Semi-Automated Training Field Extraction and Analysis for Efficient Digital Image Classification

*Martin P. Buchheim** and *Thomas M. Lillesand*
Environmental Remote Sensing Center, University of Wisconsin-Madison, Madison, WI 53706

ABSTRACT: Two non-traditional approaches to digital automated training data extraction and analysis were developed in response to limitations experienced with traditional manual training techniques as applied to the spatially and spectrally complex digital data available from second generation spaceborne sensing systems (Landsat Thematic Mapper [TM], SPOT). The first of these was a Semi-Automated Training Field Extraction (SATFE) procedure which used a least-variance training field growth strategy to extract training data from digital imagery. Data extraction was based only upon the coordinates of a single seed pixel for each field and several training data extraction parameters associated with each field. The second was a companion technique developed to conduct Semi-Automated Training Field Analysis (SATFA). In this procedure, training field data were analyzed and refined for use in image classification. Testing of the techniques consisted of a comparison of their accuracy and efficiency with that from traditional manual training data extraction and analysis procedures. Satellite imagery covering four test areas in Wisconsin (two TM, two SPOT) served as the base for testing. Results indicated that the two techniques were capable of significantly improving efficiency in the training phase of per-point classification, including reductions in both analyst and digitization time requirements, without incurring losses in classification accuracy, as compared to traditional procedures.

INTRODUCTION

MOST OF THE AUTOMATED IMAGE classification procedures in widespread use today can be considered "first generation approaches" to image classification in that they were developed primarily for use with Landsat Multispectral Scanner (MSS) data during the 1970s. Recent research indicates that these first generation image processing procedures are often incapable of fully exploiting the information content of the "second generation data" currently available from the Landsat Thematic Mapper (TM) and the SPOT High Resolution Visible (HRV) sensors (Hopkins *et al.*, 1988; Haack *et al.*, 1987; Parks *et al.*, 1987, Williams *et al.*, 1987; Woodcock and Strahler, 1987; Buchheim *et al.*, 1985; Irons *et al.*, 1985; Latty *et al.*, 1985; Acevedo *et al.*, 1984; Latty, 1984; Markham and Townshend, 1981). This suggests that "traditional" image analysis and classification approaches must be modified if their application to second generation data sets is to result in both accurate *and* efficient image classification. This paper focuses on two such new methodologies, namely, a technique for Semi-Automated Training Field Extraction (SATFE) and a companion procedure for performing Semi-Automated Training Field Analysis (SATFA). Used in combination, these techniques permit the extraction and analysis (refinement through merger or deletion) of training field data with a minimum of human intervention. Below, we describe the nature and function of these techniques and their accuracy and efficiency when applied to two SPOT and two TM test images.

SEMI-AUTOMATED TRAINING FIELD EXTRACTION: SATFE

The impetus for the development of a non-traditional approach to training field extraction grew out of growing evidence of the inadequacy and inefficiency of traditional training approaches when applied to second generation data, in general, and out of dissatisfaction with conventional training procedures, and unsupervised clustering techniques, in particular. The tremendous spatial and spectral complexity of second generation data can make it extremely difficult and tedious both to locate and delineate *supervised* training fields having statistical properties appropriate for maximum likelihood classification. This is compounded by the need to ensure that all important sub-classes within the high-variability data have been adequately characterized spectrally, or even sampled at all. Similarly, with many traditional *unsupervised* techniques, the impact of increased data complexity may be manifest as a loss of efficiency in the post-classification labeling stage. Furthermore, the clustering algorithms and parameters of many widely used unsupervised techniques were designed specifically with MSS data in mind, and thus are not universally applicable to data with fundamentally different characteristics. Because of these data-induced changes, the application of traditional supervised and unsupervised techniques to second generation digital satellite imagery often requires a much greater investment of time and expertise in order to produce accurate and repeatable classification results than has typically been the case with coarser, first generation data.

The new technique (SATFE) was envisioned as a means for improving both the efficiency and consistency of training field extraction. It was designed to increase the degree of automation in training field extraction, thereby significantly reducing analyst time requirements, while maintaining acceptable levels of classification accuracy. The procedure is still fundamentally grounded within a "per-point" (as opposed to "per-field") classification framework and, therefore, carries with it all the problems inherent in such an approach (i.e., no spatial, textural, or contextual information is used in classification). The intended purpose of the new technique, however, was to provide a tool whereby the utility of a per-point approach, limited as it may be for certain second generation data applications, could be improved. It was also viewed as a potential complement to other non-traditional image analysis techniques and as a possible efficient conduit for integration of remotely sensed data and geographic information systems (GISs).

DESCRIPTION

SATFE is designed to automatically delineate training fields within an image using only the coordinates of an initial "seed pixel" for each field and a set of user-specified extraction parameters associated with each seed pixel. The technique may function in either a supervised mode, wherein analyst-designated seeds representing known cover categories are used, or an

*Presently with Information Services International, 100 International Drive, Mt. Olive, NJ 07828–1808.

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 55, No. 9, September 1989, pp. 1347–1355

0099-1112/89/5509–1347$02.25/0

unsupervised mode, in which the cover category of a seed is unknown at the time of clustering. A training field (supervised mode) or spectral cluster (unsupervised mode) is "grown" around each seed pixel based upon the characteristics of neighboring pixels and the values of the field extraction parameters.

Supervised Mode. In the supervised implementation of SATFE, an image analyst identifies seed pixels in a manner similar to that used with traditional manual training field delineation, except that only one coordinate (that of the seed pixel) per training field need be identified, and parameters controlling the training data extraction process for each field must be specified. These parameters are typically specified uniquely for each seed pixel, but a common set of parameters may also be used. Supervised training fields are then grown, sequentially, around the designated seed pixels according to one of two possible strategies: "linear" growth or "concentric" growth.

