



Article

Evaluation of an Airborne Remote Sensing Platform Consisting of Two Consumer-Grade Cameras for Crop Identification

Jian Zhang ^{1,2}, Chenghai Yang ^{2,*}, Huaibo Song ^{2,3}, Wesley Clint Hoffmann ², Dongyan Zhang ^{2,4} and Guozhong Zhang ⁵

¹ College of Resource and Environment, Huazhong Agricultural University, 1 Shizishan Street, Wuhan 430070, China; jz@mail.hzau.edu.cn

² USDA-Agricultural Research Service, Aerial Application Technology Research Unit, 3103 F & B Road, College Station, TX 77845, USA; songyangfeifei@163.com (H.S.); clint.hoffmann@ars.usda.gov (W.C.H.); hello-lion@hotmail.com (D.Z.)

³ College of Mechanical and Electronic Engineering, Northwest A&F University, 22 Xinong Road, Yangling 712100, China

⁴ Anhui Engineering Laboratory of Agro-Ecological Big Data, Anhui University, 111 Jiulong Road, Hefei 230601, China

⁵ College of Engineering, Huazhong Agricultural University, 1 Shizishan Street, Wuhan 430070, China; zhanggz@mail.hzau.edu.cn

* Correspondence: chenghai.yang@ars.usda.gov; Tel.: +1-979-571-5811; Fax: +1-979-260-5169

Academic Editors: Mutlu Ozdogan, Clement Atzberger and Prasad S. Thenkabail

Received: 21 January 2016; Accepted: 11 March 2016; Published: 18 March 2016

Abstract: Remote sensing systems based on consumer-grade cameras have been increasingly used in scientific research and remote sensing applications because of their low cost and ease of use. However, the performance of consumer-grade cameras for practical applications has not been well documented in related studies. The objective of this research was to apply three commonly-used classification methods (unsupervised, supervised, and object-based) to three-band imagery with RGB (red, green, and blue bands) and four-band imagery with RGB and near-infrared (NIR) bands to evaluate the performance of a dual-camera imaging system for crop identification. Airborne images were acquired from a cropping area in Texas and mosaicked and georeferenced. The mosaicked imagery was classified using the three classification methods to assess the usefulness of NIR imagery for crop identification and to evaluate performance differences between the object-based and pixel-based methods. Image classification and accuracy assessment showed that the additional NIR band imagery improved crop classification accuracy over the RGB imagery and that the object-based method achieved better results with additional non-spectral image features. The results from this study indicate that the airborne imaging system based on two consumer-grade cameras used in this study can be useful for crop identification and other agricultural applications.

Keywords: consumer-grade camera; crop identification; RGB; near-infrared; pixel-based classification; object-based classification

1. Introduction

Remote sensing has played a key role in precision agriculture and other agricultural applications [1]. It provides a very efficient and convenient way to capture and analyze agricultural information. As early as 1972, the Multispectral Scanner System (MSS) sensors were used for accurate identification of agricultural crops [2]. Since then, numerous commercial satellite and custom-built airborne imaging systems have been developed for remote sensing applications with agriculture being

one of the major application fields. Remote sensing is very mysterious for most people who usually perceive it as complex designs and sophisticated sensors on satellites and other platforms. This is true for quite a long time, especially for the scientific-grade remote sensing systems on satellites and manned aircraft. However, with the advances of electronic imaging technology, remote sensing sensors have significant breakthroughs, including digitalization, miniaturization, ease of use, high resolution, and affordability. Recently, more and more imaging systems based on consumer-grade cameras have been designed as remote sensing platforms [3–6].

As early as in the 2000s, consumer-grade cameras with 35 mm film were mounted on a unmanned aerial vehicle (UAV) platform to acquire imagery in a small area for range and resource managers [7]. Remote sensing systems with consumer-grade cameras have such advantages over scientific-grade platforms as low cost with high spatial resolution [8,9] and easy deployment [7]. There are two main types of consumer-grade cameras: non-modified and modified. The non-modified cameras capturing visual light with R, G, and B channels have been commonly used since the early periods of consumer-grade cameras for remote sensing [7,10,11]. Additionally, this type of cameras have been commonly used for aerial photogrammetric surveys [12]. However, most scientific-grade remote sensing sensors include near-infrared (NIR) detection capabilities due to its sensitivity for plants, water and other cover types [13]. Therefore, using modified cameras to capture NIR light with different modifying methods is becoming popular.

Modified NIR cameras are generally obtained by replacing the NIR-blocking filter in front of the complementary metal-oxide-semiconductor (CMOS) or the charge-coupled device (CCD) with a long-pass NIR filter. A dual-camera imaging system is a common choice with one camera for normal color imaging and the other for NIR imaging [14,15]. With a dual-camera configuration, four-band imagery with RGB and NIR sensitivities can be captured simultaneously. Some consumer-grade cameras have been modified to capture this imagery by this method [10]. The other method is to remove the internal NIR-blocking filter and replace it with a blue-blocking [16] or a red-blocking filter [17]. This method can capture three-band imagery, including the NIR band and two visible bands, with just one single camera.

Both methods can be implemented, but both still have some issues [16]. The separate RGB and NIR images from dual cameras need to be registered for dual cameras. The NIR light may influence the other two bands for the single camera when capturing the NIR band. In other words, extensive post-processing is required to get the desired imagery. Therefore, it is necessary to decide if a simple unmodified RGB camera or a relatively complex dual-camera system with the NIR band should be selected. Many remote sensing imaging systems based on unmodified single cameras have been used for some agricultural applications to achieve satisfying results [9,14,18]. Consumer-grade cameras cannot capture very detail spectral information, but most of them can obtain very high or ultra-high ground resolution because these cameras can be carried on such low-flying platforms as UAV, airplanes and balloons. Moreover, with the rapid advancement of object-based image analysis (OBIA) methods, imagery with abundant shape, texture and context information could usually improve classification performance for practical applications [19,20]. Therefore, to choose pixel-based or object-based image processing methods is another question that needs to be considered.

Despite some shortcomings, consumer-grade cameras with the attributes of low cost and easy deployment have been gradually used in different application fields in the last decade, in particular for agricultural applications, such as crop identification [10], monitoring [11,14,21], mapping [9], pest detection [18], species invasion [17] and crop phenotyping [6,22]. It is becoming a reality for farmers to use these consumer-grade cameras for crop production and protection. Therefore, it is important to evaluate and compare the performances of normal RGB cameras and modified NIR cameras using different image processing methods.

Scientists at the Aerial Application Technology Research Unit at the U.S. Department of Agriculture-Agricultural Research Service's Southern Plains Agricultural Research Center in College Station, Texas has assembled a dual-camera imaging system with both visible and NIR sensitivities

using two consumer-grade cameras. The overall goal of the study was to evaluate the performance of RGB imagery alone and RGB combined with NIR imagery for crop identification. The specific objectives were to: (1) compare the differences between the three-band imagery and four-band imagery for crop identification; (2) assess the performance on image classification between the object-based and pixel-based analysis methods; and (3) make recommendations for the selection of imaging systems and imagery processing methods for practical applications.

2. Materials and Methods

2.1. Study Area

This study was conducted at the Brazos River Valley near College Station, Texas, in July 2015 within a cropping area of 21.5 km² (Figure 1). This area is located at the lower part of the Brazos River (the “Brazos Bottom”) with humid subtropical climate.

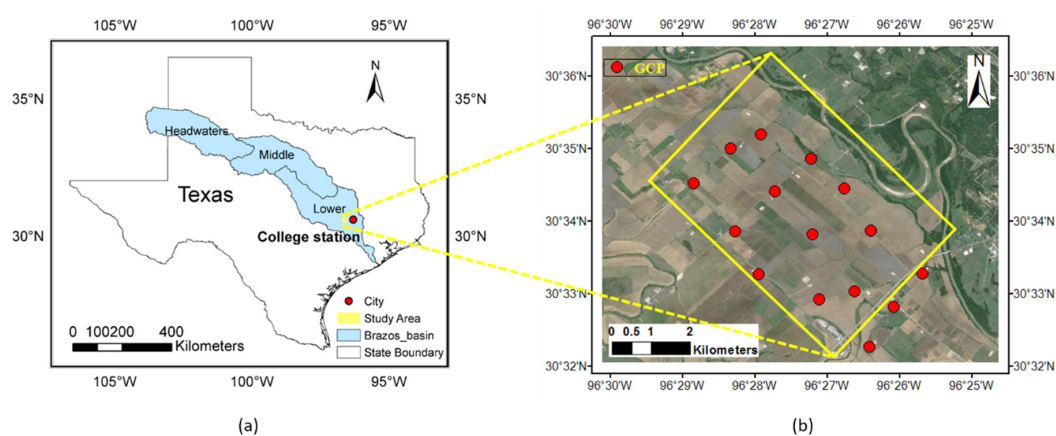


Figure 1. The study area: (a) Geographic map of the study area, which is located in the bottom of Brazos basin and near College Station, Texas; and (b) image map of the study area from Google Earth.

In this area, the main crops were cotton, corn, sorghum, soybean and watermelon in the 2015 growing season. Cotton was the main crop with the largest cultivated area, and it had very diverse growing conditions due to different planting dates and management conditions. Most cornfields were near physiological maturity with very few green leaves, and most of sorghum fields were in the generative phase reflected by beginning senescence at the imaging time. Especially corn was drying in fields for harvest. Soybean and watermelon were at the vegetative growth stages. Due to cloudy and rainy weather in much of May and June, aerial imagery was not acquired during the optimum period of crop discrimination, based on the crop calendars for this study area. However, this type of weather conditions is probably a common dilemma for agricultural remote sensing.

2.2. Imaging System and Airborne Image Acquisition

2.2.1. Imaging System and Platform

The dual-camera imaging system used in this study consisted primarily of two Nikon D90 digital CMOS cameras with Nikon AF Nikkor 24mm f/2.8D lenses (Nikon, Inc., Melville, NY, USA). One camera was used to capture three-band RGB images. The other camera was modified to capture NIR images after the infrared-blocking filter installed in front of the CMOS of the camera was replaced by a 720 nm long-pass filter (Life Pixel Infrared, Mukilteo, WA, USA). The other components of the system included two GPS receivers, a video monitor and a wireless remote trigger as shown in Figure 2. The detailed description of this system can be found in a single-camera imaging system described by Yang *et al.* [18]. The difference between the two imaging systems was that the single-camera system contained only one Nikon D90 camera for taking RGB images, while the dual-camera system had a the

RGB camera and a modified camera for NIR imaging necessary for this study. This dual-camera imaging system was attached via a camera mount box on to an Air Tractor AT-402B as shown in Figure 2.

