Strategies for Integrating Information from Multiple Spatial Resolutions into Land-Use/Land-Cover Classification Routines

DongMei Chen and Douglas Stow

Abstract

With the development of new remote sensing systems, very-high spatial and spectral resolution images now provide a source for detailed and continuous sampling of the Earth's surface from local to regional scales. This paper presents three strategies for selecting and integrating information from different spatial resolutions into classification routines. One strategy is to combine layers of images of varying resolution. A second strategy involves comparing the a posteriori *probabilities of each class at different resolutions. Another strategy is based on a top-down approach starting with the coarsest resolution.*

The multiresolution strategies are tested using simulated multiresolution images for a portion of the rural-urban fringe of the San Diego Metropolitan Area. The classification accuracy obtained from using three multiple strategies was greater when compared with that from a conventional single-resolution approach. Among the three strategies, the top-down approach resulted in the highest classification accuracy with a Kappa value of 0.648, compared to a Kappa of 0.566 for the conventional classifier.

Introduction

Accurate and timely land-use/land-cover information is essential to many government and private organizations at local, regional, national, and global levels for different applications, such as environmental monitoring and planning, land-use/land-cover change modeling, transportation planning, urban development planning, and urban modeling. Remotely sensed data have been the major sources of generating land-use/land-cover maps.

Spatial resolution is one of the fundamental considerations when using remotely sensed images. The simplest measurement of spatial resolution is ground resolved distance (GRD), defined as the dimensions of the smallest objects recorded on an image. For a satellite remote sensing system, the spatial resolution is the dimension of the ground-projected instantaneous-field-of-view (IFOV) recorded in an image (e.g., 30 meters per pixel for TM imagery or 10 meters per pixel for SPOT panchromatic imagery). Often the ground sampling distance (*pixel size*) is used to represent the spatial resolution. But ground sampling distance of an image can be different from the spatial resolution of the sensor that records the image after resampling. For this paper, spatial resolution is used as the ground sampling distance of an image.

Traditionally, land-use/land-cover mapping with remotely sensed data is conducted at a single resolution by visual interpretation and/or different semiautomated image classification algorithms and strategies, including supervised, unsupervised, and hybrid training approaches (Richards and Jia, 1999); parametric and nonparametric classifiers; segmentation (Conner *et al.*, 1984); artificial neural networks (ANN) (Civco, 1993); fuzzy sets (Wang, 1990; Foody, 1996); and knowledge-based systems (Kontoes and Rokos, 1996). With the development of new remote sensing systems, very high spatial resolution images provide a set of continuous samples of the Earth's surface from local to regional scales. The spatial resolution of various satellite sensors presently ranges from 0.6 to 25,000 m. Furthermore, high resolution airborne data acquisition technology has developed rapidly in recent years. As an increasing number of high resolution data sets become available, such as Digital Globe (Quickbird), Space Imaging (Ikonos), Orbimage, Indian Remote Sensing (IRS), Digital Ortho Quarter Quadrangle (DOQQ), etc., there is an increasing need for more efficient approaches to process and analyze these data sets.

A considerable amount of previous research has been devoted to exploring the magnitude and impact of spatial resolution on image analysis when shifting scale from coarse to fine resolutions (Gong and Howarth, 1992; Cihlar, 2000). Due to the more heterogeneous spectral-radiometric characteristics within land-use/land-cover units portrayed in high resolution images, many applications of traditional single-resolution classification approaches have not led to satisfactory results (Barnsley and Barr, 1996). In general, traditional single-resolution classification procedures are inadequate for discriminating between land-use/land-cover classes where spectral/spatial features and spatial patterns vary as a function of spatial resolution.

Based on the concepts of the scene model developed by Strahler *et al.* (1986) and scale of variance of features in the image by Woodcock and Strahler (1987), scene models may be either high (H) resolution with pixels smaller than objects, or low (L) resolution with pixels larger than objects to be mapped. The ideal situation for image analysis and classification is reached when the pixel size of the image corresponds to the objects in the ground scene (Woodcock and Harward, 1992; Marceau *et al.*, 1994a; Marceau *et al.*, 1994b). Because land-cover/land-use units vary in size, the analysis scale corresponding to one object does not match the others. Therefore, it is difficult to achieve optimal classification performance at a

D.M. Chen is with the Department of Geography, Queen's University, Kingston, Ontario K7L 3N6, Canada (chendm@post.queensu.ca).

D. Stow is with the Department of Geography, San Diego State University, San Diego, CA 92115.

Photogrammetric Engineering & Remote Sensing
Vol. 69, No. 11, November 2003, pp. 1279–1287.

0099-1112/03/6911–1279/$3.00/0

single resolution. From a practical standpoint, building a framework to represent, analyze, and classify images represented by multiple resolutions is necessary in order to capture unique spatial-resolution-related information about land-use/land-cover classes.

The purpose of this paper is to present three different strategies for selecting and integrating information from multiple spatial resolutions into classification routines under a multiresolution classification framework. The research background on multiresolution analysis and classification framework is discussed. The details of three strategies are presented. The proposed strategies were applied to two Color Infrared Digital Ortho Quarter Quadrangle (DOQQ) subsets in the urban and rural fringe areas of the San Diego Metropolitan Area. The overall and individual Kappa coefficients obtained from both multiresolution strategies and single-resolution strategies were compared to assess the efficiency of the different strategies.

Research Background

The spatial resolution of an image substantially affects image classification results. Spatial variation of image brightness values is important in the classification of remotely sensed data because pixels are grouped into classes because of similarities and classes are distinguished because of dissimilarities. Most studies on spatial resolution have examined the accuracy of estimating some property on the ground with remotely sensed imagery of different spatial resolutions (Latty *et al.*, 1985; Irons *et al.*, 1985; Toll, 1985; Johnson and Howarth, 1987; Marceau *et al.*, 1994a; Marceau *et al.*, 1994b; Hlavka and Livingston, 1997). These studies are empirical and the choice of spatial resolution depends on the availability of images.