Unsupervised Mode. SATFE is equally amenable to implementation in an unsupervised mode. In this case, seed pixels may be located automatically according to a regular, random, or stratified sampling strategy, or manually through quasi-random placement of seeds by the image analyst into spectrally homogeneous objects of initially unknown cover type. Sampling may be conducted for either the entire image or an analyst-specified subset of the image. This capability permits heavier sampling in areas of either greater importance or higher spectral diversity. Processing from this point on is identical to that for the supervised mode except that, when seed pixels are located automatically, all spectral clusters will utilize the same set of extraction parameters.

Linear Training Field Growth. With the linear growth strategy, pixels are added to a training field based upon their contribution to the summed variance for the field. The summed variance for a field is simply the sum of the variances of the field pixel values for all bands of imagery used as input. Thus, during training data extraction, the pixel that is added to a field next is the pixel which increases the summed variance of the field the least or, equivalently, decreases the summed variance of the field the most. Pixels added to a field must be spatially adjacent to a pixel already included in the field. Diagonal, as well as horizontal and vertical neighbors are considered spatially adjacent. Choices among these "candidate" pixels are made based upon the summed variance that the field would have following the addition of each candidate pixel. A field grows one pixel at a time along the path of least summed variance in the imagery, with the updated composition of the field serving as the basis for growth decisions and variance calculations during the next pass through the algorithm. Linear growth is illustrated in Figure 1a.

Concentric Training Field Growth. In concentric growth, pixels are added to a field, one at a time, in ever widening concentric circles around the seed pixel, irrespective of their relative impacts on field summed variance. This method was implemented primarily to provide a faster field growth procedure for relatively large, rectilinear, homogeneous fields, and also to reduce problems with statistical bias in variance estimation that may result from the linear growth technique and the high positive spatial autocorrelation that is commonly found in remotely sensed data (Campbell, 1981) (The potential impact of this bias will be discussed in greater detail later). Concentric growth is illustrated in Figure 1b.

Training Data Extraction Parameters. Training data extraction is controlled primarily by means of three criteria which determine when field growth is terminated: (1) a threshold maximum absolute summed variance for a field; (2) a maximum size (number of pixels) for a field; or (3) a maximum relative-variance-increase-ratio for a field. The first of these, the variance threshold, is a field homogeneity criterion which uses the same measure that is used for pixel extraction in the linear growth strategy. The

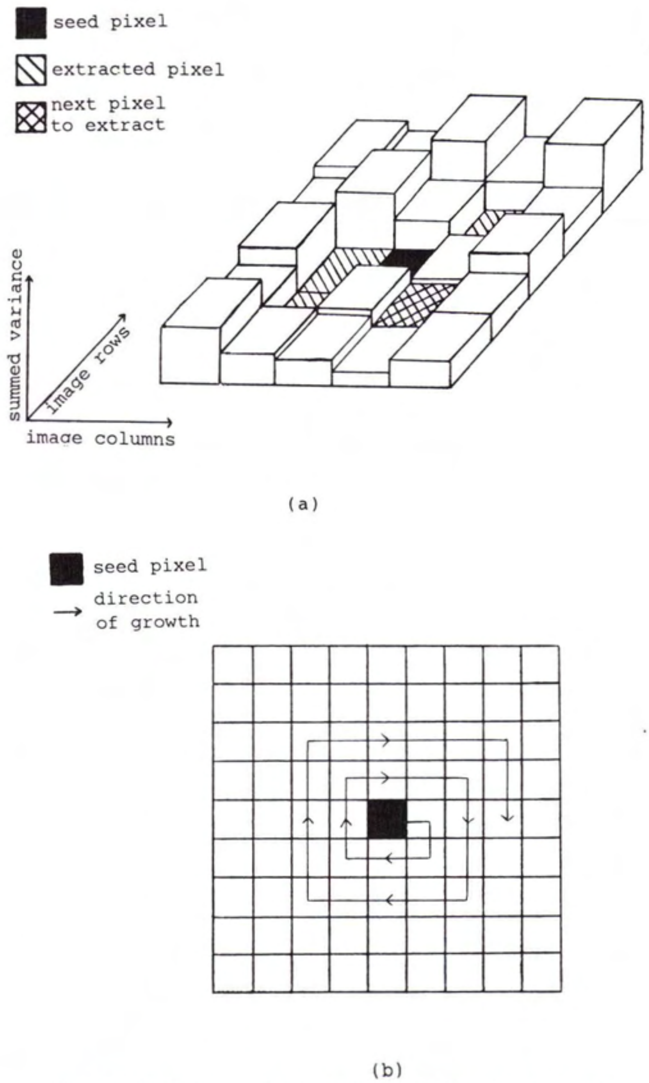

FIG. 1. Illustration of field growth strategies for the semi-automated training field extractor (SATFE): (a) the linear growth strategy; (b) the concentric growth strategy.

second criterion is simply an upper limit on the size that a field may have. The last criterion is a mathematical comparison of the new field summed variance – the variance of the field including the most recently extracted pixel – to the old field variance – the variance of the field prior to the latest pixel extraction. If this ratio is too large – if the variance of the field increases too much because of the addition of a single pixel – then processing for the field is terminated, regardless of the magnitude of the summed variance of the field at that point. Prudent use of this criterion helps to prevent the growth of fields over object boundaries and to reduce the sensitivity of field growth termination to the specification of the variance threshold parameter.

Two additional parameters are implemented to further enable the analyst to control the training field growth process. A minimum field size must be specified for each seed pixel. The method of growth – linear or concentric – must also be specified for each seed. A slight variation on this either/or choice, however, allows the analyst to select concentric growth as the initial growth strategy, with linear growth implemented only in the event that the field fails to reach minimum size using the concentric strategy. These two criteria are useful primarily to reduce spatial autocorrelation bias in variance estimates and, in the unsupervised mode, to prevent the creation of small, meaningless, high variance clusters. In the supervised mode, the analyst should know which growth strategy is most appropriate for each field based upon the spatial properties of the object upon which training is being conducted.

A final parameter available to the analyst is an automatic-variance-increase-value. This value is a percentage by which the variance threshold for a field should be increased if the field fails to reach minimum size before the variance threshold is exceeded. If this parameter is in force – and it need not be – the variance threshold for a field is increased automatically by the specified percentage in cases when field growth is terminated by the variance threshold criterion prior to the field attaining minimum size.