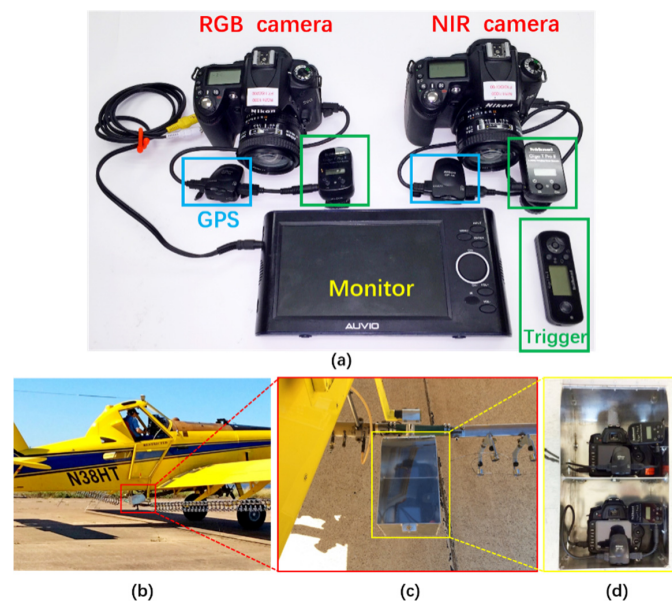


Figure 2. Imaging system and platform: (a) Components of the imaging system (two Nikon D90 cameras with Nikkor 24 mm lenses, two Nikon GP-1A GPS receiver, a 7-inch portable LCD video monitor, a wireless remote shutter release); (b) cameras mounted on the right step of an Air Tractor AT-402B; (c) close-up picture showing the custom-made camera box; and (d) close-up picture of the cameras in the box.

2.2.2. Spectral Characteristics of the Cameras

The spectral sensitivity of the two cameras was measured in the laboratory through a monochromator (Optical Building Blocks, Inc., Edison, NJ, USA) and a calibrated photodiode. The two cameras were spectrally calibrated with the lenses by taking photographs of monochromatic light from the monochromator projected onto a white panel. A calibrated photodiode was positioned at the same distance of the camera to measure the light intensity. The relative spectral response of one channel could be calculated for a given wavelength λ and a given channel (RGB) as shown in Equation (1) [23].

$$R(\lambda, n) = \frac{C(\lambda, n) - C_{dark}}{I(\lambda)}, \quad n = r, g, b; \quad \lambda = 400 - 1000 \text{ nm} \quad (1)$$

where $R(\lambda, n)$ is the spectral response of channel $n = r, g, b$ at λ wavelength. $I(\lambda)$ is the light intensity measured with the photodiode at λ wavelength. $C(\lambda, n)$ is the digital count corresponding to channel $n = r, g, b$ at λ wavelength. C_{dark} is the mean digital count of the dark background of channel $n = r, g, b$ at λ wavelength. Wavelength ranged from 400 to 1000 nm, and the measurement wavelength interval was 20 nm. The average digital count for each channel was determined for the center of the projected light beam using image-processing software (MATLAB R2015a, MathWorks, Inc., Natick, MA, USA). In addition, the images were recorded by the raw camera format.

From the normalized spectral sensitivity of the two cameras (Figure 3), the sensitivity varied from 400 to 700 nm for the non-modified camera and from 680 to 1000 nm for the modified camera. With the 720 nm long-pass filter, the spectral response rose from 0 at 680 nm to maximum at 720 nm. It can be seen that the spectral sensitivity curves have some overlaps among the channels of each camera. This is very common in consumer-grade cameras, and is also one of the reasons that this type of cameras had not been commonly used for most scientific applications in the past. For the modified camera, the red channel had a much stronger response than the other two channels (blue and green)

and the monochrome imaging mode in the NIR range. Thus, the red channel was chosen as the NIR image for remote sensing applications.

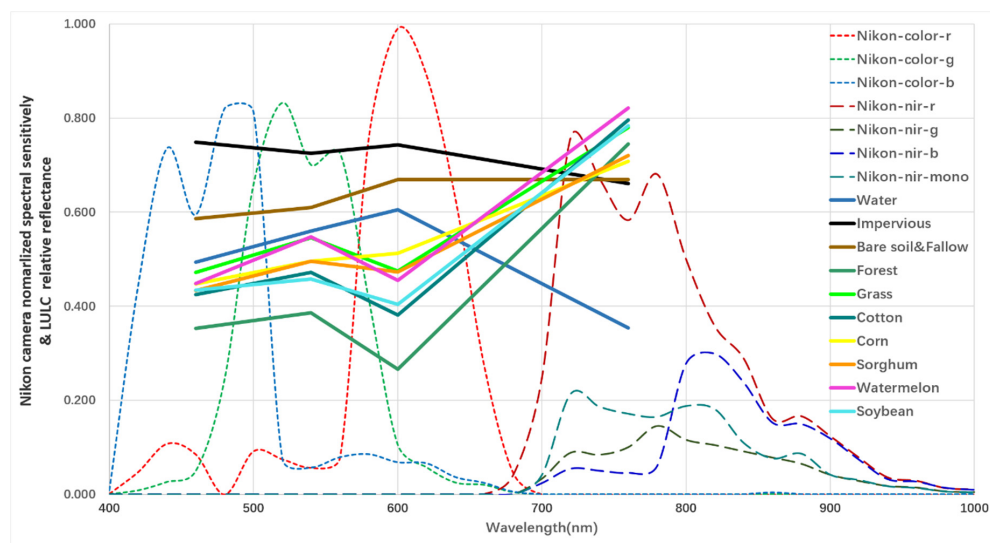


Figure 3. Normalized spectral sensitivity of two Nikon D90 cameras and relative reflectance of 10 land use and land cover (LULC) classes. The dotted lines represent different channels of the RGB camera (Nikon-color-r, Nikon-color-g, and Nikon-color-b) and the modified NIR camera (Nikon-nir-r, Nikon-nir-g, Nikon-nir-b, and Nikon-nir-mono). The solid lines represent the relative reflectance of 10 LULC classes.

2.3. Image Acquisition and Pre-Processing

2.3.1. Airborne Image Acquisition

Airborne images were taken from the study area at altitudes of 1524 m (5000 ft.) above ground level (AGL) with a ground speed of 225 km/h (140 mph) on 15 July 2015 under sunny conditions. The spatial resolution was 0.35 m at this height. In order to achieve at least 50% overlaps along and between the flight lines, images were acquired at 5-s intervals. Both cameras simultaneously and independently captured 144 images each over the study. Moreover, each image was recorded in both 12-bit RAW format for processing and JPG format for viewing and checking.

2.3.2. Image Pre-Processing

Vignetting and geometric distortion problems are the inherent issues of most cameras which usually cause some inaccuracy in image analysis results, especially for modified cameras [21,24]. Therefore, the free Capture NX-D 1.2.1 software (Nikon, Inc., Tokyo, Japan) provided with the camera manufacturer was used to correct the vignetting and geometric distortion in the images. The corrected images were saved in 16-bit TIFF format to preserve image quality.

There were 144 images to be mosaicked for each camera. The Pix4D Mapper software (Pix4D, Inc., Lausanne, Switzerland) was selected, which is a software package for automatic image mosaicking with high accuracy [25]. To improve the positional accuracy of the mosaicked image, some white plastic square panels with a side of 1 m were placed across the study area. A Trimble GPS Pathfinder ProXRT receiver (Trimble Navigation Limited, Sunnyvale, CA, USA), which provided a 0.2-m average horizontal position accuracy with the real-time OmniSTAR satellite correction, was used to collect the coordinates from these panels. Fifteen ground control points (GCP) as shown in Figure 1 were used for geo-referencing. As shown in Figure 4, the spatial resolutions were 0.399 and 0.394 m for the mosaicked RGB and NIR images. The absolute horizontal position accuracy was 0.470 and 0.701 m for the respective mosaicked images. These positional errors were well within 1 to 3 times of the ground sampling distance (GSD) or spatial resolution [26].

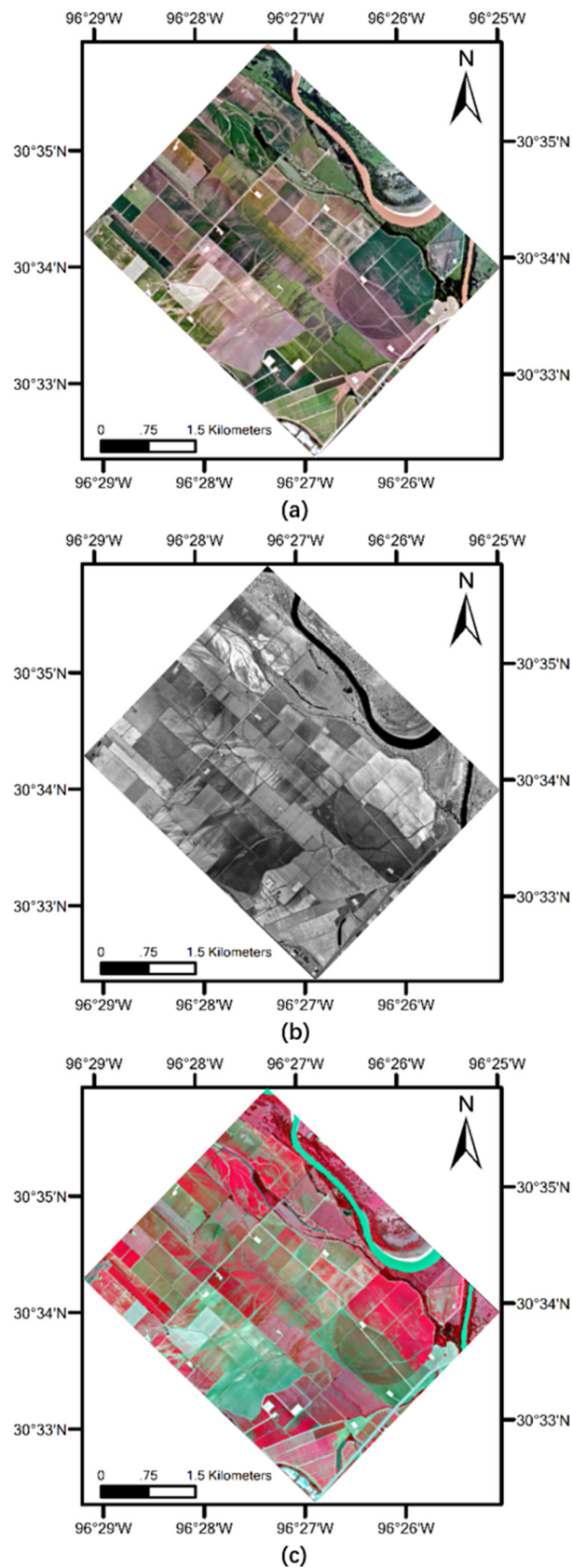


Figure 4. Mosaicked images for this study: (a) three-band RGB image, band 1 = blue, band 2 = green and band 3 = red; (b) NIR band image; and (c) CIR composite (NIR, red and green) extracted from the four-band image, band 1 = blue, band 2 = green, band 3 = red and band 4 = NIR.

