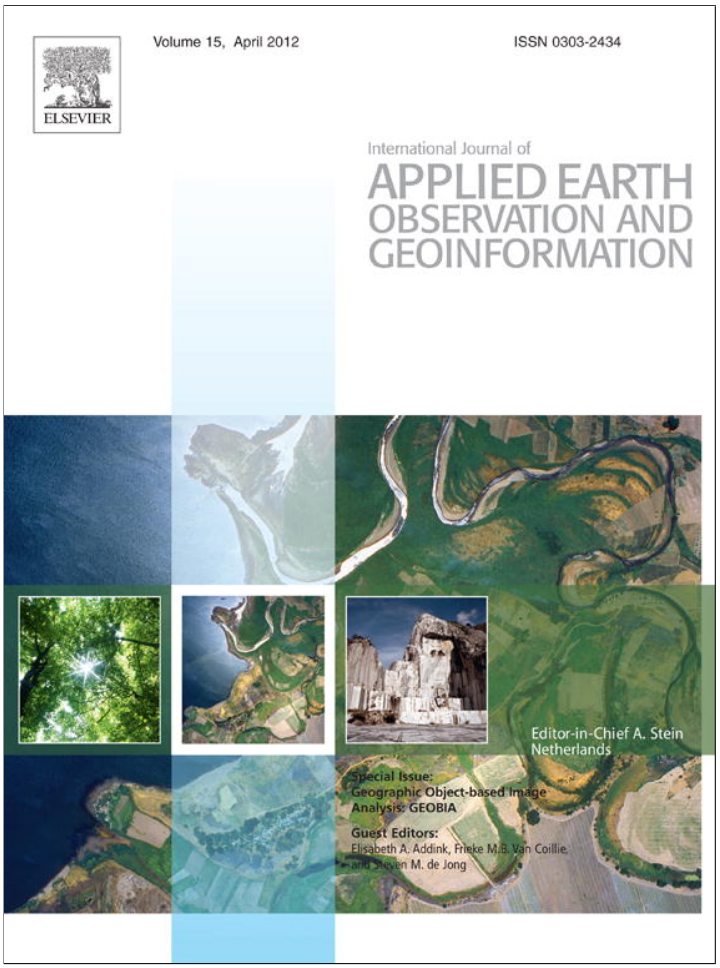

International Journal of Applied Earth Observation and Geoinformation 15 (2012) 70–78

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

A comparison of three feature selection methods for object-based classification of sub-decimeter resolution UltraCam-L imagery

A.S. Laliberte[a,*], D.M. Browning[b], A. Rango[b]

[a] Jornada Experimental Range, New Mexico State University, Las Cruces, NM 88003, USA
[b] USDA-Agricultural Research Service, Jornada Experimental Range, Las Cruces, NM 88003, USA

ARTICLE INFO

Article history:
Received 23 August 2010
Accepted 18 May 2011

Keywords:
Feature selection
Object based image analysis
Aerial imagery
Very high resolution
Classification

ABSTRACT

The availability of numerous spectral, spatial, and contextual features with object-based image analysis (OBIA) renders the selection of optimal features a time consuming and subjective process. While several feature selection methods have been used in conjunction with OBIA, a robust comparison of the utility and efficiency of approaches would facilitate broader and more effective implementation. In this study, we evaluated three feature selection methods, (1) Jeffreys–Matusita distance (JM), (2) classification tree analysis (CTA), and (3) feature space optimization (FSO) for object-based vegetation classifications with sub-decimeter digital aerial imagery in arid rangelands of the southwestern U.S. We assessed strengths, weaknesses, and best uses for each method using the criteria of ease of use, ability to rank and/or reduce input features, and classification accuracies. For the five sites tested, JM resulted in the highest overall classification accuracies for three sites, while CTA yielded highest accuracies for two sites. FSO resulted in the lowest accuracies. CTA offered ease of use and ability to rank and reduce features, while JM had the advantage of assessing class separation distances. FSO allowed for determining features relatively quickly, because it operates within the OBIA software used in this analysis (eCognition). However, the feature ranking in FSO is not transparent and accuracies were relatively low. While all methods offered an objective approach for determining suitable features for classifications of sub-decimeter resolution aerial imagery, we concluded that CTA was best suited for this particular application. We explore the limitations, assumptions, and appropriate uses for this and other datasets.

1. Introduction

The selection of appropriate spectral bands and/or image features is a crucial step in any image analysis process. Using a set of optimal features ensures that the classes in question are discriminated effectively and with sufficiently high accuracy, and that the dimensionality is reduced for efficient use of training samples (Jensen, 2005). With increased use of high resolution aerial and satellite imagery, object-based image analysis (OBIA) has become more commonplace in recent years due to its ability to extract meaningful image objects by segmentation, and to bridge remote sensing and GIS (Blaschke, 2010; Blaschke et al., 2008; Hay et al., 2005). The ability to incorporate elements used traditionally in aerial photo interpretation (color, size, shape, texture, pattern, and contextual information), is one of the strengths of OBIA. However, the availability of hundreds on spectral, spatial, and contextual features for each image object can make the determination of optimal features a time consuming or subjective process. In this paper, the term "feature" describes spectral (image bands, band ratios), spatial (area, compactness, etc.), contextual (difference to neighbor), and texture properties in an OBIA context.

Feature selection techniques range from graphic methods to statistical approaches involving separation distances for image classes. Several feature selection methods have been used in conjunction with OBIA. Herold et al. (2003) and Carleer and Wolff (2006) used the Bhattacharyya distance, while Nussbaum et al. (2006), Marpu et al. (2008), and Zhang et al. (2010) employed the Jeffreys–Matusita distance for feature selection. Van Coillie et al. (2007) used a genetic algorithm, and Johansen et al. (2009) evaluated feature space plots, box plots, band histograms, and feature space optimization. Classification tree analysis for selection of optimal features was successfully applied by Chubey et al. (2006), Yu et al. (2006), Laliberte et al. (2007), and Addink et al. (2010).

The above-mentioned studies used satellite images (QuickBird, Ikonos, SPOT) or aerial photography at resolutions ranging from 0.3 to 1.25 m. In recent years, the use of digital mapping cameras has greatly increased, and examples for use of these images include mapping benthic habitats (Green and Lopez, 2007), land use/land cover mapping (Rosso et al., 2008), border monitoring (Coulter and Stow, 2008), and mapping changes in shrub cover

* Corresponding author. Tel.: +1 575 646 3557.
E-mail address: alaliber@nmsu.edu (A.S. Laliberte).

0303-2434/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jag.2011.05.011

(Stow et al., 2008). Digital airborne imagery can now be acquired at sub-decimeter resolution while maintaining sufficient image overlap for photogrammetric processing, allowing for creation of orthophotos and digital surface models at very high resolution (Wiechert and Gruber, 2010). This imagery exhibits great potential for vegetation mapping at very high resolution, despite multiple challenges such as high spatial frequency, the effect of shadows, viewing geometry, and illumination (Laliberte et al., 2010a).

