An Empirical Investigation of Image Resampling Effects Upon the Spectral and Textural Supervised Classification of a High Spatial Resolution Multispectral Image

O. Dikshit and D.P. Roy

Abstract

An empirical examination of the effects of resampling upon the spectral and textural classification of a high spatial resolution multispectral image is described. Spectral and textural maximum-likelihood classifications are performed upon unresampled and upon bilinear and cubic convolution resampled versions of the image. The texture classifications use spectral training data and additional texture training data calculated using the grey-level difference histogram algorithm. The resampling algorithms increase the overall classification accuracies in a statistically significant manner which vary between individual classes. These results are explained by consideration of the interaction between the local smoothing properties of the resampling algorithms and the grey-level structure of the image. The results indicate that spectral and textural classification procedures may be applied to images, after they have been resampled, without a reduction in the classification accuracy, and that textural classification procedures should use class training statistics collected from the resampled image.

Introduction

Automated image classification procedures are required to routinely process the increasingly large remotely sensed data sets that are becoming available. The classified images produced by these procedures must be georeferenced to allow them to be compared temporally and to produce synoptic land-cover maps. Conventional classification procedures assume that surface features located in the image can be discriminated systematically by examination of their spectral signatures. However, because of sensitivity to externally induced spectral variations (e.g., atmospheric and viewing geometry effects) and because of intrinsic variations associated with surface features (e.g., vegetation varies with maturity, canopy closure, and moisture stress) (Wharton, 1989), it is necessary to redefine the spectral signatures (i.e., class training statistics) each time a new image is classified. Under an automated classification scheme, the class training statistics may be collected by the simultaneous examination of the image and coregistered ground reference data that define the locations of representative land-cover units. The image and ground reference data may be coregistered most conveniently by georeferencing the image into the same coordinate system as the ground reference data. This is because ground reference data are conventionally collected in an Earth-based coordinate system and because well established image georeferencing techniques may be used. This approach also ensures that the resultant classified image will be georeferenced.

Image georeferencing is performed by mapping a regular grid of Earth-based coordinates into the remotely sensed image. The coordinates of the regular grid define the locations of the pixel centers in the georeferenced image and may be mapped into the sensed image using non-parametric techniques (i.e., ground control collected from the image and from a map) or using parametric techniques (i.e., a model of the sensing geometry) (Baker *et al.*, 1975). The grey-level values of each georeferenced image pixel are then interpolated from the grey-level values of a local neighborhood of pixels located in the sensed image. This interpolation process is termed image resampling and is known to modify the grey-level structure and, therefore, the appearance of the image. Resampling effects are most obvious where neighboring image grey-level values vary rapidly, for instance, within texturally diverse regions and along boundaries between spectrally distinct regions of the image. For these reasons, it is expected that image classification results may be different when georeferenced (i.e., resampled) images are used. This may be particularly likely when image texture information is incorporated into the classification procedure, because resampling operations have been shown to modify the values of texture measures in a non-uniform manner (Roy and Dikshit, 1994).

The objective of this paper is to investigate the impact of conventional image resampling techniques upon the spectral and textural classification of a remotely sensed image. Past workers have observed only insignificant differences between the classification accuracies of satellite images resampled using conventional techniques (e.g., Ferneyhough and Niblack, 1977; Etheridge and Nelson, 1979; Logan and Strahler, 1979; Smith and Kovalick, 1985). It has been suggested that this may have been due to a failure to classify images that con-

O. Dikshit is with the Indian Institute of Technology, Department of Civil Engineering, Survey, Photogrammetry and Remote Sensing, Kanpur, U.P. 208016, India.

D.P. Roy was with the NERC Unit for Thematic Information Systems, Department of Geography, University of Reading, Reading RG6 2AB, U.K. He is presently with the Department of Geography, University of Maryland and NASA Goddard Space Flight Center, Code 920.2, Greenbelt, MD 20771.

Photogrammetric Engineering & Remote Sensing, Vol. 62, No. 9, September 1996, pp. 1085–1092.

0099-1112/96/6209-1085$3.00/0

tain many boundaries between spectrally distinct classes where the resampling effects will be most evident (Atkinson, 1988) and because of insufficiently detailed ground truth information (Shlien, 1979). In accordance with these suggestions, a high spatial resolution test image that contains many class boundaries and a ground reference map depicting a corresponding level of detail are used in the experiments described.