To generate a mosaicked four-band image, the mosaicked RGB and NIR images were registered to each other using the AutoSync module in ERDAS Imagine (Intergraph Corporation, Madison, AL, USA). The RGB image was chosen as the reference image as it had better image quality than the NIR image. Several tie control points were chosen manually before the automatic registration. Thousands of tie points were generated by AutoSync and a third-order polynomial geometric model as recommended with the number of tie points was used [27]. The root mean square (RMS) error for the registration was 0.49 pixels or 0.2 m. The combined image was resampled to 0.4-m spatial resolution. The color-infrared (CIR) composite of the four-band image is shown in Figure 4.

2.4. Crop Identification

Selection of different band combinations and classification methods generally influence classification results. To quantify and analyze these effects on crop identification results, three typical and general image classification methods (unsupervised, supervised and object-based) were selected. Meanwhile, to examine how numbers of LULC classes affect the classification results, six different class groupings were defined as shown in Table 1. It should be noted that the three-band or four-band image was first classified into 10 classes and then the classification results were regrouped into six, five, four, three and two classes. For the ten-class grouping, the impervious class mainly included solid roads and buildings. Bare soil and fallow were grouped in one class and the water class included river, ponds, and pools. Considering soybeans and watermelon accounted for only a small portion of the study area, they were treated as non-crop vegetation with grass and forest in the five-class grouping and as non-crop in the four- and three-class groupings.

Table 1. Definition of six different class groupings for image classification.

Ten-Class	Six-Class	Five-Class	Four-Class	Three-class	Two-Class
impervious	non-crop	non-crop (non-vegetation)	non-crop (with soybean and watermelon)	non-crop (with soybean and watermelon)	non-crop
bare soil and fallow					
water		non-crop (vegetation)			
grass					
forest					
soybean	soybean	crop	crop	crop	
watermelon	watermelon				
corn	corn				
sorghum	sorghum				
cotton	cotton				

2.4.1. Pixel-Based Classification

The unsupervised Iterative Self-Organizing Data (ISODATA) and the supervised maximum likelihood classification were chosen as pixel-based methods in this study. Given the diversity and complexity of the land cover in the study area, the number of clusters was set to ten times of the number of land cover classes. The number of maximum interactions was set to 20 and the convergence threshold to 0.95. Then all of the clusters were assigned to the 10 land cover classes. For the supervised maximum likelihood classification, each class was further divided into 3 to 10 subclasses due to the variability within each of the land cover classes. After supervised classification, these subclasses were merged. For each subclass, 5 to 15 training samples were selected and the total number of training samples was almost equal to the number of clusters in ISODATA. The same training samples were used for supervised classification for both the three-band and four-band images.

2.4.2. Object-Based Classification

OBIA has been recognized to have outstanding classification performance for high-resolution imagery. Segmentation and definition of classification rules are the main steps of object-based

classification. In order to show a transparent process and obtain an objective result, the estimation of scale parameter (ESP) tool was used for choosing segmentation parameters [28] and the classifier known as classification and regression tree (CART) was used for generating classification rules.

Segmentation for object-based classification was performed using the commercial software eCognition Developer (Trimble Inc., Munich, Germany). The segmentation processing that integrates spectral, shape and compactness factors is very important for the subsequent classification [29], but standardized or widely accepted methods are lacking to determine the optimal scale for different types of imagery or applications [30]. To minimize the influence of contrived factors in this step, some reference segmentation scales can be estimated by the estimation of scale parameter (ESP) tool [28], which evaluates variation in heterogeneity of image objects that are iteratively generated at multiple scale levels to obtain the most appropriate scales. For this study, a scale step of 50 was set to find some optimal segmentation scales from 0 to 10,000 with the ESP tool, and three segmentation parameters (SP) (1900, 4550 and 9200) had been estimated. To simplify the processing, the SP 4550 was used for image segmentation, which is suitable for most of land cover classes without over-segmentation or under-segmentation. To further improve the segmentation results, spectral difference segmentation with a scale of 1000 was performed to merge neighboring objects with similar spectral values. The three-band and four-band image segmentation produced 970 and 950 image objects, respectively, as shown in Figure 5.

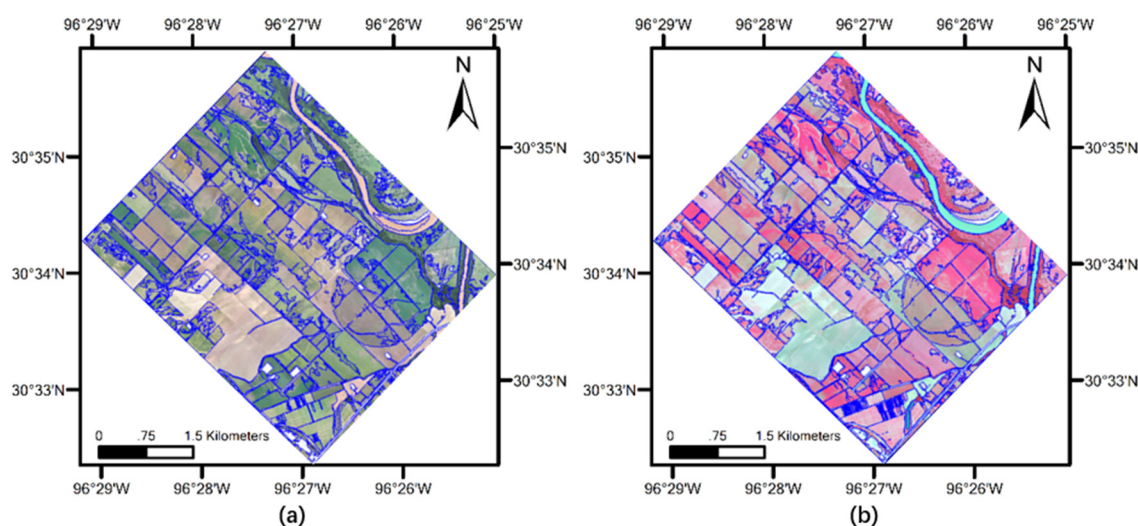


Figure 5. Image segmentation results: (a) three-band image segmentation with 970 image objects; and (b) four-band image segmentation with 950 image objects.

The classification pattern of object-based classification like eCognition is mainly based on a series of rules from several features. User knowledge and past experience could be transferred to some constraint rules for classification [31]. However, it is very unreliable and highly individualized. Therefore, the CART algorithm was used for the training of object-based classification rules [32]. Because it is a non-parametric rule-based classifier and has a “white box” workflow [30], the structure and terminal nodes of a decision tree is easy to interpret, allowing the user to know the mechanism of the object-based classification method and evaluate it.

The CART classifier included in eCognition could create the decision-tree model based on some features from training samples. In order to minimize the impact by the selection of different sample sets, the sample sets used in the supervised classification was also used. The difference was that the samples were turned into image objects. Then these image objects containing the class information were used as the training samples for the object-based classification. There were three main feature

types used for modeling: layer spectral features including some vegetation indices (VIs), shape features, and texture features (Table 2) [33–45].

Table 2. List of object features for decision tree modeling.

Source	Feature Types	Feature Name	For Three-Band	For Four-Band
References	VI	(1) ¹ Normalized Difference Vegetation index (NDVI) = (NIR − R)/(NIR + R) [33]	×	✓
		(2) Ratio Vegetation index (RVI) = NIR/R [34]	×	✓
		(3) Difference Vegetation index (DVI) = NIR − R [35]	×	✓
		(4) Renormalized Difference Vegetation Index(RDVI) = $\sqrt{\text{NDVI} \times \text{DVI}}$ [36]	×	✓
		(5) NDWI = (G − NIR)/(G + NIR) [37]	×	✓
		(6) Optimization of Soil-adjusted Vegetation Index (OSAVI) = (NIR − R) / (NIR + R + 0.16) [38]	×	✓
		(7) Soil Adjusted Vegetation Index (SAVI) = $1.5 \times (\text{NIR} - \text{R}) / (\text{NIR} + \text{R} + 0.5)$ [39]	×	✓
		(8) Soil Brightness Index (SBI) = $\sqrt{\text{NIR}^2 + \text{R}^2}$ [40]	×	✓
		(9) B* = B/(B + G + R), (10) G* = G/(B + G + R), (11) R* = R/(B + G + R)		
		(12) Excess Green (ExG) = 2G* − R* − B* [41], (13) Excess Red (ExR) = 1.4R* − G* [42],	✓	✓
		(14) ExG − ExR [43]		
		(15) CIVE = 0.441R − 0.811G + 0.385B + 18.78745 [44]	✓	✓
		(16) Normalized Difference index (NDI) = (G − R)/(G + R) [45]	✓	✓
		Mean of (17) B,(18) G,(19) R and (20) Brightness	✓	✓
		Mean of (21) NIR	✓	✓
		Standard deviation of (22) B,(23) G,(24) R	✓	✓
eCognition	Layer	(25) Standard deviation of NIR	✓	✓
		HIS((26) Hue, (27) Saturation, (28) Intensity)	✓	✓
		(29) Area, (30) Border length	✓	✓
	Geometry	(31) Asymmetry, (32) Compactness, (33) Density, (34) Shape index	✓	✓
		GLMC ((35) Homogeneity, (36) Contract, (37) Dissimilarity, (38) Entropy, (39) Ang.2nd moment, (40) Mean, (41) StdDev, (42) Correlation)	✓	✓
	Texture			
Total Number of Features			32	42

¹ (n) is the number for each feature, ranging from (1) to (42).