The ability to map vegetation at fine resolution allows for comparison of ground- and image-based measurements of cover and species composition at the plot level, and subsequent extrapolation to the landscape level. This approach has the potential to be implemented for monitoring purposes by land management agencies (i.e., U.S. Bureau of Land Management (BLM) and U.S. Natural Resources Conservation Service (NRCS)) that are mandated to monitor and assess millions of acres of rangelands. Broad-scale implementation has been demonstrated to be well suited for very high resolution remote sensing (Ehlers et al., 2003, 2006; Laliberte et al., 2010b). This study is part of a larger research effort focused on developing and implementing novel remote sensing acquisition and analysis techniques suitable for potential integration into the National Resources Inventory (Nusser and Goebel, 1997). This study builds on previous research examining the use of OBIA for estimating vegetation cover in very high resolution digital imagery (Laliberte et al., 2010a) by extending the analysis to classification at the species level. In order to evaluate species separability in very high resolution imagery, suitable features for classification have to be determined.

Optimal features for classification may be scale dependent, and features used in the analysis of moderate resolution imagery (i.e., satellite or aerial photography) may not be applicable to finer resolution data. Determination of appropriate features for very high resolution imagery, and a robust comparison of the utility and efficiency of various feature selection methods could facilitate broader use of sub-decimeter aerial imagery for vegetation mapping.

We focus on assessing three commonly used feature selection approaches used in conjunction with OBIA rather than presenting an exhaustive investigation of feature selection methods. The objectives of this study were to (1) determine the optimal features for fine-scale vegetation mapping, and (2) evaluate three feature selection methods (i.e., Jeffreys–Matusita distance (JM), classification tree analysis (CTA), and feature space optimization (FSO)), in the context of object-based classification of rangeland vegetation with digital aerial imagery with a 6 cm ground resolved distance. Evaluation criteria for the feature selection methods were efficiency and ease of use, ability to rank and reduce features, and classification accuracies.

2. Methods

2.1. Study area and images

The study sites were located on the Jornada Experimental Range (JER) and the Chihuahuan Desert Rangeland Research Center (CDRRC) in southwestern New Mexico, USA (32°34′11″W, 106°49′44″N) (Fig. 1). Average elevation is ca. 1200 m, and rainfall amounts and distribution are highly variable, with more than 50% of the mean annual precipitation of 245 mm occurring during July, August, and September (Wainwright, 2006). Much of the historic semi-desert grassland has experienced marked increase in shrub abundance and distribution over the last century, although some grass-dominated areas remain (Gibbens et al., 2005). For this study, imagery was acquired over five 150 m × 150 m plots in five vegetation communities. The five sites are a subset of fifteen plots established as long-term research plots as part of the Jornada Basin long-term ecological research (LTER) program and represent five plant communities with a wide range in vegetation structure, primary productivity, and species composition. Of the five plant communities, two are dominated by grass – upland grasslands (GIBPE), low-lying areas (i.e., playas) (PCOLL), and three are dominated by shrubs with co-occurring mixed grasses – creosote (*Larrea tridentata*) (CGRAV), tarbush (*Flourensia cernua*) (TEAST) and mesquite (*Prosopis glandulosa*) (MWELL) (Peters and Gibbens, 2006).

The images were acquired 13 September, 2009 with an UltraCam-L digital mapping camera at a flying height of approximately 480 m above ground. The camera acquires multispectral data in the red (580–700 nm), green (480–630 nm), blue

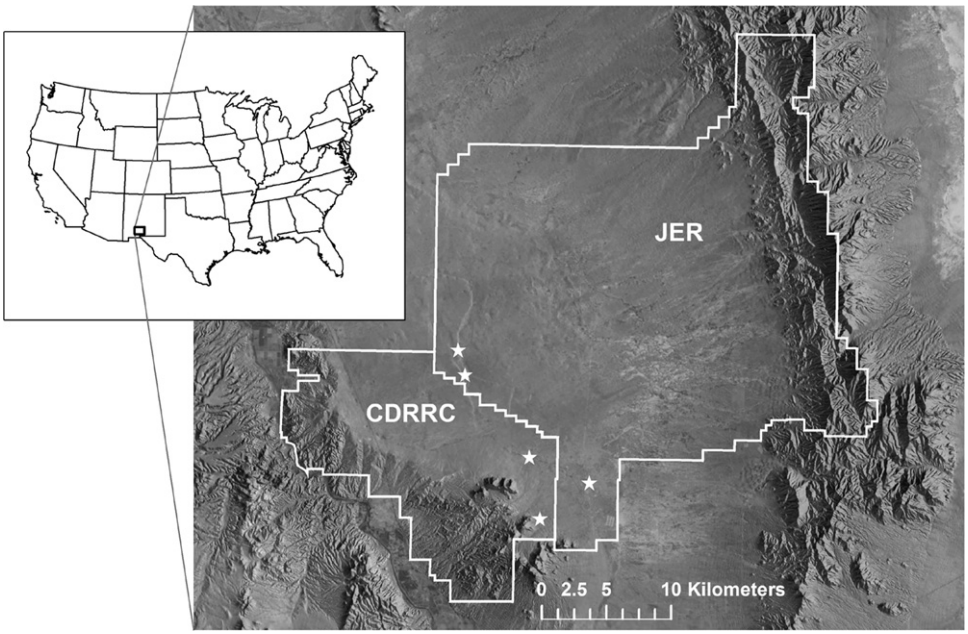

Fig. 1. Study area in southwestern New Mexico, USA, showing the five plots over which the UltraCam-L imagery was acquired at the Jornada Experimental Range (JER) and the Chihuahuan Desert Rangeland Research Center (CDRRC).

Table 1
Four-letter abbreviations, scientific names, common names, and growth forms of species mapped in five plots at the JER and CDRRC.

Abbreviation	Scientific name	Common name	Growth form
ATCA	*Atriplex canescens*	Four-wing saltbush	Shrub
EPTR	*Ephedra trifurca*	Longleaf ephedra	Shrub
FAPA	*Fallugia pardoxa*	Apache plume	Shrub
FLCE	*Flourensia cernua*	Tarbush	Shrub
LATR	*Larrea tridentata*	Creosote bush	Shrub
LYBE	*Lycium berlandieri*	Wolfberry	Shrub
PRGL	*Prosopis glandulosa*	Honey mesquite	Shrub
RHMI	*Rhus microphylla*	Littleleaf sumac	Shrub
YUBA	*Yucca bacata*	Banana yucca	Shrub
YUEL	*Yucca elata*	Soaptree yucca	Shrub
GUSA	*Gutierrezia sarothrae*	Broom snakeweed	Sub-shrub
ARPU	*Aristida purpurea*	Purple threeawn	Grass
BOER	*Bouteloua eriopoda*	Black grama	Grass
DAPU	*Dasyochloa pulchella*	Fluffgrass	Grass
MUPO	*Muhlenbergia porteri*	Bush muhly	Grass
PAOB	*Panicum obtusum*	Vine mesquite	Grass
PLMU	*Pleuraphis mutica*	Tobosa grass	Grass
SCBR	*Scleropogon brevifolius*	Burrograss	Grass
SPFL	*Sporobolus flexuosus*	Mesa dropseed	Grass
GRSQ	*Grindelia squarrosa*	Curlycup gumweed	Forb
HECI	*Helianthus ciliaris*	Blueweed	Forb
HYOD	*Hymenoxys oderata*	Bitterweed	Forb
LACO	*Laennecia coulteri*	Coulter conyza	Forb
LEVI	*Leptochloa viscida*	Gum sprangletop	Forb
SONU	*Sophora nuttalliana*	Silky sophora	Forb
ZIAC	*Zinnia acerosa*	Desert zinnia	Forb
OP spp.	*Opuntia*	Prickly pear	Cactus

