# An Evaluation of an Object-Oriented Paradigm for Land Use/Land Cover Classification*

**Rutherford V. Platt and Lauren Rapoza**
*Gettysburg College*

Object-oriented image classification has tremendous potential to improve classification accuracies of land use and land cover (LULC), yet its benefits have only been minimally tested in peer-reviewed studies. We aim to quantify the benefits of an object-oriented method over a traditional pixel-based method for the mixed urban–suburban–agricultural landscape surrounding Gettysburg, Pennsylvania. To do so, we compared a traditional pixel-based classification using maximum likelihood to the object-oriented image classification paradigm embedded in eCognition Professional 4.0 software. This object-oriented paradigm has at least four components not typically used in pixel-based classification: (1) the segmentation procedure, (2) nearest neighbor classifier, (3) the integration of expert knowledge, and (4) feature space optimization. We evaluated each of these components individually to determine the source of any improvement in classification accuracy. We found that the combination of segmentation into image objects, the nearest neighbor classifier, and integration of expert knowledge yields substantially improved classification accuracy for the scene compared to a traditional pixel-based method. However, with the exception of feature space optimization, little or no improvement in classification accuracy is achieved by each of these strategies individually. **Key Words: image classification, land cover, land use, object-oriented.**

面向对象的图像分类法 对于提高土地利用和土地覆盖的分类准确度，备有巨大的潜能，但其好处较少在同行审评的研究被测试。我们的目的是要将面向对象的方法，用于围绕在宾州葛底斯堡城市-郊区-农业混合的景观，来量化其方法的优点。比起以像素为基础的传统方法，该方法确实有它优越之处。为了达到设定目标，我们比较了由极大似然分类器取得的传统像素为基础的分类和嵌入在易康eCognition专业4.0软件之中的面向对象的地物分类范式。其范式有至少四个组件是不经常使用在以像素为基础的分类上的：（1）分割程序，（2）就近邻分类，（3）整合的专业知识，以及（4）特征空间优化。我们评估了各别的组件以确定分类精度改善的来源。我们发现比起以像素为基础的传统方法，分割图像物体，近邻分类器以及整合专业知识的综合能提高场景的分类精度。然而，除了空间的优化，每一个上述策略很少能够单独地取得分类精度的改善。**关键词：图像分类，土地覆盖，土地利用，面向对象。**

La clasificación de imágenes orientadas a objetos tiene un enorme potencial para mejorar la precisión en la clasificación de los usos y coberturas del suelo (land use and land cover, LULC); sin embargo, sus beneficios solo se han probado mínimamente en estudios revisados por expertos en el campo. Nuestro objetivo es cuantificar los beneficios de un método orientado a objetos en comparación con un método tradicional basado en pixeles en la región mixta urbana-suburbana-agrícola que circunda a Gettysburg, Pensilvania. Para hacerlo, comparamos la clasificación tradicional basada en pixeles usando la máxima probabilidad con el paradigma de la clasificación por imágenes orientadas a objetos integrada en el programa eCognition Professional 4.0. Este paradigma orientado a objetos tiene al menos cuatro componentes que típicamente no se usan en la clasificación basada en pixeles: (1) el procedimiento de segmentación, (2) el clasificador por el vecino más cercano, (3) la integración de conocimiento experto, y (4) la optimización del espacio de trabajo. Evaluamos individualmente cada uno de estos componentes para determinar la fuente de cualquier mejora en la precisión de la clasificación. Encontramos que la combinación de la segmentación en objetos imágenes, el clasificador

*Thanks to Gettysburg College for supporting this research through a Research and Professional Development Grant.

The Professional Geographer, 60(1) 2008, pages 87–100 © Copyright 2008 by Association of American Geographers.
Initial submission, August 2006; revised submissions, May and June 2007; final acceptance, June 2007.
Published by Taylor & Francis, LLC.

por el vecino más cercano y la integración de conocimiento experto dan como resultado una precisión en la clasificación substancialmente mejorada para la escena cuando se compara con el método basado en pixeles. Sin embargo, con excepción de la optimización del espacio de trabajo, se logra poca o ninguna mejora en la precisión de la clasificación con cada una de estas estrategias implementadas de manera individual. **Palabras clave: clasificación de imágenes, cobertura del suelo, uso del suelo, orientado a objetos.**



Object-oriented image classification at its simplest level is the classification of homogeneous image primitives, or objects, rather than individual pixels. Object-oriented image classification has numerous potential advantages. If carefully derived, image objects are closely related to real-world objects. Once these objects are derived, topological relationships with other objects (e.g., adjacent to, contains, is contained by, etc.), statistical summaries of spectral and textural values, and shape characteristics can all be employed in the classification procedures (Benz et al. 2004). Despite many advances, object-oriented methods are still computationally intensive and the improvement in classification accuracy over traditional methods is not always clear.

The idea of object-based image analysis has been around since the early 1970s (de Kok, Schneider, and Ammer 1999), but implementation lagged due to lack of computing power. On a limited basis, specialized object-oriented software packages were employed in the 1980s to extract roads and other linear features (McKeown 1988; Quegan et al. 1988). These methods of analysis were difficult to employ and inefficient compared to pixel-based methods, which began to employ advanced techniques such as fuzzy sets, neural networks, and textural measurements. Since the 1990s, computing power has increased and high spatial resolution imagery has become common, prompting a new emphasis on object-oriented techniques (Franklin et al. 2003).

Recently, object-oriented image classification has been successfully used to identify logging and other forest management activities using Landsat ETM+ imagery (Flanders, Hall-Beyeer, and Pereverzoff 2003); to map shrub encroachment using QuickBird imagery (Laliberte et al. 2004); to quantify landscape structure using imagery from Landsat ETM+, QuickBird, and aerial photography (Ivits et al. 2005); to map fuel types using Landsat TM and Ikonos imagery (Giakoumakis, Gitas, and San-Miguel 2002); and to detect changes in land use from imagery in the German national topographic and cartographic database (Walter 2004).