As an increasing number of sources of multispectral remotely sensed data become available, the selection among different spatial resolutions becomes difficult. To select an image with appropriate spatial resolution for an image classification, one should examine the characteristics of scene content, especially the changing pattern of a scene as a function of changes in analytical scale and resolution. For example, what is the relationship between the ground sampling distance (GSD) (e.g., 1, 5, 10, or 100 m) of image data relative to the statistical parameters used in classifiers and the result of classification in land-use/land-cover classification using remotely sensed imagery?

Several techniques have been employed to assess appropriate (or optimal) spatial resolutions (such as Atkinson and Curran (1997)). Although a particular classification approach may achieve the best result for a single resolution and a particular class, there is no single resolution at which optimal results would be achieved for all classes (Marceau *et al.*, 1994b). Clearly landscape objects (e.g., land-cover/land-use polygons) and their internal material elements are not the same size; therefore, the analysis scale corresponding to one object may not be applicable to the others.

Several researchers have emphasized the importance of multiscale approaches. For example, De Cola (1994; 1997) proposed multiscale visualization as a component of exploration and analysis of image data. Friedl (1997) introduced a multiscale simulation model of spatial variation and used it in an environmental study using satellite data. Couloigner and Ranchin (2000) developed a multiscale approach for automatic road extraction and indicated the potential of the approach. Some more recent work using a hidden Markov model in multi-resolution classification is reported in Li *et al.* (2000).

The question of how one might choose scales that best represent each land-use/land-cover type is the basis for the multiresolution classification approaches proposed by Chen *et al.* (2002). The following briefly illustrates the multiresolution classification framework.

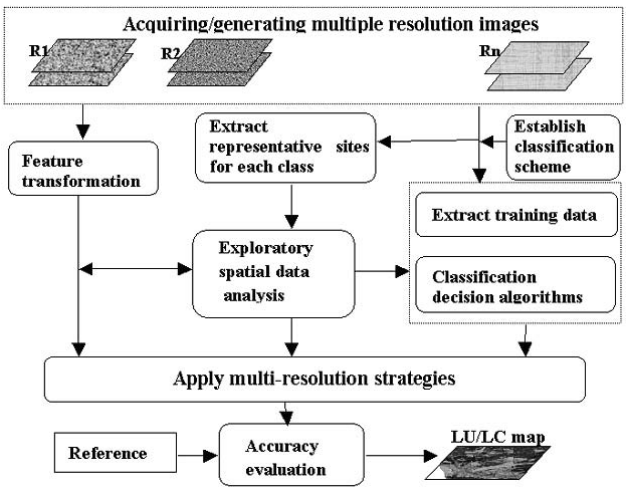

Figure 1. A proposed multiresolution image analysis and classification framework. Rn stands for variable image spatial resolution.

Multiresolution Image Analysis and Classification Framework

A proposed multiresolution image analysis and classification framework for integrating information from multiple resolutions is illustrated in Figure 1. Central to the approach are the following assumptions and hypotheses:

- There is no single spatial resolution which is optimal for classifying land-use/land-cover types, but an appropriate spatial resolution range exists for specific land-use/land-cover types and a given classification routine and mapping approach;
- Information obtained for a range of spatial resolutions is useful for discriminating land-cover/land-use classes, but improvements in accuracy are a function of the complexity of the environment, analysis methods, and land-cover/land-use types;
- Choosing the appropriate resolution range can be aided by information on the spatial characteristics of image brightness using spatial analytical methods; and
- Different land-use/land-cover types can be identified based on their spectral digital numbers and their spatial features extracted from different resolutions.

The framework proposed here focuses on the examination of image pattern/autocorrelation using different spatial analytical techniques in order to select appropriate methods in different stages of classification such as training strategy, feature extraction, scene models, and classification accuracy. Multiresolution images can be generated through aggregation methods. Spatial analysis techniques for measuring the pattern size and degree of autocorrelation are computed for each training class in order to determine whether they can guide selection of training data and the range of spatial resolutions used for classification. The results obtained from single resolutions are used as the benchmark to evaluate the efficiency of multiresolution methods. The different strategies for incorporating information from different resolutions are described below.

Different Strategies for Incorporating Information from Multiple Resolutions

Three strategies are developed to exploit information obtained from different resolutions and, thus, to improve the classification results.

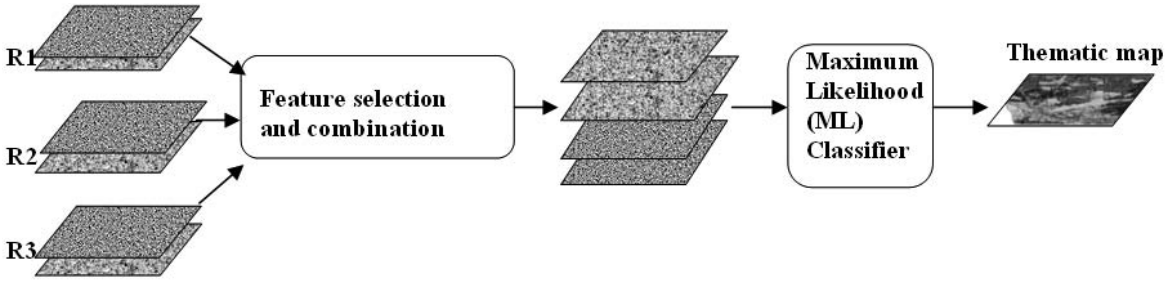

Figure 2. Flowchart for Strategy I (R1, R2, and R3 represent the different resolution levels). It combines layers of images of varying resolutions and then classifies them.

Of the many classification approaches available, the Gaussian maximum-likelihood classifier (GMLC) is utilized here for developing the multiresolution strategies, although the multiple resolution approach is applicable to other classifiers. The GMLC was selected because it is one of the most common classification rules used for supervised classification. The GMLC is robust and utilizes means, variances, and covariances of training site statistics, whereas most decision rules are based on simpler statistics (Jensen, 1996).

The GMLC is based on an estimated probability density function for each of the classes under consideration. The class statistics are obtained from the training data. Pixels are assigned to the most likely class of membership.