(410–570 nm), and near infrared (690–1000 nm) bands. Five overlapping images were acquired over each site, orthorectified using interior and exterior orientation values using a direct georeferencing procedure (no ground control points), and mosaicked using Leica Photogrammetric Suite (Erdas, 2009). The image mosaic (6 cm ground resolved distance) was subsequently co-registered to an orthorectified QuickBird image using Erdas Imagine AutoSync® with an average of 230 automatically generated tie points and an average RMS error of 5 cm. Images were clipped to the plot boundaries for further analysis.

2.2. Field data

At each of the five plots, training and test samples for the dominant vegetation species were collected in polygon format. Due to the relatively small size of the plots (150 m × 150 m), training samples were collected by systematically walking the plot area. On average, 520 species-level samples were collected per plot, and classes per plot ranged from 6 to 10. A total of 27 species-specific classes were mapped (Table 1). Species with fewer than 20 samples per class were omitted from the analysis, because CTA is sensitive to large differences in sample size among classes.

Training and test samples were collected by digitizing vegetation boundaries over the displayed image using ArcPad on a tablet PC. We chose this approach rather than delineating vegetation boundaries with a GPS due to the discrepancy between image resolution and GPS positional error. At 6 cm image resolution and sub-meter accuracy of differentially corrected GPS data, the error of the vegetation boundaries would be a minimum of 10 pixels, making it difficult to determine the precise location of a vegetation patch. Half of the samples in each plot were randomly selected as training sites, and half were retained for accuracy assessment.

2.3. Image analysis

2.3.1. Segmentation and rule-based classification

We used eCognition Developer 8 (Definiens, 2009) for the object-based image analysis.

The images were segmented at two scales, a fine scale multiresolution segmentation with scale parameter 100, and a coarser scale spectral difference segmentation with a maximum spectral difference of 1500. A spectral difference segmentation allows for combining adjacent image objects with similar spectral properties into larger objects while maintaining small spectrally distinct objects. In this manner, small patches of vegetation can be maintained within larger bare soil areas and vice versa, while simultaneously reducing the number of objects. All bands were weighted equally for the segmentation; color/shape was set to 0.9/0.1, and smoothness/compactness was set to 0.5/0.5.

The segmentation parameters were determined based on expert judgment and visual interpretation. While more objective segmentation approaches have been developed recently (Esch et al., 2008; Dragut et al., 2010), we wanted to ensure that our segmentation approach (multiresolution and spectral difference segmentation) matched the one used over the same plots with 4-cm resolution digital imagery (Laliberte et al., 2010a). In addition, evaluating different segmentation approaches was not an objective in this study.

All classifications were done at the coarser segmentation scale. A rule-based approach was used to classify the images into shadow, bare ground, and vegetation. In the tarbush, playa, and grassland plots, a sparse vegetation class was added, consisting of widely spaced tufts of sparse vegetation and litter, which were not of interest in this study. A process tree, a collection of rules for segmentation and classification, was developed on the first image and applied to the remaining images for consistency in the analysis. The threshold parameters for the rule-based classification of shadow, bare ground, and vegetation were changed for each plot according to differences in vegetation. A detailed assessment of threshold values in a related study using 4-cm digital mapping camera (DMC) imagery over the same plots resulted in broad guidelines for these threshold values, and demonstrated that small changes in values did not affect the transferability of the rule to other plots (Laliberte et al., 2010a).

2.3.2. Feature selection methods and nearest neighbor classification

We tested the three feature selection methods for the species-level classification only. We started the analysis with 31 spectral, spatial, and texture features. The initial features were selected based on previous OBIA projects using aerial imagery in the Jornada Basin (Laliberte and Rango, 2009; Laliberte et al., 2010a,b). The dominant grasses in the grass-dominant plots GIBPE and PCOLL were black grama (*Bouteloua eriopoda*) and tobosa (*Pleuraphis mutica*), respectively, and both species were relatively unique spectrally. Other mixed grasses in both grass- and shrub-dominated plots, and several shrub species were more difficult to differentiate unambiguously with visual analysis alone. The initial features included a sufficiently wide range of spectral, spatial, and texture features to determine suitable features for each species.

Spearman's rank correlation analysis was used to eliminate features with correlation coefficients >0.9 to reduce data dimensionality. As a result, 12–18 features remained per plot (Table 2). Three feature selection methods were tested, (1) JM, (2), CTA, and (3) FSO. Inputs for all three methods consisted of the feature values of the training objects for each class. In this paper, we use the abbreviation JM for the feature selection method based on the Jeffreys–Matusita distance, and the abbreviation JMdist for the distance measure.

The Jeffreys–Matusita distance, JMdist, is a pair-wise measure of class separability based on the probability distributions of two classes and is calculated as:

$$\mathrm{JMdist} = \sqrt{2(1 - e^{-B})}, \quad (1)$$

Table 2
Features selected from 31 input features using three feature selection methods (JM, CTA, FSO) for five plots. Uncorrelated features per plot have correlation coefficients <0.9. Numbers in columns represent feature ranks for the JM method by largest JM distance, for CTA by primary splitter in decision tree, and for FSO by order of selection in the FSO tool.

Plots	CGRAV			MWELL			TEAST			GIBPE			PCOLL		
Method	JM	CTA	FSO	JM	CTA	FSO	JM	CTA	FSO	JM	CTA	FSO	JM	CTA	FSO
Total uncorrelated	12	16	17	17	18										
Total selected	7	7	7	7	9	13	8	5	16	7	6	15	10	4	15
Features															
NDVI	2	2		1	1	7	3	2	11	2		13	2	3	14
Mean blue															
Mean green				6	5	12									
Mean NIR	4	4			3		6		16	6	3	10	6		13
Mean red	1	1	6	3	8	10	2	1	3	4	1	6	5	1	
Max difference					6	13	8		13	7	4	15	7		6
Standard deviation blue															
Standard deviation green						1									
Standard deviation NIR			4			4			6			1			2
Standard deviation red		6	2						15			5			9
Ratio blue	3	5				2	7	5	1			4			
Ratio green				7	2	5		4	7			11	3		1
Ratio NIR							4		8						
Ratio red							5		2	1	2		1	2	8
Mean diff. to neighb. blue													8	4	5
Mean diff. to neighb. green				5	9	6									
Mean Diff. to neighb. NIR			1										9		11
Mean diff. to neighb. red			7						10		5	8			
Area				2	4		1	3	5	3	6		4		
Compactness												14			7
Density			3			8			4			2			3
Roundness					7	9			9						
Shape index															
GLCM Homogeneity						11			14			7			12
GLCM contrast												9			
GLCM dissimilarity									12						10
GLCM entropy	7		5												15
GLCM std. dev.															
GLCM correlation						3						3			4
GLCM ang. 2nd moment	5	7								5		12	10		
GLCM mean	6	3		4											

where B is the Bhattacharyya distance:

$$B = \frac{1}{8}(\mu_1 - \mu_2)^2 \left(\frac{C_1 + C_2}{2}\right)^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \left(\frac{|((C_1 + C_2)/2)|}{\sqrt{(|C_1||C_2)}} \right) \quad (2)$$

and μ_1, μ_2 are the means and C_1, C_2 the covariance matrices for two classes, 1 and 2 (Ferro and Warner, 2002). Compared to the Bhattacharyya distance, the Jeffreys–Matusita distance has a finite dynamic range. This characteristic allows for easier comparison of class separability between images.