2.5. Accuracy Assessment

For accuracy assessment, 200 random points were generated and assigned to each class in a stratified random pattern based on each classification map. At least 10 points were generated for each class. For this study, three classification methods were applied to two types of images, so there were six classification maps. A total of 1200 points were used for accuracy assessment of the six classification maps [30]. The number of points and percentages by class type are given in Table 3. Ground verification of all the points for the LULC classes was performed shortly after image acquisition for this study area. If one or more points fell within a field, the field was checked. Overall accuracy [46], confusion matrix [47], and kappa coefficient [48] were calculated. In order to evaluate the performance of the image types and methods, average kappa coefficients were calculated by class and by method.

Table 3. Count and percentage by class type for 1200 reference points.

Class Type	Count	Percentage	Class Type	Count	Percentage
Impervious	55	4.6%	Soybean	29	2.4%
Bare Soil and Fallow	186	15.6%	Watermelon	49	4.1%
Grass	162	13.5%	Corn	100	8.3%
Forest	106	8.8%	Sorghum	115	9.6%
Water	69	5.8%	Cotton	329	27.3%

3. Results

3.1. Classification Results

Figure 6 shows the ten-class classification maps based on the three methods applied to the three-band and four-band images, including unsupervised classification for the three-band image (3US), unsupervised classification for the four-band image (4US), supervised classification for the

three-band image (3S), supervised classification for the four-band image (4S), object-based classification for the three-band image (3OB), and object-based classification for the four-band image (4OB). To compare the actual differences between the pixel-based and object-based methods directly [29], no such post-processing operations as clump, sieve, and eliminate were performed for the pixel-based classification maps and no generalization was applied to the object-based classification maps either.

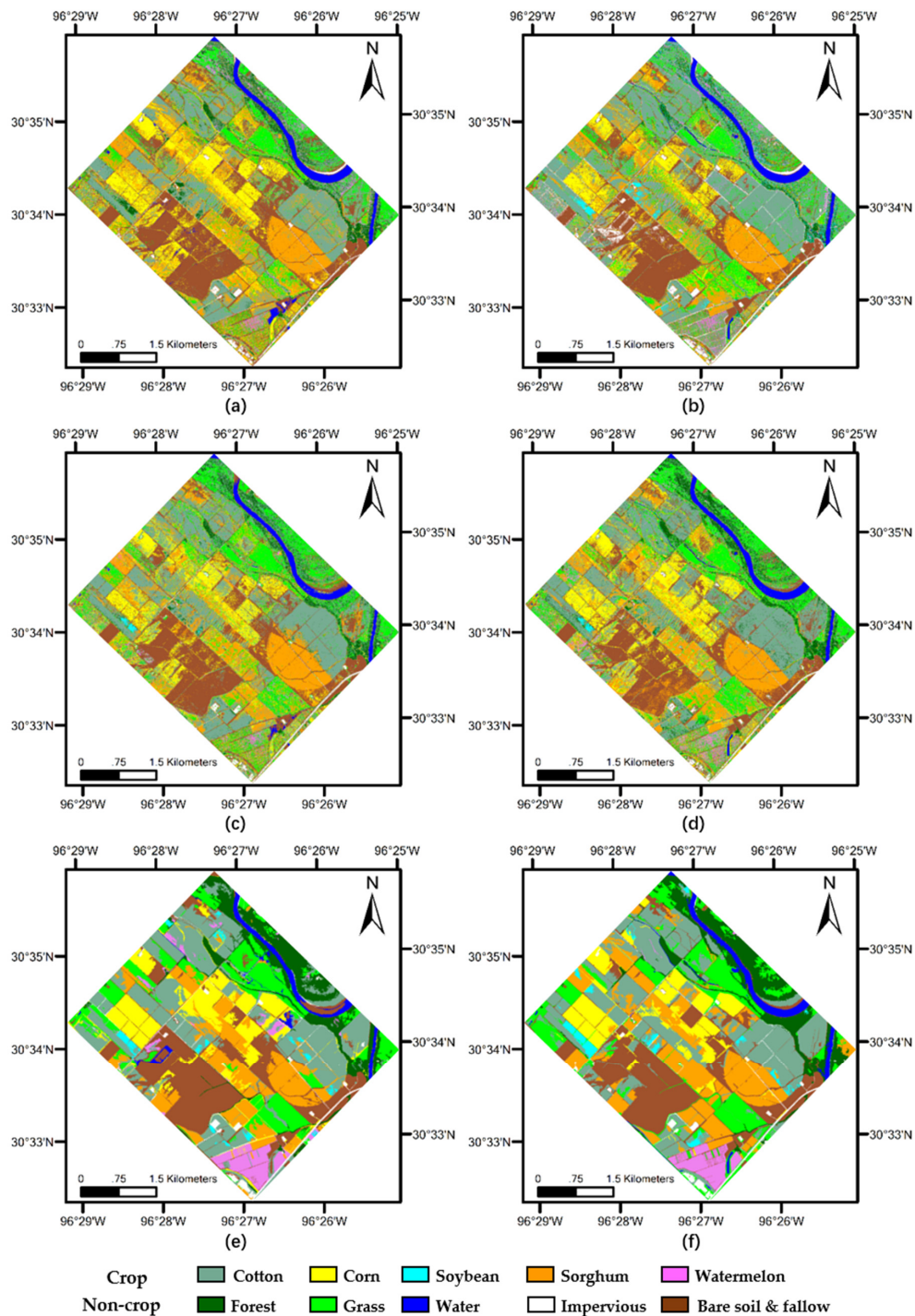


Figure 6. Image classification results (ten-class): (a) unsupervised classification for three-band image (3US); (b) unsupervised classification for four-band image (4US); (c) supervised classification for three-band image (3S); (d) supervised classification for four-band image (4S); (e) object-based classification for three-band image (3OB); and (f) object-based classification for three-band image (4OB).

Most of the classification maps appear to distinguish different land cover types reasonably well. From a visual perspective, the “salt-and-pepper” effect on the pixel-based maps is the obvious difference with the object-based maps. The object-based maps present a very good visual effect as different cover types are shown by the homogenous image objects. Without considering the accuracy of the maps, the object-based classification maps look cleaner and more appropriate to produce thematic maps. Visually, it is difficult to evaluate the differences between the unsupervised and supervised methods or between the three-band and four-band images.

Specifically, the classification results of water and impervious had high consistence. Because of the lack of NIR band, some water areas in the three-band image was classified as bare soil and fallow for all the methods. Sorghum and corn were difficult to distinguish because both crops were at their late growth stages with reduced green leaf area. Corn was close to physiological maturity and above ground biomass was fully senescent, whereas sorghum was in the generative phase and started senescence, but still had significant green leaf material. Although late growth stages caused a reduction in canopy NDVI values for both corn and sorghum, the background weeds and soil exposure also affected the overall NDVI values. All crops and cover types show varying degrees of confusion among themselves. This problem also occurred in the object-based maps, but it does not appear to be as obvious as in the pixel-based maps.

3.2. Accuracy Assessment

Table 4 summarizes the accuracy assessment results for the ten-class and two-class classification maps for the three methods applied to the two images. The accuracy assessment results for the other groupings are discussed and compared with the ten-class and two-class results in Section 4.3. Overall accuracy for the ten-class classification maps ranged from 58% for 3US to 78% for 4OB and overall kappa from 0.51 for 3US to 0.74 for 4OB. As expected, the overall accuracy and kappa were higher for the four-band image than for the three-band image for all the three methods. Among the three methods, the object-based method performed better than the two pixel-based methods, and the supervised method was slightly better than the unsupervised method.

For the individual classes, the non-plant classes such as water, impervious, and bare soil and fallow had better and stable accuracy results for all the six scenarios with an average kappa of 0.85, 0.82 and 0.74, respectively. Due to variable growing stages and management conditions, the main crop class cotton had a relatively low accuracy with an average kappa of 0.47 for all the scenarios. Although at later growing stages, sorghum and corn had a relatively good accuracy with an average kappa of 0.67 and 0.62, respectively. The main reason was that both crops were at senescence and had less green leaf material, so they could easily be distinguished with other vegetation. Soybean and watermelon had unstable accuracy results among the six scenarios, but their differentiation was significantly improved with the object-based method. The grass and forest in the study area were difficult to distinguish using the pixel-based methods, but they were more accurately separated with the object-based method.

For the two broad classes (crop and non-crop), overall accuracy ranged from 76% for 3US to 91% for 4OB and overall kappa from 0.51 for 3US to 0.82 for 4OB. Clearly, overall accuracy and kappa were generally higher for the two-class maps than for the ten-class maps. The two-class classification maps will be useful for some applications when total cropping area information is needed.

Table 4. (a–f) Accuracy assessment results.

(a) Unsupervised Classification for Three-Band Image (3US)																	
	CD ⁴	Ten-class ²										Pa ⁵ %	Ua ⁶ %	Kp ⁷	Two-class		
		RD ³													Pa	Ua	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT				%	%	%
Non-crop	IM	44 ¹	11	0	0	0	0	0	0	0	0	80	80	0.79	75	75	0.51
	BF	9	143	12	1	3	0	5	13	7	49	77	59	0.71			
	GA	0	2	71	7	1	0	9	0	9	17	44	61	0.38			
	FE	0	0	1	57	1	24	0	0	0	13	54	59	0.50			
	WA	0	9	0	0	61	0	1	0	0	1	88	85	0.88			
	SB	0	0	0	0	0	0	0	0	0	0	0	NaN	0.0	76	77	0.51
	WM	0	0	14	9	0	0	9	0	0	0	18	28	0.16			
	CO	0	16	7	0	2	1	6	72	15	29	72	49	0.68			
	SG	0	2	35	4	0	2	17	13	79	62	69	37	0.62			
	CT	2	3	22	28	1	2	2	2	5	158	48	70	0.36			
Overall kappa=0.51 Overall accuracy=58%															Overall kappa=0.51 Overall accuracy=76%		
(b) Supervised Classification for Three-Band Image (3S)																	
	CD	Ten-class										Pa	Ua	Kp	Two-class		
		RD													Pa	Ua	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT				%	%	%
Non-crop	IM	37	3	0	0	0	0	0	0	0	0	67	93	0.66	77	80	0.58
	BF	12	152	7	1	6	1	2	8	5	32	82	67	0.77			
	GA	4	4	83	20	0	0	19	4	12	14	51	52	0.44			
	FE	1	0	5	51	1	0	0	0	0	11	48	74	0.45			
	WA	0	4	0	0	56	0	1	0	1	0	81	90	0.80			
crop	SB	0	0	0	1	0	19	0	0	0	3	66	83	0.65	82	80	0.62
	WM	0	3	10	3	0	0	9	1	2	2	18	30	0.16			
	CO	0	9	10	0	4	0	6	63	13	35	63	45	0.58			
	SG	1	8	19	0	0	1	9	12	69	48	60	41	0.54			
	CT	0	3	28	30	2	8	3	12	13	184	56	65	0.42			
Overall kappa=0.53 Overall accuracy=60%															Overall kappa=0.60 Overall accuracy=80%		

Table 4. Cont.