This last parameter was implemented to further reduce the sensitivity of the extraction technique to the specification of the variance threshold criterion and to allow the analyst greater flexibility in regulating field growth. An analyst can choose to rely heavily on accurate specification of the variance threshold, if knowledge of that entity is abundant or the analyst is very comfortable with its use. Alternatively, the analyst can place more emphasis on field size parameters during growth, using the automatic-variance-increase value as a means for achieving a spectrally homogeneous field of appropriate size.

Parameter specification is often a problem with automated image analysis and classification techniques. In many cases, the parameters are sufficiently arcane measurement entities so as to make their specification difficult when substantial previous experience with their use is lacking. In the case of SATFE, a conscious effort was made to ensure that the necessary field extraction parameters were familiar, intuitive, or easily calculated entities.

Inputs and Outputs Summary

Inputs to the SATFE procedure are simply the multispectral image data upon which the training is to be conducted, the coordinates of the seed pixels for each training field, and the associated extraction parameters for each training field.

Outputs from the procedure are a digital file containing the raw training field data extracted during the processing for each field and a separate output image data file which can be used to record and display the locations of the individual training field pixels generated by the process.

SEMI-AUTOMATED TRAINING FIELD ANALYSIS: SATFA

The process of moving from a group of initial training fields to a set of refined summary statistics suitable for use in classification (herein referred to as "training field analysis") has historically been a difficult one. The process typically involves more art than science in terms of the reasoning behind the merger and deletion decisions required to develop final training statistics. The task has also typically involved a significant investment of expert analyst time to conduct because the decisions made during the process have not lent themselves to automation or delegation to untrained personnel.

Experience has suggested that these training field analysis problems may be even more acute in the case of the more spectrally complex second generation data analyses – so much so that the problems may become a severe impediment to the operational utilization of second generation data for automated, per-point classification. As a result, a procedure for automating this labor intensive task of training field merging and deleting was developed. The procedure, SATFA, was designed specifically as a companion program for SATFE in the hope that it would increase the consistency of training field analysis procedures and further increase the automation of the training phase of per-point classification. It was also envisioned as a means for creating a capability for analyzing the large number of training fields which may be required to adequately spectrally characterize second generation imagery (and which can very easily be produced as output from SATFE, particularly in its unsupervised implementation). The process of analyzing large numbers of training fields is time-consuming and tedious under any circumstances, but it is acutely so when the informational relationships of the fields are not known *a priori* (i.e., in analysis of unsupervised clusters).

Training field analysis generally involves two types of decisions: whether or not to merge the data from two or more training fields if they are similar spectrally, and whether or not to delete a training field if its characteristics are sub-optimal in the context of the classification algorithm that will ultimately be based on the training data. In SATFA, two parameters have been developed to facilitate automated training field merging, and two procedures have been implemented to aid in training field deletion.

Merging Parameters

Merging decisions with SATFA are based upon two user-specified transformed divergence criteria which attempt to mimic the procedures employed in previous manual training data analyses. Transformed divergence is a familiar concept in remote sensing whose use as a measure of spectral similarity is well documented elsewhere (e.g., Horler and Ahern, 1986; Swain, 1978).

The first of the merging criteria is a *maximum* pairwise transformed divergence value. Fields having pairwise divergences greater than this value are discounted as candidates for merger. When divergences are less than this value, a training field is merged with another (candidate) training field provided that (1) the merger represents the best available "match" for the current field – the field has its lowest transformed divergence (greatest spectral similarity) with the candidate field – or (2) the two fields are compatible in terms of their similarity relationships with all other fields – the fields which satisfy the maximum divergence threshold with the current field are identical to those meeting the threshold for the candidate field.

The second criterion is a *minimum* divergence value below which two fields are merged regardless of their relationships with other fields. Using this second criterion, a field is merged with all currently unmerged fields which satisfy this threshold. In cases of confusion where a non-transitive spectral similarity relationship exists among fields (i.e., field 1 is similar to field 2, field 2 is similar to field 3, but field 3 is not similar to field 1), priority is given to field pairs with the lowest transformed divergence (greatest spectral similarity).

We believe the use of these two criteria and their associated compatibility evaluations and "best-match" determinations produces a more globally correct merging result than is often the case with other field merging approaches such as those embedded in many unsupervised clustering procedures. In many such approaches, mergers are done "on-the-fly," before all clusters have been identified, creating an order-dependence in the merging process. The merging procedures designed herein are neither affected by the order in which the fields are identified in the imagery nor the order in which they are stored in the digital training field data file.

Deletion Procedures

Automated handling of training field deletions proved to be a much more difficult process. As implemented for the evaluations described herein, SATFA relies heavily on manual input concerning fields which should be deleted – the analyst may pre-specify fields for deletion. The only other procedure for deletion is a "similarity percentage" value. This value is a measure of the number of fields that a given field is "similar to," based upon the maximum transformed divergence criterion. Samples similar to more than a user-specificed percentage of the total number of fields are deleted. This is a less than ideal approach to automated field deletion, but it is useful at least in deleting fields which, because of inordinately large variances, are spectrally similar to a relatively large proportion of the total number of training fields.

With SATFA, a conscious effort was again made to use parameters which are familiar to image analysts (transformed divergence) or easily understood and calculated (deletion percentage). Also, to allow greater flexibility and control in merging decisions, a merging constraint option was included in the procedure. This option allows an analyst to constrain training field mergers to within or among a particular group (or groups) of training fields. Thus, an analyst may use logic and informational relationships derived from a supervised approach to restrict mergers to spectral sub-classes within the same information class, thereby avoiding the problematic possibility of producing a training field consisting of samples from different information classes. Similarly, an analyst may specify, during either a supervised or unsupervised approach, that certain high quality training fields not be "contaminated" by merger with other, presumably less suitable, sets of fields.

