhang, B.; Tong, Q.X. Research on crop-weed discrimination using a field imaging spectrometer. *Spectrosc. Spectr. Anal.* **2010**, *30*, 1830–1833.
30. Zhang, D.; Qin, Z.; Wei, D.; Xiu, W. Using hyperspectral sensor to recognize monocotyledonous/dicotyledonous weeds and crops. *Infrared Laser Eng.* **2013**, *42*, 208–213.
31. Zwiggelaar, R. A review of spectral properties of plants and their potential use for crop/weed discrimination in row-crops. *Crop Prot.* **1998**, *17*, 189–206. [CrossRef]