Few direct comparisons of object-oriented and traditional pixel-based methods have been published, and these primarily appear in conference proceedings. One study found that object-oriented methods yield similar classification accuracy to traditional methods, but that segmentation of pixels into objects makes the classification more "map-like" by reducing the number of small disconnected patches (Willhauck 2000). Another study compared four change detection methods: traditional post-classification, cross-correlation analysis, neural networks, and an object-oriented method (Civco et al. 2002). The study found that there was no single best way to perform change analysis, and suggested that future object-oriented methods based on "multitemporal objects" could improve on current results. A third study compared pixel-based and object-oriented land use classification for a scene in the Black Sea region of Turkey and found that the object-oriented classification method had a substantial advantage over the pixel-based parallelepiped, minimum distance, and maximum-likelihood classifiers (Oruc, Marangoz, and Buyuksalih 2004). A comparison of land cover classification in northern Australia found that object-oriented classification yielded improved classification accuracy, 78 percent versus 69.1 percent (Whiteside and Ahmad 2005). Overall, the literature suggests either that object-oriented methods have a real advantage or that they demonstrate little advantage but much promise.

These studies all use various versions of the eCognition software (recently renamed Definiens), a specialized image classification software package that integrates hierarchical object-oriented image classification, fuzzy logic, and other strategies to improve classification accuracy. At the time of this research, eCognition Professional 4.0 was the most fully developed object-oriented classification software available. The object-oriented paradigm in eCognition has at least four components not typically used in traditional pixel-based classification methods: (1) the segmentation procedure, (2) nearest neighbor classifier, (3) the integration of expert knowledge, and (4) feature space optimization. We evaluate the object-oriented paradigm in eCognition for classification of an urban–suburban–agricultural landscape surrounding Gettysburg, Pennsylvania.

Our research questions are as follows: First, to what extent does the eCognition object-oriented paradigm increase land use/land cover (LULC) classification accuracy over a traditional pixel-based method for this scene? Second, how much of the increased accuracy, if any, is due to segmentation of image objects, the classifier, expert knowledge, or feature space optimization? This study is distinct from previous studies because it independently evaluates these four elements of the eCognition object-oriented paradigm to determine their effect on classification accuracy.

## Methods

### *Study Area and Data*

The study area is 148 km$^2$ in size, and located in a rural area northwest of the Baltimore and Washington, DC, metropolitan areas (Figure 1). It encompasses the borough of Gettysburg, the Gettysburg National Military Park, and surrounding areas. The area is located in the southwest corner of the Newark-Gettysburg basin, which is situated between the South Mountain (an extension of the Blue Ridge Mountains) to the west and the Susquehanna River to the east. The area comprises extensive agriculture divided by narrow wooded bands and, increasingly, low-density residential development. We acquired a georectified IKONOS satellite image of the study area taken 25 July 2003. The image has a spatial resolution of 4 m and has four spectral bands: blue (480.3 nm), green (550.7 nm), red (664.8 nm), and near infrared (805.0 nm).

### *Classification Models*

The image was classified into seven LULC classes: forest, fallow, water, recreational grasses, commercial/industrial/transportation, cultivated, and residential (Table 1). These classes represent the most common and important land uses and covers in the area. We did not use a more detailed classification scheme such as the one used for the National Land Cover data set (NLCD; Vogelmann et al. 2001) because many of the NLCD classes are missing from the area (e.g., ice/snow, shrublands), rare (e.g., high-density development), or not distinguishable using a single date image. Two to four image objects were selected as training samples for each class for use in the classification procedures described in the next section.

A total of eight image classifications were constructed to compare the object-oriented paradigm to a traditional classification (Table 2). Model 1 represents the best object-oriented model, model 2 represents the traditional pixel-based classification, and models 3 through 8 fall somewhere in between. We compared pairs of models that differed in one key respect; either in terms of analysis level (pixel or object), classifier (maximum likelihood or nearest neighbor), expert knowledge (used or not), or feature space optimization (used or not). These terms and the specific differences between models are discussed in the following sections.

### *Analysis Level: Pixel versus Object*

Traditional image classification methods classify individual pixels, whereas object-oriented classification methods classify homogeneous regions, or image objects. The process of aggregating pixels into image objects is known as image segmentation. For the models operating at the object level, the image used in this study was segmented using the fractal net evolution approach (FNEA), which is a multifractal approach implemented in the eCognition image processing software (Baatz and Schaepe 2000; Baatz et al. 2004). FNEA is a pairwise clustering process that finds areas of minimum spectral and spatial heterogeneity given a set of scale, color, and shape parameters (Benz et al. 2004).

The size of the image objects is determined by the scale parameter, a unitless number related to the image resolution that describes the maximum allowable heterogeneity of image objects. As the scale parameter increases, the size of the image objects also increases (Benz et al. 2004). The color and shape parameters are weights between zero and one that determine the contribution of spectral heterogeneity (in this case red, green, blue, and near infrared)