For JM, we used the SEATH tool (Nussbaum et al., 2006; Marpu et al., 2008) due to its compatibility with eCognition's exported object statistics. Based on the input sample data, a probability distribution for each class is estimated based on the means and variances of the two classes in question. The thresholds are determined by fitting a Gaussian probability mixture model to the frequency distribution of a feature for the two classes (Nussbaum et al., 2006). The class separation between two classes for each feature is measured on a scale of 0–2, with 2 indicating complete separability. SEATH calculates the class separability and threshold for every feature and two-class combination and outputs individual text files for each two-class comparison. The data were compiled to determine the largest average JMdist (meaning the average of all two-class combinations) for every possible 4–10 feature combination. Combinations with fewer features always had lower average JMdist values than combinations with more features, and selecting the lowest average JMdist from the seven candidates (best 4 features through best 10 features) would not have been appropriate. Therefore, we selected the feature combination that resulted in the largest JMdist for the least separable pair of classes, a common strategy in feature selection when multiple classes have to be considered (Swain and Davis, 1978).

For CTA, we used CART® software (Salford Systems) which outputs decision trees based on the algorithm developed by Breiman et al. (1984). CTA is a nonparametric statistical technique, whereby a dataset is successively split into increasingly homogenous subsets. At each node in the tree, the splitting rule is defined by the Gini index, a measure of heterogeneity. The Gini index at node t is defined as

$$g(t) = \sum p(j)p(i) \quad (3)$$

where $p(j)$ and $p(i)$ are the probability of class j and i at node t, with a range of $g(t)$ from 0 to 1 (Steinberg and Colla, 1997). To prevent overfitting of the tree, splitting was stopped when a terminal node had less than ten classes. The optimal tree was determined by 10-fold cross validation. The selection of optimum features and the ranking was based on the variable importance scores of the primary splitters in the decision tree. The scores reflect the contribution of each feature in predicting the output class, with scores ranging from 0 to 100 (Steinberg and Colla, 1997).

FSO is a tool available in eCognition and it calculates an optimum feature combination based on class samples. FSO evaluates the Euclidean distance in feature space between the samples of all classes and selects a feature combination resulting in the best class separation distance, which is defined as the largest of the minimum distances between the least separable classes (Definiens, 2009).

The broad vegetation class from the initial rule-based classification was further refined into species-level classes using a nearest

neighbor classification with the features obtained from the three feature selection methods. A nearest neighbor method is simple to implement and generally has good results with carefully chosen features, although it is sensitive to irrelevant predictors. Because we used feature selection methods in this study, a nearest neighbor classification approach was deemed appropriate. We determined classification accuracies for the three output maps for the species-level classes only by creating an error matrix to determine user's, producer's, and overall classification accuracies, and Kappa statistics (Congalton and Green, 2009). Statistical differences between classifications were assessed with McNemar's test, a nonparametric test of contingency tables (Foody, 2004).

3. Results

3.1. Feature selection

The features selected with the three methods are shown in Table 2. CTA selected the lowest number of features in three of the five plots. FSO consistently selected the largest number of features, in two plots twice as many features as JM, and in two plots three times as many features as CART, with one noted exception being the CGRAV plot. We observed more consistency in feature selection between JM and CTA than between FSO and either of the other two methods. For example, in CGRAV, FSO did not select NDVI or Mean NIR, while both features were selected by JM and CTA (Table 2). A similar observation was apparent with ranked features, with greater similarity between JM and CTA, and less similarity between FSO and either JM or CTA. For JM, the ranking is based on the average JMdist of all two-class comparisons for a particular feature space. For CTA, the ranking was obtained using the variable importance scores of the primary splitters in the decision tree. The FSO ranking is based on the order of selection within the FSO tool. Overall, spectral features were more likely to be selected than spatial or texture features, and on average the highest ranking features were mean red, NDVI, and ratio red.

3.2. Classification accuracy

The highest overall classification accuracies and Kappa values were obtained using JM for three plots (MWELL, GIBPE, CGRAV), followed by CTA for two plots (PCOLL, TEAST). Classifications with features from FSO had the lowest overall accuracies and Kappa values (Fig. 2). p-Values from McNemar's tests for all two-method comparisons were <0.001 except for the CTA vs. FSO comparisons in MWELL ($p = 0.052$) and GIBPE ($p = 0.029$). In general, Kappa values for individual classes showed more consistency across methods if the class was highly separable from other classes (based on JMdist). Conversely, less separable classes had greater variation in Kappa values for the three methods (Fig. 3). For example, at the MWELL site, ATCA (Fourwing saltbush) and YUEL (Yucca), which were confused with each other, showed a large variation in Kappa values for the methods, while highly separable species such as GUSA (Broom snakeweed) and PRGL (Honey mesquite) had comparable Kappa values for all methods. However, there were exceptions to this observation related to the size class of certain shrubs. While large PRGL (avg. diameter >0.6 m) were highly separable from other species and had large Kappa values (as in plot MWELL), small PRGL (avg. diameter <0.6 m) were often confused with other small shrubs due to their higher reflectance, resulting in Kappa values around 0.5 for PRGL in CGRAV (Fig. 3).

The species-level producer's and user's accuracies summarized by plot based on the feature selection method resulting in the highest overall classification accuracy were relatively high, with a mean, standard deviation, and median of 70%, 25%, 72%, respectively, for producer's accuracy, and 77%, 18%, 82%, respectively, for user's accuracy (Table 3). The variability in accuracy for a given species was highly dependent on the size class of the species in question, and the occurrence of species with similar spectral properties in the same plot. The classification of plot TEAST shows the fine detail achievable with this imagery (Fig. 4). In TEAST, three shrub species, five grass species, and one forb were classified with the nearest neighbor classification in addition to the classes bare, shadow, and sparse vegetation obtained with the rule-based classification. Overall accuracy for TEAST was 80%.