TESTING METHODOLOGY

The foundation for testing the SATFE and SATFA algorithms was a direct comparison of conducting training field extraction, training field analysis, and image classification using the new procedures versus execution of the same tasks using conventional methods. Specifically, the potential utility of the semi-automated procedures was assessed by evaluating the comparative accuracy and efficiency of the two approaches, traditional and non-traditional.

ACCURACY

The accuracy of the two approaches was compared primarily using the KHAT classification accuracy measure and its associated tests for statistical significance (Lathrop *et al.*, 1987; Rosenfield and Fitzpatrick-Lins, 1986; Congalton *et al.*, 1983; Fleiss *et al.*, 1969; Cohen, 1960). Ground reference information was derived from field visits, aerial photography, and available maps. Accuracy values were determined for a random sample of points (pixels), rather than fields, located in the test imagery. Because of problems with expense, accessibility, and registration of samples, however, two restrictions were placed on the location of these "random" samples within the test imagery: (1) samples were taken only in areas for which ground information was available, and (2) samples were limited to field interior (non-edge) pixels. A minimum sample size of 50 pixels per class was taken, when possible.

The *Null Hypothesis* of this aspect of the testing was that no significant differences in accuracy would be observed between the two approaches. The hope was that this hypothesis would not be rejected – that the use of a more automated approach to training field extraction and analysis would not result in a loss in classification accuracy as compared to more traditional, manual procedures.

EFFICIENCY

The comparative efficiency of the semi-automated versus the manual approaches was determined by measuring the time required to complete each approach (both analyst time and processing time) and by qualitatively evaluating the relative ease of implementation of each approach. The number of coordinates digitized during training field delineation was also recorded. As with the accuracy tests, the *Null Hypothesis* for this aspect of the study was that no difference in the efficiency of the two procedures would be detected. In this case, however, the hope was that the new procedures would be noticeably more efficient than the traditional, manual procedures.

ADDITIONAL TESTING

Two additional tests were conducted specifically on the SATFA procedure. First, the two growth strategies, linear and concentric, were evaluated using image output for several example training fields. Second, the summed variance of training field data produced by the two approaches was compared for several test fields.

ANALYSIS AND CLASSIFICATION METHODS

Study Sites. Testing of the new procedures was accomplished using SPOT and TM data from four test sites in Wisconsin (Figure 2). The sites were chosen to represent a range of physiographic conditions and land-use practices, and to allow direct comparison of the possible differential effects of the procedures on SPOT and TM data.

Study site #1 was an agricultural area covering a portion of south central Wisconsin including the villages of Lone Rock and Spring Green. A subset of 26 August 1984 Landsat-TM image was extracted for use in this area. Study site #2 consisted of an urbanized area in northern Dane County, Wisconsin including a portion of the Madison metropolitan area. SPOT multispectral data from 3 June 1986 were used for this study site. The third study site was a heavily forested area located in northwestern Wisconsin near the city of Minong. Data for study site #3 were extracted from a 29 July 1986 TM image. Study site #4 was selected to cover approximately the same area as that for site #3 – Minong and vicinity. For this site, however, SPOT multispectral data from a 9 August 1986 overpass were used.

FIG. 2. Map depicting the four study sites used for testing SATFE and SATFA (study sites 3 and 4 consisted of TM and SPOT data, respectively, covering approximately the same area).

For each study site, approximately a 500- by 500-pixel subset of the appropriate data set was extracted for use in analysis (approximately 225 km^2 for the TM scenes and 100 km^2 for the SPOT scenes). Comparisons between the two approaches were conducted for all four study sites using a supervised approach to classification. Testing of SATFE's unsupervised mode was less extensive and rigorous, consisting of a comparison of its results to those of the supervised approaches solely for study site #1.

Approaches to Training Data Extraction and Analysis

For the traditional approach, training areas for each study site were delineated manually using a Gould DeAnza (model FD5000) color graphics display. An iterative procedure was utilized to develop the final set of summary statistics used as input to the classifier. It involved the creation, analysis, and refinement of these statistics using a variety of statistical and graphical analyses (transformed divergence, training area classification accuracy assessment, histograms, and scatter plots of training field data).

Training field extraction using SATFE consisted of manual delineation, on the same display hardware, of a single seed coordinate for each training field and the specification of extraction parameters for each seed. Specifications of the parameters for each seed were based upon various combinations of the following sources of information:

- general knowledge of the spectral variability of the cover type being represented, the spectral variability of the image data used for training, and the spectral bands used for training;
- visual assessment of the size, shape, and orientation of the object in which the seed was placed and assessment of the homogeneity of the image data for that object; and
- preliminary calculation of the summed variance for objects in a small number of evaluation areas.

The number and location of the initial set of training fields for each study site were identical for both the traditional and non-traditional procedures (the seed pixel for a field in the non-traditional approach was contained within the multi-coordinate polygon for that field in the traditional approach). Additionally, in order to reduce the possibility of order- or analyst-related bias in the testing methodology, the order of training field delineation — traditional first or semi-automated first — was altered as each of the four study areas was processed, and the same image analyst was employed in training data extraction and analysis procedures for all four study sites.

Analysis of the raw training field data produced by SATFE was carried out using the SATFA procedure. Here again, the entire analysis scenario was an iterative one, with SATFA executed repetitively until a satisfactory set of summary statistics was produced.

The summary statistics produced as output for both approaches for each study site were used as input to a modified maximum likelihood classifier (patterned after Addington (1975)) implemented on an IBM-AT microcomputer available at the University of Wisconsin-Madison's Environmental Remote Sensing Center (ERSC). With the two SPOT scenes, all three multispectral bands were used in training data extraction and image classification. For the TM data, on the other hand, processing was limited to bands 1, 3, 4, 5, and 7 in an attempt to reduce data dimensionality without significantly impacting information content. The level of detail extracted during classification of each test image was based upon scene-specific physiography and cover type differentiability. The information classes utilized for each site are listed in Table 1.