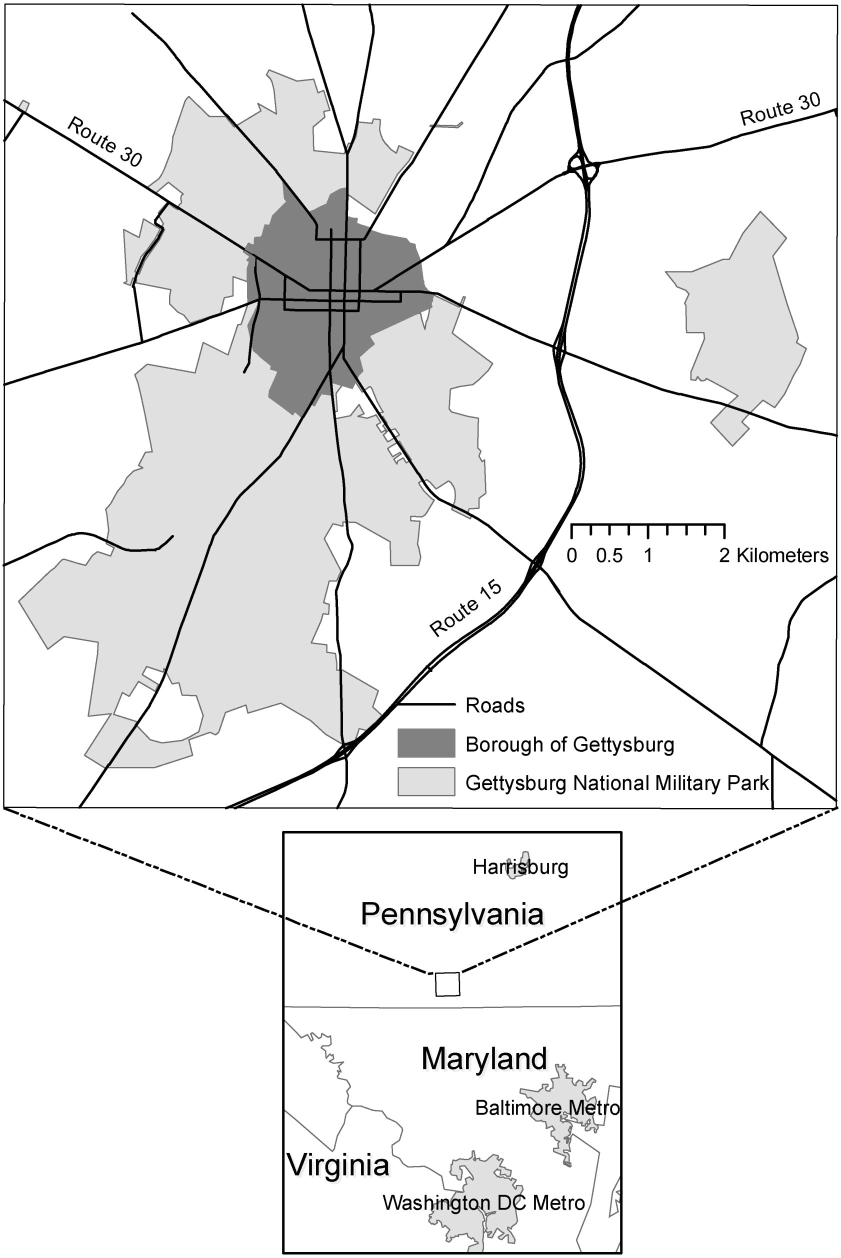


**Figure 1** *Study area.*

and shape to the overall heterogeneity that is to be minimized. The smoothness and compactness parameters are additional weights between zero and one that determine how shape is calculated.

Spectral heterogeneity is defined as the sum of standard deviations of each image band. Minimizing only spectral heterogeneity results in objects that are spectrally similar, but that might have fractally shaped borders or many

**Table 1** *Land use/land cover classes*

| Class | Description |
|---|---|
| Forest | Closed canopy forest |
| Fallow | Field with little or no vegetation |
| Water | Stream, lake, or pond |
| Recreational grasses | Mowed grass in a suburban or urban context, might have scattered trees |
| Commercial/industrial/transportation | Parking lots, industrial sites, strip malls, and associated infrastructure |
| Cultivated | Cropland, pasture, and unmowed grasses |
| Residential | Single or multifamily housing |

branched segments (Baatz et al. 2004). To address this issue, the segmentation process can also incorporate shape in terms of compactness or smoothness. Compactness is defined as the ratio of the border length and the square root of the number of object pixels. Smoothness is defined as the ratio of the border length and the shortest possible border length derived from the bounding box of an image object (Baatz et al. 2004).

The "best" compactness and smoothness parameters depend on the size and types of objects to be extracted. For example, an object representing an agricultural field would ideally have high smoothness and compactness, whereas an object representing a riparian area along a stream would ideally have low smoothness and compactness.

It is important to note that there is no such thing as optimal parameters for image segmentation (Benz et al. 2004). For object-level models, the normal procedure is simply to iteratively try different parameters until the resulting objects are appropriately sized and shaped for the particular task. After testing many possible parameters, we used the following: scale: 50, color: 0.7, shape: 0.3, smoothness: 0.5, compactness: 0.5. The resulting objects closely corresponded to the boundaries of fields, woodlots, strip malls, and other elements of interest in the image (Figure 2). For pixel-level models, we segmented the image into single-pixel objects.

### *Image Classifier*

The models are divided according to what classifier they use: maximum likelihood or nearest neighbor. The maximum likelihood classifier calculates the probability that a pixel or object belongs to each class and then assigns the pixel or object to the class with the highest probability (Richards 1999). It is one of the most commonly used classifiers because of its simplicity and robustness (Platt and Goetz 2004). The Environment for Visualizing Images (ENVI) image processing software was used for maximum likelihood classification. The nearest neighbor classifier is a part of the eCognition object-oriented paradigm and assigns each object to the class closest to it in feature space.

### *Expert Knowledge*

The models are also divided according to whether or not they employ expert knowledge (i.e., user-developed classification rules) in addition to the classifier (nearest neighbor classifier). Traditional classification methods typically do not employ expert knowledge, whereas the eCognition object-oriented

**Table 2** *Summary of classification models*

| Model | Analysis level | Classifier | Expert knowledge | Feature space |
|---|---|---|---|---|
| 1 | Object | Nearest neighbor | Yes | Spectral |
| 2 | Pixel | Maximum likelihood | No | Spectral |
| 3 | Object | Nearest neighbor | No | Spectral |
| 4 | Pixel | Nearest neighbor | No | Spectral |
| 5 | Object | Maximum likelihood | No | Spectral |
| 6 | Pixel | Nearest neighbor | Yes | Spectral |
| 7 | Object | Nearest neighbor | No | Optimized |
| 8 | Pixel | Nearest neighbor | No | Optimized |

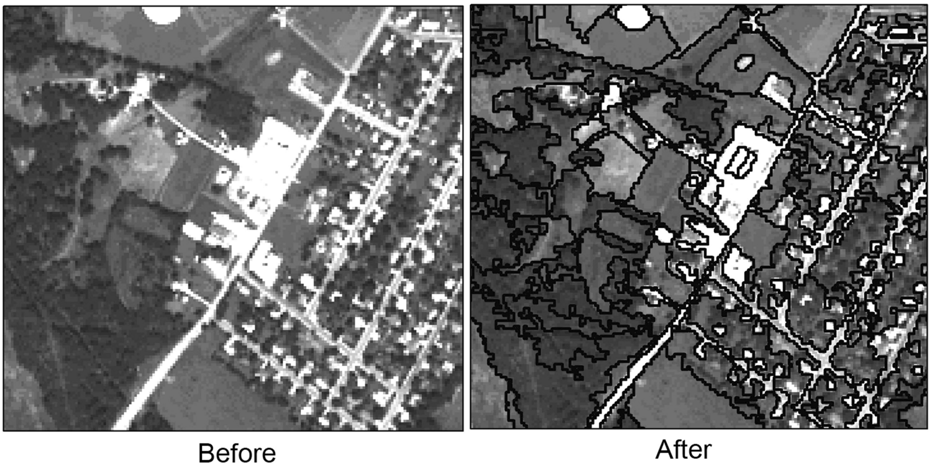


**Figure 2** *Before and after segmentation.*

paradigm integrates membership functions with the nearest neighbor classifier. Membership functions allow objects to be classified using simple rules related to spectral, shape, and textural characteristics or on relationships between neighboring objects. For example, we might observe that recreational grasses are typically surrounded by developed areas like commercial/industrial/transportation or residential. Knowing this, a membership function can be constructed that assigns each object a value between zero and one depending on the percentage of that object that is surrounded by an object defined as developed (Figure 3). For each object, the lowest of the two probabilities (the membership function probability and the nearest neighbor probability) is selected as the actual probability of belonging to a class. The closer this number is to one, the more likely the object will be classified as recreational grasses. For models that use expert knowledge, membership functions were defined for every class except for water (Table 3).