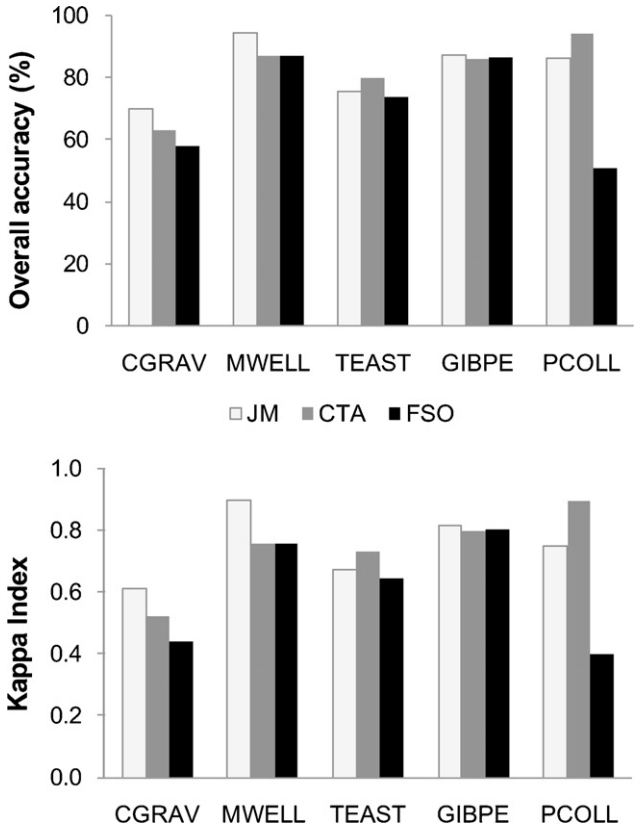

Fig. 2. Overall classification accuracies and Kappa indices for classifications of five plots with three feature selection methods (JM, CTA, FSO). p-Values from McNemar's tests for all two-method comparisons were <0.001 except for the CTA vs. FSO comparisons in MWELL ($p = 0.052$) and GIBPE ($p = 0.029$).

3.3. Class separation distances

Both FSO and JM provide class separation distances, and while a detailed comparison of the separation distances is beyond the scope of this paper, we compared the values obtained with FSO and JM for all two-class combinations for each plot. In contrast to the JMdist calculation, the separation distances obtained from FSO are more simplified. Distances between classes are calculated by determining for each sample of class a the sample of class b with the smallest Euclidean distance to it. The process is repeated for samples of class b compared to class a, and the Euclidean distances are finally averaged over all samples. The correlations for the two measures were relatively high with the exception of one plot (GIBPE), and p-values were <0.001 for four plots (Table 4).

3.4. Processing times

All processing was performed on a computer with 4 GB of RAM and two dual-core 2.6 GHz processors. The three methods for feature selection resulted in a range in processing times due to the

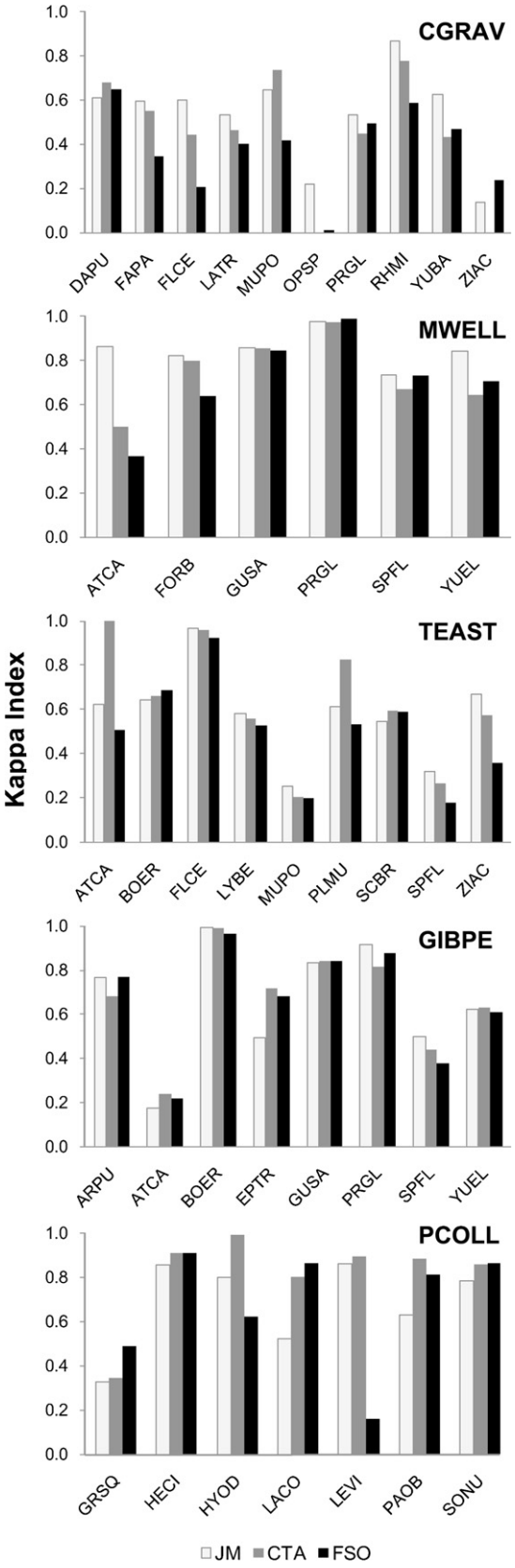

Fig. 3. Kappa indices for individual classes for three classification methods (JM, CTA, FSO) in five plots Species are listed with four-letter abbreviations, details are in Table 1.

Table 3
Percent vegetation cover and producer's and user's accuracies for species-level classes in five plots. For each plot, the values are based on classifications based on the feature selection method resulting in the highest overall classification accuracy: JM for CGRAV, MWELL, and GIBPE, and CTA for TEAST and PCOLL.

Species	Plot	Vegetation cover (%)	Producer's accuracy (%)	User's accuracy (%)
DAPU	CGRAV	7.5	61	85
FAPA	CGRAV	0.6	62	73
FLCE	CGRAV	0.3	63	43
LATR	CGRAV	9.0	65	65
MUPO	CGRAV	7.4	66	62
OP spp	CGRAV	0.8	22	94
PRGL	CGRAV	3.3	71	68
RHMI	CGRAV	10.9	89	88
YUBA	CGRAV	3.8	65	84
ZIAC	CGRAV	2.0	57	68
ARPU	GIBPE	1.2	77	92
ATCA	GIBPE	1.2	20	30
BOER	GIBPE	41.1	100	91
EPTR	GIBPE	3.6	51	59
GUSA	GIBPE	1.7	84	85
PRGL	GIBPE	14.0	94	93
YUEL	GIBPE	2.3	66	82
ATCA	MWELL	5.8	88	98
FORB	MWELL	8.2	82	93
GUSA	MWELL	9.2	87	93
PRGL	MWELL	0.5	99	95
SPFL	MWELL	1.7	74	93
YUEL	MWELL	12.2	87	90
GRSQ	PCOLL	0.3	35	67
HECI	PCOLL	0.3	91	64
HYOD	PCOLL	21.7	99	99
LACO	PCOLL	7.3	81	42
LEVI	PCOLL	23.2	96	100
PAOB	PCOLL	4.8	91	100
SONU	PCOLL	0.5	86	64
ATCA	TEAST	1.5	100	83
BOER	TEAST	30.6	70	60
FLCE	TEAST	9.6	98	90
LYBE	TEAST	1.8	56	81
MUPO	TEAST	1.6	22	53
PLMU	TEAST	10.3	87	71
SCBR	TEAST	0.9	64	96
SPFL	TEAST	0.7	50	80
SPFL	TEAST	5.0	27	45
ZIAC	TEAST	3.8	14	79

different steps required for each method. For all plots, CTA was faster than JM, which was faster than FSO for obtaining an optimum feature combination. However, the speed of FSO was entirely dependent on how many texture features were included, because texture determination is CPU demanding. In this analysis, between one and four texture features were selected by the FSO tool for each plot, slowing down the operation. For example, running the FSO tool took on average 10 s for a test using 10 features without texture. With four texture features, this task required up to 2 h. If no texture features are included, FSO has the potential to be the fastest feature selection method.

