ELSEVIER

Integrating Contextual Information with per-Pixel Classification for Improved Land Cover Classification

J. Stuckens, P. R. Coppin* and M. E. Bauer†*

A hybrid segmentation procedure to integrate contextual information with per-pixel classification in a metropolitan area land cover classification project is described and evaluated. It is presented as a flexible tool within a commercially available image processing environment, allowing components to be adapted or replaced according to the user's needs, the image type, and the availability of state-of-the-art algorithms. In the case of the Twin Cities metropolitan area of Minnesota, the combination of the Shen and Castan edge detection operator with iterative centroid linkage region growing/merging based on Student's t*-tests proved optimal when compared to other more common contextual approaches, such as majority filtering and the Extraction and Classification of Homogeneous Objects classifier. Postclassification sorting further improved the results by reducing residual confusion between urban and bare soil categories. Overall accuracy of the optimal classification technique was 91.4% for a level II classification (10 classes) with a* K_e *of 90.5%. The incorporation of contextual information in the classification process improved accuracy by 5.8% and* K_e *by 6.5%. As expected, classification accuracy for a simplified level I classification (five classes) was higher with 95.4% and 94.3% for* K_e*. A second important advantage of the technique is the reduced occurrence of smaller mapping units, resulting in a more attractive classification map compared to traditional per-pixel maximum likelihood classification results.* ©*Elsevier Science Inc., 2000*

INTRODUCTION

Several methods and algorithms have been developed to maximize extraction of information from digital satellite imagery. The recent scientific literature documents the role of multitemporal data sets (Coppin and Bauer, 1996), the development of more sophisticated classifiers (e.g., neural networks) (Paola and Schowengerdt, 1995), guided clustering (Bauer et al., 1994), and fuzzy sets (Bastin, 1997), sub-pixel classification approaches (Foody and Cox, 1994), the integration of textural information (Ryherd and Woodcock, 1996), and the incorporation of contextual techniques (e.g., segmentation) (Lobo, 1997). In addition, use has been made of ancillary data in stratification (Fuller et al., 1994), and in postclassification correction or filtering routines (Harris and Ventura, 1995). Many if not most of the studies on context have focused on smaller data sets, not on large metropolitan areas with their typical pattern of intricately mixed land use ranging from pure urban to rural. They all nevertheless testify to the fact that classification schemes taking into account context result in improved thematic classification accuracy.

The data source for land cover classification often consists of satellite imagery. Currently operational optical satellite sensors such as the Landsat Thematic Mapper (TM) and the French SPOT system deliver medium resolution products in which the pixel size is smaller than the general extent of landscape objects. All images therefore exhibit a high degree of spatial autocorrelation. In other words, the knowledge that a pixel belongs to a certain class increases the probability that its neighboring pixels belong to the same class. While traditional per-

*Department of Land Management, Faculty of Agricultural and Applied Biological Sciences, Katholieke Universiteit Leuven, Leuven, Belgium

†Department of Forest Resources, College of Natural Resources, University of Minnesota

Address correspondence to P. R. Coppin, Department of Land Management, Katholieke Universiteit Leuven, Vital Decosterstraat 102, B-3000 Leuven, Belgium. E-mail: pol.coppin@agr.kuleuven.ac.be

Received 14 September 1998; revised 10 July 1999.

REMOTE SENS. ENVIRON. 71:282–296 (2000)

655 Avenue of the Americas, New York, NY 10010

0034-4257/00/$–see front matter
PII S0034-4257(99)00083-8

pixel classifiers ignore this, contextual classifiers specifically try to incorporate this additional information into the classification routine. It is evident that the higher the degree of spatial autocorrelation is, the better the results of contextual classifiers will be compared to per-pixel classifiers.

However, one must be aware of the underlying processes when interpreting images generalized through context. All per-pixel classifiers try to assign every pixel to its spectrally closest class. For example, a wet area in a crop field may be classified as a wetland class, which is correct from the physical point of view. Yet after generalization through context (implementation of a contextual technique), the whole field would be categorized as cropland, which is a more correct assignment from the human interpretation perspective. The relationship between spectral response and information class is therefore often many-to-one. Where this phenomenon occurs for isolated pixels or smaller pixel groups, contextual information can assist in the derivation of the correct information class.

Extraction of Contextual Information

Two general approaches exist by which context can be calculated. One uses a moving window (ERDAS Inc., 1995); the other is based on image segmentation. With moving windows, contextual information is extracted from the pixel's immediate neighborhood by imposing a search window, computing a contextual parameter within that window, and assigning that value to the original pixel. While a 3×3 pixel kernel is probably the most widely applied window size, larger windows are appropriate for small pixel sizes and/or large land cover entities. Contextual data can be incorporated, either during the classification routine itself or as a postclassification step. In the former case, each pixel is assigned to the class with the highest probability, as in the case with the per-pixel maximum-likelihood-classifier—however, with probability defined as a measure not only of spectral similarity, but also of context (Kartikeyan et al., 1994; Jhung and Swain, 1996). Postclassification context integration generally involves filtering with a majority filter, whereby each pixel is recoded to the majority class of a neighborhood defined by the filter. More refined methods have been proposed by Barnsley and Barr (1996), Kim (1996), and Groom et al. (1996). Not only does this operation reduce the "salt-and-pepper" effect typical for per-pixel classifiers, it also results in larger classification units that might adhere more to the human perception of land cover. However, final classification accuracies do not necessarily improve dramatically. Thunnissen et al. (1992), for example, report an improvement of only 2% compared to a pure per-pixel classifier.

Segmentation operations, on the other hand, divide the image into contiguous clumps of pixels called "segments" or "regions." This division is based on spectral similarity or dissimilarity measures. All pixels of a segment or region are then assumed to belong to the same information class. Segmentation algorithms are based on either region growing/merging (Kettig and Landgrebe, 1976), boundary detection (Marr and Hildreth, 1980), or a combination of both (Tilton, 1996). Haralick and Shapiro (1985) also describe image segmentation routines that are based on measurement space-guided spatial clustering, on spatial clustering, and on split-and-merge techniques, but their application seems to be very limited in the earth observation domain. More recently, segmentation algorithms based on statistical models have been proposed (e.g., the PICO regression model) (Acton, 1996) and several Markov Field models (e.g., Bauman and Shapiro, 1994).

In region growing/merging, neighboring pixels or small segments that have similar spectral properties are assumed to belong to the same larger segment and are therefore merged. During further processing, segments are then often represented by their mean spectral vector. Both single-pass and iterative procedures have been developed. The computationally easiest, but least adequate, region growing algorithms are single linkage segmentations where pixels are simply linked to neighboring pixels and segments are connected components of linked pixels (Haralick and Shapiro, 1985). The same authors stated, however, that this algorithm is not likely to perform well on complex (such as metropolitan area imagery) or noisy data because of the ease with which unwanted region chaining can occur. An improvement was proposed by Bader and Jájá (1995) to suppress noise by incorporating various preprocessing steps. Centroid linkage region growing algorithms merge pixels with already existing adjacent core segments or seed pixels. Different ways are available to initiate the region growing process using either seed pixels (Latty, 1984), seed areas (Adams and Bischof, 1994), or initial edge detection-based segmentations (Egawa and Kusaka, 1988; Tilton, 1996). Merging then groups all processes where adjacent segments are combined into larger classification units based on statistical similarity rules. Growing and merging techniques are often aggregated into single segmentation algorithms. Examples of region growing approaches that have been successfully applied in a variety of settings are the Extraction and Classification of Homogeneous Objects or Extraction and Classification of Homogeneous Objects (ECHO) algorithm (Kettig and Landgrebe, 1976), the Hierarchical Stepwise Optimization (HSWO) algorithm (Beaulieu and Goldberg, 1989), the Woodcock-Haward algorithm (Woodcock and Haward, 1992), and the Iterative Mutually Optimum Region Merging (IMORM) algorithm (Lobo, 1997).