(c) Object-Based Classification For Three-Band Image (3OB)																	
	CD	Ten-class										Two-class					
		RD										Pa	Ua	Kp	Pa	Ua	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT	%	%		%	%	
Non-crop	IM	51	7	0	0	1	0	0	0	1	0	93	85	0.92	87	87	0.75
	BF	2	142	2	1	10	1	0	3	7	28	76	72	0.72			
	GA	1	3	91	5	2	2	0	0	13	15	56	69	0.51			
	FE	0	2	25	96	3	1	0	0	1	1	91	74	0.89			
	WA	0	7	1	0	50	0	0	0	0	1	72	85	0.71			
	SB	0	0	0	0	0	21	0	0	0	21	72	50	0.71	88	88	0.75
	WM	0	7	5	0	0	0	46	0	0	9	94	69	0.94			
	CO	1	14	3	1	3	0	0	84	9	17	84	64	0.82			
	SG	0	1	5	0	0	0	0	13	80	29	70	63	0.66			
	CT	0	3	30	3	0	4	3	0	4	208	63	82	0.53			
Overall kappa=0.68 Overall accuracy=72%															Overall kappa=0.75 Overall accuracy=88%		
(d) Unsupervised Classification for Four-Band Image (4US)																	
	CD	Ten-class										Two-class					
		RD										Pa	Ua	Kp	Pa	Ua	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT	%	%		%	%	
Non-crop	IM	50	28	0	0	1	0	0	0	0	0	91	63	0.90	70	79	0.48
	BF	4	141	11	1	1	0	6	12	6	34	76	65	0.71			
	GA	0	1	56	2	0	1	7	5	17	17	35	53	0.28			
	FE	0	0	1	44	1	0	0	0	0	0	42	96	0.39			
	WA	0	0	0	0	64	0	0	0	0	0	93	100	0.92			
	SB	0	0	0	11	0	22	0	0	0	11	76	50	0.75	83	75	0.60
	WM	0	0	6	13	1	0	15	0	0	12	31	32	0.28			
	CO	0	4	6	0	0	0	3	58	9	17	58	60	0.54			
	SG	1	4	22	2	1	2	1	15	62	38	54	42	0.47			
	CT	0	8	60	33	0	4	17	10	21	200	61	57	0.44			
Overall kappa=0.52 Overall accuracy=59%															Overall kappa=0.54 Overall accuracy= 77%		

Table 4. Cont.

(e) Supervised Classification for Four-Band Image (4S)																	
	CD	Ten-class										Pa %	Ua %	Kp	Two-class		
		RD													Pa %	Ua %	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT						
Non-crop	IM	38	2	0	0	0	0	0	0	0	0	69	95	0.68	79	82	0.61
	BF	10	141	10	2	1	0	3	7	5	36	76	66	0.71			
	GA	2	8	86	18	3	0	6	1	4	20	53	58	0.46			
	FE	1	0	3	67	1	1	0	0	0	18	63	74	0.60			
	WA	0	0	0	0	64	0	0	0	0	0	93	100	0.92			
	SB	0	0	0	0	0	19	0	0	0	2	66	90	0.65	84	81	0.65
	WM	0	5	15	4	0	0	23	0	4	1	47	44	0.45			
	CO	0	14	15	0	0	1	4	69	16	27	69	47	0.65			
	SG	0	10	13	1	0	1	2	12	71	29	62	51	0.57			
	CT	4	6	20	14	0	7	11	11	15	196	60	69	0.47			
Overall kappa =0.59 Overall accuracy=65%															Overall kappa=0.63 Overall accuracy=82%		
(f) Object-Based Classification for Four-Band Image (4OB)																	
	CD	Ten-class										Pa %	Ua %	Kp	Two-class		
		RD													Pa %	Ua %	Kp
		IM	BF	GA	FE	WA	SB	WM	CO	SG	CT						
Non-crop	IM	51	4	0	0	0	0	0	0	0	0	93	93	0.92	90	91	0.80
	BF	1	153	2	0	4	1	2	3	3	21	82	81	0.79			
	GA	1	8	103	7	0	1	1	2	2	13	64	75	0.59			
	FE	1	0	23	92	3	0	0	0	0	1	87	77	0.85			
	WA	1	0	0	3	61	0	0	0	0	0	88	94	0.88			
	SB	0	0	0	0	0	23	0	0	0	17	79	58	0.79	92	91	0.83
	WM	0	6	0	0	0	0	44	0	0	0	90	88	0.89			
	CO	0	7	3	1	0	0	1	74	7	11	74	71	0.72			
	SG	0	0	9	0	0	1	0	20	103	38	90	60	0.88			
	CT	0	8	22	3	1	3	1	1	0	228	69	85	0.61			
Overall kappa=0.74 Overall accuracy=78%															Overall kappa=0.82 Overall accuracy=91%		

¹ Bold values correspond to number of points correctly classified. ² IM = impervious, BF = bare soil and fallow, GA = grass, FE = forest, WA = water, SB = soybean, WM = watermelon, CO = corn, SG = sorghum, CT = cotton. ³ RD = reference data. ⁴ CD = classification data. ⁵ Pa = producer's accuracy, ⁶ Ua = user's accuracy. ⁷ Kp = kappa coefficient.

4. Discussion

4.1. Importance of NIR Band

To analyze the importance of the NIR band, some kappa coefficients from Table 4 were rearranged and the average coefficients by image (AKp1) and by method (AKp2) were calculated (Table 5).

Table 5. Results of kappa analysis by three methods and two images for (a) Crop and (b) Non-crop.

(a)	Crop	Three-Band (Kappa)				Four-Band (Kappa)			
		3US	3S	3OB	AKp1	4US	4S	4OB	AKp1
	SB ¹	0.00	0.65	0.71	0.45	0.75	0.65	0.79	0.73
	WM	0.16	0.16	0.94	0.42	0.28	0.45	0.89	0.54
	CO	0.68	0.58	0.82	0.69	0.54	0.65	0.72	0.64
	SG	0.62	0.54	0.66	0.61	0.47	0.57	0.88	0.64
	CT	0.36	0.42	0.53	0.44	0.44	0.47	0.61	0.51
	AKp2	0.36	0.47	0.73		0.50	0.56	0.78	

(b)	Non-crop	Three-band (Kappa)				Four-band (Kappa)			
		3US	3S	3OB	AKp1	4US	4S	4OB	AKp1
	IM	0.79	0.66	0.92	0.79	0.90	0.68	0.92	0.83
	BF	0.71	0.77	0.72	0.73	0.71	0.71	0.79	0.74
	GA	0.38	0.44	0.51	0.44	0.28	0.46	0.59	0.44
	FE	0.50	0.45	0.89	0.61	0.39	0.60	0.85	0.61
	WA	0.88	0.80	0.71	0.80	0.92	0.92	0.88	0.91
	AKp2	0.65	0.62	0.75		0.64	0.68	0.81	

¹ IM = impervious, BF = bare soil and fallow, GA = grass, FE = forest, WA = water, SB = soybean, WM = watermelon, CO = corn, SG = sorghum, CT = cotton, AKp1 = Average kappa coefficient among the three methods for each class with the same image, and AKp2 = Average kappa coefficient among the crop classes or non-crop classes with the same classification methods.

It can be seen from Table 5 that the NIR band improved the kappa coefficients for four of the five crops and for three of the five non-crop classes. The net increases in AKp1 for the four crops were 0.28 for soybean, 0.12 for watermelon, 0.07 for cotton and 0.03 for sorghum, while the decrease in AKp1 for corn was 0.05. Although the classification for soybean was greatly improved, soybean only accounted for a very small portion of the study area which was less than 2.5%. Due to its small area and misclassification, there were unstable classification results for soybean as shown by the unique zero kappa value in Table 4. The contribution of the improvement for watermelon was mainly due to the object-based classification method. The classification for corn got worse mainly due to its later growth stage. Corn had low chlorophyll contents as shown by its flat RGB and reduced water contents as indicated by the relatively low NIR reflectance compared to the other vegetation classes. These observations could be confirmed by the spectral curves shown in Figure 3, which were derived by calculating the average spectral values from each class using the training samples for the supervised classification. The spectral curve of corn had the lowest reflectance at the NIR band among the vegetation classes. In other words, the NIR band was not sensitive to corn at this stage, which had a similar NIR response to the bare soil and fallow fields. In fact, the bare soil and fallow class was one of the main classes for misclassification with corn as shown in Table 4.

For the non-crop classes, the NIR band improved the classification for the water, impervious and bare soil classes. This result conforms with the general knowledge that NIR is effective at distinguishing water and impervious. The classes of grass and forest also benefited from the NIR band with the supervised method.

To compare the differences between the three-band and four-band images by classification method, average kappa coefficients (AKp2) increased for each of the three methods for the combined crop class and for two of the three methods for the combined no-crop class. If AKp3 is the average of the

AKp2 values for the three methods, AKp3 increased from 0.52 for the three-band image to 0.61 for the four-band image for the crop class, and from 0.67 for the three-band image to 0.71 for the four-band image for the non-crop class. The crop class benefited from the NIR band more than the non-crop class.

If AKp4 is the average of the AKp3 values for the two general classes, AKp4 increased from 0.6 for the three-band image to 0.66 for the four-band image. Therefore, the addition of NIR improved the classification results over the normal RGB image.