Both CTA and JM required exporting the object statistics from eCognition to another program for determining the optimum features. CTA was more time efficient, because CART® processing times

Table 4
Correlations between JM distances and class separation distances obtained with FSO tool for all two-class comparisons for five plots.

Plot	Correlation coefficient	*p*-Value	*n*
CGRAV	0.88	<0.001	45
MWELL	0.81	<0.001	15
TEAST	0.82	<0.001	36
GIBPE	0.41	0.03	28
PCOLL	0.76	<0.001	21

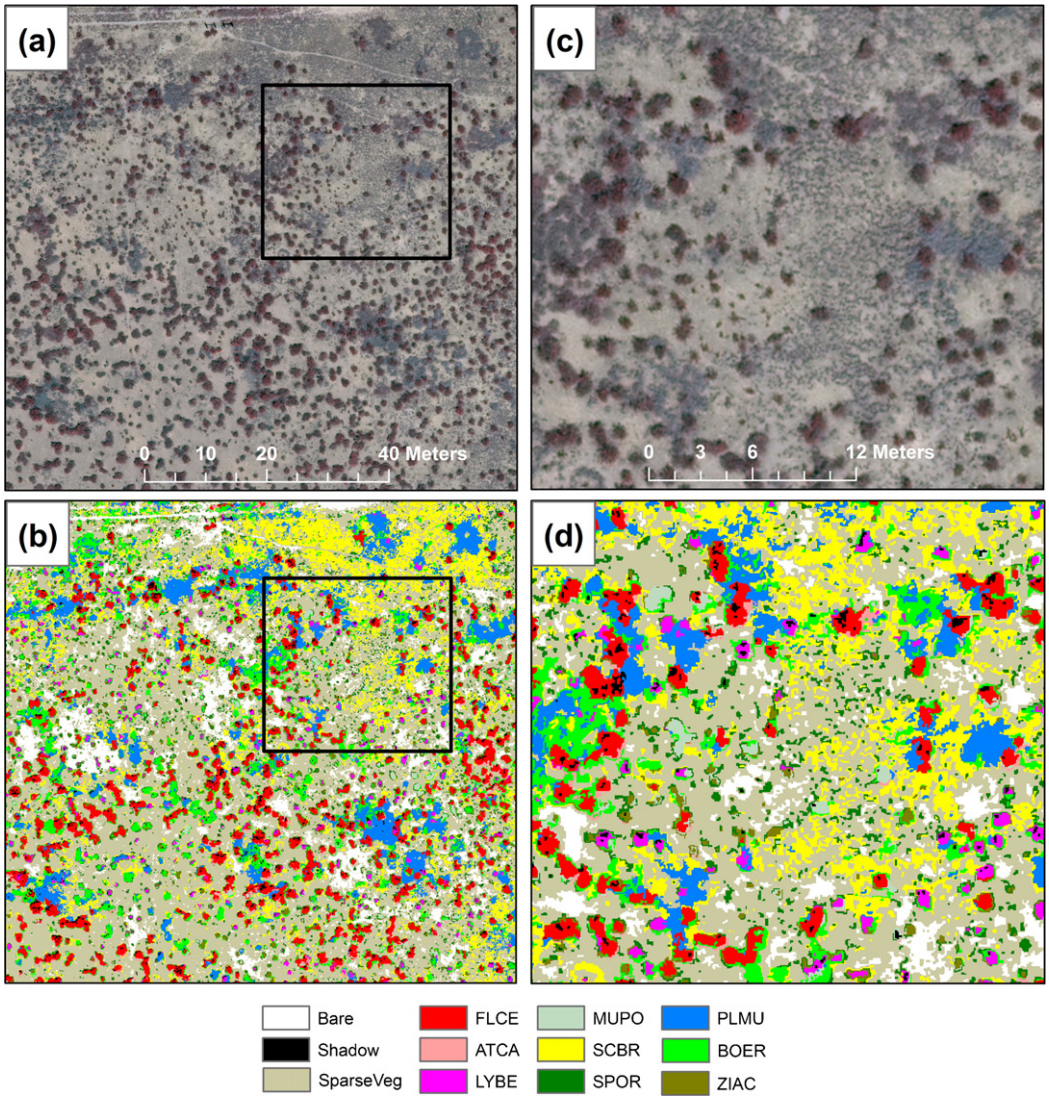

Fig. 4. UltraCam-L image with 6 cm ground resolved distance (a) and classification (b) for a portion of plot TEAST. The black box delineates the enlarged area in (c) and (d). Classification was performed using CTA. Species are listed with four-letter abbreviations, details are in Table 1.

from import to outputs of results required only seconds. Feature rankings based on variable importance scores of the primary splitters in the tree were easy to interpret. JM required the most data manipulation, because several steps were required for analyzing object statistics using the SEATH tool, compiling individual output files for each two-class combination (up to 45), and obtaining average separation distances for every possible 4–10 feature combination before determining the optimum feature combination for the least separable pair of classes. Classification times in eCognition were also highly dependent on the number of texture features included. Plots without texture features were classified in minutes, while plots with four texture features took up to 9 h to classify.

4. Discussion

Each feature selection method assessed in this study had advantages and disadvantages. In order to assess the suitability of a particular feature selection approach, classification accuracy is an important criterion, because ultimately the analyst desires a result with the highest achievable accuracy. However, other criteria have to be considered as well, especially if multiple high resolution images have to be analyzed. Ease of use, efficiency, and processing times are equally important, and an efficient workflow may be preferable to higher overall accuracy, especially if the difference in accuracy consists of a few percentage points. The ability to clearly interpret the results of a feature selection method also has to be taken into consideration. Finally, a robust feature selection technique has to be capable of ranking and reducing a potentially large number of input features. In most situations, training samples are costly and therefore limited. Using a reduced number of features for a given set of training samples reduces dimensionality and prevents what is known as the Hughes phenomenon, the deterioration of classification accuracy due to the addition of unnecessary features (Kim and Landgrebe, 1991).