Boundary detection starts from the assumption that two neighboring pixels that exhibit a large difference in their spectral-radiometric vectors are members of differ-

ent segments. Consequently an edge or boundary can be drawn between them, and a segment is defined as the aggregation of the connected nonedge pixels. The edge pixels can then be either merged with the most similar segment, initiate their own segment, or be left unassigned (to be classified with a per-pixel classifier after segmentation). Very often, edges are extracted from gradient images, produced using local variance, Sobel gradient (Richards, 1993), or more sophisticated filters (Latty, 1984; Egewa and Kusaka, 1988; Shen and Castan, 1992). They can be the local maxima of the gradient image or may result from thresholding the gradient image. The latter process results in what is frequently referred to as an "edge-nonedge" or "ENE" image. Though an ENE image doesn't necessarily need to have all boundaries closed, several pre- and postprocessing procedures have been developed to improve the quality and consistency of ENE images. Haralick and Dinstein (1975) performed a region growing of the edge pixels to close gaps. Latty (1984) computed edge magnitude as well as edge orientation where the latter is used to project the intersection between interrupted edges, while Egewa and Kusaka (1988) defined algorithms for edge thinning and elimination of noisy edges.

Hybrid segmentation schemes may combine boundary detection and region growing/merging in various ways, implementing existing algorithms in flexible structures so as to fit classification requirements and imagery characteristics. Such systems must be seen as frameworks rather than as ready-to-use algorithms, and can be optimized according to predefined objectives. They are characterized by an open-structure sequential data flow, allowing for the replacement of algorithms with faster, less complicated, better adapted, state-of-the-art versions when appropriate. There are two major ways to classify images thus segmented: individual segments can be labeled using per-segment statistics, or the segmented layer can be combined with per-pixel classifier results via an overlay operation and per-segment majority rules. While the latter approach is independent of the training phase of the classifier, the former may require specific training data collection (Lobo and Chic, 1996). On the other hand, the integration of segmentation and classification steps in the former may result in better algorithm performance. Hybrid segmentation can be particularly useful in landscapes such as metropolitan areas, where land cover is spatially discontinuous with abrupt changes rather than gradual transitions. This is because it can be initiated from seed regions derived from edge detection, and because abrupt changes can be well represented by a step-edge model such as the one developed by Shen and Castan (1992).

The overall objective of this research was to integrate contextual information into a land cover classification of a large metropolitan area and to investigate to what extent this additional information contributed to overall classification accuracy. A major difference with previous, mainly smaller scope, studies resides in the fact that the latter often had full-coverage reference or training data available or very large data subsets. In a setting such as the land cover classification or LCC of the metropolitan area of the Twin Cities, Minnesota, USA, the 1% of the total area we had at our disposal for training and verification purposes (9,000 ha) was a proportionally small subset, although in real terms large with respect to ground data collection. A realistic assessment of classifier performance could therefore be expected.

STUDY AREA, IMAGERY, AND REFERENCE DATA

This research was carried out in the metropolitan area of Minneapolis-St. Paul, Minnesota, USA, totaling around 8,650 km^2 (Fig. 1). The central part (the Twin Cities of Minneapolis and St. Paul and the surrounding suburbs) is characterized by high-density urban, industrial, and residential land uses, but also includes lakes, small forest patches, grass fields, and lawns. It is surrounded by a seven-county area that is predominantly agricultural to the south, and composed of a patchwork of croplands, forests, wetlands, and lakes to the north.

The imagery, available for this research project, consisted of spectral bands 1 to 5 of two Landsat-5 TM scenes acquired June 16 and September 4, 1991. Both scenes provided full coverage of the seven-county study area. The June image allowed for good separation between major land cover classes (agriculture, forest, wetland, water, urban) while the September image displayed late-season characteristics of the landscape that provided greater specificity of classes. The multitemporal data also aided in the discrimination between row crops and other agricultural crops and the discrimination of forests and wetlands from other vegetation types. Both images had already been rectified and registered to USGS 7.5-minute quadrangle maps in UTM zone 15 using a second-order polynomial function with 40 ground control points and nearest neighbor resampling. The respective RMS errors were 0.37 pixels for June and 0.49 pixels for September. The two 5-band dates of imagery were treated as a single 10-band image (or a stacked data vector) in all digital processing.

A systematic stratified sample of 72 sections, derived in a previous research project (Bauer et al., 1998) was used as reference data for classifier training and evaluation. Cover type identifications were determined from small-format (35 mm) aerial color photography flown in July/August of 1991 and field data from the U.S. Department of Agriculture, Agricultural Stabilization and Conservation Service. Field information was obtained from ASCS records rather than relying on the small scale, color slides; the photos were used to locate field boundaries. Forty of the 72 reference sections were used for

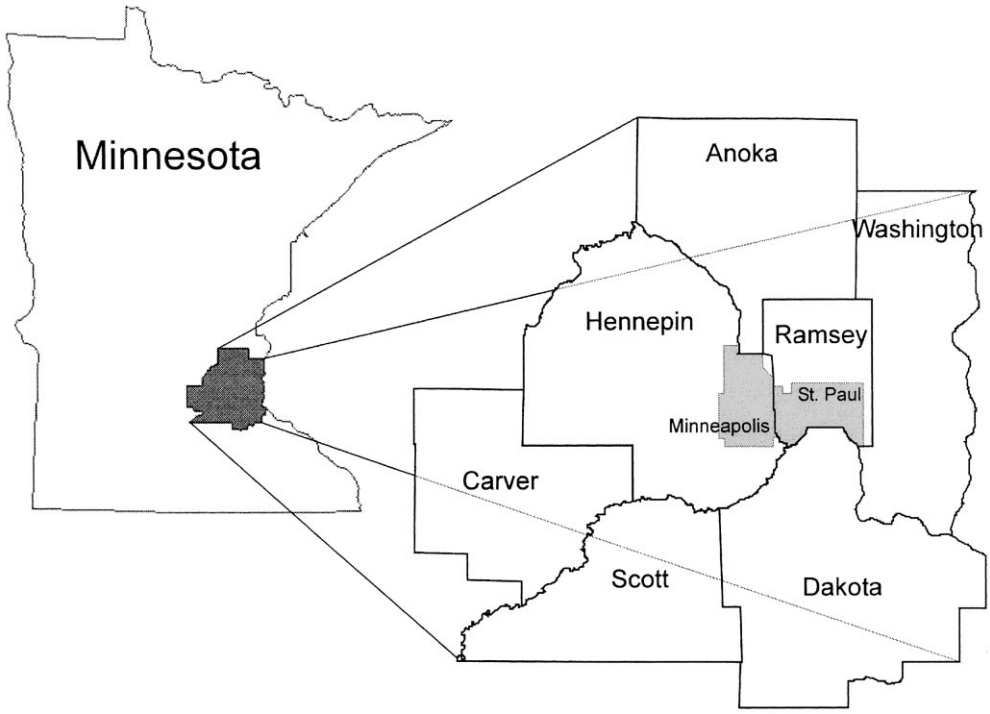

Figure 1. Twin Cities metropolitan study area.

classifier training purposes, while the remaining were kept for independent accuracy assessment. Additional reference data for nonagricultural classes were established from 1990 Metropolitan Council land use maps and National Wetland Inventory data. The land cover classification scheme included the 10 (somewhat modified) Anderson level II categories (Anderson et al., 1976) listed in Table 1.

METHODOLOGY

We implemented a phased methodological approach as depicted in Fig. 2. All analyses were executed in an ERDAS Imagine 8.2 software environment (ERDAS Inc., 1995) except for the edge detection, where Khoros 2.2 was used (Khoral Research Inc., 1997), and the ECHO routine, which was incorporated in the MultiSpec image processing system (Landgrebe and Biehl, 1995).