### *Feature Space Optimization*

Finally, models were divided into those that use feature space optimization and those that do not. Typically classification is conducted using the spectral or textural bands chosen by the analyst. In the feature space optimization procedure, the user first specifies which feature space bands should be included in the analysis. For objects, the feature space bands might number in the hundreds and include spectral data, hierarchical data, shape data, and textural data. For pixels, the feature space is considerably smaller and is limited to spectral and textural data. Feature space optimization then calculates all the bands and sorts them based on how well they separate the classes.

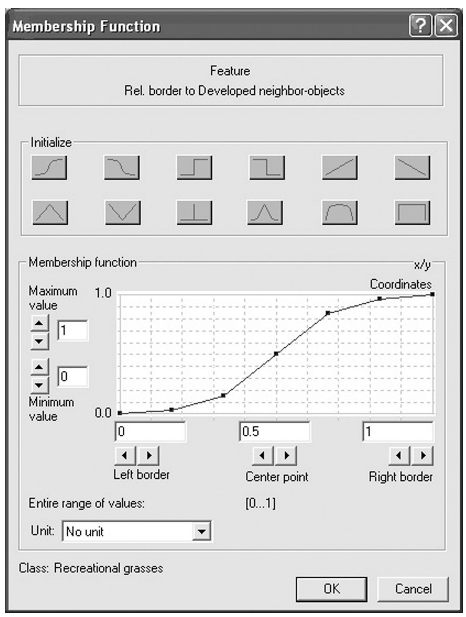


**Figure 3** *Example of a membership function.*

***Table 3*** *Membership functions*

| Class | Factors promoting membership | Factors limiting membership |
|---|---|---|
| Commercial/industrial/transportation | — | Relative border to residential |
| Recreational grasses | Relative border to residential or commercial/industrial/ transportation | Relative border to residential or commercial/industrial/transportation |
| Residential | Adjacent to recreational grasses, low homogeneity | Relative border to commer- cial/industrial/transportation, cultivated, or fallow |
| Cultivated | Rectangular fit | Relative border to residential |
| Fallow | Rectangular fit | Relative border to residential |
| Forested | — | Relative border to residential |
| Water | — | — |

The thirty feature space bands that best separated the LULC classes were used to classify the image for models that used feature space optimization. The bands beyond these thirty did little to help separate classes and so were not used in the classification. Textural data were computed following Haralick, Shanmugam, and Dinstein (1973). The grey level co-occurrence matrix describes how different combinations of pixel values occur within an object (Baatz et al. 2004). The grey level difference vector, another way to measure texture, is the sum of the diagonals of the grey level co-occurrence matrix.

At the object level (model 7), the following six bands best separated the classes according to the feature space optimization:

- Grey level co-occurrence matrix correlation of band 3 (red). This measures the correlation between the values of neighboring pixels in the red band.
- Shape index: perimeter divided by four times the square root of the area of an object.
- Degree of skeleton branching: describes the highest order of branching of skeletons, defined as lines that connect the midpoints of a Delaunay triangulation of an object. High values indicate a complex geometrical structure of the object (Benz et al. 2004).
- Maximum pixel value of band 4 (near infrared).
- Grey level difference vector entropy of band 3 (red). This measures whether pixels have similar brightness levels in the red band.
- Grey level co-occurrence matrix entropy of band 1 (blue). This measures whether pixels have similar brightness levels in the blue band.

At the object level (model 8), the following six bands best separated the classes according to the feature space optimization:

- Grey level co-occurrence matrix mean of band 1 (blue). This measures the mean frequency of pixel values in combination with neighboring pixel values in the blue band.
- Grey level co-occurrence matrix contrast of band 4 (near infrared). This measures the amount of variation in the near infrared band within an object.
- Grey level co-occurrence matrix variance of band 4 (near infrared). Similar to grey level co-occurrence matrix contrast, this measures the dispersion of variation in the near infrared band within an object.
- Grey level co-occurrence matrix mean of band 4 (near infrared). This measures the mean frequency of pixel values in combination with neighboring pixel values in the near infrared band.
- Mean difference to neighbor, band 4 (near infrared). This measures the mean difference between a pixel value and its neighbor in the near infrared band.
- Grey level co-occurrence matrix homogeneity of band 1 (blue). This measures degree that the object displays a lack of variation in the blue band.

One possible issue with feature space optimization is overfitting—fitting a statistical model with too many parameters such that the

***Table 4*** *Confusion matrix, model 1*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Object | Nearest neighbor | Yes | Spectral | 72% | 71% | 78% | 13% | 9% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 94 | 0 | 4 | 1 | 0 | 8 | 0 | 107 | 88% |
| Cultivated | 3 | 101 | 2 | 18 | 12 | 3 | 0 | 139 | 73% |
| Fallow | 14 | 1 | 14 | 2 | 0 | 7 | 0 | 38 | 37% |
| Forested | 0 | 4 | 1 | 85 | 6 | 5 | 0 | 101 | 84% |
| Recreational grass | 3 | 18 | 0 | 7 | 54 | 17 | 0 | 99 | 55% |
| Residential | 8 | 0 | 2 | 1 | 0 | 23 | 0 | 34 | 68% |
| Water | 1 | 0 | 0 | 0 | 0 | 2 | 28 | 31 | 90% |
| Total | 123 | 124 | 23 | 114 | 72 | 65 | 28 | 549 | 71% |
| User's acc. | 76% | 81% | 61% | 75% | 75% | 35% | 100% | 72% | |

*Note: CIT = Commercial/industrial/transportation.*

model fits the training data much better than the validation data. To minimize overfitting, we excluded bands with nonnormalized units of length or area; objects of any size can belong to any class.