We determined that CTA was best suited for this particular study based on ability to rank and reduce features, relatively low processing times, and relatively high accuracy. However, we acknowledge that either of the other methods may be preferable for different research objectives. CTA proved to be an excellent feature reduction and ranking tool; it required few steps, and the results were easy to interpret. CTA also has the advantage of being a non-parametric approach. In terms of efficiency, if many classes and/or many features are involved, JM can be time-consuming, because the analysis is essentially based on multiple two-class comparisons. In cases where the assumption of normality underlying the JM distance may not be met due to limited sample sizes, CTA would be preferable.

Table 5
Strengths and weaknesses of, and best uses for three feature selection methods for object-based image analysis of very high resolution digital aerial photography. The feature selection methods are Jeffreys–Matusita distance (JM), applied using the SEATH tool, classification tree analysis (CTA), and feature space optimization (FSO).

	JM	CTA	FSO
Strengths	(1) Feature ranking for 2-class comparisons, (2) JM distances and rules, (3) compatible with eCognition export	(1) Feature reduction and ranking, (2) non-parametric, (3) can obtain features or specific rules, (4) fast analysis	(1) Feature reduction within eCognition, (2) class separation distances, (3) fast analysis without texture features
Weaknesses	(1) No initial feature reduction, (2) requires multiple steps for feature selection, (3) assumes normality	(1) No class separation distances, (2) Potential for overfitting decision tree	(1) "Black box" feature selection approach, (2) unclear feature ranking, (3) no rules
Best uses	(1) Feature ranking for NN or rule-based classification when separation distances are needed, (2) for limited number of classes and features	(1) Feature reduction and ranking for NN or rule-based classification, (2) for many classes and/or features, (3) for features with non-normal distributions	(1) Feature reduction for NN classification

On the other hand, the JM method has the advantage of quantitative class separation distances. We found that class separation distances offered new insights into the influence of size classes of certain species. While larger mesquite shrubs were clearly separable from other species of shrubs, small mesquites had a tendency to be confused with similarly sized individuals of other species. FSO also provides class separation distances, but has certain limits. The JMdist has the advantage of being scaled from 0 to 2, which makes is convenient for comparing separability of the same classes in different plots. This was not possible using the FSO-derived class separation distances, because the maximum distances varied greatly from plot to plot. However, the correlation between the two distance measures was relatively high, indicating that either measure provides useful insight into class separation. FSO-derived class separation distances have been shown to be a reasonable indicator of classification accuracy in object based image analysis (Laliberte and Rango, 2009).

When all factors were included, FSO was the lowest ranking method in this study. While it can be applied directly in eCognition and can potentially be the fastest approach without texture features, it consistently selected the largest number of features of all three methods, thereby achieving the lowest relative feature reduction capability. In addition, FSO appeared to be a "black box" without clear transparent mechanisms that provided the lowest classification accuracies.

We summarized strengths, weaknesses, and suggested best uses for the three feature selection methods, demonstrating that classification objectives and logistical constraints are best considered a priori (Table 5). We suggest that CTA is best suited for cases with numerous image classes and many features due to its excellent feature reduction ability, ease of use, and easy interpretability. JM is appropriate if class separation distances are of interest, and if the analyst desires to rank input features for two-class comparisons. In this study, we only tested nearest neighbor classification using training samples. However, two of the methods, CTA and JM, also provide threshold values that can be used as inputs for rule-based classification. While CTA outputs thresholds based on the entire classification tree, providing easy transfer to a rule-based classification (Laliberte et al., 2007), the SEATH tool used for determination of the JM distance provides threshold values for separating only two classes, and the analyst has to choose how many features to use per class. Nussbaum et al. (2006) and Gao et al. (2007) used this approach, choosing first the two features with the highest separability for each class combination, and then limiting the inputs to the final rule-base for each class to the features with the maximum separability.

Classification accuracies in this study varied for the five plots, although we did not interpret this as a lack of robustness of the method. Because the plots were located in five distinct vegetation communities with different size classes of the same species, different accuracy results were not deemed unusual. We expect the method to be relatively robust if transferred to imagery taken at another time of day or year. Due to different viewing geometry, illumination, and the resulting effect on shadows, adjustments in the thresholds would be required for the classes shadow, bare ground, and vegetation. However, based on previous research (Laliberte et al., 2010a), we expect that those adjustments would be minor and would not affect the transferability to a large degree.

Our main objective was to evaluate the feature selection methods, but the classification accuracy results offered other valuable insights. For example, classes with relatively low producer's accuracies should be aggregated into higher parent classes, if the rule base were to be applied to a larger image. Another aspect relates to potential segmentation inaccuracies and their effect on classification results. As in any OBIA study, there are always some under- or over-segmented objects of interest. Our quality control consisted of expert judgment and visual interpretation, which was feasible on the small images we analyzed. For transferring the approach to larger images, we would suggest the use of more objective segmentation approaches (Esch et al., 2008; Dragut et al., 2010). Because our comparison of feature selection methods was based on the same segmentation, it was reasonable to assume that the classification accuracies could be attributed to the feature selection methods.

Very high resolution digital aerial imagery has the potential to be used for monitoring purposes by land management agencies, and good correlations between ground- and image-based measures of cover have been achieved at the structure group level (grass, shrub, etc.) (Laliberte et al., 2010a,b). This study shows that relatively high classification accuracies can be achieved with sub-decimeter digital aerial imagery for detailed species-level classification with the use of training samples and an appropriate feature selection approach. Lower resolution data (10–15 cm) would offer advantages in processing times and might result in comparable classification accuracies, although a direct comparison with coarser resolution data is needed to evaluate tradeoffs between resolution, efficiency, and accuracy.

5. Conclusions

In this study, we assessed three feature selection methods for object-based classification of rangeland vegetation using sub-decimeter resolution digital aerial imagery. All methods offered an objective approach for feature selection. JM resulted in the highest classification accuracies for three of the five sites; however, CTA was deemed to be the best method for this particular imagery, because of the efficient workflow, the ability to both rank and significantly reduce input features, and the lack of parametric assumptions for normality. Even though CTA ranked second in terms of accuracy, we determined that it was the best method if multiple images were to be analyzed. FSO was ranked third based on feature reduction capability, the "black box" nature, and lowest

accuracies. JM is an attractive method if class separation distances are of interest. FSO also provided class separation distance, but they are not scaled for easy comparisons. Further studies will investigate the validity of these findings when the methods are applied to classifications of larger areas.

Acknowledgements

This research was funded by the USDA Agricultural Research Service and the National Science Foundation Long-Term Ecological Research Program, Jornada Basin V: Landscape Linkages in Arid and Semiarid Ecosystems, and the USDA Natural Resources Conservation Service in support of the Conservation Effects Assessment Project. The authors would like to thank Debra Peters and John Anderson for access to the LTER NPP sites, and Peg Gronemeyer, Amy Slaughter and Connie Maxwell for field data collection.

References

Addink, E.A., de Jong, S.M., Davis, S.A., Dubyanskiy, V., Burdelow, L.A., Leirs, H., 2010. The use of high-resolution remote sensing for plague surveillance in Kazakhstan. Remote Sens. Environ. 114, 674–681.

Blaschke, T., Lang, S., Hay, G.J. (Eds.), 2008. Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. Springer-Verlag, Berlin, p. 817.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. 65, 2–16.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth International Group, Belmont, CA.