Stratification and Training

Before classification the imagery was stratified according to physiography and training data sets were developed separately for each physiographic stratum. The main purpose of this preprocessing step was to subdivide the rather large study area into physiographically homogeneous zones to avoid confusion between training signatures across the region, and thus improve final classification accuracy. The Twin Cities metropolitan area was divided into four physiographic regions: the Anoka Sand Plain, the Rochester Till Plain, the Owatonna Moraine, and the Eastern St. Croix Moraine. A fifth physiographic region, the Minnesota River Valley, was too small to have

Table 1. Land Cover Classification Scheme

Class	*Description*
Grass	Lawns, golf courses, sod fields
High-Density Development	Commercial, industrial and multifamily residential, with high percentage of impervious surfaces and little to no vegetation
Low-Density Development	Residential with lower percentage of impervious surfaces interspersed with vegetation such as lawns and trees
Row Crop	Corn, soybeans
Small Grain Crop	Wheat, oats, rye
Nonrow Crop	Alfalfa, bromegrass, etc. used for hay and pasture, Conservation Reserve Program, and diverted acres
Forest	Hardwood, conifer, mixed forest types
Wetland	Undifferentiated wetland types, including palustrine and lacustrine with emergent vegetation
Water	Permanent open water; lakes and rivers
Bare soil	Bare areas, not cropped in summer or fall

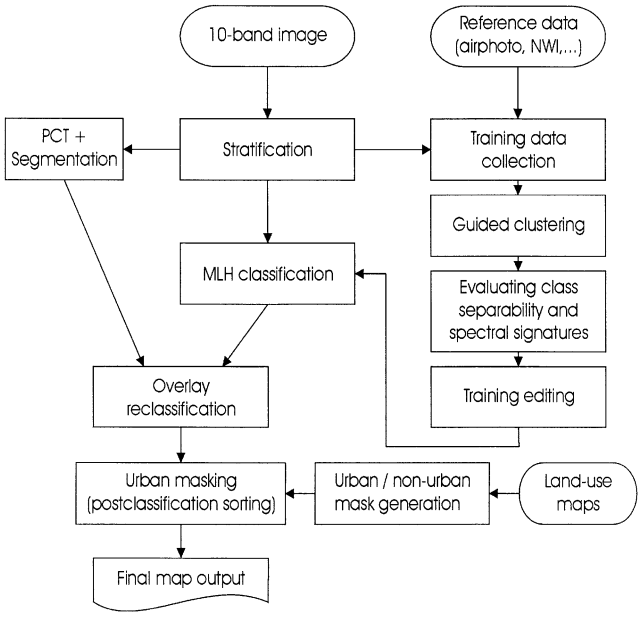

Figure 2. Flow chart of the phased approach.

enough reference data and was merged with the neighboring regions according to closest distance.

Because only general information was available in the reference data (e.g., one general class for hardwood forest) and because the inherent variability of such classes is too great for these to be of use in spectral-radiometric classification, it was necessary to break them up into subclasses (e.g., 10 sub- or cluster-classes for hardwood forest). These subclasses were then regrouped under their original general class name after image classification. The process is referred to as guided clustering (Bauer et al., 1994), and it provided consistently superior results to any of the other methods tested. It is highly automated and thus ideal for large-area applications, combining supervised (reference data classes) and unsupervised (cluster classes) training approaches.

The number of land cover subclasses for guided clustering was user-defined between 4 and 10, depending on the estimated spectral variability and the sample size (number of training polygons available). Too few subclasses would not cover all modes in the histogram; too many would result in small classes that could not encompass all data points in the multidimensional feature space. The latter fact is certainly a constraint where data dimensionality is high. The performance of the maximum likelihood classifier (MLH) was therefore checked for the Hughes phenomenon (Hughes, 1968) prior to any classification. It was found that all 10 bands could be used in the classification without risking loss of accuracy.

Next the training sets were evaluated for separability and exhaustiveness (inclusion of all spectral signatures present in the n-dimensional image). Training signatures were merged, added, or deleted until each set met both criteria. Class separability was evaluated using separability matrices composed of transformed divergences as well as contingency matrices derived from classifications of the training data sets. With respect to exhaustiveness, signatures not present in the original training set were identified using a Mahanalobis distance image derived from a provisional classification (Richards, 1993). Such an image is composed of pixel values that represent the Mahanalobis distance between the pixel and the subclass to which it is assigned. High pixel values thus represent spectral signatures that lack training data and likewise have a higher probability of being erroneously classified. Additional training data were created where these pixels tended to cluster and represented identifiable land cover entities. For example, the reference data set was supplemented with temporal flood plain signatures that were not included in the original reference data set.

Classification Approaches

Three approaches for the integration of contextual information were investigated. One approach (ECHO) classifies independently while imposing its own segmentation algorithm. Another approach, hybrid segmentation, is based on edge detection and centroid linkage region growing/merging. A third approach, majority filtering, is a simple postclassification procedure entailing the use of a 3×3 majority filter kernel, and as such will not be discussed further. In addition, residual confusion, typical for metropolitan areas, required the implementation of a postclassification sorting routine.

ECHO

The ECHO single-pass segmentation (Kettig and Landgrebe, 1976) works in two stages: first cell selection and then cell annexation. In the first stage, cells ($n \times n$ pixel squares) are selected as initial regions according to a homogeneity criterion. In the second stage, statistically similar cells that are adjacent are merged into segments. Leftover nonhomogeneous pixel groups are classified on a pixel basis. The parameters for the ECHO algorithm are actually threshold values for statistical testing; they have no direct physical meaning. For a detailed explanation we refer to Kettig and Landgrebe (1976). Our ECHO parameters were set at 2×2 pixels for cell size, at 2 for the cell annexation criterion, and at 165 for the cell selection criterion after optimization. These optimized parameter values for the ECHO process were derived from a small pilot experiment and the training statistics used for ECHO were the same as those for the MLH classification (including guided clustering).

HYBRID SEGMENTATION

The hybrid segmentation scheme proposed in general here (and implemented for this research in particular) is

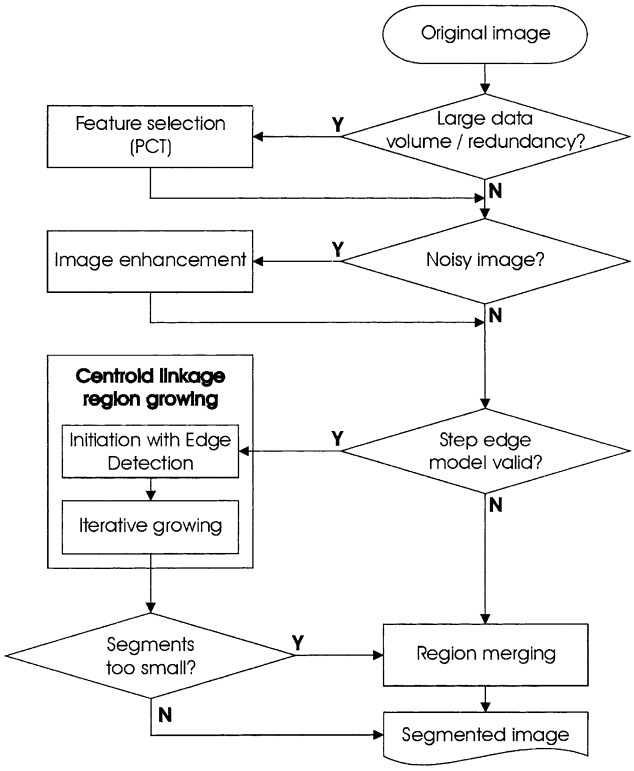

Figure 3. Hybrid segmentation scheme.

based on edge detection and centroid linkage region growing/merging followed by classification via overlay and per-segment majority rules. As can be seen from Fig. 3, it incorporates several optional routines, some of which may only be necessary for particular types of imagery and applications.

Edge detection was conducted based on Shen and Castan's "optimal linear operator for step edge detection" (Shen and Castan, 1992), a filter the same authors had found to be rather insensitive to noise and more precise in edge location than Gaussian (Marr and Hildreth, 1980) or Canny (Canny, 1983) filters, and that was successfully used in earth observation applications by Tilton (1996). The procedure encompassed two phases. First the input image (first three principal components of the 10-band imagery) was smoothed using an infinite symmetric exponential filter or ISEF (Shen and Castan, 1992). The theoretical design of the filter has an infinite window size whereby pixel weight factors diminish exponentially from the window center. In practice, a 7×7 window is applied and the contribution of pixels farther away is neglected. Although bigger windows better mimic the theoretical ISEF idea, they seem to have little effect on the ISEF result. Second, step edge detection was carried out using first derivative adaptive thresholding for second derivative zero crossing. The first derivative can be seen as a type of gradient image as used in many other edge detection algorithms. Implementation was in the Khoros software environment (Khoral Research Inc., 1997) using the "vsdef" algorithm with the following parameter settings for all input bands: smoothing parameter $a=0.5$, minimum length of edge segment $l=5$ pixels or 150 m, lower threshold $t_1=1$, and upper threshold $t_2=3$. These settings resulted from a pilot experiment on a subset of 200×200 pixels where visual edges and homogeneous polygons were manually digitized as vector entities and subsequently cross-referenced against the algorithm outcome with different parameter values. To insure closed boundaries around core segments (and thus avoid excessive undersegmentation), an edge-closing algorithm was applied to the ENE image of each separate input band. The algorithm followed the maxima of the equivalent Sobel gradient image (that was also made available) from one open end to another. Finally, the three ENE images were integrated via Boolean logic: a pixel is part of an edge on the composite ENE image if it is an edge on either of the three input ENE images.