### *Classification Evaluation*

To evaluate classification accuracy of the eight models, a random sample of 300 points was generated across the image in areas that are not training sites. This random sample was used to determine the relative proportion of each class within the image. Following Congalton and Green (1999), the random sample was supplemented by a stratified random sample of 250 points. This ensured that each class in the classified image contained at least thirty validation points. All 550 validation points were overlain on the original imagery and on 0.6 m georectified aerial photography taken in spring 2003. An image analyst identified the LULC class at the object level and the pixel level, always deferring to the IKONOS imagery for classes that might have changed between the times the two images were taken. A second image analyst revisited any points that were tagged as “questionable” by the first analyst. If the two analysts disagreed, the point was visited in the field by the first analyst to make a final determination of the class. A total of thirty-five points were visited in the field.

Using all 550 points, a confusion matrix was generated to compare the predicted LULC classes to the actual LULC classes. The diagonal of the matrix shows the number of objects or pixels where the predicted class is the same as the actual class, whereas the off-diagonal values show the number of objects or pixels where the predicted class is different from the actual class. For each model, user's accuracy (probability that a site in the classified image actually represents that class on the ground) and producer's accuracies (probability that a site on the ground was classified correctly) were calculated for each class. The average user's accuracy, average producer's accuracy, and weighted agreement (percent correct weighted by the actual occurrence in the landscape, as estimated by the 300-point random sample) were also reported.

Note that our accuracy reporting does not include Kappa, which attempts to correct for chance agreement, because this measure is difficult to interpret, has an arbitrary definition of chance agreement, and conflates different sources of error (Pontius 2000). Instead, we reported weighted disagreement due to location (percentage of objects or pixels incorrectly classified because the predicted location was incorrect) and weighted disagreement due to quantity (percentage of objects or pixels incorrectly classified because the predicted quantities of classes were incorrect). Both percentages are weighted by prevalence of the classes in the image.

## Results and Discussion

Representing the eCognition object-oriented paradigm, model 1 (Table 4) operates on the object level, uses the nearest neighbor classifier, incorporates expert knowledge, and uses only the spectral feature space. The weighted agreement (percent correct weighted by prevalence of the class) was 78 percent, the highest of all the models (Table 4). The user's accuracy was over 70 percent for all classes except fallow and residential. Because residential is a rare class that is spectrally similar to many other classes, areas known to be residential were only 68 percent likely to be classified correctly. Producer's accuracy was over 70 percent for all classes except fallow, residential, and recreational grasses. Sites known to be fallow were misclassified 37 percent of the time because of the spectral similarity to commercial/industrial/transportation. The classes of commercial/industrial/transportation, cultivated, forested, and water were over 70 percent accurate from both the user's and producer's perspective. The weighted disagreement due to location is 13 percent and the weighted disagreement due to quantity is 9 percent, showing that neither type of error predominates.

Representing pixel-based image classification, model 2 operates on the pixel level, uses the maximum likelihood classifier, does not incorporate expert knowledge, and uses spectral feature space (Table 5). The weighted agreement is 64 percent, which is 14 percent lower than model 1. Model 2 has a lower classification accuracy than model 1 for every class except fallow, which has a higher user's and producer's accuracy than model 1. Whereas weighted disagreement due to quantity is similar between model 1 and 2 (9 percent versus 12 percent), the weighted disagreement due to location is clearly superior in model 1 compared to model 2 (13 percent versus 24 percent).

### Analysis Level: Pixel versus Object

To evaluate the effect of analysis level we compared model 3 (Table 6) to model 4 (Table 7). Both models use the nearest neighbor classifier, omit expert knowledge, and use spectral feature space, but model 3 operates at the object level whereas model 4 operates at the pixel level. The weighted agreement is similar for both models—60 percent for model 3 and 61 percent for model 4. Weighted disagreement due to location is identical (17 percent) and similar for weighted disagreement due to quantity (23 percent versus 22 percent). The accuracy within classes varies, however. These results suggest that classifying objects rather than pixels does not necessarily improve classification accuracy.

### Classifier

To evaluate the effect of the classifier we compared model 3 (Table 6) to model 5 (Table 8). Both models operate at the object level, omit expert knowledge, and use spectral feature space. They differ only in terms of the classifier: model 3 uses the nearest neighbor classifier, whereas model 5 uses the maximum likelihood classifier. Model 3 has a lower weighted agreement than model 5 (60 percent versus 71 percent). Model 3 also has a similar weighted disagreement due to location (17 percent versus 18 percent) and a higher weighted disagreement due to quantity (23 percent versus 11 percent). This suggests that maximum likelihood does a better job in predicting the quantity of classes, but is comparable to nearest

**Table 5** *Confusion matrix, model 2*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 2 | Pixel | Maximum likelihood | No | Spectral | 63% | 53% | 64% | 24% | 12% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 77 | 1 | 2 | 0 | 1 | 26 | 0 | 107 | 72% |
| Cultivated | 0 | 64 | 8 | 0 | 51 | 16 | 0 | 139 | 46% |
| Fallow | 1 | 0 | 30 | 0 | 0 | 7 | 0 | 38 | 79% |
| Forested | 0 | 10 | 2 | 53 | 26 | 10 | 0 | 101 | 52% |
| Recreational grass | 1 | 26 | 3 | 7 | 42 | 20 | 0 | 99 | 42% |
| Residential | 16 | 0 | 2 | 0 | 2 | 14 | 0 | 34 | 41% |
| Water | 7 | 0 | 0 | 0 | 0 | 12 | 12 | 31 | 39% |
| Total | 102 | 101 | 47 | 60 | 122 | 105 | 12 | 549 | 53% |
| User's acc. | 75% | 63% | 64% | 88% | 34% | 13% | 100% | 63% | |

*Note: CIT = Commercial/industrial/transportation.*

**Table 6** *Confusion matrix, model 3*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 3 | Object | Nearest neighbor | No | Spectral | 62% | 61% | 60% | 17% | 23% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 62 | 0 | 24 | 0 | 0 | 21 | 0 | 107 | 58% |
| Cultivated | 0 | 73 | 3 | 1 | 47 | 15 | 0 | 139 | 53% |
| Fallow | 1 | 2 | 32 | 0 | 0 | 3 | 0 | 38 | 84% |
| Forested | 0 | 10 | 2 | 68 | 18 | 3 | 0 | 101 | 67% |
| Recreational grass | 1 | 51 | 7 | 2 | 20 | 18 | 0 | 99 | 20% |
| Residential | 3 | 1 | 11 | 0 | 0 | 19 | 0 | 34 | 56% |
| Water | 2 | 0 | 1 | 0 | 0 | 0 | 28 | 31 | 90% |
| Total | 66 | 137 | 80 | 71 | 85 | 79 | 28 | 549 | 61% |
| User's acc. | 94% | 53% | 40% | 96% | 24% | 24% | 100% | 62% | |

*Note: CIT = Commercial/industrial/transportation.*

neighbor in predicting the location. However, because several less common classes (water and residential) are poorly classified in model 5, the average producer's accuracy for this model (58 percent) is lower than for model 3 (61 percent).