Maximum-likelihood classifications of the test image and resampled versions of the test image are performed using spectral class training data and additional image texture information. As resampling effects are expected to be a function of the image grey-level structure, the classifications are performed using a number of texturally distinct classes. Because there may be situations when it is more convenient to classify resampled images using training statistics collected from the unprocessed image, the impact of using training statistics collected from the original test image and from resampled versions of the test image are examined. The accuracy of the classifications are compared and the significance of any differences observed between them are assessed using test statistics derived from kappa analyses.

Overview of Commonly Used Resampling Techniques

The most commonly used resampling techniques are cubic convolution, bilinear, and nearest-neighbor. Generally, cubic convolution resampling has the potential to reconstruct exactly any second-degree polynomial, whereas bilinear can exactly reconstruct at most a first-degree polynomial, and nearest-neighbor can only produce exact reconstruction when the image has a constant grey-level value (Keys, 1981). Nearest-neighbor and bilinear resampling techniques are heuristic, whereas cubic convolution is a limited span approximation of the theoretically optimal sinc resampling function which, in practice, cannot be used as it requires an infinitely large pixel neighborhood (Shlien, 1979).

Nearest-neighbor resampled pixels are allocated a grey-level value equal to the grey-level value of the nearest pixel in the original image. Nearest-neighbor resampling is often used because it is computationally simple and because it does not introduce new grey-level values into the resampled image. The disadvantage of the technique is that it introduces pixel-level geometric discontinuities (up to a maximum of $\sqrt{2}/2$ of a pixel), changing the textural properties of the resampled imagery, making it appear visually "blocky" or "dappled" (Ferneyhough and Niblack, 1977). For this reason, it is recommended that nearest-neighbor resampling is not used for applications where the textural properties of the image are important. Nearest-neighbor resampling is not considered further in this paper.

Bilinear resampling fits a hyperbolic paraboloid through four neighboring pixel values in the original image to estimate the resampled pixel value (Castleman, 1979). Bilinear resampling has the effect of a low-pass filter and gives a visually smooth image without the geometric discontinuities of nearest-neighbor resampling.

Cubic convolution was originally developed for resampling of Landsat MSS imagery (Riffman and McKinnon, 1974) and is a truncated approximation of the sinc function (Foley *et al.*, 1990). The cubic convolution resampler approximates the sinc function over a 4- by 4-pixel neighborhood and produces a visually smooth image except in regions of high grey-level contrast where the contrast may be enhanced (Park and Schowengerdt, 1983). Park and Schowengerdt (1983) formulated the cubic convolution parametrically and demonstrated that the mean square error between the grey-level values of the original and the resampled image is minimized using the cubic convolution, followed by the bilinear and then the nearest-neighbor resamplers, respectively. Similar results were observed by Shlien (1979) and Keys (1981).

Texture Calculating Algorithm

Incorporation of texture information has been advocated as a means of improving spectrally based classification accuracies (Marceau *et al.*, 1990; Wang and He, 1990), especially in high spatial resolution imagery where classification accuracies remain generally low (e.g., Cushnie, 1987; Jensen and Hodgson, 1987). Image texture may be described qualitatively in many ways but is most often described over a fine to coarse continuum. As spatial patterns become more definitive in an image and extend over many pixels, the texture will become coarse (Haralick *et al.*, 1983). Texture algorithms extract quantitative measures of the image texture measured over a small window located within the image (e.g., Haralick, 1979; Gool *et al.*, 1985). The suitability of a particular texture measure for image classification and image segmentation procedures is dependent upon the grey-level structure of the image (e.g., Sali and Wolfson, 1992). However, recent work has shown that the one-dimensional grey level difference histogram (GLDH) algorithm yields classification accuracies similar to computationally more expensive second-order algorithms (such as neighboring grey-level dependence matrices) when applied to airborne remotely sensed terrain images (Dikshit, 1992).

The GLDH algorithm (Conners and Harlow, 1980) computes a histogram of absolute grey-level difference values along a specified direction and then converts the histogram into a probability distribution. Texture information is extracted from the probability distribution to give five texture features: inertia, mean, entropy, energy, and inverse difference moment. Of these, the energy texture feature was used in the experiments described in this paper because it gave the highest overall classification accuracy when applied to the experimental test data.

The grey-level difference histogram is found by counting the frequency of occurrence of differences in grey-level values for a given pixel displacement: i.e.,

$$\delta(x, y) = |f(x, y) - f(x + \Delta x, y + \Delta y)|$$

where $\delta(x, y)$ is the grey-level difference, $f(x,y)$ is the grey-level value of the digital image at (x,y), and Δx, Δy is the pixel displacement.