After edge detection, connected component labeling (or clumping) was applied to the nonedge pixels of the composite ENE image under a four-connectivity rule (to avoid connecting adjacent segments across diagonal edges), resulting in core segment pixels that were all labeled to the segment's unique tag value. Initial core segments, smaller than three pixels, were merged into the background together with the edge pixels, a process referred to as sieving. Tests had shown that the elimination of these small core segments resulted in larger entities after region growing without affecting the performance of the iterative region growing algorithm.

Regions were grown from the initial core segments via an iterative region growing algorithm that was governed by a distance measure and growing threshold, the latter experimentally defined to be optimal at a value of 20 8-bit digital numbers. The growing threshold designated the maximum Euclidean distance between the feature vectors of candidate member pixels and the means of the adjacent region, whereby feature vectors were defined in the three-dimensional feature space formed by the three first PCs. Upon every pass, all background pixels (edge and small-segment) adjacent to a core region were evaluated as candidate members. A pixel was accepted if its Euclidean distance to a region's mean vector was smaller than a set maximum value (derived from a pilot experiment through visual interpretation) and smaller than all its Euclidean distances to other adjacent core regions, if any. Iterations continued until less than 0.5% of unlabeled pixels were assigned a region on a single pass. Convergence generally occurred after four iterations.

Region growing was followed by a visual inspection of the output image. If it was determined that the average segment or region size was too small to adequately represent land cover features, region merging was performed. As was the case with region growing, the proce-

dure was implemented iteratively until convergence was attained, requiring a segment distance measure and a merging threshold often expressed as a confidence level. With each iteration, and assuming normal distributions, the multivariate normalized differences between the mean vectors of all candidate pairs of adjacent segments were computed as shown in Eq. (1):

$$d_{j,k}=\sum_{i=1}^{3}\frac{|m_{i,j}-m_{i,k}|}{\sqrt{\frac{s_{i,j}^2}{n_j-1}+\frac{s_{i,k}^2}{n_k-1}}} \quad (1)$$

where

$d_{j,k}$ = multivariate distance over all three principal component bands ($i=3$)
$m_{i,j}$ = mean vector for the jth segment in the ith band
$m_{i,k}$ = mean vector for the kth segment in the ith band
$s_{i,j}^2$ = the variances for the jth segment in the ith band
$s_{i,k}^2$ = the variances for the kth segment in the ith band
n_j = the segment size in pixels of the jth segment
n_k = the segment size in pixels of the kth segment

Where segment sizes fell below 30 pixels, the formula was corrected for small sample sizes. Each candidate pair (j,k) was allowed one merge on each pass if two conditions are fulfilled. First, there were no other matches possible with a smaller multivariate distance (they are each other's mutual optimum in terms of spectral distance), and second, that multivariate distance did not exceed a preset maximum threshold τ. As the values of $d_{i,j,k}$ are distributed according to a Student's t-distribution with (n_j+n_k-2) degrees of freedom, a general confidence limit can be set from which τ can be derived for every pair of segments. Merging is thus performed if (with a confidence limit a) the null hypothesis that both samples come from the same distribution is not rejected in either of the three principal component bands. For this study, a was set at 0.01. A more extensive description of the merging algorithm can be found in Lobo (1997).

If a step-edge model is not appropriate, for example when regions are very small, then merging should occur immediately following image enhancement and should be carried out by connected components labeling of pixels with identical spectral vectors.

Classification of the hybrid segmentation results via an overlay technique of per-pixel-classified and segmented images was preferred here because of the robustness of this method against undersegmentation errors. The method itself was straightforward: for every segment, the majority MLH class and fraction were calculated. If more than 50% of the pixels of a particular segment were categorized in one class (majority fraction>0.5), then the segment was considered appropriate both with respect to existence and labeling (majority class). If not, then the segment was evaluated as a product of undersegmentation and its pixels were classified according to the 3×3 majority-filtered MLH output. Apart from forcing per-pixel classification where segments lack homogeneity, this technique also partially assists in overcoming problems caused by incomplete edge detection and by situations where adjacent segments with different land cover types were spectrally very similar. It is also interesting to note that the use of a stricter majority rule (for example, 66%) did not result in significantly different classifications accuracies.

POSTCLASSIFICATION SORTING

All four classification approaches that were tested here (MLH, ECHO, majority filter, and hybrid segmentation) exhibited spectral confusion between high-density urban and bare soil classes. Since this might be typical for metropolitan LCC endeavors, a postclassification sorting routine was developed based on ancillary data. Since the main objective of the project framework, within which this research was carried out, was production of "the best possible" land cover map of the metropolitan area of Minneapolis-St. Paul, postclassification sorting was applied solely to the "best" classification result, the hybrid segmentation. Results of hybrid segmentation with postclassification sorting should therefore be compared exclusively with hybrid segmentation without postclassification sorting, not with the other contextual classifiers. Metropolitan Council digital land use maps were converted into binary urban/nonurban (rural) masks. All pixels categorized as high-density urban but falling within the nonurban land use mask were recoded as bare soil. The inverse operation could not be validly applied because "bare soil" does indeed occur within urban areas.

Accuracy Assessment

Visual information content (map readability) was assessed qualitatively as well as quantitatively. The first operation involved a visual cross-comparison of the classification outputs and false-color composites (FCC) (TM bands 5, 4, and 3) of the single-date TM images. Although Landsat TM FCCs are commonly composed of bands 4, 3, and 2, we selected bands 5, 4, and 3 because band 5 provided additional information for distinguishing vegetation cover types. To arrive at a numeric measure of information content, for each classification method the size distribution of the mapping units (segments, polygons, regions) within a 400×400 pixel subset was computed and graphically plotted as a proportional cumulative curve (after conversion to hectares). This was done for a northern subset (mosaic of natural land cover

types), a central subset (urban St. Paul), and a southern subset (agricultural land cover types). The rationale was that visual information content is positively affected when fewer small patches are present, which makes it easier to interpret the map.

To allow for a comparative accuracy assessment, contingency tables and accuracy statistics were computed separately for each combination of physiographic stratum and classification method. Additional evaluations for the entire seven-county metropolitan area were subsequently carried out for each method (irrespective of physiographic stratum) via simple table summation and recalculation of the overall statistics. The accuracy assessment was exclusively based on the reference data sections that were not used in the classifier training phases.

The number of reference pixels per physiographic stratum for accuracy assessment was a function of the size of that stratum. An average of 860 reference pixels per class were available, with an absolute minimum of 82. To provide enough reference data for all 10 classes, less frequently occurring land cover types were somewhat oversampled, risking the introduction of a bias in the accuracy statistics. To avoid this, contingency table columns were weighted for total area in their class, so that column totals became proportional to effective ground area, as is the case in random sampling. Corrected land cover type estimates were obtained from classified-correct percentages in the tables using inverse calibration techniques (Czaplewski and Catts, 1992).

Comparative accuracy statistics that were computed include the overall percentage correct (P_c), user and producer accuracies, and Kappa with random chance agreement (K_e) (Foody, 1992) [see Eq. (2)]:

$$K_e = \frac{P_c - \frac{1}{q}}{1 - \frac{1}{q}} \quad \text{with } q = \text{number of classes} \tag{2}$$

Preference was given to the K_e coefficient over the standard Kappa because the adjustment for chance agreement in the latter assumes that row-marginal proportions are *a priori* specified, and that final maps are forced to have exact proportions (Stehman, 1997). As this is clearly not the case, $1/q$ as an adjustment term for chance agreement was deemed more appropriate than the standard Kappa adjustment term [see Eq. (3)]:

$$\sum_{k=1}^{q} p_{k+}p_{+k} \quad \text{with } p_{k+} \text{ and } p_{+k} \text{ as row and column marginals} \tag{3}$$

Another advantage of K_e is that it is simply a linear rescaling of P_c, so that both K_e and P_c use the same test statistics for hypothesis testing.

The P_c and K_e parameters of the four classification schemes were mutually compared, and for each pair the difference was computed and checked for statistical difference. Because more than one comparison between accuracy parameters was made, the original z statistics as described by Ma and Redmond (1995) were modified for Tukey multiple comparison procedures and confidence intervals were established for both K_e and P_c. In addition, the same procedures were implemented to verify whether postclassification significantly improved final classification accuracy.