To illustrate the classification results and explain the misclassification between some classes, spectral separability between any two classes in terms of Euclidean distance was calculated by ERDAS. To facilitate discussion, the Euclidean distance was normalized by the following formula:

$$x' = (x - x_0) / \max |(x - x_0)| \quad (2)$$

where x is the absolute Euclidean distance any two classes based on the training samples, and x_0 is the average of all the two-class Euclidean distances for either the three-band or four-band image, and x' is the normalized spectral distance ranging from -1 for the worse separability to 1 for the best separability.

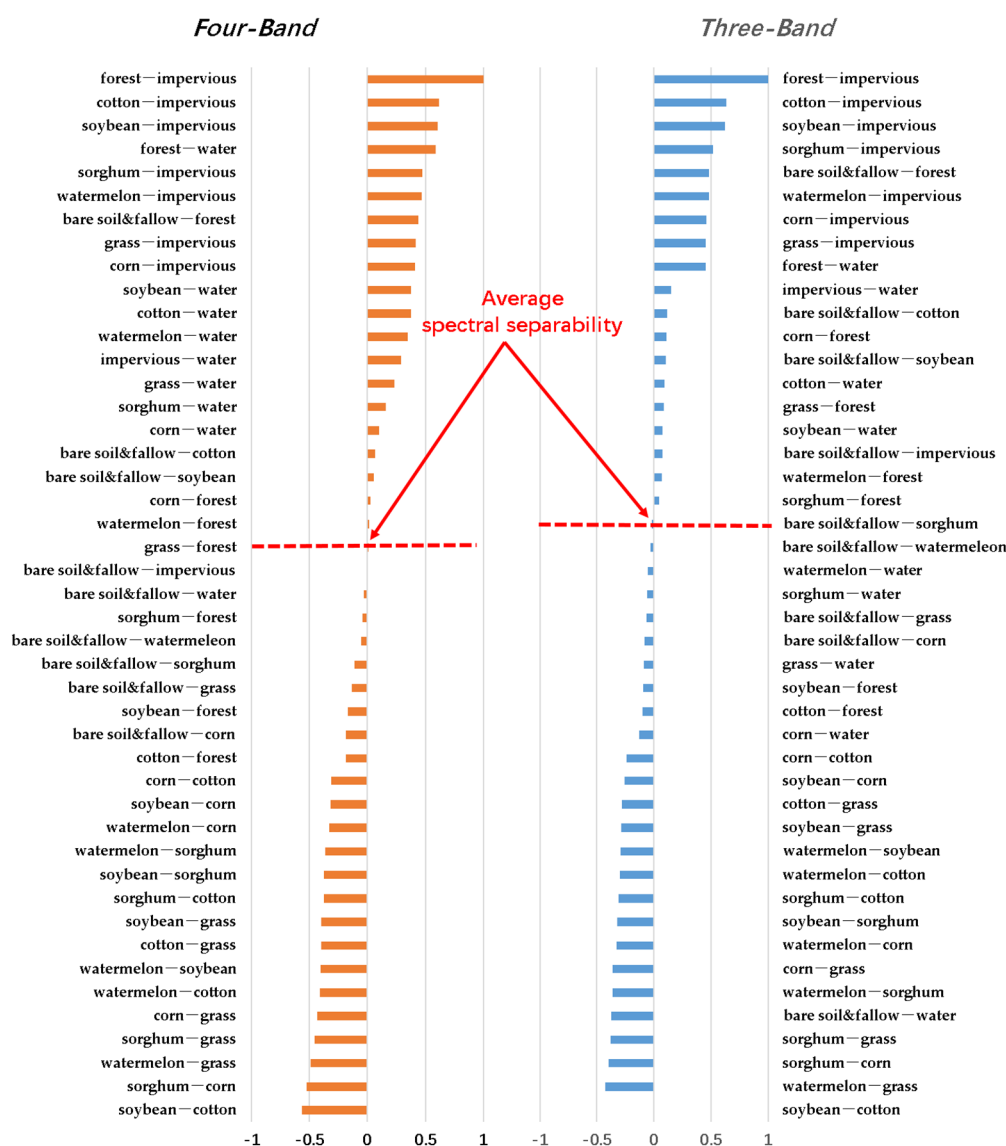


Figure 7. Spectral separability between any two classes for the three-band and four-band images.

From the normalized Euclidean distance results shown in Figure 7, the forest and impervious classes had the best separation, while soybean and cotton had the worse separation for both the three-band and four-band images. These results should clearly explain why some of the classes had higher classification accuracy and kappa values than others. In general, the non-crop classes such as forest, water and impervious had high separability with crop classes, while the crop classes had relatively low separability among themselves. Since corn and sorghum are near the bottom of the list, it explains in another way why they were difficult to separate. There are more class pairs above the average spectral separability for the four-band than for the three-band, indicating that the NIR band is a useful for crop identification, especially for plants at their vegetative growth periods.

4.2. Importance of Object-Based Method

As can be seen from Tables 4 and 5 the selection of the classification methods had a great effect on classification results. To clearly see this effect, the kappa analysis results were rearranged by the classification methods as shown in Table 6.

The average kappa coefficients between the three-band and four-band images (AKp5) were calculated for all the crop and non-crop classes for each of the three classification methods. For all the crop classes, the object-based method performed best with the highest AKp5 values, followed by the supervised and unsupervised methods. Moreover, the object-based method performed better than the other two methods for all the no-crop classes except for water, for which the unsupervised method was the best. Similarly, if AKp6 is the average of the AKp5 values for the five crop classes, the AKp6 values for the crop class were 0.43, 0.51 and 0.76 for the unsupervised, supervised and object-based methods, respectively. The AKp6 values for the non-crop class were 0.65, 0.65 and 0.78 for the three respective classification methods.

Clearly, the object-based method was superior to the pixel-based methods. This was because the object-based method used many shape and texture features as shown in Table 2 to create homogeneous image objects as the processing units during the classification, while the pixel-based classification methods only used spectral information in each pixel during the classification. Figure 8 shows the decision trees and the number of features involved in the image classification process using the object-based method, which was created automatically by eCognition.

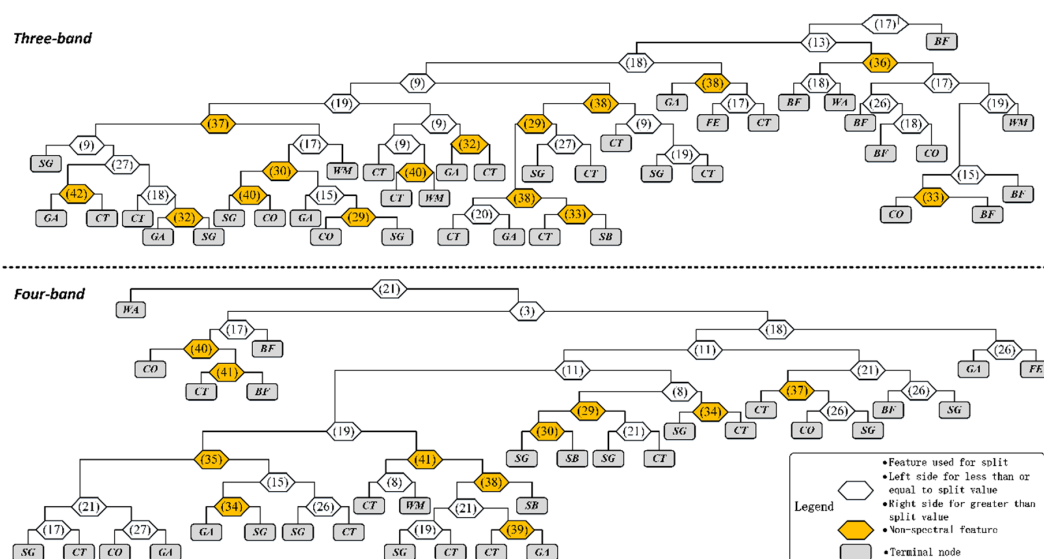


Figure 8. Decision tree models for object-based classification. Abbreviations for the 10 classes: IM=impervious, BF=bare soil and fallow, GA=grass, FE=forest, WA=water, SB=soybean, WM=watermelon, CO=corn, SG=sorghum, and CT=cotton. ¹ (n) is the number ID of each feature, ranging from (1) to (42), which is described in Table 2.

Figure 9 shows the average kappa coefficients and differences for the crop and non-crop classes for the three classification methods. The difference in AKp6 between the crop and non-crop classes reduced from 0.22 for the unsupervised to 0.14 for the supervised and to 0.02 for the object-based method. Evidently, non-crop had a better average kappa coefficient than crop for the pixel-based methods because most of the non-crop classes such as water, impervious, and bare soil and fallow classes had better spectral separability than the other classes. However, both crop and non-crop had essentially the same average kappa coefficient for the object-based classification method.

Table 6. Kappa analysis results arranged by classification method for (a) Crop and (b) Non-crop.

(a)	Crop	Unsupervised			Supervised			Object-Orient		
		3US	4US	AKp5	3S	4S	AKp5	3OB	4OB	AKp5
	SB ¹	0.00	0.75	0.38	0.65	0.65	0.65	0.71	0.79	0.75
	WM	0.16	0.28	0.22	0.16	0.45	0.31	0.94	0.89	0.92
	CO	0.68	0.54	0.61	0.58	0.65	0.62	0.82	0.72	0.77
	SG	0.62	0.47	0.55	0.54	0.57	0.56	0.66	0.88	0.77
	CT	0.36	0.44	0.40	0.42	0.47	0.45	0.53	0.61	0.57
	AKp6	0.43			0.51			0.76		
(b)	Non-crop	Unsupervised			Supervised			Object-orient		
		3US	4US	AKp5	3S	4S	AKp5	3OB	4OB	AKp5
	IM	0.79	0.90	0.85	0.66	0.68	0.67	0.92	0.92	0.92
	BF	0.71	0.71	0.71	0.77	0.71	0.74	0.72	0.79	0.76
	GA	0.38	0.28	0.33	0.44	0.46	0.45	0.51	0.59	0.55
	FE	0.50	0.39	0.45	0.45	0.60	0.53	0.89	0.85	0.87
	WA	0.88	0.92	0.90	0.80	0.92	0.86	0.71	0.88	0.80
	AKp6	0.65			0.65			0.78		

¹ IM=impervious, BF=bare soil and fallow, GA=grass, FE=forest, WA=water, SB=soybean, WM=watermelon, CO=corn, SG=sorghum, CT=cotton, AKp5=Average kappa coefficient for each class between the three-band and four-band images, and AKp6=Average of the AKp5 values for the crop or non-crop classes.