Let p be the probability density function (PDF) of $\delta(x, y)$ calculated by normalizing the frequencies of the difference histogram by dividing each frequency by the total number of occurrences. If the number of grey levels in the image is n_g, then the PDF takes the form of an n_g-dimensional vector whose i^{th} component is the probability that $\delta(x, y)$ will have a grey-level difference value i. The GLDH energy texture feature is defined (Conners and Harlow, 1980) as

$$\text{Energy} = \sum_{i=0}^{n_g-1} p(i)^2.$$

Generally, heterogeneous, structured, or visually coarse images are indicated by low energy values and homogeneous, unstructured, or visually fine images are associated with higher energy values.

Bilinear and cubic-convolution resampling operations have been shown empirically to modify the values of GLDH texture features extracted from a high spatial resolution image (Roy and Dikshit, 1994). The degree of modification was found to be dependent upon the GLDH texture feature measure, the resampling algorithm, and the textural properties of the image. It was found that GLDH texture features that measure heterogeneity had reduced average values, whereas the GLDH texture features that measure homogeneity, such as the energy texture feature measure, had increased average values after application of bilinear and cubic-convolution resamplers. This was attributed to the local smoothing properties

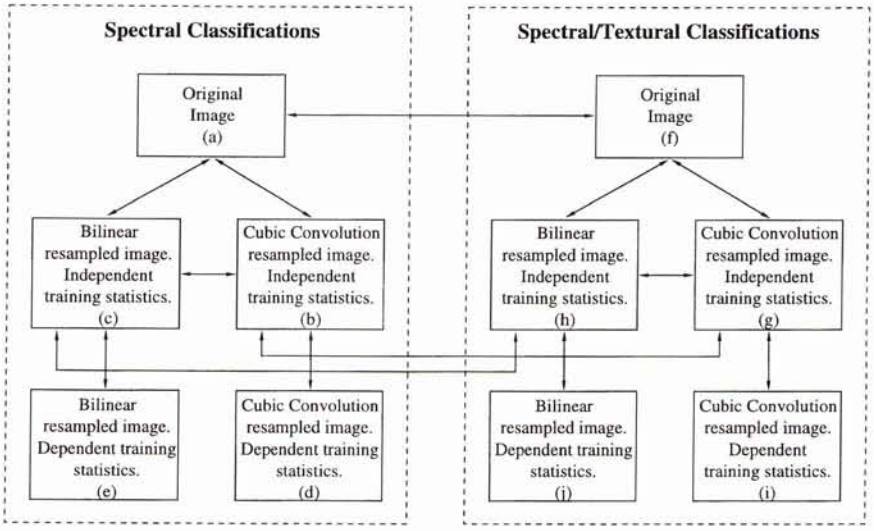

Figure 1. Spectral and spectral/textural classifications and the pairwise inter-classification significance tests performed between them.

of the resamplers. Changes in the GLDH texture feature values were generally observed to be more pronounced in those parts of the image characterized by texturally fine cover types and in images resampled using the bilinear resampler.

Study Site and Test Data

The test image was sensed by a Daedalus AADS1268 Airborne Thematic Mapper (ATM) over Woodwalton Fen, Cambridgeshire, U.K. This site is flat and characterized by diverse semi-natural cover types which exhibit complex spatial distributions and provide visually discrete textural classes. The image was sensed under calm flying conditions with the ATM S-bend correction facility activated to reduce across track tangential scale distortions (Callison *et al.*, 1987). ATM image bands 5 (0.63 to 0.69 μm), 7 (0.76 to 0.90 μm), and 9 (1.55 to 1.75 μm) were used for classification purposes. This band combination is known to have low inter-band correlation (Morris and Barnsley, 1989) and is established as being suitable for vegetation cover mapping (e.g., Jensen and Hodgson, 1987; Curran and Pedley, 1989). The dimensions of the test image were 220 by 1022 pixels with an approximate nadir pixel dimension of $1.25m^2$.

A ground reference map, produced by large-scale photo-interpretation and frequent ground based survey (Fuller *et al.*, 1986), was manually coregistered with the test image and then rasterized so that it contained the same number of pixels as the test image. The resultant ground reference image was used to define the class training statistics and to assess the classification results. Eight texturally distinct classes were used. The classes, listed in subjectively judged order of increasing coarseness, were water, wet grass mown and grazed (MG), wet grass, fen herbs, sallow scrub, fen reed, fen *phalaris*, and woodland.

Experimental Methodology

Spectral and combined spectral/textural maximum-likelihood classifications were performed on the original and resampled versions of the test image. In total, five spectral and five spectral/textural classifications were performed (Figure 1). In order to examine the impact of resampling upon training statistic collection practices, each resampled image classification was performed twice using *independent* and *dependent* types of training statistics. Independent training statistics were collected from the resampled bands and dependent training statistics were collected from the original unresampled bands.