Constraints on the Classification Comparisons. It should be emphasized that comparisons of the results of the two approaches were conducted with only the original sets of training data available for use in classification. Re-training and the multiple iterations of a full supervised approach were not conducted so as to limit the amount of subjectivity involved in the testing procedures. Thus, the results do not indicate the maximum classification accuracy possible with *either* approach. Rather, they indicate the classification accuracy attained with each approach using the same source of training data.

Table 1. List of the Information Classes Used in the Classification of Each Study Site.

STUDY SITE:	CLASSES:
#1	urban, corn, continuous cover agriculture, bare soil, upland hardwood, red pine, jack pine, wetland, lowland hardwood, water, barren
#2	barren, urban/commercial, urban/residential, upland forest, wetland, water, continuous cover agriculture, bare soil
#3	wetland, water, jack pine, red pine, seedlings, swamp conifer, hardwoods, agriculture, urban
#4	same as study site #3

RESULTS

Accuracy

Results from the accuracy comparisons of the traditional versus the non-traditional training procedures are presented in Table 2. For the supervised approach, no significant differences in classification accuracy (as measured by KHAT) were observed for study sites #2, 3, and 4: standardized Z-scores for the KHAT comparisons were 0.90, 0.35, and 0.77, respectively. For study site #1, however, KHAT for the non-traditional procedures was actually significantly greater than that for the traditional procedures ($Z = -2.91$).

Comparisons of the non-traditional unsupervised procedure with the traditional supervised procedure again showed no significant difference in classification accuracy ($Z = 0.19$). Of interest, however, is the fact that the respective accuracies of the supervised and unsupervised procedures for the non-traditional approach were significantly different ($Z = 3.06$). A detailed evaluation of results from the unsupervised procedure indicated that, while it functioned relatively well in terms of overall classification accuracy (at least in comparison to the traditional supervised approach), it failed to accurately discriminate classes occurring infrequently in the imagery, presumably because of its sampling approach to cluster extraction. Thus, the average by-class accuracy of the unsupervised approach was noticeably lower that that of either supervised procedure.

Because of this sampling problem, a hybrid supervised-unsupervised approach was examined. In this hybrid approach, 12 supervised training fields (out of a total of more than 75) representing infrequently occurring classes were imported from the supervised procedure, combined with the unsupervised clusters, and used in image classification. Accuracy for the augmented classes increased, as did overall accuracy — to the point that it was no longer significantly different from that obtained with the non-traditional supervised procedure ($Z = 0.00$).

Efficiency

Results from the efficiency comparisons of the two procedures are presented in Table 3. The number of points digitized during training field delineation and the time requirements for the

TABLE 2. SUMMARY OF THE STATISTICAL DIFFERENCES IN ACCURACIES OF THE TRADITIONAL AND NON-TRADITIONAL (SAFTE/SATFA) APPROACHES FOR EACH STUDY SITE (SPV: SUPERVISED; UNS: UNSUPERVISED; HYB: HYBRID).

Study Site/ Approach	Traditional (K_1)		Non-Traditional (K_2)			
	% Correct	KHAT	% Correct	KHAT	Z-Score	δ KHAT*
#1/SPV	70	0.66	77	0.74	−2.91[a]	±0.052
#2/SPV	75	0.71	73	0.68	0.90	±0.060
#3/SPV	84	0.81	83	0.80	0.35	±0.047
#4/SPV	76	0.72	74	0.70	0.77	±0.054
#1/UNS	70	0.66	71	0.65	0.19	±0.054
#1/HYB	70	0.66	77	0.73	−2.91[a]	±0.052

*KHAT difference (K_1 - K_2) necessary to indicate significance at the 95 percent level of confidence
[a]indicates a significant difference at the 95 percent level of confidence (Z-scores greater than 1.96 or less than −1.96 indicate significant differences at this level of confidence)

TABLE 3. COMPARISONS OF THE EFFICIENCY OF THE TRADITIONAL AND NON-TRADITIONAL (SATFA/SATFE) SUPERVISED TRAINING DATA EXTRACTION AND ANALYSIS APPROACHES FOR ALL STUDY SITES.

Study Site	Traditional:			Non-Traditional:		
	# Points Digitized	Training Time	Analysis Time	# Points Digitized	Training Time	Analysis Time
#1	506	60 min.	75 min.	77	30 min.	45 min.
#2	562	60 min.	75 min.	71	30 min.	45 min.
#3	531	45 min.	90 min.	81	30 min.	30 min.
#4	676	135 min.	75 min.	90	45 min.	60 min.

training data extraction and analysis phases of the procedures were markedly reduced with the non-traditional approach for all four study sites. Ratios of the number of points digitized in the two procedures were on the order of 6:1 in favor of SATFE, while time savings of from 20 to 60 percent were achieved with the SATFE/SATFA approach as a whole. The relative simplicity of implementing the non-traditional approach was a further advantage in this context. Use of the new procedures was straightforward, requiring very little prior knowledge of the techniques for parameter specification and permitting substantial reductions in the number of repetitive decisions required of the analyst.

ADDITIONAL TESTING

Figures 3a and 3b are "before" and "after" images, respectively, illustrating image output for the linear and concentric training field growth strategies for several training fields in study site #1. Training data extraction ranging from intricate linear objects to large homogeneous objects was accomplished using these two growth strategies.

Concerns about the possible bias due to spatial autocorrelation in the semi-automated field growth strategies motivated the comparison presented in Table 4. Listed herein are the summed variances (in Digital Number [DN] units) for several traditionally and non-traditionally derived training fields from study site #1. In almost every case, the SATFE-derived summed variances are lower. This indicates the presence of additional autocorrelation bias in the SATFE-derived fields, the presence of multiple spectral sub-classes in the traditionally derived fields, or a combination of both.