RESULTS AND DISCUSSION

Segmentation performance was assessed separately from classification accuracy. The latter is a function of both training and classifier performance, encompassing a certain generalization of the information content and the thematic map accuracy. As computation time was never a limiting factor, no comparative tests of it were incorporated in the analysis. It must be pointed out immediately that due to the semantic nature of the segmentation algorithm parameters (which were fitted to the data, and thus were data-driven), their values as chosen for this study should not necessarily be extended to other types of imagery or other image analysis objectives.

Segmentation Characteristics

One way of cross-checking different segmentation algorithms is the calculation of their respective regions-per-pixel (R/P) ratios (Ryherd and Woodcock, 1996). The R/P value for the hybrid segmentation was 0.03, corresponding to an average region size of 34 pixels or about 3 ha. Average segment size increased progressively throughout the segmentation scheme: 10 pixels before sieving, 18 at the start of the growing process, 27 at the end of the growing process, and 34 after region merging. Because relatively few regions were merged, indicating that the edge detection algorithm produced very few oversegmentations, one may consider omitting the region merging step in similar future projects since oversegmentations are generally of less importance than undersegmentations in metropolitan land cover assessment. ECHO gave a higher R/P ratio of 0.036, with an average segment being made up of 28 pixels or about 2.5 ha.

The R/P values given above concern solely the image partitions delineated as segments. Although they may also be considered as image partitions, any remaining single-pixel elements were excluded from the R/P computation because they were classified via per-pixel pattern recognition routines. If one compares the segmentation results proportionally, 65% of the total image area was classified with segment majority rules by the hybrid segmentation scheme, compared to 75% for ECHO. Since the hybrid segmentation scheme ultimately became the preferred method (see discussion of classification accuracy below), we also looked at the contribution

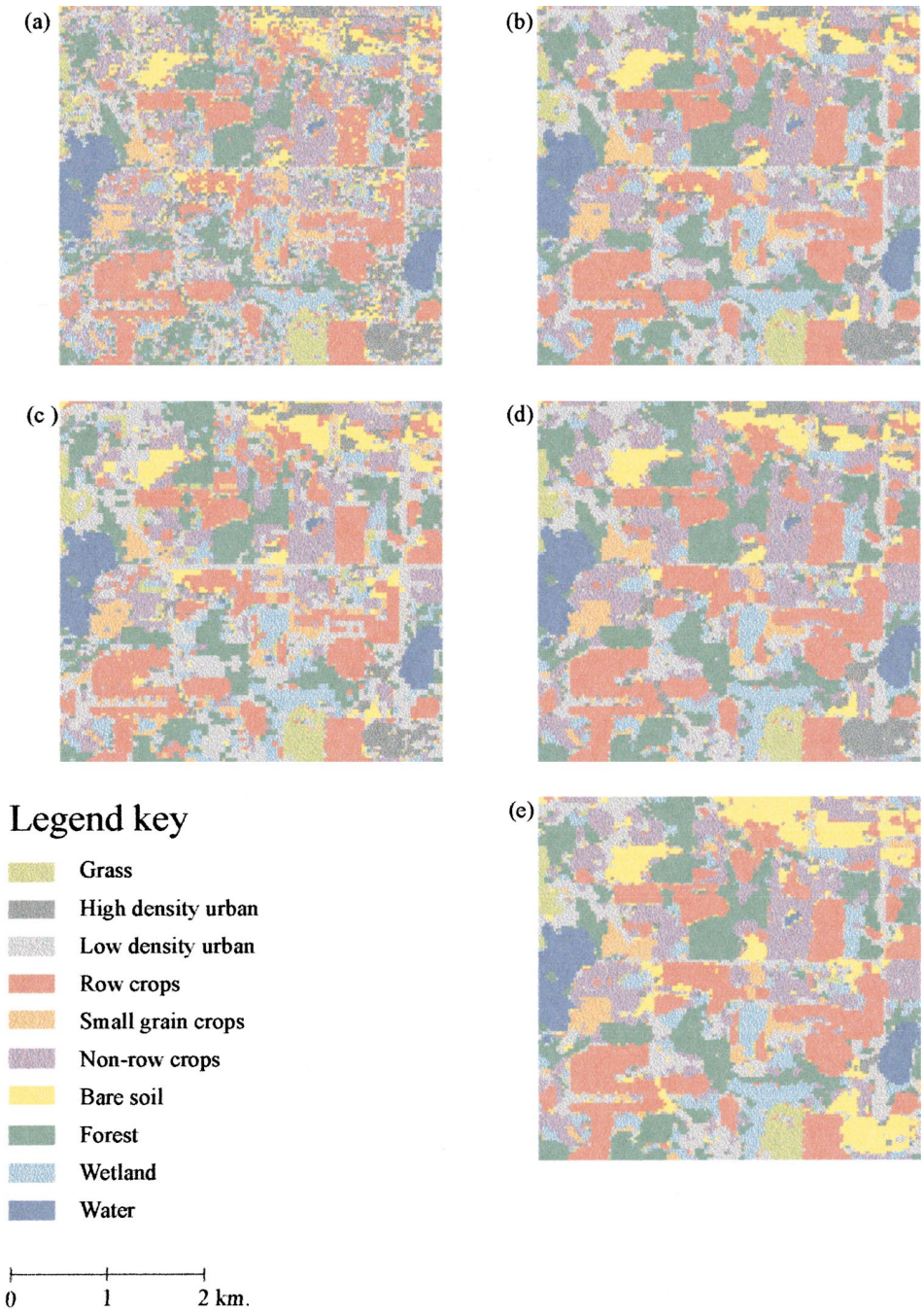

Figure 4. Comparison of visual information content between (a) maximum likelihood classification, (b) majority filtering classification, (c) ECHO classification, (d) hybrid segmentation classification, and (e) hybrid segmentation plus postclassification sorting.

of its component processes. After edge detection and sieving, 45% of the image was clipped to edge. Region growing reduced this to 17%. The final overlay operation rejected 22% of the area delineated as segments (83%) as nonhomogeneous (or not homogeneous enough), leading to the 65% figure indicated above.

Visual Information Content

Comparative visual information content of the tested metropolitan area land cover classification schemes is illustrated in Fig. 4, where 130×130 pixel subsets are shown for each of the four classification approaches (MLH, majority, ECHO, and hybrid segmentation) and, in addition, for hybrid segmentation followed by postclassification sorting. It is immediately evident that the salt-and-pepper effect (individual pixels being differently classified than their surrounding neighborhoods) decreases from MLH (a) to ECHO (c) to majority filter (b) and to hybrid segmentation (d). While the MLH algorithm gives evidence of quite a bit of confusion at the polygon edges, presumably due to residual misregistration and mixed boundary pixels, the three contextual methods result in much more regular polygon edges and an absence of small pixel clumps. This can be explained

by the fact that they inherently incorporate smoothing features such as filters and region growing/merging procedures. In the case of larger pixel clumps, high-resolution visual cross-verification against the small-format aerial photography used for reference data generation and the Landsat FCC used for visual accuracy assessment demonstrated that the hybrid segmentation approach generalized them more correctly than the two other contextual classifiers. In other words, the hybrid technique gave a visually more accurate representation of, for example, lake shores and parcel boundaries. Though there is definitely a subjective dimension in this assessment, this contention is validated by the accuracy assessment results (see Table 2), where the hybrid technique clearly outperforms the other contextual classifiers. As a consequence of its segmentation algorithm using $n \times n$ pixel kernels, ECHO has the additional drawback that its classification product appears unrealistically blocky, and that edges and shapes were preserved better for the north-south and/or west-east oriented agricultural parcels than for more naturally zoned and diagonally oriented objects.

As is the case in all land cover assessment exercises, only landscape elements that exceeded the minimum mapping unit (MMU) were classified. Linear features such as roads often appear broken up in multiple segments once they surpassed the MMU threshold, and these segments were then, although inconsistently, categorized as high- or low-density urban areas. The problem did not occur for waterways: water is a distinct spectral class and rivers in the Metropolitan Area are rather large (e.g., the Mississippi, Minnesota, and St. Croix Rivers). In residential areas, individual houses were dissolved into the urban matrix. In rural zones, individual large buildings were eliminated via generalization through context. It is obvious that for these types of landscape elements, small residual misregistrations and spectral class mixing could also be a sources of error.