Carleer, A.P., Wolff, E., 2006. Urban land cover multi-level region-based classification of VHR data by selecting relevant features. Int. J. Remote Sens. 27, 1035–1051.

Chubey, M.S., Franklin, S.E., Wulder, M.A., 2006. Object-based analysis of Ikonos-2 imagery for extraction of forest inventory parameters. Photogramm. Eng. Remote Sens. 72, 383–394.

Congalton, R.G., Green, K., 2009. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, 2nd ed. Taylor and Francis, Boca Raton, FL.

Coulter, L.L., Stow, D.A., 2008. Assessment of the spatial co-registration of multitemporal imagery from large format digital cameras in the context of detailed change detection. Sensors 8, 2161–2173.

Definiens, 2009. eCognition Developer 8.0 User Guide. Definiens AG, Munich, Germany.

Dragut, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. Int. J. Geogr. Inf. Sci. 24 (6), 859–871.

Erdas, 2009. Leica Photogrammetric Suite User Guide. Erdas Inc., Norcross, GA.

Ehlers, M., Gaehler, M., Janowsky, R., 2003. Automated analysis of ultra high resolution remote sensing data for biotope type mapping: new possibilities and challenges. ISPRS J. Photogramm. 57, 315–326.

Ehlers, M., Gaehler, M., Janowsky, R., 2006. Automated techniques for environmental monitoring and change analysis for ultra high resolution remote sensing data. Photogramm. Eng. Remote Sens. 72 (7), 835–844.

Esch, T., Thiel, M., Bock, M., Roth, A., Dech, S., 2008. Improvement of image segmentation accuracy based on multiscale optimization procedure. IEEE Geosci. Remote Sens. 5 (3), 463–467.

Ferro, C.J.S., Warner, T.A., 2002. Scale and texture in digital image classification. Photogramm. Eng. Remote Sens. 68, 51–63.

Foody, G.M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogramm. Eng. Remote Sens. 70, 627–633.

Gao, Y., Mas, J.F., Niemeyer, I., Marpu, P.R., Palacio, J.L., 2007. Object-based image analysis for mapping land-cover in a forest area. In: 5th International Symposium: Spatial Data Quality , Enschede, The Netherlands, 13–15 June, 2007, p. 4.

Gibbens, R.P., McNeely, R.P., Havstad, K.M., Beck, R.F., Nolen, B., 2005. Vegetation changes in the Jornada Basin from 1858 to 1998. J. Arid Environ. 61, 651–668.

Green, K., Lopez, C., 2007. Using object-oriented classification of ADS40 to map benthic habitats of the state of Texas. Photogramm. Eng. Remote Sens. 73, 861–865.

Hay, G., Castilla, G., Wulder, M., Ruiz, J.R., 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. Int. J. Appl. Earth Obs. 7, 339–359.

Herold, M., Liu, X., Clarke, K.C., 2003. Spatial metrics and image texture for mapping urban land use. Photogramm. Eng. Remote Sens. 69, 991–1001.

Jensen, J.R., 2005. Introductory Digital Image Processing: A Remote Sensing Perspective. Prentice-Hall, Inc., Upper Saddle River, NJ.

Johansen, K., Phinn, S., Witte, C., Philip, S., Newton, L., 2009. Mapping banana plantations from object-oriented classification of SPOT-5 imagery. Photogramm. Eng. Remote Sens. 75, 1069–1081.

Kim, B., Landgrebe, A., 1991. Hierarchical classifier design in high-dimensional numerous class cases. IEEE T. Geosci. Remote 29, 518–528.

Laliberte, A.S., Fredrickson, E.L., Rango, A., 2007. Combining decision trees with hierarchical object-oriented image analysis for mapping arid rangelands. Photogramm. Eng. Remote Sens. 73, 197–207.

Laliberte, A.S., Rango, A., 2009. Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery. IEEE T. Geosci. Remote 47, 761–770.

Laliberte, A.S., Browning, D.M., Herrick, J.E., Gronemeyer, P., 2010a. Hierarchical object-based classification of ultra high resolution digital mapping camera (DMC) imagery for rangeland mapping and assessment. J. Spat. Sci. 55 (1), 101–115.

Laliberte, A.S., Herrick, J.E., Rango, A., Winters, C., 2010b. Acquisition, orthorectification, and object-based classification of unmanned aerial vehicle (UAV) imagery for rangeland monitoring. Photogramm. Eng. Remote Sens. 76, 661–672.

Marpu, P.R., Niemeyer, I., Nussbaum, S., Gloaguen, R., 2008. A procedure for automatic object-based classification. In: Blaschke, T., Lang, S., Hay, G.J. (Eds.), Object-based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. Springer, Berlin, pp. 169–184.

Nussbaum, S., Niemeyer, I., Canty, M.J., 2006. SEATH – a new tool for automated feature extraction in the context of object-based image analysis. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVI-4/C42, 6.

Nusser, S.M., Goebel, J.J., 1997. The National Resources Inventory: a long term multi resource monitoring programme. Environ. Ecol. Stat. 4, 181–204.

Peters, D.C., Gibbens, R.P., 2006. Plant communities in the Jornada Basin: the dynamic landscape. In: Havstad, K.M., Huennecke, L.F., Schlesinger, W.H. (Eds.), Structure and Function of a Chihuahuan Desert Ecosystem. The Jornada Basin Long-Term Ecological Research Site. Oxford University Press, Oxford, England, pp. 211–231.

Rosso, P.H., Klonus, S., Ehlers, M., Tschach, E., 2008. Comparative properties of four airborne sensors and their applicability to land surface interpretation. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXVII (Part B1), 545–550.

Steinberg, D., Colla, P., 1997. CART – Classification and Regression Trees. Salford Systems, San Diego, CA.

Stow, D., Hamada, Y., Coulter, L., Anguelova, Z., 2008. Monitoring shrubland habitat changes through object-based change identification with airborne multispectral imagery. Remote Sens. Environ. 112, 1051–1061.

Swain, H., Davis, S.M., 1978. Remote Sensing: The Quantitative Approach. McGraw-Hill, New York, NY.

Van Coillie, F.M.B., Verbeke, L.P.C., De Wulf, R.R., 2007. Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. Remote Sens. Environ. 110, 476–487.

Wainwright, J., 2006. Climate and climatological variations in the Jornada Basin. In: Havstad, K.M., Huennecke, L.F., Schlesinger, W.H. (Eds.), Structure and Function of a Chihuahuan Desert Ecosystem. The Jornada Basin Long-Term Ecological Research Site. Oxford University Press, Oxford, England, pp. 44–80.

Wiechert, A., Gruber, M., 2010. DSM and true ortho generation with the UltraCam-L – a case study. In: Proceedings of the American Society for Photogrammetry and Remote Sensing Annual Conference , San Diego, CA, 26–30 April, 2010, p. 8.

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., Schirokauer, D., 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. Photogramm. Eng. Remote Sens. 72, 799–811.

Zhang, X., Feng, X., Jiang, H., 2010. Object-oriented method for urban vegetation mapping using IKONOS imagery. Int. J. Remote Sens. 31, 177–196.