DISCUSSION

ADVANTAGES

By virtue of its design and implementation, the non-traditional procedures have the potential to solve several of the problems encountered with the use of the traditional forms of supervised and unsupervised image analysis with second generation satellite imagery:

- Because only one pixel must be located in the image, SATFE reduces the potential for ground-to-image misregistration and may produce dramatic reductions in the amount of analyst time required for training field delineation (the program, rather than the analyst, delineates the fields); thus, SATFE, used in conjunction with SATFA, may eliminate much of the complexity involved in training on spectrally and spatially complex imagery.
- Actual image data values for each field or cluster are retained in the output from SATFE, not just their statistical summary, so the statistical properties of the fields can be examined and their spectral homogeneity ensured.
- In contrast to the algorithms in many traditional unsupervised procedures which cluster using rectangular windows of image data, cluster extraction using SATFE's linear growth strategy is not biased toward the exclusion of objects of a particular shape or orientation: fields or clusters may grow out along linear features as easily as within rectangular features; the importance of adequately characterizing such features spectrally has increased with the improved spatial resolution of second generation data.
- In the supervised mode, and, if desired, in the unsupervised mode, each seed pixel specified in SATFE may possess a unique variance threshold (and other extraction parameters) such that, during training data extraction, the procedure can be sensitive to differences in within-class variation in spectral response that are often observed among cover types (e.g., water versus urban) in remotely sensed data.
- The actual geographic locations of the spectral clusters selected during unsupervised processing can be identified within the remotely sensed data by means of the output image data file; this is of particular importance during the process of post-classification labeling.
- SATFA may be used very effectively to analyze the large numbers of training fields which may be required to adequately characterize the spectral variability of second generation satellite data; manual analysis of such data can be exceedingly tedious.

DISADVANTAGES

The semi-automated procedures also have several potential limitations:

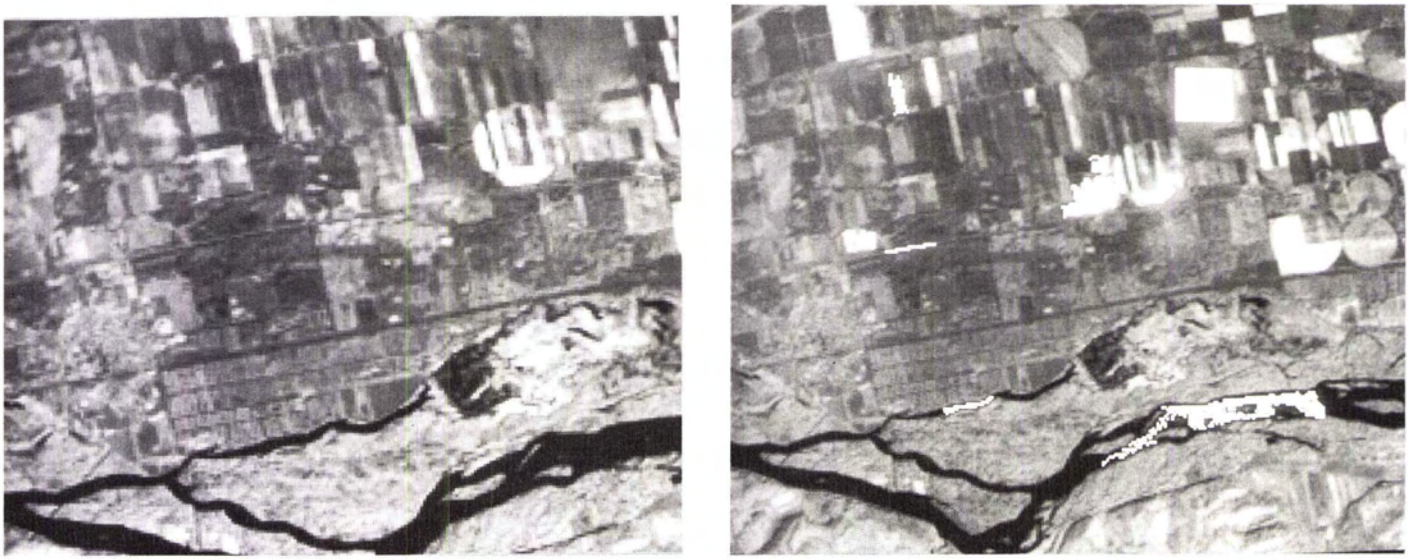

FIG. 3. Illustration of the two growth strategies for the semi-automated training field extractor (SATFE) showing actual image output for example fields for each strategy: (a) TM band-4 for a portion of study site #1 without training field locations; (b) same image as in (a) with training fields overlayed as white onto the original imagery.

TABLE 4. COMPARISON OF THE SUMMED VARIANCES FOR SEVERAL TEST FIELDS EXTRACTED FROM STUDY SITE #1 USING THE TRADITIONAL AND NON-TRADITIONAL APPROACHES (DN UNITS).

	Traditional:		Non-Traditional:	
Class	# Pixels	Variance	# Pixels	Variance
urban	24	266	40	84
corn	38	23	48	19
continuous cover	25	29	31	24
bare soil	28	126	25	65
upland hardwood	59	28	100	20
red pine	70	13	125	22
wetland	18	68	16	39
water	106	10	250	4
barren	9	432	10	109

- Even when implemented in the unsupervised mode, SATFE does not provide for complete sampling of an image; thus, some potentially important spectral information may be overlooked during processing; however, this is a limitation of traditional supervised approaches, as well.
- The data merging criteria implemented in SATFA, while generally applicable to training data analysis, may not constitute the best possible data analysis approach in every situation.
- The per-point framework of the entire approach limits its ability to fully exploit the additional spatial and contextual information that is available from second generation satellite imagery.
- Because of the nature of the pixel extraction techniques in SATFE, the variance of a field or cluster may be underestimated, particularly with rectangularly shaped objects.

SPATIAL AUTOCORRELATION

The latter disadvantage is potentially the most serious technical problem with the SATFE procedure. Underestimation of field variance may result from the fact that contiguous, non-independent (positively autocorrelated) pixels are used for estimation (Campbell, 1981; Cliff and Ord, 1981; Basu and Odell, 1974). Further, depending upon the size and shape of the object (and the specified maximum field size), only the most similar of those contiguous pixels may be used for estimation (in the linear growth case). Within linearly or irregularly shaped objects, the bias introduced as a result of using this approach is probably not much different from that associated with traditional contiguous supervised training because the procedure would likely select nearly the same pixels that an image analyst would select (given sufficient time and patience to do so). With rectangular or other regularly shaped objects, however, this approach may introduce additional bias into the variance estimation process because only the most similar pixels within an object might be included in the field.