Strategy I: Adding Feature Layers from Different Resolutions as Additional Layers in Classification

A simple means for using information from multiple resolution images is to incorporate them simultaneously in a classification routine. In this way feature measures obtained at various resolutions are merged (see Figure 2). Then the ML classifier is applied. This approach is simple, and no other algorithms are needed to organize the data. All features from coarse resolutions need to be mapped back to the finest resolution using pixel replication. The major drawback is that computation cost may be high.

Strategy II: Comparing *a Posteriori* Probabilities from Multiple Resolutions

The flowchart for Strategy II is illustrated in Figure 3. For this approach, the classifier is applied at each resolution to obtain the probability $P(k|i)$ for each pixel k as a member of class i ($i = 1, 2, \ldots, m$ possible classes). The probabilities are then converted to *a posteriori* probabilities of class membership, which are assessed as the probability density of a case for a class relative to the sum of the densities (Jensen, 1996). The *a posteriori* probability of a pixel k belonging to a class i, $L(i|k)$, is determined by the following equation:

$$L(i|k) = \frac{a_i P(k|i)}{\sum_{i=1}^{m} a_i P(k|i)} \quad (1)$$

where $P(k|i)$ is the probability for a pixel k as a member of class i; a_i is the *a priori* probability of membership of class i, and m is the total number of classes.

For each pixel, the *a posteriori* probabilities sum to 1.0. At each resolution, the highest *a posteriori* probability and its related class are output for each pixel. $L_l(i|k)$ represents the maximum *a posteriori* probability of a pixel k belonging to class i at resolution level l, $L_l(i|k)$ derived from all resolutions, and k is assigned to the class with the highest maximum *a posteriori* probability. Thus, k is in class c if, and only if,

$$L(c|k) \geq L_l(i|k) \quad (2)$$

where $i = 1, 2, 3, \ldots, m$ possible classes, $l = 4m, 8m, \ldots,$ possible resolutions.

With the above approach, the feature layers obtained from coarse resolutions do not have to map back to the finest resolution specified and the computation cost is lower than that in the first strategy. However, there is a requirement for classifying the entire image at all resolutions, so this method may still not be very efficient. Pixels having the highest confidence at one resolution can be removed (e.g., masked out) at subsequent finer resolutions and this leads to the third strategy, a top-down procedure.

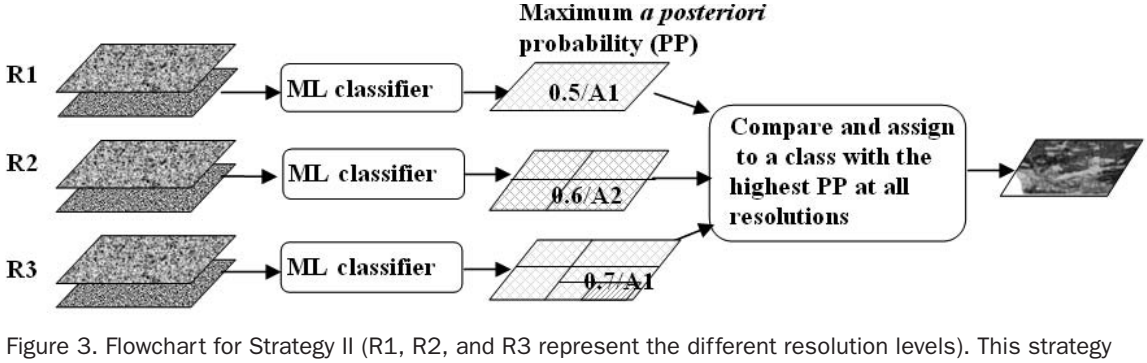

Figure 3. Flowchart for Strategy II (R1, R2, and R3 represent the different resolution levels). This strategy compares the *a posteriori* probabilities of each class at different resolutions and assign to the class with the highest *a posteriori* probability.

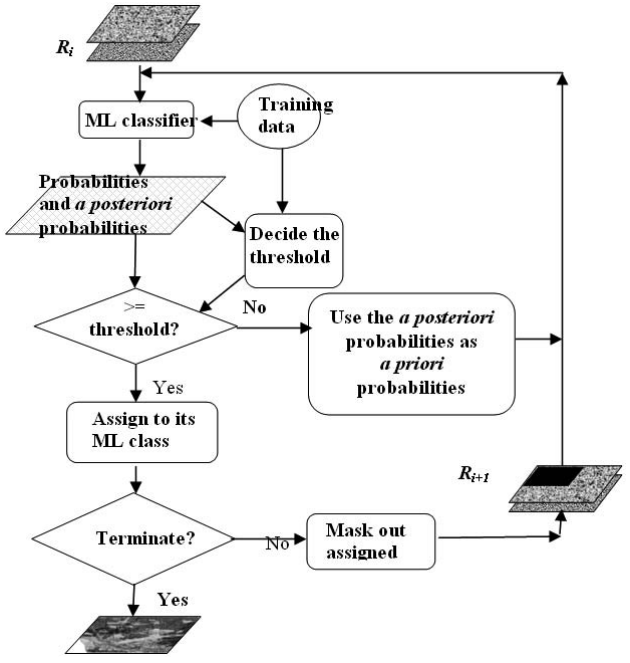

Figure 4. Flowchart of basic steps involved in Strategy III top-down procedure (R_i, R_{i+1} represent the different resolution levels).

Strategy III: Top-Down Filtering Approach

The basic steps involved in the top-down, multiple-resolution classification process are illustrated in Figure 4. This procedure starts with the coarsest resolution image. The finer resolution images are used only when necessary. Previous studies show that classification accuracy obtained using per-pixel classification in urban land-use/land-cover classification usually decreases as a function of resolution (Martin *et al.*, 1988; Gong and Howard, 1990). Therefore, it may be appropriate to utilize coarser resolutions and incorporate finer resolution data as necessary. Classes whose differentiation depends on finer detail can be handled later at some finer resolution level using smaller pixel sizes. A mixed pixel composed of several classes can be separated at finer resolution levels. In this way both the spatial details and classification accuracy will be improved. A possible disadvantage is that the class decision may be made in the absence of fine-resolution information, which could provide additional benefits to the classification decision.