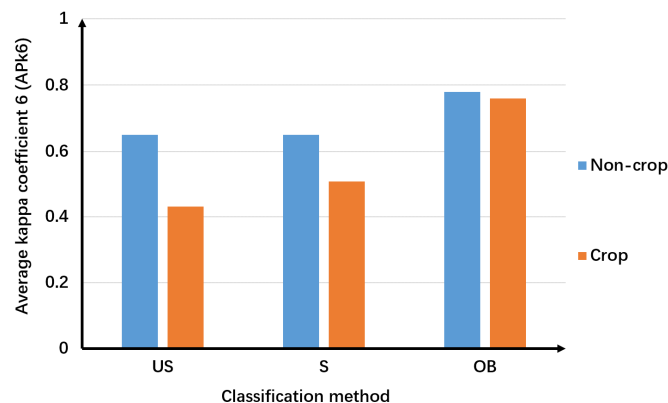


Figure 9. Average kappa coefficient difference for crop and non-crop by three classification methods (APk6) and difference of them. Unsupervised classification method (US), supervised classification method (S), object-based classification method (OB).

To explain the reason for this, the statistical results for the decision tree models used in the object-based classification method are summarized in Table 7. It can be seen from Figure 8 and Table 7, the three-band image used more non-spectral features at a higher frequency than the four-band image, which could compensate for the lacking of the NIR band in the normal RGB image. Most of the branches of three-band or four-band image decision tree models for classification used the shape and texture features (95% for the three-band and 82% for the four-band). These features were used more than one time with an average of 1.62 times for the three-band and 1.15 times for the four-band image.

All these showed the importance and advantage of the non-spectral features for image classification. The non-spectral features are particularly important when there is no sufficient spectral information.

As shown in Table 6, the pixel-based methods performed better than object-based method for distinguish water. This is because the spectral information was enough to distinguish water and non-spectral features could cause a worse result with the object-based method. Thus, the four-band image with the pixel-based methods achieved better classification results for water.

Table 7. Statistical results of decision tree models for object-based classification.

Properties	Three-Band	Four-Band
Number of end nodes (number of branches)	39	33
Maximum number of tree levels	10	10
First level to use non-spectral features	3	4
Number of branches that used non-spectral features	63	38
Average times non-spectral features were used for each branch	1.62	1.15
Ratio of branches that used non-spectral features (%)	95	82

4.3. Importance of Classification Groupings

Thus far, only the ten-class and two-class classification results shown in Table 4 have been discussed. Figure 10 shows the overall accuracy and overall kappa for the six class groupings defined in Table 1 based on the six classification types.

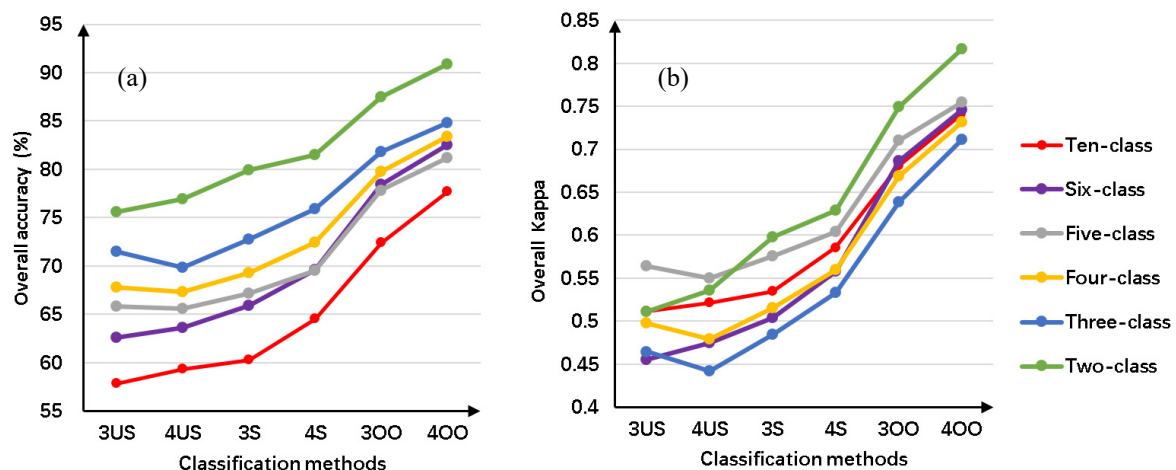


Figure 10. Overall accuracy (a) and overall kappa (b) for six class groupings based on six classification types. Classification methods were represented as unsupervised classification for three-band image (3US), unsupervised classification for four-band image (4US), supervised classification for three-band image (3S), supervised classification for four-band image (4S), object-based classification for three-band image (3OB), and object-based classification for four-band image (4OB).

The overall accuracy generally increased as the number of classes decreased. However, this was not necessarily the case for the overall kappa. The two-class, five-class and ten-class classifications had higher kappa values than the three-class, four-class and six-class classifications except that the two-class classification for the object-based method had slightly higher kappa values than the ten-class classifications. Overall classification accuracy simply considers the probability of image pixels being correctly identified in the classification map. Kappa coefficient, by contrast, considers not only the correct classification but also the effect of omission and commission errors. By using the spectral separability shown in Figures 3 and 7 it could be found that class groupings with poor spectral separability between subclasses generally had higher kappa values. For example, the five-class classifications achieved the second highest kappa value for both the supervised and object-based

methods. This particular class grouping combined four vegetation classes (grass, forest, soybean and watermelon) with similar spectral characteristics into one class. Approximately two-thirds of the spectral separability values between any two of the four classes were below the average level and these classes were very easy to become confused during the classification process. This confusion was eliminated when these classes were grouped into one class. Therefore, depending on the requirements of particular applications, all available classes can be regrouping based on their spectral characteristics into appropriate classes to improve classification results. With such a regrouping, the agronomical use of the classification map is practically reduced for relevant crops, but it could still be used for a LULC census.

4.4. Implications for Selection of Imaging Platform and Classification Method

From the above analysis, the additional NIR band and the object-based method both could improve the performance of image classification for crop identification. The imaging system used in this study included a modified camera to capture NIR information. The camera along with the lens, GPS, and modification fees was about \$1300. Moreover, the images from the two cameras need to be aligned for analysis. The object-based classification method performed better than the pixel-based methods. However, the object-based method involves complex and time-consuming processing such as segmentation and rule training, and requires experienced operators to use the software. Therefore, how to weigh such factors as the cost, ease of use and acceptable classification results is a real and practical issue for users, especially for those without much remote sensing knowledge and experience.

Based on the results from this study, some suggestions are provided for consideration. If users do not have much experience in image processing, a single RGB camera with pixel-based classification can be used. For users with some image processing experience, a dual-camera system with the NIR sensitivity and pixel-based classification methods may be a good combination. For users with sufficient image processing experience, either a single RGB camera or a dual-camera system in conjunction with object-based classification may be an appropriate choice. It is also possible to modify a single RGB camera to have two visible bands and one NIR band [16,49]. This will eliminate the image alignment involved with the dual-camera system.

5. Conclusions

This study addressed important and practical issues related to the use of consumer-grade RGB cameras and modified NIR cameras for crop identification, which is a common remote sensing application in agriculture. Through synthetically comparing the performance of the three commonly-used classification methods with the three-band and four-band images over a relative large cropping area, some interesting results have been found.

Firstly, the NIR image from the modified camera improved classification results from the normal RGB alone. This finding is consistent with the common knowledge and results from scientific-grade imaging systems. Moreover, the importance of the NIR band appears to be especially evident in the classification results from pixel-based methods. Since pixel-based methods usually are easy to use by users without much experience in remote sensing, imaging systems with more spectral information should be used for these users.

Secondly, many non-spectral features such as shape and texture can be obtained from the image to improve the accuracy of image classification. However, object-based methods are more complex and time-consuming and require a better understanding of the classification process, so only advanced users with much experience in image processing could use object-based methods to obtain good results even with RGB images. Moreover, appropriately grouping classes with similar spectral response can improve classification results if these classes do not need to be separated. All in all, the selection of imaging systems, image processing methods, and class groupings needs to consider the budget, application requirements and operating personnel's experience. The results from this study have demonstrated that the dual-camera imaging system is useful for crop identification and has the

potential for other agricultural applications. More research is needed to evaluate this type of imaging systems for crop monitoring and pest detection.

Acknowledgments: This project was conducted as part of a visiting scholar research program, and the first author was financially supported by the National Natural Science Foundation of China (Grant No. 41201364 and 31501222) and the China Scholarship Council (201308420447). The author wishes to thank Jennifer Marshall and Nicholas Mondrik of the College of Science of Texas A & M University, College Station, Texas, for allowing us to use their monochromator and assisting in the measurements of the spectral response of the cameras. Thanks are also extended to Fred Gomez and Lee Denham of USDA-ARS in College Station, Texas, for acquiring the images for this study.

Author Contributions: Jian Zhang designed and conducted the experiment, processed and analyzed the imagery, and wrote the manuscript. Chenghai Yang guided the study design, participated in camera testing and image collection, advised in data analysis, and revised the manuscript. Huaibo Song, Wesley Clint Hoffmann, Dongyan Zhang, and Guozhong Zhang were involved in the process of the experiment and ground data collection. All authors reviewed and approved the final manuscript.