### *Expert Knowledge*

To evaluate the effect of expert knowledge integrated into membership functions, we compared model 1 (Table 4) to model 3 (Table 6) and model 6 (Table 9) to model 4 (Table 7). Within each pair, the models differ only in terms of whether they use membership functions in addition to the nearest neighbor classifier (models 1 and 6 do, models 3 and 4 do not). The first pair of models operates at the object level and the second set operates at the pixel level.

At the object level, membership functions improve classification accuracy considerably. Model 1 is higher than model 3 in terms of weighted agreement (78 percent versus 60 percent), average user's accuracy (72 percent versus 62 percent), and average producer's accuracy (71 percent versus 61 percent). In particular, cultivated, residential, and recreational grasses are improved from both the user's and producer's perspective when membership functions are used. The classification accuracy of water remains the same because membership functions are not used in either model. Model 1 is superior to model 3 in terms of weighted agreement due to location (13 percent versus 17 percent) and quantity (9 percent versus 23 percent). This shows that membership rules especially help predict the quantity of objects in each class.

At the pixel level, there is no improvement in classification accuracy when membership functions are used. Models 4 and 6 are similar in weighted agreement (61 percent versus 60 percent). Model 6 has a higher average user's

**Table 7** *Confusion matrix, model 4*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 4 | Pixel | Nearest neighbor | No | Spectral | 60% | 62% | 61% | 17% | 22% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 71 | 3 | 11 | 1 | 0 | 19 | 2 | 107 | 66% |
| Cultivated | 0 | 80 | 10 | 2 | 37 | 10 | 0 | 139 | 58% |
| Fallow | 1 | 3 | 31 | 0 | 0 | 3 | 0 | 38 | 82% |
| Forested | 0 | 17 | 5 | 69 | 9 | 0 | 1 | 101 | 68% |
| Recreational grass | 1 | 42 | 8 | 10 | 28 | 10 | 0 | 99 | 28% |
| Residential | 10 | 0 | 9 | 1 | 0 | 14 | 0 | 34 | 41% |
| Water | 1 | 0 | 1 | 0 | 0 | 0 | 29 | 31 | 94% |
| Total | 84 | 145 | 75 | 83 | 74 | 56 | 32 | 549 | 62% |
| User's acc. | 85% | 55% | 41% | 83% | 38% | 25% | 91% | 60% | |

*Note: CIT = Commercial/industrial/transportation.*

***Table 8*** *Confusion matrix, model 5*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 5 | Object | Maximum likelihood | No | Spectral | 69% | 58% | 71% | 18% | 11% |
| | **CIT** | **Cultivated** | **Fallow** | **Forested** | **Recreational grass** | **Residential** | **Water** | **Total** | **Prod. acc.** |
| CIT | 84 | 0 | 1 | 0 | 0 | 22 | 0 | 107 | 79% |
| Cultivated | 0 | 79 | 10 | 0 | 41 | 9 | 0 | 139 | 57% |
| Fallow | 1 | 0 | 31 | 0 | 0 | 6 | 0 | 38 | 82% |
| Forested | 0 | 6 | 0 | 61 | 32 | 2 | 0 | 101 | 60% |
| Recreational grass | 2 | 25 | 1 | 2 | 52 | 17 | 0 | 99 | 53% |
| Residential | 18 | 1 | 2 | 0 | 0 | 13 | 0 | 34 | 38% |
| Water | 10 | 0 | 0 | 0 | 0 | 10 | 11 | 31 | 35% |
| Total | 115 | 111 | 45 | 63 | 79 | 125 | 11 | 549 | 58% |
| User's acc. | 73% | 71% | 69% | 97% | 66% | 10% | 100% | 69% | |

*Note: CIT = Commercial/industrial/transportation.*

accuracy than model 4 (73 percent versus 60 percent), but a lower average user's accuracy (60 percent versus 62 percent). An important difference between the two models is the weighted disagreement due to location and quantity. For model 6, the weighted disagreement due to location is 7 percent, whereas the weighted disagreement due to quantity is 33 percent. Classes like residential and recreational grasses are severely underpredicted by model 6. In contrast, for model 4, the weighted disagreement due to location is 17 percent, whereas the weighted disagreement due to quantity is 22 percent. Compared to model 6, model 4 does a better job predicting the number of pixels in each class, but a worse job predicting the locations of these pixels. Overall the results show that these membership functions do not improve classification at the pixel level.

### *Feature Space Optimization*

To evaluate the effect of feature space optimization, we compared model 3 (Table 6) to model 7 (Table 10) and model 4 (Table 7) to model 8 (Table 11). The first pair operates on the object level, whereas the second pair operates on the pixel level. Models 3 and 7 use spectral feature space only, whereas models 7 and 8 use feature space optimization. Feature space optimization benefits classification accuracy at both the object and pixel level. At the object level, model 7 has a weighted agreement of 71 percent, whereas model 3 has a weighted agreement of 60 percent. At the pixel level, model 8 has a weighted agreement of 67 percent, whereas model 4 has a weighted agreement of 61 percent. In neither case do the models outperform model 1, which uses expert knowledge but not feature space optimization.