The basic steps involved in the top-down, per-pixel classification process can be stated explicitly as follows:

(1) At each level of resolution l, select groups of training pixels for each class through the image. Use $T_l(i)$ to denote the training data set for class i at resolution level l. These training data are used for calibrating the classification routine.

(2) Classify the image at the coarsest resolution R with the maximum-likelihood classifier (MLC). For each pixel k, calculate $P(k|i)$ and $L(i|k)$, which are the probability and *a posteriori* probability as a member of class i, respectively. Using Equation 1, obtain the maximum probability $MaxP(k|i)$ and maximum *a posteriori* probability $MaxL(i|k)$.

(3) Calculate the mean and standard deviation of maximum probabilities ($Mean(P(k|i))$, $Std.(P(k|i))$ and maximum *a posteriori* probability ($Mean(L(i|k))$, $Std.(L(i|k))$ for each class i based on the training data $T_R(i)$. If the training data for each class are unbiased samples, a pixel k is assigned to a class c with high confidence if

(i) $P(k|c) \geq P(k|i)$ where $i = 1, 2, \ldots, m$ possible classes, and

$$Mean(P(k|c) - Std.(P(k|c) \leq P(k|c) \quad (3)$$

(ii) $L(c|k) \geq L(i|k)$ where $i = 1, 2, 3, \ldots, m$ possible classes; and

$$Mean(L(c|k)) - Std.(L(c|k)) \leq L(c|k) \quad (4)$$

where $k \in T_R(c)$.

Then Equations 3 and 4 are used as thresholds for deciding whether a pixel is assigned to the maximum-likelihood class.

(4) All pixels with maximum probabilities not satisfying Equations 3 and 4 are regarded as mixed pixels, or pixels which do not have identical signatures and cannot be identified at this resolution level. For these pixels, the *a posteriori* probability $L_R(i|k)$ is calculated using Equation 1, where $L_R(i|k)$ is the *a posteriori* probability $L(i|k)$ at resolution level R for pixel k as a member of class i.

(5) All pixels that are already assigned to a class are excluded (masked) for subsequent processing. For other pixels, the process goes to finer resolution images. Repeat step 2. Use $L_R(i|k)$ as *a priori* probability a_i in Equation 1 to calculate the *a posteriori* probability for pixel k.

(6) Repeat steps 3 through 5.

The sequential process above stops once (a) all the pixels are assigned to a class, or (b) the finest resolution is reached. In case (b), if there are still pixels unassigned to a class, the rule in Equation 2 is used. That rule is to compare the *a posteriori* probability at all resolution levels, and a pixel is assigned to the class with the highest *a posteriori* probability.

Case Study

The three strategies were tested using simulated multiresolution images derived from 1-m USGS color infrared Digital Ortho Quarter Quads (DOQQ) data for a portion of the rural-urban fringe of Del Mar, San Diego County, California. The image data have a spatial resolution of 1 m with three spectral bands (green, red, and NIR). Two study areas, each covering about 3.5 Km^2, were selected. Each subset contains several major land-use/land-cover types so that a variety of land-use/land-cover types are covered.

DOQQ images with 1-m resolution were aggregated progressively into five nominal resolution levels (4 m, 8 m, 12 m, 16 m, and 20 m) by an averaging method. Eight land-use/land-cover classes were used, including single-family residential, multi-family residential, industrial/commercial, irrigated grassland, high-density vegetation, cleared land, undeveloped land, and agriculture land. Figure 5 shows a DOQQ subset of one study area.

Training data were selected by visually identifying and manually digitizing blocks of pixels. As a general rule, the length and width of small blocks for each class were close to the range obtained from the semi-variogram, so that each block was big enough to represent the spectral and spatial properties of each class. Thus, the heterogeneity or autocorrelation within each class was included in the training data. The distance between any two blocks was greater than or equal to the range of the semi-variogram, so that the pixels in one block were correlated, but not spatially autocorrelated with those in another block. The detailed information on training selection can be found in Chen and Stow (2002).

Both single-resolution and multiresolution classifications were conducted. The single-resolution classification was used as a benchmark for evaluating various multiresolution approaches.

The results were evaluated and analyzed based on classification accuracy for eight land-use/land-cover classes. Reference data were randomly selected from the whole study area, and their categories were identified with the aid of an extant land-use GIS layer, aerial photographs, and field reconnaissance. In order to use one set of reference data at several

Figure 5. Subset of a USGS color-infrared DOQQ (NIR band) for an area of Sorrento Valley, California. The black lines mask out the major roads.

resolution levels, boundary or mixed pixels which could not be clearly identified as the same classes at both 4-m and 8-m resolutions were deleted from the samples. A total of 1867 randomly selected samples were identified for the study area. The overall and individual Kappa coefficients (Jensen, 1996) were reported for the study area for a series of classification maps in order to evaluate the agreement between the classification results and the reference data.

To determine the difference between two kappa coefficients, the significance test proposed by Cohen (1960) for comparing two classification results was adopted. With this method, the difference between two Kappa coefficients resulting from two classifications was first obtained. The square root of the sum of the variances V_K between the two classifications was then calculated. A z-value is determined by dividing the difference by the square root. A z-value above 1.96 indicates that two classification results are significantly different at the 0.95 confidence level.

Results and Discussion

Table 1 shows the overall Kappa values obtained from single resolutions and Table 2 shows the overall Kappa values obtained from classification of several multiple-resolution combinations using Strategy I. If not all bands at all resolutions were used, the band combinations with the best average separability at that number of bands were selected.