Because large amounts of small mapping units negatively affect the visual interpretability of a land cover map, generalization was to be maximized without negatively impacting information content and accuracy. Figure 5 summarizes the cumulative size distributions of the cover maps of a natural land cover subzone, an urban subzone, and an agricultural land cover subzone, each produced via the MLH, the majority filtering, the ECHO, and the hybrid segmentation algorithms. Note that the X-axis of Fig. 5 follows a logarithmic scale. While the entire graphs may be of interest to express different levels of landscape scale, we were especially interested in the size distributions of the smallest mapping units. Mapping units below the 1 ha (or about 11 pixels) threshold represented, respectively, 27% (northern subset, natural land cover), 24% (southern subset, agricultural land cover), and 12% (central subset, urban cover) of the total number of pixels. The contextual methods reduced these totals to half to a third of those values.

Table 2. MLH Accuracies and K_e and Increases from Incorporation of Contextual Information

	Class Accuracies (%)											
Classifier	*Nonrow Crop*	*Row Crop*	*Small Grain Crop*	*Bare Soil*	*Forest*	*Wetland*	*Water*	*Grass*	*High-Density Development*	*Low-Density Development*	*Overall Accuracy (%)*	*Kappa* e *(%)*
MLH	62.6	93.0	72.9	70.0	93.3	76.6	98.6	93.4	91.5	89.5	85.6	84.0
Change relative to MLH												
Majority Filter	+4.5	+3.7	+2.4	+3.3	+1.1	+5.4	+0.6	+1.4	+4.4	+5.0	+3.4	+3.7
ECHO	+3.3	+1.8	+1.0	+12.0	−1.6	+0.3	−0.1	+1.0	−8.5	+3.6	+1.5	+1.7
Hybrid	+7.1	+4.5	+7.3	+3.0	+1.6	+8.7	+0.8	+1.4	+6.1	+5.3	+4.5	+4.9
Hybrid+												
Postclassification Sort	+7.1	+4.5	+7.3	+24.7	+1.6	+8.7	+0.8	+1.4	−6.2	+5.3	+5.8	+6.5

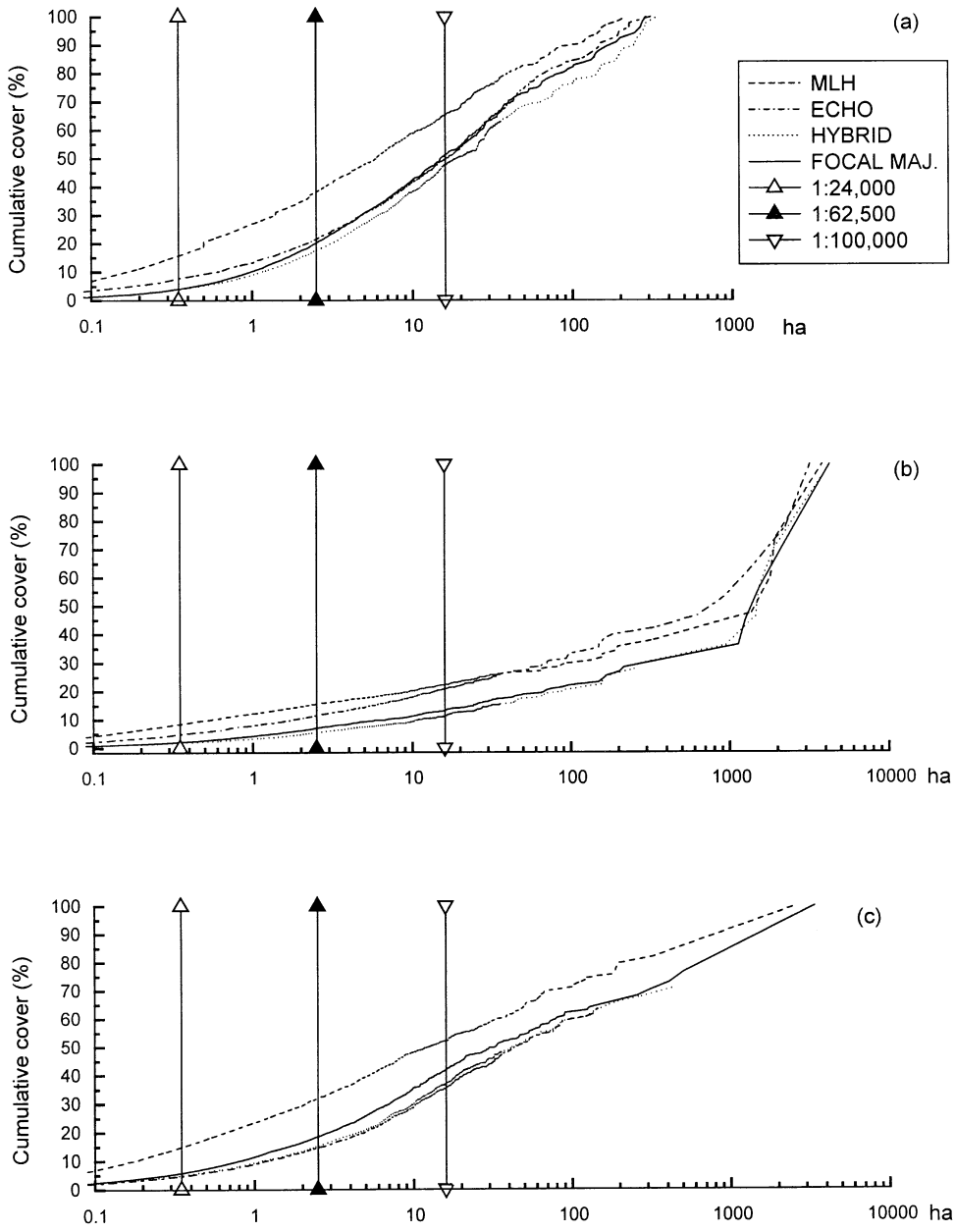

Figure 5. Cumulative size distribution graphs: (a) northern subset—natural land cover types; (b) central subset—urban St. Paul; (c) southern subset—agricultural land cover types.

The hybrid segmentation method consistently produced the maps with the lowest number of small mapping units. From the same perspective, ECHO proved better than the majority contextual filter in the southern agricultural subzone (many large rectangular fields) while the contrary was true for the urban and natural land cover subzones (more irregular patterns). From the same figure it can be inferred that identical conclusions can be drawn for the MMUs of the standard United States Geological Survey survey maps at 1:24,000 (0.35 ha), 1:62,500 (2.5 ha), and 1:100,000 (16 ha) scales.

Thematic Accuracy

Table 2 cross-references the producer's (class), the overall (overall percentage correct) accuracies, and the K_e coefficients computed from the respective contingency tables after correction to unbiased class cover areas for the entire seven-county area. The majority filtering and hybrid segmentation algorithm improved upon the standard MLH approach over the whole line (all land cover types), with increases in overall accuracy of up to 4.5% and K_e increases of up to 4.9% relative to the MLH outcome. The ECHO classifier was not always significantly

Table 3. Tukey Multiple Comparisons for K_e at 95% Confidence Intervals (All Numbers in %)

Comparison	*Improvement*	*Lower CI Limit*	*Upper CI Limit*
Majority Filter vs. MLH	+3.4	+2.6	+4.1
ECHO vs. MLH	+1.5	+0.8	+2.3
Hybrid vs. MLH	+4.5	+3.7	+5.2
ECHO vs. Majority Filter	−1.8	−1.1	−2.6
Hybrid vs. Majority Filter	+1.1	+0.4	+1.8
Hybrid vs. ECHO	+2.9	+2.2	+3.6
Hybrid + Postclassification Sort vs. Hybrid	+1.4	+0.7	+2.0

better than the MLH for forest, water, and high-density urban classes, for reasons that have to do with the dimensions and shapes of these landscape elements in relation to the $n \times n$ pixel kernel implemented by the ECHO algorithm. The overall ECHO results were nevertheless superior by 1.5% and 1.7%. The greatest increases in classification accuracy due to the contextual classifiers (as compared to per-pixel classifiers) were the wetlands, nonrow and grain crop agricultural classes, and the urban categories. It is striking that even the rather rudimentary majority filter performed 3.4% better than the MLH. The urban/nonurban postclassification sorting process applied on the hybrid segmentation resulted in an additional gain of 1.3% in overall accuracy (total of +5.8%) and 1.6% for the K_e value (total of +6.5%), with a major increase in class accuracy of 24.7% for the bare soil class, a decrease of 6.2% for the high-density urban class, and the other class gains remaining unchanged.