***Table 9*** *Confusion matrix, model 6*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 6 | Pixel | Nearest neighbor | Yes | Spectral | 73% | 60% | 60% | 7% | 33% |
| | **CIT** | **Cultivated** | **Fallow** | **Forested** | **Recreational grass** | **Residential** | **Water** | **Total** | **Prod. acc.** |
| CIT | 88 | 4 | 12 | 1 | 0 | 0 | 2 | 107 | 82% |
| Cultivated | 0 | 125 | 11 | 3 | 0 | 0 | 0 | 139 | 90% |
| Fallow | 3 | 3 | 32 | 0 | 0 | 0 | 0 | 38 | 84% |
| Forested | 0 | 24 | 5 | 71 | 0 | 0 | 1 | 101 | 70% |
| Recreational grass | 1 | 73 | 12 | 12 | 1 | 0 | 0 | 99 | 1% |
| Residential | 20 | 0 | 13 | 1 | 0 | 0 | 0 | 34 | 0% |
| Water | 1 | 0 | 1 | 0 | 0 | 0 | 29 | 31 | 94% |
| Total | 113 | 229 | 86 | 88 | 1 | 0 | 32 | 549 | 60% |
| User's acc. | 78% | 55% | 37% | 81% | 100% | 0% | 91% | 63% | |

*Note: CIT = Commercial/industrial/transportation.*

***Table 10*** *Confusion matrix, model 7*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 7 | Object | Nearest neighbor | No | Optimized | 57% | 56% | 71% | 20% | 9% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 31 | 0 | 34 | 0 | 0 | 40 | 2 | 107 | 29% |
| Cultivated | 0 | 72 | 24 | 8 | 22 | 10 | 3 | 139 | 52% |
| Fallow | 1 | 0 | 33 | 0 | 0 | 4 | 0 | 38 | 87% |
| Forested | 0 | 11 | 5 | 47 | 19 | 19 | 0 | 101 | 47% |
| Recreational grass | 2 | 6 | 22 | 0 | 23 | 43 | 3 | 99 | 23% |
| Residential | 3 | 0 | 6 | 0 | 0 | 24 | 1 | 34 | 71% |
| Water | 2 | 0 | 1 | 0 | 0 | 1 | 27 | 31 | 87% |
| Total | 39 | 89 | 125 | 55 | 64 | 141 | 36 | 549 | 56% |
| User's acc. | 79% | 81% | 26% | 85% | 36% | 17% | 75% | 57% | |

*Note: CIT = Commercial/industrial/transportation.*

### *Limitations*

This analysis has several limitations. First, the results should not automatically be extended to other contexts; they are partially a function of the spatial and spectral resolution of the image, the composition of the scene, the classification system, and the segmentation parameters. However, the results are consistent with the growing body of literature that shows an advantage for object-oriented classification methods (Willhauck 2000; Oruc, Marangoz, and Buyuksalih 2004; Whiteside and Ahmad 2005). Second, although we tested several aspects of the object-oriented image classification paradigm in eCognition, we cannot be sure that we maximized the software's potential. For example, the segmentation parameters and membership functions we used are not necessarily optimal.

A third limitation is that, despite the simple classification system, many of the classes might be difficult to distinguish spectrally. Part of the potential difficulty in separating these LULC classes is due to the spatial and spectral resolution of the imagery. High spatial resolution imagery, like the imagery used in this study, can be difficult to classify due to the high spectral heterogeneity within classes (Woodcock and Strahler 1987; Donnay 1999; Laliberte et al. 2004). Furthermore, multispectral IKONOS imagery contains only four spectral bands and thus has poorer spectral resolution than many other common multispectral sensors such as Landsat 7.

## Conclusions

This study started with two research questions: (1) to what extent object-oriented image classification increases LULC classification accuracy over a traditional pixel-based method

***Table 11*** *Confusion matrix, model 8*

| Model | Analysis level | Classifier | Expert knowledge | Feature space | Average user's acc. | Average prod. acc. | Weighted agreement | Weighted disagree location | Weighted disagree quantity |
|---|---|---|---|---|---|---|---|---|---|
| 8 | Pixel | Nearest neighbor | No | Optimized | 59% | 63% | 67% | 22% | 11% |

| | CIT | Cultivated | Fallow | Forested | Recreational grass | Residential | Water | Total | Prod. acc. |
|---|---|---|---|---|---|---|---|---|---|
| CIT | 65 | 0 | 7 | 0 | 0 | 35 | 0 | 107 | 61% |
| Cultivated | 0 | 71 | 16 | 4 | 40 | 8 | 0 | 139 | 51% |
| Fallow | 1 | 3 | 31 | 0 | 1 | 2 | 0 | 38 | 82% |
| Forested | 0 | 15 | 4 | 61 | 8 | 13 | 0 | 101 | 60% |
| Recreational grass | 1 | 18 | 8 | 6 | 26 | 40 | 0 | 99 | 26% |
| Residential | 5 | 0 | 3 | 0 | 0 | 26 | 0 | 34 | 76% |
| Water | 1 | 0 | 0 | 0 | 0 | 4 | 26 | 31 | 84% |
| Total | 73 | 145 | 75 | 83 | 74 | 56 | 32 | 549 | 63% |
| User's acc. | 89% | 49% | 41% | 73% | 35% | 46% | 81% | 59% | |

*Note: CIT = Commercial/industrial/transportation.*

for this scene, and (2) how much of the increased accuracy, if any, is due to segmentation, the classifier, expert knowledge, or feature space optimization. The answer to the first question is clear: we found that the object-oriented paradigm implemented in eCognition yields a considerable improvement in classification accuracy over a traditional method for this scene. Weighted agreement of the best object-oriented method (model 1) was 78 percent, which is 14 percent higher than the model representing the traditional pixel-based classification (model 2). That said, the object-oriented paradigm is no magic bullet: when classes overlap spectrally as in this study, high classification accuracy is still difficult to achieve.

The answer to the second question is that the combination of the object analysis level, the nearest neighbor classifier, and expert knowledge yields the highest classification accuracy. Therefore, for the most part, we cannot attribute the advantage of the object-oriented paradigm to any one of these methods in isolation. For example, we found that the classification of image objects rather than pixels by itself yields no increase in weighted agreement (~60 percent weighted agreement in each case for models 3 and 4). Furthermore, using the nearest neighbor classifier rather than the maximum likelihood classifier actually decreases weighted agreement (60 percent for model 3, 71 percent for model 5) and does a worse job predicting the quantity of classes (weighted disagreement of 23 percent for model 3 and 14 percent for model 5).

When paired with a nearest neighbor classifier, membership functions yielded a large improvement in classification accuracy at the object level (weighted agreement of 78 percent for model 1 compared to 60 percent for model 3). At the pixel level, there is no improvement in classification when membership functions are used (weighted agreement of ~60 percent for both models 4 and 6). It is important to underscore membership functions are not exclusive to object-oriented classification; they can be implemented for pixel-based classification as well. In this study, most of the membership functions are related to relative border and adjacency, which can be applied at both the pixel and object level. However, membership functions are more broadly applicable at the object level because certain variables commonly used in membership functions, such as shape, are only meaningful for objects. Feature space optimization led to an increase in classification accuracy at both the object and pixel level, but feature space optimization models still performed poorer than model 1, which integrated expert knowledge.