The Kappa values obtained from classification using multiple resolutions are greater when compared to those from single-resolution image input. This was the case for combinations of two, three, four, and five resolutions. One exception is using four spectral bands with the best average separability from the stacked images of 12 m, 16 m, and 20 m. Classification accuracy improvements are significant at the 0.95 confidence levels for one two-resolution, two three-resolution, and one five-resolution combinations relative to comparable results from all single-resolution classifications. In addition, there are seven combinations at which the classification accuracy improvements are significant at 0.95 confidence levels, when compared with the result with the lowest accuracy among the resolutions involved. The highest classification accuracy result, 0.601, was achieved using the five bands with the best separability from the image with stacked layers from 4 m, 8 m, and 16 m. The combined image from 4 m and 20 m yielded the second highest classification accuracy of 0.591. This was an increase of 0.102 as compared with the accuracy from 4 m alone.

It is not efficient to use all spectral bands from all resolutions considering the computation cost increase and classification accuracy obtained. The overall Kappa value for classification results from the combined image with 15 input channels of five resolutions (4 m, 8 m, 12 m, 16 m, and 20 m) for all three spectral bands is 0.589. Although the increase in Kappa value is significant at the 0.05 level, this combination did not yield the greatest increase in accuracy. Two combinations with 12 input channels (four resolutions and three bands) did not lead to improvements at the 0.95 confidence level compared with all results from a single resolution. Combining all possible bands and resolutions was not necessary for obtaining a more accurate classification. Several large improvements were achieved using five or six input channels with two or three different resolutions.

Although tests of Strategy I did not include all possible combinations of bands and resolutions, the results suggest

TABLE 1. OVERALL CLASSIFICATION ACCURACIES DERIVED FROM SINGLE-RESOLUTION. ACCURACY IS EXPRESSED AS KAPPA VALUES

Spatial Resolutions	4 m	8 m	12 m	16 m	20 m
Kappa Values	0.4789	0.5181	0.5435	0.5526	0.5664

TABLE 2. CLASSIFICATION ACCURACIES USING STRATEGY I OF COMBINING LAYERS OF IMAGES OF DIFFERENT SPATIAL RESOLUTIONS. CLASSIFICATION ACCURACY IS EXPRESSED AS KAPPA VALUES

The Number of Resolutions	Resolutions Used (m)	Bands Used	Kappa Value
Two	4 and 8	6	0.5596
	4 and 16	6	0.5753
	4 and 20	6	0.5908*
	4 and 12	6	0.5453
	8 and 16	6	0.569
	8 and 12	6	0.5768†
	8 and 20	6	0.5647
	12 and 20	6	0.5697
	12 and 16	6	0.5661
	16 and 20	6	0.5583
Three	4, 8, and 12	3	0.558
	4, 8, and 16	5	0.601*
	4, 12, and 20	9	0.5959*
	8, 12, and 20	9	0.5889†
	8, 16, and 20	9	0.589†
	12, 16, and 20	4	0.5145
Four	4, 8, 12, and 16	12	0.5623†
	4, 12, 16, and 20	5	0.5459
	8, 12, 16, and 20	6	0.5688
	8, 12, 16, and 20	12	0.5664†
Five	4, 8, 12, 16, and 20	9	0.5754†
	4, 8, 12, 16, and 20	8	0.5688†
	4, 8, 12, 16, and 20	7	0.5575
	4, 8, 12, 16, and 20	15	0.589†

*Indicates result was significantly higher at the 0.95 confidence levels for all single resolutions.

†Indicates result was significantly higher at the 0.95 confidence levels for at least one resolution involved.

that the selection of bands is important. Some bands may add extraneous or redundant information and increase confusion in classification. The average separability (as quantified by the transformed divergence statistics) does not have an exact correspondence to the classification accuracy, as was suggested in the results of the previous section and from other studies (Gong and Howarth, 1990). This caused some difficulties in selecting the combination of resolutions and bands that was needed to generate the most accurate classification map.

Table 3 lists the overall Kappa values obtained from classification of those combinations using Strategy II. When compared to the accuracy obtained from single-resolution images, the Kappa values obtained from classifications using multiple resolutions are always higher. This was the case for the results obtained from comparing the *a posteriori* probabilities of two, three, four, and five resolutions. Classification accuracy improvements are significant at the 0.95 confidence levels for five two-resolution, two three-resolution, and one four-resolution combinations relative to comparable results from all single-resolution classifications. The highest Kappa value, 0.6321, was achieved using Strategy II by comparing the *a posteriori* probabilities of classified images from 12 m and 16 m. This strategy resulted in an increase of 0.15 over the Kappa value from 12 m and 16 m. The combined image from 16 m and 20 m yielded the second highest Kappa value of 0.6244.

The inclusion of more images of different resolutions was not necessary to achieve significant improvements in classification results. The overall Kappa value for classification results from comparing the *a posteriori* probabilities of classified images of all five resolutions (4, 8, 12, 16, and 20) was 0.5662, which was not significantly different at the 0.05 level when compared with any results from single resolutions. The greatest improvements were achieved with two resolutions.

The improvement of the overall classification accuracy is minimal when the *a posteriori* probabilities obtained from fine resolutions (such as 4 m) were used. None of the combinations with a 4-m resolution in Table 3 yielded significant improvements relative to comparable products from all single-resolution classifications. The smallest improvement was achieved at a four-resolution combination of 4 m, 8 m, 12 m, and 16 m, followed by a combination of 4 m and 8 m. This may be explained by the low probability resulting from the non-normal distribution of the training data at fine resolutions. The wide range of pixel values in the training data at fine resolutions yielded overlaps in the feature space between different classes. Thus, the probability of a class being confused with another was high, resulting in misclassification. *A posteriori* probabilities with high likelihood of misclassification at fine resolution integrated with those obtained from lower resolutions were not effective in improving the classification accuracy.

With the data and classification scheme used in this study, a high Kappa value was achieved with the combination of coarse resolutions. This is different from the results obtained from using Strategy I. With Strategy I the combination of a fine resolution and a coarse resolution resulted in classification results with high Kappa values.