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

Acknowledgments: Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

References

1. Mulla, D.J. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosyst. Eng.* **2013**, *114*, 358–371. [[CrossRef](#)]
2. Bauer, M.E.; Cipra, J.E. Identification of agricultural crops by computer processing of ERTS MSS data. In Proceedings of the Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, New Carrollton, IN, USA, 5–9 March 1973; pp. 205–212.
3. Colomina, I.; Molina, P. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2014**, *92*, 79–97. [[CrossRef](#)]
4. Zhang, C.; Kovacs, J.M. The application of small unmanned aerial systems for precision agriculture: A review. *Precis. Agric.* **2012**, *13*, 693–712. [[CrossRef](#)]
5. Watts, A.C.; Ambrosia, V.G.; Hinkley, E.A. Unmanned aircraft systems in remote sensing and scientific research: Classification and considerations of use. *Remote Sens.* **2012**, *4*, 1671–1692. [[CrossRef](#)]
6. Araus, J.L.; Cairns, J.E. Field high-throughput phenotyping: The new crop breeding frontier. *Trends Plant Sci.* **2014**, *19*, 52–61. [[CrossRef](#)] [[PubMed](#)]
7. Quilter, M.C.; Anderson, V.J. Low altitude/large scale aerial photographs: A tool for range and resource managers. *Rangel. Arch.* **2000**, *22*, 13–17.
8. Hunt, E., Jr.; Daughtry, C.; McMurtrey, J.; Walthall, C.; Baker, J.; Schroeder, J.; Liang, S. Comparison of remote sensing imagery for nitrogen management. In Proceedings of the Sixth International Conference on Precision Agriculture and Other Precision Resources Management, Minneapolis, MN, USA, 14–17 July 2002; pp. 1480–1485.
9. Wellens, J.; Midekor, A.; Traore, F.; Tychon, B. An easy and low-cost method for preprocessing and matching small-scale amateur aerial photography for assessing agricultural land use in burkina faso. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *23*, 273–278. [[CrossRef](#)]
10. Oberthür, T.; Cock, J.; Andersson, M.S.; Naranjo, R.N.; Castañeda, D.; Blair, M. Acquisition of low altitude digital imagery for local monitoring and management of genetic resources. *Comput. Electron. Agric.* **2007**, *58*, 60–77. [[CrossRef](#)]
11. Hunt, E.R., Jr.; Cavigelli, M.; Daughtry, C.S.; McMurtrey, J.E., III; Walthall, C.L. Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* **2005**, *6*, 359–378. [[CrossRef](#)]
12. Chiabrando, F.; Nex, F.; Piatti, D.; Rinaudo, F. UAV and RPV systems for photogrammetric surveys in archaeological areas: Two tests in the piedmont region (Italy). *J. Archaeol. Sci.* **2011**, *38*, 697–710. [[CrossRef](#)]
13. Nijland, W.; de Jong, R.; de Jong, S.M.; Wulder, M.A.; Bater, C.W.; Coops, N.C. Monitoring plant condition and phenology using infrared sensitive consumer grade digital cameras. *Agric. For. Meteorol.* **2014**, *184*, 98–106. [[CrossRef](#)]

14. Murakami, T.; Idezawa, F. Growth survey of crisphead lettuce (*Lactuca sativa* L.) in fertilizer trial by low-altitude small-balloon sensing. *Soil Sci. Plant Nutr.* **2013**, *59*, 410–418. [CrossRef]
15. Yang, C.; Westbrook, J.K.; Suh, C.P.C.; Martin, D.E.; Hoffmann, W.C.; Lan, Y.; Fritz, B.K.; Goolsby, J.A. An airborne multispectral imaging system based on two consumer-grade cameras for agricultural remote sensing. *Remote Sens.* **2014**, *6*, 5257–5278. [CrossRef]
16. Miller, C.D.; Fox-Rabinovitz, J.R.; Allen, N.F.; Carr, J.L.; Kratochvil, R.J.; Forrestal, P.J.; Daughtry, C.S.T.; McCarty, G.W.; Hively, W.D.; Hunt, E.R. Nir-green-blue high-resolution digital images for assessment of winter cover crop biomass. *GISci. Remote Sens.* **2011**, *48*, 86–98.
17. Artigas, F.; Pechmann, I.C. Balloon imagery verification of remotely sensed phragmites australis expansion in an urban estuary of New Jersey, USA. *Landsc. Urban Plan.* **2010**, *95*, 105–112. [CrossRef]
18. Yang, C.; Hoffmann, W.C. Low-cost single-camera imaging system for aerial applicators. *J. Appl. Remote Sens.* **2015**, *9*, 096064. [CrossRef]
19. Wang, Q.; Zhang, X.; Wang, Y.; Chen, G.; Dan, F. The design and development of object-oriented uav image change detection system. In *Geo-Informatics in Resource Management and Sustainable Ecosystem*; Springer-Verlag: Berlin/Heidelberg, Germany, 2013; pp. 33–42.
20. Diaz-Varela, R.; Zarco-Tejada, P.J.; Angileri, V.; Loudjani, P. Automatic identification of agricultural terraces through object-oriented analysis of very high resolution dsms and multispectral imagery obtained from an unmanned aerial vehicle. *J. Environ. Manag.* **2014**, *134*, 117–126. [CrossRef] [PubMed]
21. Lelong, C.C.D. Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. *Sensors* **2008**, *8*, 3557–3585. [CrossRef]
22. Liebisch, F.; Kirchgeßner, N.; Schneider, D.; Walter, A.; Hund, A. Remote, aerial phenotyping of maize traits with a mobile multi-sensor approach. *Plant Methods* **2015**, *11*. [CrossRef] [PubMed]
23. Labbé, S.; Roux, B.; Bégué, A.; Lebourgeois, V.; Mallavan, B. An operational solution to acquire multispectral images with standard light cameras: Characterization and acquisition guidelines. In Proceedings of the International Society of Photogrammetry and Remote Sensing Workshop, Newcastle, UK, 11–14 September 2007; pp. TS10:1–TS10:6.
24. Lebourgeois, V.; Bégué, A.; Labbé, S.; Mallavan, B.; Prévot, L.; Roux, B. Can commercial digital cameras be used as multispectral sensors? A crop monitoring test. *Sensors* **2008**, *8*, 7300–7322. [CrossRef]
25. Visockiene, J.S.; Brucas, D.; Ragauskas, U. Comparison of UAV images processing softwares. *J. Meas. Eng.* **2014**, *2*, 111–121.
26. Pix4D. Getting GCPs in the Field or through Other Sources. Available online: <https://support.pix4d.com/hc/en-us/articles/202557489-Step-1-Before-Starting-a-Project-4-Getting-GCPs-in-the-Field-or-Through-Other-Sources-optional-but-recommended-> (accessed on 26 December 2015).
27. Hexagon Geospatial. ERDAS IMAGINE Help. Automatic Tie Point Generation Properties. Available online: <https://hexagongeospatial.fluidtopics.net/book#!book;uri=fb4350968f8f8b57984ae66bba04c48d;breadcrumb=23b983d2543add5a03fc94e16648eae7-b70c75349e3a9ef9f26c972ab994b6b6-0041500e600a77ca32299b66f1a5dc2d-e31c18c1c9d04a194fdbf4253e73afaa-19c6de565456ea493203fd34d5239705> (accessed on 26 December 2015).
28. Drăguț, L.; Tiede, D.; Levick, S.R. Esp: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [CrossRef]
29. Duro, D.C.; Franklin, S.E.; Dubé, M.G. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using Spot-5 HRG imagery. *Remote Sens. Environ.* **2012**, *118*, 259–272. [CrossRef]
30. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q. Per-pixel vs. Object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [CrossRef]
31. Laliberte, A.S.; Fredrickson, E.L.; Rango, A. Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. *Photogramm. Eng. Remote Sens.* **2007**, *73*, 197–207. [CrossRef]
32. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; Wadsworth, Inc.: Monterey, CA, USA, 1984.
33. Rouse, J.W., Jr.; Haas, R.; Schell, J.; Deering, D. Monitoring vegetation systems in the great plains with ERTS. In Proceedings of the Third ERTS-1 Symposium NASA, NASA SP-351, Washington, DC, USA, 10–14 December 1973; pp. 309–317.

34. Jordan, C.F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* **1969**, *50*, 663–666. [[CrossRef](#)]
35. Richardson, A.J.; Everitt, J.H. Using spectral vegetation indices to estimate rangeland productivity. *Geocarto Int.* **1992**, *7*, 63–69. [[CrossRef](#)]
36. Roujean, J.L.; Breon, F.M. Estimating par absorbed by vegetation from bidirectional reflectance measurements. *Remote Sens. Environ.* **1995**, *51*, 375–384. [[CrossRef](#)]
37. McFeeters, S. The use of the normalized difference water index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
38. Rondeaux, G.; Steven, M.; Baret, F. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* **1996**, *55*, 95–107. [[CrossRef](#)]
39. Huete, A.R. A soil-adjusted vegetation index (SAVI). *Remote Sens. Environ.* **1988**, *25*, 295–309. [[CrossRef](#)]
40. Kauth, R.J.; Thomas, G. The tasselled cap—A graphic description of the spectral-temporal development of agricultural crops as seen by landsat. In Proceedings of the Symposia on Machine Processing of Remotely Sensed Data, West Lafayette, IN, USA, 29 June–1 July 1976; pp. 41–51.
41. Woebbecke, D.; Meyer, G.; Von Bargen, K.; Mortensen, D. Color indices for weed identification under various soil, residue, and lighting conditions. *Trans. ASAE* **1995**, *38*, 259–269. [[CrossRef](#)]
42. Meyer, G.E.; Hindman, T.W.; Laksmi, K. Machine vision detection parameters for plant species identification. *Proc. SPIE* **1999**, 3543. [[CrossRef](#)]
43. Camargo Neto, J. A Combined Statistical-Soft Computing Approach for Classification and Mapping Weed Species in Minimum-Tillage Systems. Ph.D. Thesis, University of Nebraska, Lincoln, NE, USA, 2004.
44. Kataoka, T.; Kaneko, T.; Okamoto, H. Crop growth estimation system using machine vision. In Proceedings of the 2003 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Kobe, Japan, 20–24 July 2003; pp. 1079–1083.
45. Woebbecke, D.M.; Meyer, G.E.; Von Bargen, K.; Mortensen, D.A. Plant species identification, size, and enumeration using machine vision techniques on near-binary images. *Proc. SPIE* **1993**, 1836, 208–219.
46. Story, M.; Congalton, R.G. Accuracy assessment—a user’s perspective. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 397–399.
47. Stehman, S.V.; Czaplewski, R.L. Design and analysis for thematic map accuracy assessment: Fundamental principles. *Remote Sens. Environ.* **1998**, *64*, 331–344. [[CrossRef](#)]
48. Rosenfield, G.H.; Fitzpatrick-Lins, K. A coefficient of agreement as a measure of thematic classification accuracy. *Photogramm. Eng. Remote Sens.* **1986**, *52*, 223–227.
49. Hunt, E.R.; Hively, W.D.; Fujikawa, S.J.; Linden, D.S.; Daughtry, C.S.; McCarty, G.W. Acquisition of NIR-green-blue digital photographs from unmanned aircraft for crop monitoring. *Remote Sens.* **2010**, *2*, 290–305. [[CrossRef](#)]