Knowledge of this potential source of bias is important for appropriate and judicious use of the SATFE procedure. Quantitative compensation for this bias, however, is not necessary in order for the procedure to be useful. Acceptably accurate results can be obtained with the new procedure, consistently and repeatedly. Hence, the bias in variance estimation appears to be of little or no consequence relative to the dramatic improvements in overall classification efficiency realized. The success of the non-traditional approach in three relatively diverse areas (an agricultural area, an urban area, and a forested area) and using data from two sensors (SPOT and TM) supports this contention.

OVERALL ASSESSMENT OF ADDED UTILITY

The improved spatial resolution of second generation imagery has increased the need for extracting representative training data for many small, linear, or irregularly shaped objects which were not even detected by earlier, first generation satellite sensors. Even relatively large objects, such as certain agricultural fields, can no longer necessarily be considered as spectrally uniform for purposes of training field delineation. The spectral sub-components of these objects may be resolved and may contain meaningful spectral information which might only serve to confuse any attempt at spectral characterization of the object as a whole. With SATFE, extraction of such detailed information may be less prone to error because extraction is driven by the data, according to its contribution to a field's variance, rather than by the error-prone digitization of an analyst.

Thus, with second generation imagery, it may no longer be possible to simply "blindly" digitize training fields in digital imagery based upon specifications of the ground coordinates for those fields. Depending upon the cover categories involved and the physiography of the area under study, these areas may in fact contain several of the resolved spectral sub-components of the identified ground categories, and thus might not provide an appropriate spectral characterization of the object as originally identified. SATFE is useful under these circumstances, both for improving the homogeneity of training data (because field growth is restricted to the spectral sub-class in which a seed pixel is located) and for reducing geometric registration problems between ground and image coordinates (because only one coordinate needs to be registered).

Training field delineation with the traditional approach was also a much more tedious and time-consuming process than with the SATFE-based approach. This was particularly the case for intricate, linear or irregularly shaped objects. Parameter specification within SATFE required some time and thought; however, the flexibility built into the algorithm's parameterization process reduced the importance of the specification of any single parameter. Had the procedure been heavily reliant on the accurate specification of a single, complex parameter, much of the efficiency of the technique would undoubtedly have been lost. Furthermore, systematic testing of the sensitivity of SATFE to changes in field extraction parameters indicated that judicious use of the relative-variance-increase ratio and the automatic-variance-increase-value permit fairly consistent results to be obtained even while varying the variance threshold and maximum field size parameters (Buchheim, 1988).

The potential improvements in processing efficiency with the non-traditional approach thus appear to be substantial. However, processing beyond the level conducted herein would likely be required to maximize the accuracy of the entire per-point image classification process. This would involve additional training field delineation and training field analysis. The advantage of the non-traditional approach for such a complete analysis and classification procedure, however, would likely only be increased because additional analyst and processing time savings would accrue with each iteration.

Depending on the nature of the study area and the application, the number of spectral classes required to adequately characterize the spectral variability of a second generation satellite data set may be substantially greater than was ever necessary previously (Malila, 1985; Price, 1984). In such situations, the non-traditional approach described herein, which allows both relatively quick and easy extraction of training data and simple and efficient analysis of these data, provides a substantial advantage in utility compared to traditional, more manual approaches.

Improved consistency and repeatability in training data extraction and analysis is yet another issue of relevance. SATFE potentially can provide a more objective, consistent, and thereby repeatable means for extracting training data. With SATFE's linear growth strategy, training sample extraction is based upon a pixel's spatial location and contribution to field variance (and other relatively objective measures). An analyst's subjective determination of a pixel's class identity, which may be based entirely upon evaluation of a particular band subset or image enhancement, need no longer serve as the basis for training field extraction decisions. With the new approach, there is certainly still some "art" involved in locating seed pixels and specifying extraction parameters; however, it is potentially a much less subjective procedure than that currently in use.

Several of the charcteristics of SATFE's unsupervised mode render it expecially useful for clustering second generation image data. No rectangular structure is imposed onto the clustering process in SATFE. Seeds located near object boundaries generally produce spectral information for the object in which they are located, rather than information representing some unknown combination of spectral responses from objects on either side of the boundary. In addition, maintenance of the raw data comprising each cluster aids in determining cluster spectral homogeneity. Examination of the spatial locations of the clusters by means of the output image data file is also an extremely useful characteristic – the tedium of post-classification labeling is greatly reduced as a result.

FUTURE DIRECTIONS

Although not specifically demonstrated in this research, one of the biggest advantages of SATFE may be its potential for integration with existing GIS processing capabilities and databases. Seed pixel coordinates, and conceivably even certain training data extraction parameters, could be derived from existing GIS data layers (e.g., the sizes and broad cover categories of potential training areas could be extracted from GIS data and used to set the maximum field size and variance threshold parameters for SATFE). Using this GIS input, SATFE could be used either for general land-cover classification for an entire region or to statistically characterize spectral information at specific locations within a region for change detection studies or special purpose land-cover inventories.

The SATFE and SATFA procedures can also be thought of as a step toward the establishment of an expert system for training data extraction and analysis. They certainly were not designed as such and have no capability for maintenance of any form of knowledge base, yet the principles upon which they were designed may be sufficiently fundamental (in terms of their global applicability to training data extraction and analysis) that they may represent a useful step in this direction. They certainly demonstrated an ability to increase automation, improve consistency, and decrease analyst time demands for these procedures.

ACKNOWLEDGMENTS

This research was supported primarily through funding provided by the Water Resources Division of the United States Geological Survey and the Wisconsin Department of Transportation for "Use of Landsat Thematic Mapper Data to Develop Land Cover Parameters for Improved Hydrologic Modeling of Wisconsin's Driftless Area." Additional support was provided by the College of Agricultural and Life Sciences, School of Natural Resources, and the Institute for Environmental Studies, University of Wisconsin-Madison, and by a grant from the William and Flora Hewlett Foundation.