Tukey multiple comparisons between the K_e statistics (Table 3) allowed for the ranking of the different land cover classification schemes. All differences were significant at a confidence level of $\alpha = 0.01$. Hybrid segmentation outperformed the majority filter, which in turn outperformed ECHO, with the latter outperforming the MLH. Postclassification sorting after hybrid segmentation again proved beneficial. One must note, however, that the nature of the reference data set used for accuracy cross-verification (pure homogeneous polygons) may have led to somewhat narrower confidence intervals and somewhat higher significance levels than may be expected with not-so-homogeneous polygons that are more typical for the metropolitan landscape.

The accuracy assessment of the final metropolitan level II land cover map, generated by hybrid segmentation plus postclassification sorting, is summarized in Table 4. When combined to five level I general land cover classes (Table 5), overall classification accuracy and K_e increased, respectively, from 91.4% to 95.4% and from 90.5% to 94.3%. Because the change from the level II to a level I classification entailed a grouping of "similar" land cover classes, one may conclude that error rates of commission and omission were relatively more prevalent in the thematically more related and subsequently combined level II classes. Results differ somewhat when the four physiographic strata were analyzed separately. Overall classification accuracies range from 79% in the southern Rochester till plain to 89% in the eastern St. Croix moraine. This may be partially explained by the differing rates of occurrence of certain land cover types that proved troublesome to delineate (e.g., nonrow crops), by the spectral overlap of some of these cover types (e.g., row and nonrow crops in the southern stratum), and by a lack of training data for some classes in the smaller strata.

Since most reference data polygons were derived from entire fields or stands (i.e., homogeneous areas), mixed pixels are somewhat underrepresented in this accuracy assessment. If more mixed pixels form part of the reference data set, the accuracy figures may be assumed to be slightly lower than those calculated and presented here. This assumption is reinforced by Coppin and Bauer (1995), who reported a 5% drop in K coefficient values when mixed boundary pixels were not excluded from the cross-referencing, due mainly to residual positional inaccuracies at the pixel level and the subsequent uncertainty in their class assignment.

Though a justified critique may be that for some minor classes (e.g., wetland) not enough reference data were available to carry out a thorough validation process with respect to user's and producer's accuracy, some general results of the proposed metropolitan land cover classification scheme are striking and not always expected. The low-density and high-density urban classes, the water bodies, the forests, the grasslands, and the row crops were consistently delineated with high accuracy. The excellent identification of the residential areas and the suburban fringe was not expected, since this is generally seen as problematic due to the mixed occurrence of houses, trees, and lawns (Møller-Jensen, 1990; Ryherd and Woodcock, 1996). The major confusion seems to have arisen in the delineation of forests vs. wetlands and of nonrow crops vs. other cover types. It is known, however, that in this area many forests are found on wet

Table 4. Contingency Table and Summary Statistics for Final 10-Class Land Cover Classification

	Reference Data											
Classified Data	*Nonrow Crop*	*Row Crop*	*Small Grain Crop*	*Bare Soil*	*Forest*	*Wetland*	*Water*	*Grass*	*High-Density Development*	*Low-Density Development*	*Total*	*User's Accuracy (%)*
Nonrow Crop	**2,991**	57	46	12	41	64	0	10	0	1	3,223	92.8
Row Crop	459	**7,248**	1	1	37	17	0	1	1	3	7,768	93.3
Small Grain Crop	153	50	**604**	2	7	35	0	0	0	0	851	71.0
Bare Soil	86	18	0	**1,471**	0	1	0	0	55	8	1,639	89.7
Forest	72	18	9	0	**5,360**	174	0	0	0	1	5,634	99.9
Wetland	292	31	27	0	145	**2,312**	9	27	0	0	2,841	81.4
Water	0	0	0	0	31	13	**1,980**	0	0	0	2,023	97.9
Grass	5	2	24	0	1	29	0	**762**	0	0	824	92.5
High-Density Development	58	4	2	9	0	0	0	0	**730**	384	1,187	61.5
Low-Density Development	175	7	39	11	28	65	2	4	12	**7,247**	7,589	95.5
Total	4,291	7,435	752	1,506	5,650	2,710	1,991	804	798	7,644	33,597	
Producer's Accuracy (%)	69.7	97.5	80.3	97.7	94.9	85.3	99.4	94.8	91.5	94.8		

Overall classification accuracy=91.4%. Overall Kappa e coefficient=90.5%. Bold type indicates agreement numbers.

soils, and that the nonrow crop class includes a wide range of cover types instead of one single ground cover.

CONCLUSIONS

Segmentation approaches have been popular for use in extensive and/or homogeneous areas with large and regularly shaped landscape elements such as those prevalent in the agricultural zone of the U.S. Midwest. They may also be very applicable to the forested landscape of the northern United States. The reasons are the obvious ease with which landscape elements can be detected and divided into segments (agriculture) and the fact that the individual pixel responses cannot be considered representative for a forest stand (forestry). Because metropolitan areas often exhibit more diverse and complex patterns that are further subdivided into smaller irregularly shaped segments, they sometimes appear less-likely subjects for segmentation, although the combination with contextual filters may offer an attractive alternative.

Given a certain pixel-based thematic accuracy, any segmentation approach, hybrid or not, can only contribute so much of the "missing" accuracy. Error sources such as image quality problems, residual misregistrations, inconsistencies in and incompleteness of reference data, and similarities in spectral response for particular land cover types account for the remainder. Parameter setting and thus user interaction, however, remains critical. We propose that although computer power is generally no longer a limiting factor, the level of operator understanding is, and prior knowledge of the properties and potential of segmentation can assist in deciding whether its use is warranted. As shown here, this potential could be expressed as a difference in accuracy between segment-based and pixel-per-pixel classification results of an identical image subset. This approach is also valid to check the intrinsic qualities of a segmentation method for a particular land cover mixture.

Because of the flexible design of the hybrid segmentation scheme that has been presented here, several alternative solutions are possible for many of the steps. Selection of a composite set of procedures must be objective and data-driven, however, taking into account hardware and software constraints and operational limitations (e.g., time budgets and training level of operators). Examples of alternative solutions that have been presented in the scientific literature since the completion of this research are the implementation of neural network or fuzzy set classifiers replacing the classic MLH algorithm (Mannan et al., 1998), new edge detection algorithms that are continuously being developed in other application fields (Paillou, 1997), and new similarity measures for region growing (Baraldi and Parmiggiani, 1996). The option also exists to replace the principal component linear data transformation by another feature selection approach such as canonical analysis (Lobo, 1997), al-

Table 5. Contingency Table and Summary Statistics for the Final Five-Class Land Cover Classification

	Reference Data						
Classified Data	*Agriculture*	*Forest*	*Wetland*	*Water*	*Developed*	*Total*	*User's Accuracy (%)*
Agriculture	**13,201**	85	117	0	79	13,482	97.9
Forest	98	**5,360**	174	0	1	5,633	95.1
Wetland	349	145	**2,312**	9	27	2,842	81.3
Water	0	31	13	**1,980**	0	2,024	97.9
Developed	336	29	94	2	**9,139**	9,600	95.2
Total	13,984	5,650	2,710	1,991	9,246	33,581	
Producer's Accuracy (%)	94.4	94.9	85.3	99.5	98.8		

Overall classification accuracy=95.3%. Overall Kappa *e* coefficient=94.3%. Bold type indicates agreement numbers.

though the drawback is that such an approach would make segmentation dependent on the training data. Texture features (Ryherd and Woodcock, 1996) could also be included in the analysis and may be a partial solution to the high-density urban/bare soil confusion. Histogram mode-seeking clustering as proposed by Narendra and Goldberg (1977) could replace the ISODATA algorithm in the guided clustering process, requiring one general parameter instead of predefining the number of subclasses for each class.

Although in this case, specifically applied to metropolitan land cover mapping with Landsat TM imagery, the hybrid segmentation described could also be implemented on, among others, similar imagery for vegetation index extraction (Schoenmakers, 1995), radar data for speckle removal, and so forth. It is equally expected that when sensor spatial resolution is further improved, per-pixel classifiers will no longer perform adequately, and segmentation or other contextual tools will become the standard in land cover mapping with remotely sensed imagery.