The study shows that, for this image and classification system, the object-oriented paradigm improves classification accuracy considerably. Unlike previous studies, we investigated the source of this advantage and found that much of the benefit is derived from the ability to integrate expert knowledge through membership functions, which is only effective at the object level. To simplify the comparison with traditional image classification methods, we used a flat classification scheme and a single level of image objects. A next step in the study will be to develop a nested hierarchy of image objects and classify land cover at each level using information from subobjects (e.g., trees, grass) and superobjects (e.g., city park, forest). This would reduce direct comparability with pixel-based methods, but allow us to take full advantage of multiresolution capabilities of eCognition and potentially further boost classification accuracy. ■

## Literature Cited


Baatz, M., U. Benz, S. Dehghani, M. Heynen, A. Höltje, P. Hofmann, I. Lingenfelder, et al. 2004. *eCognition professional user guide, version 4.0. Definiens Imaging GmbH*. München, Germany: Definiens.

Baatz, M., and A. Schaepe. 2000. Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII* [*Applied Geographic Information Processing*], ed. J. Strobl, T. Blaschke, and G. Griesebner, 12–23. Heidelberg, Germany: Wichmann-Verlag.

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heyen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 239–58.

Civco, D. L., J. D. Hurd, E. H. Wilson, M. Song, and Z. Zhang. 2002. A comparison of land use and land cover change detection methods. Paper presented at the annual conference of the ASPRS-ACSM, Washington, DC.

Congalton, R., and K. Green. 1999. Assessing the accuracy of remotely sensed data: Principles and practices. Boca Raton, FL: CRC/Lewis Press.

De Kok, R., T. Schneider, and U. Ammer. 1999. Object based classification and applications in the Alpine forest environment. In *Fusion of sensor data, knowledge sources and algorithms: Proceedings of the joint ISPRS/EARSel Workship, 3-4 June 1999, Valladolid, Spain. International Archives of Photogrammetry and Remote Sensing* 32:7-4-3 W6.

Donnay, J. P. 1999. Use of remote sensing information in planning. In *Geographical information and planning*, ed. J. Stillwell, S. Geertman, and & S. Openshaw, 242–60. Berlin: Springer-Verlag.

eCognition. *eCognition Professional*, version 4.0, Munich, Germany: Definiens.

ENVI, version 4.1. Boulder, CO: ITT Visual Information Solutions.

Flanders, D., M. Hall-Beyeer, and J. Pereverzoff. 2003. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. *Canadian Journal of Remote Sensing* 29 (4): 441–52.

Franklin, J., S. R. Phinn, C. E. Woodcock, and J. Rogan. 2003. Rationale and conceptual framework for classification approaches to assess forest resources and properties. In *Methods and applications for remote sensing of forests: Concepts and case studies*, ed. M. Wulder and S. E. Franklin, 279–300. New York: Kluwer.

Giakoumakis, M. N., I. Z. Gitas, and J. San-Miguel. 2002. Object-oriented classification modeling for fuel type mapping in the Mediterranean, using LANDSAT TM and IKONOS imagery-preliminary results. In *Forest fire research & wildfire safety*, ed. X. Viegas, 1–13. Rotterdam, Netherlands: Millpress.

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. Textural features for image classification. *IEEE TransActions on Systems, Man, and Cybernetics* 3:610–21.

Ivits, E., B. Koch, T. Blascheke, M. Jochum, and P. Adler. 2005. Landscape structure assessment with image grey-values and object-based classification at three spatial resolutions. *International Journal of Remote Sensing* 26 (14): 2975–93.

Laliberte, A. S., A. Rango, K. M. Havstad, J. F. Paris, R. F. Beck, R. McNeedly, and A. L. Gonzalez. 2004. Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. *Remote Sensing of Environment* 93:198–210.

McKeown, D. 1988. Building knowledge-based systems for detecting man-made structure from remotely sensed imagery. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical and Physical Sciences* 324 (1579): 423–35.

Oruc, M., A. M. Marangoz, and G. Buyuksalih. 2004. Comparison of pixel-based and object-oriented classification approaches using Landsat-7 ETM spectral bands. Paper presented at the conference of the ISPRS, Istanbul, Turkey.

Platt, R. V., and A. F. H. Goetz. 2004. A comparison of AVIRIS and synthetic Landsat data for land use classification at the urban fringe. *Photogrammetric Engineering and Remote Sensing* 70:813–19.

Pontius, R. G., Jr. 2000. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing* 66:1011–16.

Quegan, S., A. Rye, A. Hendry, J. Skingley, and C. Oddy. 1988. Automatic interpretation strategies for synthetic aperture radar images. *Philosophical Transactions of the Royal Society of London, Series A: Mathematical and Physical Sciences* 324 (1579): 409–20.

Richards, J. A. 1999. *Remote sensing digital image analysis*. New York: Springer Verlag.

Vogelmann, J. E., S. M. Howard, L. Yang, C. R. Larson., B. K. Wylie, and J. N. Van Driel. 2001. Completion of the 1990's National Land Cover Data Set for the conterminous United States. *Photogrammetric Engineering and Remote Sensing* 67:650–62.

Walter, V. 2004. Object-based classification of remote sensing data for change detection. *ISPRS Journal of Photogrammetry and Remote Sensing* 58:225–38.

Whiteside, T., and W. Ahmad. 2005. A comparison of object-oriented and pixel-based classification methods for mapping land cover in northern Australia. In *Proceedings of SSC2005 Spatial intelligence, innovation and praxis: The national biennial conference of the Spatial Sciences Institute*, 1225–31. Melbourne, Australia: Spatial Sciences Institute.

Willhauck, G. 2000. Comparison of object-oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos. *ISPRS* XXXIII: 214–21. Amsterdam.

Woodcock, C. E., and A. H. Strahler. 1987. The factor of scale in remote sensing. *Remote Sensing of Environment* 21:311–32.

RUTHERFORD V. PLATT is an Assistant Professor in the Department of Environmental Studies at Gettysburg College, Gettysburg, PA 17325. E-mail: rplatt@gettysburg.edu. His research interests include spatial models of human–environment interactions and emerging methods in remote sensing.

LAUREN RAPOZA is a graduate of Gettysburg College and is currently an outreach specialist at the Maryland Science Center. E-mail: laurenrapoza@gmail.com.