REFERENCES

Acevedo, W., J. S. Buis, and R. C. Wrigley, 1984. Changes in Classification Accuracy Due to Varying Thematic Mapper and Multispectral Scanner Spatial, Spectral, and Radiometric Resolution. *Proc. 18th Int'l Symp. on Remote Sensing of Environ.*, Ann Arbor, Michigan, pp. 27–44.

Addington, J. D., 1975. A Hybrid Classifier Using the Parallelepiped and Bayesian Techniques. *Proc. ACSM-ASP Annual Meeting*, pp. 772–784.

Basu, J. P., and P. L. Odell, 1974. Effect of Intraclass Correlation Among Training Samples on the Misclassification Probabilities of Bayes Procedure. *Pattern Recognition* 6(1): 13–16.

Buchheim, M. P., 1988. *Non-Traditional Approaches to Classification of High Resolution Satellite Imagery in a Microcomputer Environment.* Unpublished PhD dissertation, University of Wisconsin-Madison, Madison, Wisconsin. 497 p.

Buchheim, M. P., A. L. Maclean, and T. M. Lillesand, 1985. Forest Cover Type Mapping and Spruce Budworm Defoliation Detection Using Simulated SPOT Imagery. *Photogrammetric Engineering and Remote Sensing* 51(8): 1115–1122.

Campbell, J. B., 1981. Spatial Correlation Effects Upon the Accuracy of Supervised Classification of Land Cover. *Photogrammetric Engineering and Remote Sensing* 47(3): 355–363.

Cliff, A. D., and J. K. Ord, 1981. *Spatial Processes: Models and Applications*. Pion, Limited, London. 266 p.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. *Educ. Psych. Meas.* 20: 37–46.

Congalton, R. G., R. G. Oderwald, and R. A. Mead, 1983. Assessing Landsat Classification Accuracy Using Discrete Multivariate Analysis Statistical Techniques. *Photogrammetric Engineering and Remote Sensing* 49(12): 1671–1678.

Fleiss, J. L., J. Cohen, and B. S. Everitt, 1969. Large Sample Standard Errors of Kappa and Weighted Kappa. *Psychological Bulletin* 72(5): 323–327.

Haack, B., N. Bryant, and S. Adams, 1987. An Assessment of Landsat MSS and TM Data for Urban and Near-Urban Land-Cover Digital Classification. *Remote Sensing of Environ.* 21(2): 201–213.

Hopkins, P. F., A. L. Maclean, and T. M. Lillesand, 1988. Assessment of Thematic Mapper Imagery for Forestry Applications Under Lake States Conditions. *Photogrammetric Engineering and Remote Sensing* 54(1): 61–68.

Horler, D. N. H., and F. J. Ahern, 1986. Forestry Information Content of Thematic Mapper Data. *Int'l J. of Remote Sensing* 7(3): 405–428.

Irons, J. R., B. L. Markham, R. F. Nelson, D. L. Toll, D. L. Williams, R. S. Latty, and M. L. Stauffer, 1985. The Effects of Spatial Resolution on the Classification of Thematic Mapper Data. *Int'l J. of Remote Sensing* 6(8): 1385–1403.

Lathrop, R. G., Jr., T. M. Lillesand, and B. S. Yandell, 1987. An Evaluation of Thematic Mapper Data for Forest Cover Mapping in Northern Wisconsin. *Proc. Pecora XI Symp.*, Sioux Falls, South Dakota, pp. 386–395.

Latty, R. S., 1984. Scene Segmentation Through Region Growing. *Proc. 10th Symp. on Machine Processing of Remotely Sensed Data*, West Lafayette, Indiana, pp. 305–314.

Latty, R. S., R. Nelson, B. Markham, D. Williams, D. Toll, and J. Irons, 1985. Performance Characteristics between Information Extraction Techniques Using Variable Spatial Resolution Data. *Photogrammetric Engineering and Remote Sensing* 51(9): 1459–1470.

Malila, W. A., 1985. Comparison of the Information Contents of Landsat TM and MSS Data. *Photogrammetric Engineering and Remote Sensing* 51(9): 1449–1457.

Markham, B. L., and J. R. G. Townshend, 1981. Land Cover Classification Accuracy as a Function of Sensor Spatial Resolution. *Proc. 15th Int'l Symposium on Remote Sensing of Environ.*, Ann Arbor, Michigan, pp. 1075–1091.

Parks, N. F., G. W. Petersen, and G. M. Baumer, 1987. High Resolution Remote Sensing of Spatially and Spectrally Complex Coal Surface Mines of Central Pennsylvania: A Comparison between Simulated SPOT MSS and Landsat-5 Thematic Mapper. *Photogrammetric Engineering and Remote Sensing* 53(4): 415–420.

Price, J. C., 1984. Comparison of the Information Content of Data from the Landsat-4 Thematic Mapper and the Multispectral Scanner. *IEEE Trans. Geo. Remote Sensing* GE-22(3): 272–281.

Rosenfield, G. H., and K. Fitzpatrick-Lins, 1986. A Coefficient of Agreement as a Measure of Thematic Classification Accuracy. *Photogrammetric Engineering and Remote Sensing* 52(2): 223–227.

Swain, P. H., 1978. The Quantitative Approach: Concept and Rationale. *Remote Sensing: The Quantitative Approach* (P. H. Swain and S. H. Davis, eds.). McGraw-Hill, New York, N.Y. pp. 1–20.

Williams, V. L., W. R. Philipson, and W. D. Philpot, 1987. Identifying Vegetable Crops with Landsat Thematic Mapper Data. *Photogrammetric Engineering and Remote Sensing* 53(2): 187–191.

Woodcock, C. E., and A. H. Strahler, 1987. The Factor of Scale in Remote Sensing. *Remote Sensing of Environ.* 21(3): 311–332.