Table 4 lists the overall Kappa values obtained from classification of those combinations with only spectral bands as inputs. Similar to results from the previous two multiresolution strategies, the Kappa values obtained from classification using multiple resolutions were always higher compared to the accuracies obtained from single-resolution images. This was the case for the results obtained from top-down approaches when using two, three, four, and five resolutions. Improvements in classification accuracy are significant at the 0.95 confidence levels for most of the combinations tested relative to comparable results from all single-resolution classifications. The highest classification accuracy result, 0.6483, was achieved using Strategy III at 12 m and 16 m. This was an increase of 0.16 as compared with the Kappa value at 12 m and 16 m. The

TABLE 3. CLASSIFICATION ACCURACIES USING STRATEGY II OF COMPARING THE *A POSTERIORI* PROBABILITIES OF EACH CLASS AT DIFFERENT RESOLUTIONS. CLASSIFICATION ACCURACY IS EXPRESSED AS KAPPA VALUES

The Number of Resolutions	Resolutions Used (m)	Kappa Value
Two	4 and 8	0.5464
	4 and 16	0.5495
	4 and 20	0.5608†
	4 and 12	0.5519
	8 and 12	0.6061*
	8 and 16	0.5972†
	8 and 20	0.6192*
	12 and 16	0.6342*
	12 and 20	0.6157*
	16 and 20	0.6226*
Three	8, 16, and 20	0.5977*
	4, 8, and 12	0.5567
	4, 8, and 16	0.5387
	8, 12, and 16	0.576
	8, 12, and 20	0.5877*
	12, 16, and 20	0.5955
Four	4, 8, 12, and 16	0.5407
	8, 12, 16, and 24	0.5984*
Five	4, 8, 12, 16, and 24	0.5662

*Indicates result was significantly higher at the 0.95 confidence level compared to results obtained from all single resolutions involved.

†Indicates result was significantly higher at the 0.95 confidence level for at least one resolution involved.

TABLE 4. CLASSIFICATION ACCURACIES USING STRATEGY III TOP-DOWN APPROACH. CLASSIFICATION ACCURACY IS EXPRESSED AS KAPPA VALUES

The Number of Resolutions	Resolutions Used (m)	Kappa Value
One	4	0.4789
	8	0.5181
	12	0.5435
	16	0.5526
	20	0.5664
Two	4 and 8	0.5557†
	4 and 16	0.5581
	4 and 20	0.5620†
	4 and 12	0.5607†
	8 and 12	0.6295*
	8 and 16	0.6163*
	8 and 20	0.6309*
	12 and 16	0.6483*
	12 and 20	0.648*
	16 and 20	0.6333*
Three	8, 16, and 20	0.6014*
	4, 8, and 16	0.6019*
	8, 12, and 16	0.6395*
	8, 12, and 20	0.6224*
	12, 16, and 20	0.6341*
Four	8, 12, 16, and 20	0.6301*
	4, 8, 12, and 16	0.568†
Five	4, 8, 12, 16, and 20	0.6016*

*Indicates result was significantly higher at the 0.95 confidence level for all resolutions involved.

†Indicates result was significantly higher at the 0.95 confidence level for at least one resolution involved.