The first author is indebted to the Remote Sensing Laboratory, Department of Forest Resources, for facilitating his M.S. thesis research work at the University of Minnesota. Project support at the University of Minnesota was provided by the Department of Forest Resources and Agricultural Experiment Station (project 42-037).

REFERENCES

Acton, S. T. (1996), On unsupervised segmentation of remotely sensed imagery using nonlinear regressing. *Int. J. Remote Sens.* 17:1407–1415.

Adams, R., and Bischof, L. (1994), Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16:641–647.

Anderson, J., Hardy, E. E., Roach, J. T., and Witmer, R. E. (1976), *A Land Use and Land Cover Classification System for Use with Remote Sensor Data.* Professional Paper 964, United States Geological Survey, Washington, D.C.

Bader, D. A., and Jájá, J. (1995), Parallel algorithms for image histogramming and connected components with an experimental study. In *Proceedings 5th ACM SIGPLAN Symposium of Principles and Practice of Parallel Programming*, Santa Barbara, CA.

Baraldi, A., and Parmiggiani, F. (1996), Single linkage region growing algorithms based on the Vector Degree of Match. *IEEE Transact. Geosci. Remote Sens.* 34:137–147.

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

Bastin, L. (1997), Comparison of fuzzy c-means classification, linear mixture modeling and MLC probabilities as tools for unmixing coarse pixels. *Int. J. Remote Sens.* 17:3629–3648.

Bauer, M. E., Burk, T. E., Ek, A. R., Coppin, P. R., Lime, S. D., Walsh, T. A., Walters, D. K., Befort, W., and Heinzen, D. F. (1994), Satellite inventory of Minnesota forest resources. *Photogramm. Eng. Remote Sens.* 60:287–298.

Bauman, C., and Shapiro, M. (1994), A multi scale random field model for Bayesian image segmentation. *IEEE Image Processing* 3:162–177.

Beaulieu, J. M., and Goldberg, M. (1989), Hierarchy in picture segmentation: A stepwise optimization approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11:150–163.

Canny, J. F. (1983), *Finding Edges and Lines in Images.* MIT Technical Report 720, Massachusetts Institute of Technology, Boston.

Coppin, P. R., and Bauer, M. E. (1995), The potential contribution of pixel-based canopy change information to stand-based forest management in the northern U.S. *J. Environmental Management* 44:69–82.

Coppin, P. R., and Bauer, M. E. (1996), Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sens. Reviews* 13:207–234.

Czaplewski, R. L., and Catts, G. P. (1992), Calibration of remotely sensed proportion of area estimates for misclassification error. *Remote Sens. Environ.* 39:29–43.

Egawa, H., and Kusaka, T. (1988), Region extraction in SPOT data. *Geocarto International* 3:25–32.

ERDAS Inc. (1995), *ERDAS Imagine Field Guide.* Atlanta, GA.

Foody, G. M. (1992), On the compensation for chance agreement in image classification accuracy assessment. *Photogramm. Eng. Remote Sens.* 58:1459–1460.

Foody, G. M., and Cox, D. P. (1994), Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions. *Int. J. Remote Sens.* 15:619–631.

Fuller, R. M., Groom, G. B., and Jones, A. R. (1994), The land cover map of Great Britain: An automated classification of Landsat Thematic Mapper data. *Photogramm. Eng. Remote Sens.* 60:553–562.

Groom, G. B., Fuller, R. M., and Jones, A. R. (1996), Contextual correction: Techniques for improving land cover mapping from remotely sensed images. *Int. J. Remote Sens.* 17:69–89.

Haralick, R. M., and Dinstein, I. (1975), A spatial clustering procedure for multi-image data. *IEEE Circuits and Systems* 22:440–450.

Haralick, R. M., and Shapiro, L. G. (1985), Image segmentation techniques. *Computer Vision Graphics and Image Processing* 29:100–132.

Harris, P. M., and Ventura, S. J. (1995), The integration of geographic data with remotely sensed imagery to improve classification in an urban area. *Photogramm. Eng. Remote Sens.* 61:993–998.

Hughes, G. F. (1968), On the mean accuracy of statistical pattern recognizers. *IEEE Transact. Geosci. Remote Sens.* 21:438–441.

Jhung, Y., and Swain, P. H. (1996), Bayesian contextual classification based on modified *M*-estimates and Markov random fields. *IEEE Transact. Geosci. Remote Sens.* 34:67–75.

Kartikeyan, B., Gopalakrishna, B., Kalubarme, M. H., and Majumder, K. L. (1994), Contextual techniques for classification of high and low resolution remote sensing data. *Int. J. Remote Sens.* 15:1037–1051.

Kettig, R. L., and Landgrebe, D. A. (1976), Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Transactions on Geoscience Electronics* 14:19–26.

Khoral Research Inc. (1997), *Khoros Version 2.2 On-Line Manual*, Albuquerque, NM.

Kim, K. E. (1996), Adaptive majority filtering for contextual classification of remote sensing data. *Int. J. Remote Sens.* 17:1083–1087.

Landgrebe, D. A., and Biehl, L. (1995), *An Introduction to MultiSpec.* Purdue Research Foundation, West Lafayette, IN.

Latty, R. S. (1984), Scene segmentation through region growing. In *Proceedings, Tenth Symposium on Machine Processing of Remotely Sensed Data*, West Lafayette, IN, pp. 305–314.

Lobo, A. (1997), Image segmentation and discriminant analysis for the identification of land cover units in ecology. *IEEE Transact. Geosci. Remote Sens.* 35:1136–1145.

Lobo, A., and Chic, O. (1996), Classification of Mediterranean crops with multi sensor data: Per-pixel versus per-object statistics and image segmentation. *Int. J. Remote Sens.* 17:2385–2400.

Ma, Z., and Redmond, R. L. (1995), Tau coefficient for accuracy assessment of classification of remote sensing data. *Photogramm. Eng. Remote Sens.* 61:435–439.

Mannan, B., Roy, J., and Ray, A. K. (1998), Fuzzy ARTMAP supervised classification of multi-spectral remotely-sensed images. *Int. J. Remote Sens.* 19:767–774.

Marr, D., and Hildreth, E. C. (1980), Theory of edge detection. In *Proceedings, The Royal Society*, London, England, pp. 187–217.

Møller-Jensen, L. (1990), Knowledge-based classification of an urban area using texture and context information in Landsat-TM Imagery. *Photogramm. Eng. Remote Sens.* 56: 899–904.

Narendra, P. M., and Goldberg, M. (1977), A non-parametric clustering scheme for Landsat. *Pattern Recognition* 9:207–215.

Paillou, P. (1997), Detecting step edges in noisy SAR images: A new linear operator. *IEEE Transact. Geosci. Remote Sens.* 35:191–196.

Paola, J. D., and Schowengerdt, R. A. (1995), A detailed comparison of back propagation neural network and maximum-likelihood classifiers for urban land use classification. *IEEE Transact. Geosci. Remote Sens.* 33:981–996.

Richards, J. A. (1993), *Remote Sensing and Digital Image Analysis: An Introduction*, 2d ed., Springer-Verlag, Berlin, Heidelberg.

Ryherd, S., and Woodcock, C. (1996), Combining spectral and texture data in the segmentation of remotely sensed images. *Photogramm. Eng. Remote Sens.* 62:181–194.

Schoenmakers, R. P. H. M. (1995), *Integrated Methodology for Segmentation of Large Optical Satellite Images in Land Applications of Remote Sensing.* Office for Official Publications of the European Communities Luxembourg, Nijmegen, Netherlands.

Shen, J., and Castan, S. (1992), An optimal linear operator for step edge detection. *Graphical Models and Image Processing* 54:112–133.

Stehman, S. V. (1997), Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* 62:77–89.

Thunnissen, H. A. M., Jaarsma, M. N., and Schoumans, O. F. (1992), Land cover inventory in the Netherlands using remote sensing: Application in a soil and groundwater vulnerability assessment system. *Int. J. Remote Sens.* 13:1693–1708.

Tilton, J. C. (1996), Hybrid image segmentation for earth remote sensing data analysis. In *Proceedings, International Geoscience and Remote Sensing Symposium*, Lincoln, NE, pp. 703–705.

Woodcock, C. E., and Haward, V. J. (1992), Nested-hierarchical scene models and image segmentation. *Int. J. Remote Sens.* 16:3167–3187.