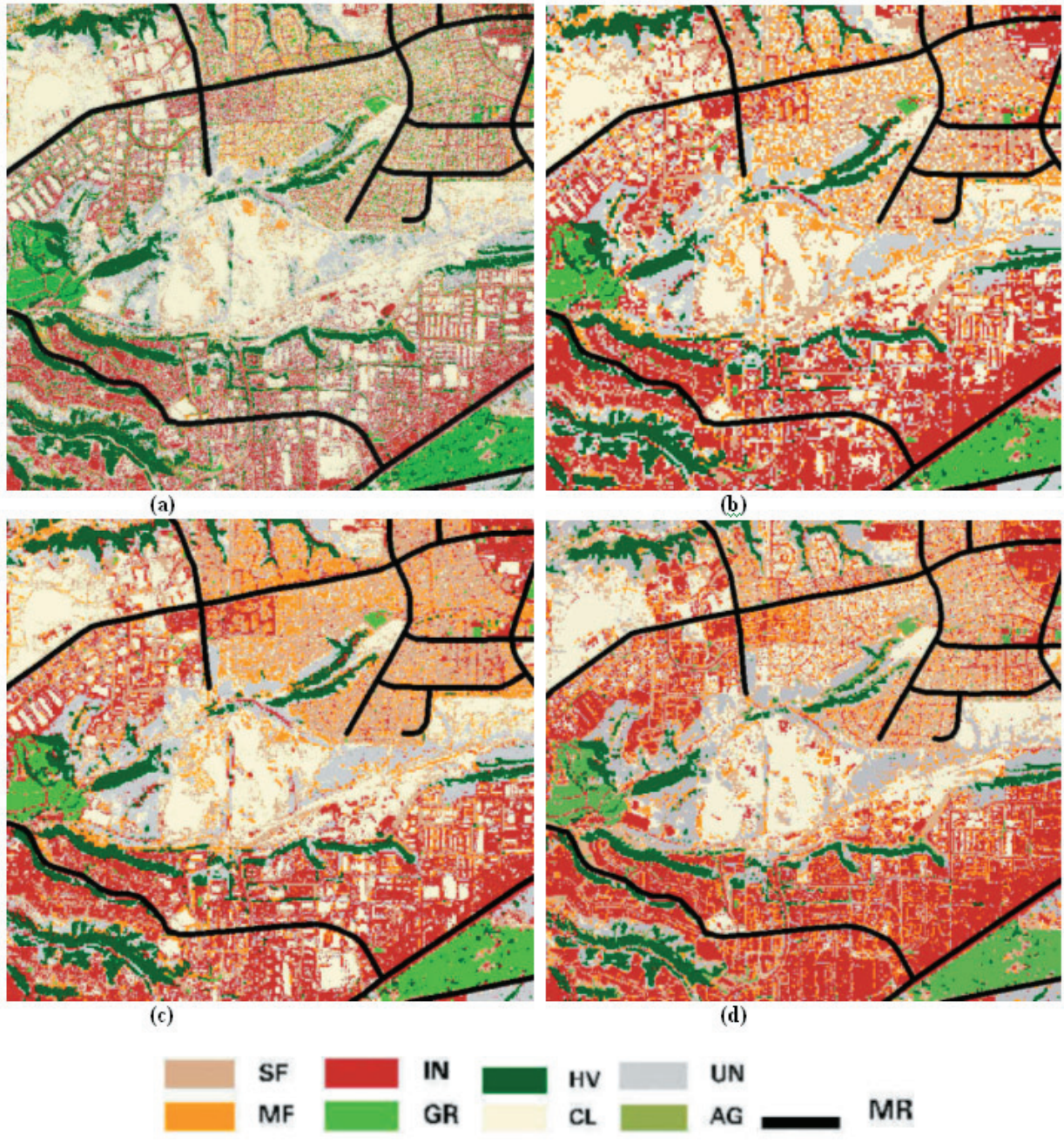

Plate 1. The classified maps obtained from (a) 4-m resolution using three spectral bands; (b) 20-m resolution using three spectral bands; (c) Strategy I of merging spectral bands of 4 m, 12 m, and 20 m; and (d) Strategy III of the top-down approach at 12 m and 16 m. (MR: masked road, SF: single-family, MF: multi-family, IN: industry, GR: irrigated grassland, HV: high-vegetation, CL: cleared land, UN: undeveloped vacant land, AG: agriculture)

combined image at 12 m and 20 m yielded a similar Kappa value of 0.648.

When compared to the accuracy results obtained from Strategy I and Strategy II, the Kappa values obtained from Strategy III are the highest when the same bands were used, except for the combinations with two resolutions in which one of them is 4 m. Kappa values from Strategy III are substantially higher than those from Strategy I and II when more than two resolutions were used. The classification accuracy improvements obtained from the six combinations of three resolutions listed in Table 4 with Strategy III were significant at 0.95 confidence levels relative to comparable results from all single-resolution classifications.

The inclusion of more than one resolution image in the top-down approach was not necessary to achieve the highest classification accuracy, particularly when the finest resolutions (e.g., 4 m and 8 m) were used in the combination. The greatest improvements (0.6483) were achieved with 12 m and 16 m resolutions. Similar to the results with Strategy II, the weaker improvements were achieved when finer resolutions

(such as 4 m and 8 m) were used in the top-down approach. However, unlike results from Strategy I and II, the Kappa values obtained from the combinations of more than two resolutions with Strategy III are similar to those from the combinations of two resolutions.

The classification products obtained from images of single resolution at 4 m and 20 m are shown in Plates 1a and 1b. Plates 1c and 1d portray the most accurate classification image obtained from Strategies I and III, respectively. The differences among these land-use/land-cover maps are readily apparent. A large portion of roofs of industrial and residential buildings were classified as cleared land at 4 m. The confusion between undeveloped vacant land and cleared land was high. "The salt-and-pepper" effect seen in Plate 1a is reduced dramatically at 20 m (Plate 1b). However, classification of the 20-m image resulted in greater confusion between undeveloped, industry, and single-family. An obvious difference between land-use/land-cover classifications is that the accuracy of residential areas and industrial areas was much higher after integrating information from 4-m and 20-m images using Strategy I (see Plate 1c). Evidence of this is seen in Plate 1d where residential, irrigated grass, and cleared land were more accurately classified.

Conclusions

Three strategies for integrating information from multiple spatial resolutions into classification procedures are developed in this paper. The first strategy is to combine layers of images of varying resolution. The second strategy involves comparing the *a posteriori* probabilities of each class at different resolutions. A pixel is assigned to the class with the highest *a posteriori* probability among all resolutions selected. The last strategy is based on a top-down approach starting with the coarsest resolution. Only when necessary to increase the maximum-likelihood probability are finer resolution images used. The *a posteriori* probabilities obtained from coarser resolutions are used as *a priori* probabilities at finer resolution levels.

The results obtained using three strategies developed in this research showed significant classification accuracy improvement when compared with those from single-resolution approaches, in terms of overall Kappa values of classified maps for the study areas that were used in this research. When compared to the accuracy results obtained from three strategies, Strategy III resulted in the highest Kappa value of 0.6483. The classification improvements from the best classification results obtained from the three strategies are significant at the 0.95 confidence level when compared with classification results obtained from a single resolution. Among the three strategies, Strategy III resulted in the highest classification accuracy while Strategy I showed the least improvement in classification accuracy. However, Strategy III requires more processing steps than do the other two strategies.

The efficiency of multiresolution strategies may be influenced by many factors during classification, including training strategy, feature extraction method and window size, characteristics of the landscape, classification algorithm, and accuracy assessment method. The impact of those factors on classification was beyond the scope of this paper. While there are certainly other methods of generating land-use/land-cover information using multiple-resolution sources of data, we found the strategies proposed in this paper reasonably efficient. The methods described in this study offer directions to those considering employing high spatial resolution data (such as Ikonos) over, or in conjunction with, the coarser spatial resolution data (such as Landsat TM) for land-use and land-cover mapping.

We view this research as an initial step towards building an integrated multiresolution classification framework for land-use/land-cover mapping using multiple spatial resolution Earth observation data. The land-use/land-cover classes selected for this study are relatively broad in spatial extent, so the value of adding higher resolution data to the analysis might be diminished. A more realistic test will include the detail land-cover types that can be discernable at higher resolutions. Further refinement, particularly of class structures and descriptors from spatial techniques, and exploration of how different spatial techniques can quantify resolution-dependent spatial characteristics of the image and can be used in the classification routine are required. More advanced classification approaches such as neural nets, fuzzy set classifiers, and expert classifier models should also be tested in the multiresolution context to further improve classification accuracy.

Acknowledgments

This research was completed in the Joint Doctoral Program in the Departments of Geography at San Diego State University and the University of California at Santa Barbara. Professors A. Getis, M. Goodchild, and R. Church provided essential guidance and support. The Authors also would like to thank the two anonymous reviewers for their insightful comments and suggestions.

References

Atkinson, P.M., and P.J. Curran, 1997. Choosing an appropriate spatial resolution for remote sensing investigations, *Photogrammetric Engineering & Remote Sensing*, 63(12):1345–1351.

Barnsley, M.J., and S.L. Barr, 1996. Inferring urban land use from satellite sensor images using kernel-based spatial reclassification, *Photogrammetric Engineering & Remote Sensing*, 62(8):949–958.

Chen, D., D. Stow, and A. Getis, 2002. Multi-resolution classification framework for improving land use/cover mapping, *Linking People, Place, and Policy: A GIScience Approach* (S. Walsh and K.A. Crews-Meyer, editors), Kluwer Academic Publisher, Boston, Massachusetts, pp. 235–262.

Chen, D., and D. Stow, 2002. The effect of training strategies on supervised classification at different spatial resolutions, *Photogrammetric Engineering & Remote Sensing*, 68(11):1155–1161.

Cihlar, J., 2000. Land cover mapping of large areas from satellites: Status and research priorities, *International Journal of Remote Sensing*, 21(6):1097–1114.

Civco, D.L., 1993. Artificial neural networks for land-cover classification and mapping, *International Journal of Geographical Information Systems*, 7(2):173–186.

Cohen, J., 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement*, 20(1):37–46.

Conners, R.W., M.M. Trivedi, and C.A. Harlow, 1984. Segmentation of a high-resolution urban scene using texture operators, *Computer Vision, Graphics, and Image Processing*, 25:273–310.

Couloigner, I., and T. Ranchin, 2000. Mapping of urban areas: A multiresolution modeling approach for semi-automatic extraction of streets, *Photogrammetric Engineering & Remote Sensing*, 66(7):867–875.

De Cola, L., 1994. Simulating and mapping spatial complexity using multi-scale techniques, *International Journal of Geographical Information Systems*, 8(5):411–427.

———, 1997. Multiresolution covariation among Landsat and AVHRR vegetation indices, *Scale in Remote Semsing and GIS* (D.A. Quattrochi, and M.F. Goodchild, editors), Lewis Publishers, Boca Raton, Florida, pp. 73–93.

Foody, G.M., 1996. Approaches for the production and evaluation of fuzzy land cover classification from remotely sensed data, *International Journal of Remote Sensing*, 17:1317–1340.

Friedl, M.A., F.W. Davis, J. Michaelsen, and M.A. Moritz, 1995. Scaling and uncertainty in the relationship between NDVI and land surface biophysical variables: An analysis using a scene simulation model and data from FIFE, *Remote Sensing of Environment*, 54:233–246.

Gong, P., and P.J. Howarth, 1990. An assessment of some factors influencing multispectral land-cover classification, *Photogrammetric Engineering & Remote Sensing*, 56(5):597–603.

_____, 1992. Frequency-based contextual classification and gray-level vector reduction for land-use identification, *Photogrammetric Engineering & Remote Sensing*, 58(4):423–437.

Hlavka, C.A., and G.P. Livingston, 1997. Statistical models of fragmented land cover and the effects of coarse spatial resolution on the estimation of area with satellite sensor imagery, *International Journal of Remote Sensing*, 18(10):2253–2259.

Irons, J.R., B.L. Markham, R.F. Nelson, D.L. Toll, D.L. Williams, S. Latty, and M.L. Staufer, 1985. The effects of spatial resolution on the classification of TM data, *International Journal of Remote Sensing*, 6(8):1385–1403.

Jensen, J.R., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective*, Prentice Hall Series in Geographic Information Science, Prentice Hall, Upper Saddle River, New Jersey, 316 p.

Johnson, D.D., and P.J. Howarth, 1987. The effects of spatial resolution on land cover/use theme extraction from airborne digital data, *Canadian Journal of Remote Sensing*, 13(2):68–74.

Kontoes, C.C., and D. Rokos, 1996. The integration of spatial context information in an experimental knowledge-based system and the supervised relaxation algorithms—two successful approaches to improving SPOT-XS classification, *International Journal of Remote Sensing*, 17(16):3093–3106.

Latty, R.S., R.F. Nelson, B.L. Markham, D.L. Williams, D.L. Toll, and J.R. Irons, 1985. Performance comparisons between information extraction techniques using variable spatial resolution data, *Photogrammetric Engineering & Remote Sensing*, 51(9):1159–1170.

Li, J., R.M. Gray, and R.A. Olshen, 2000. Multiresolution image classification by hierarchical modeling with two-dimensional hidden Markov models, *IEEE Transactions on Information Theory*, 46(5): 1826–1841.

Marceau, D.J., P.J. Howarth, and D.J. Gratton, 1994a. Remote sensing and the measurement of geographical entities in a forest environment 1: The scale and spatial aggregation problem, *Remote Sensing of Environment*, 49:93–104.

_____, 1994b. Remote sensing and the measurement of geographical entities in a forest environment 2: The optimal spatial resolution, *Remote Sensing of Environment*, 49:105–117.

Martin, L.R.G., P.J. Howarth, and G. Holder, 1988. Multispectral classification of land use at the rural-urban fringe using SPOT data, *Canadian Journal of Remote Sensing*, 14(2):72–79.

Richards, J.A., and X. Jia, 1999. *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, New York, N.Y., 384 p.

Strahler, A.H., C.E. Woodcock, and J.A. Smith, 1986. On the nature of models in remote sensing, *Remote Sensing of Environment*, 20:121–139.

Toll, D.L., 1985. Effects of Landsat TM sensor parameters on land cover classification, *Remote Sensing of Environment*, 17(2):129–140.

Wang, F., 1990. Improving remote sensing image analysis through fuzzy information representation, *Photogrammetric Engineering & Remote Sensing*, 56(8):1163–1169.

Woodcock, C.E., and A.H. Strahler, 1987. The factor of scale in remote sensing, *Remote Sensing of Environment*, 21:311–332.

Woodcock, C.E., and V.J. Harward, 1992. Nested-hierarchical scene models and image segmentation, *International Journal of Remote Sensing*, 13(16):3167–3187.

(Received 05 June 2002; accepted 30 October 2002; revised 17 December 2002)