The spectral classifications used the three image bands, and the spectral/textural classifications used an additional band of GLDH energy texture feature values. The energy texture band was calculated from the first principal component of the three image bands scaled over an 8-bit range. A square window was translated across the principal component image in steps of one pixel, and the values of the energy texture feature were calculated in the left-diagonal, the right-diagonal, and the horizontal and vertical pixel grid directions. The direction average of these four values was then assigned to the center of each window. The values of the energy texture feature were calculated using a one-pixel displacement over a window 15 pixels square. This particular combination of texture parameters was selected because it was found to give high overall classification accuracies when used in conjunction with the original test image bands (Dikshit, 1992).

When resampled classifications were performed, the spectral classifications used resampled versions of the original image bands and the spectral/textural classifications used an additional texture band derived from them. The original image bands were resampled with a translational shift of one-half pixel applied in both image axes using the cubic convolution and bilinear resampling techniques. This translational shift corresponds to the zone of maximum reconstruction error in the resampling process (Shlien, 1979) and ensures that the relative phase between the test and the resampled image pixel coordinates is uniform across the image.

Data Analysis

The impact of resampling upon the class training data was investigated prior to examination of the classification accuracies. Average transformed divergence values were computed from the training data used to perform the spectral and the spectral/textural classifications of the original and the resampled versions of the test image. The transformed divergence is a measure of the relative effectiveness of two candidate feature sets for inter-class discrimination (Swain and Davis, 1978) and the average transformed divergence is a measure of inter-class discrimination averaged over a set of classes (Thomas *et al.*, 1987). The average transformed divergence

TABLE 1. AVERAGE TRANSFORMED DIVERGENCE VALUES OF SPECTRAL AND SPECTRAL/TEXTURAL CLASS TRAINING DATA COLLECTED FROM ORIGINAL AND RESAMPLED VERSIONS OF A HIGH SPATIAL RESOLUTION TEST IMAGE (VALUES ARE SCALED OVER 0 TO 100 RANGE).

	Original Image	Cubic Convolution Resampled Image	Bilinear Resampled Image
Spectral training data (image bands 5,7,9)	87	90	92
Spectral and textural training data (image bands 5,7,9, and GLDH energy band)	97	97	98

values were scaled so that they could be interpreted over a 0 to 100 range (Mather, 1987).

The overall classification accuracy and the accuracy of the individual classes were assessed in a conventional manner by examination of classification error matrices. The classification accuracies were also assessed by kappa analyses used to compute *Khat* indices and associated asymptotic variances (Bishop *et al.*, 1975). Pairwise statistical tests were performed to assess the significance of any differences observed between two classifications using a Z statistic (Congalton and Mead, 1986): i.e.,

$$Z_{ab} = \frac{\kappa_a - \kappa_b}{\sqrt{\sigma_a^2 + \sigma_b^2}}$$

where Z_{ab} is the Z statistic for comparison of classifications a and b; κ_a, κ_b are the *Khat* indices of classifications a and b; and σ_a^2, σ_b^2 are the asymptotic variances of *Khat* indices a and b.

The difference between two classifications was considered to be significant at the 95 percent confidence level if the absolute value of the Z statistic exceeded 1.96. The significance tests are illustrated in Figure 1. Significance tests were performed between the different types of spectral classifications and between the different types of spectral/textural classifications. In these tests, the resampled image classifications performed using independent training statistics were compared with the original image classifications, and the bilinear and cubic convolution resampled image classifications were compared. Significance tests were also performed between the resampled images classified using dependent and independent training statistics to assess the classification impact of using training statistics collected from the original rather than the resampled image bands. A small number of significance tests were performed between the spectral and the spectral/textural classifications to verify that the incorporation of image texture information gave significantly improved classification accuracies.

Results

Table 1 shows the average transformed divergence values of the spectral and the spectral/textural class training data collected from the original and resampled versions of the test image. The average transformed divergences of the spectral/textural training data are greater than the average transformed divergences of the spectral training data, indicating that the incorporation of texture information has improved discrimination between the classes. The average transformed divergences of the spectral/textural training data are very high (97-98) and imply that the incorporation of additional feature sets would not provide significantly improved class discrimination. In general, the training data collected from the resampled images have higher average transformed divergence values than the training data collected from the original image. This implies that the class discrimination will be improved in the resampled images, particularly the bilinear resampled image which has the highest average transformed divergence values.

Tables 2 and 3 show the overall and individual classification accuracies and their associated *Khat* indices and *Khat* standard deviations for the spectral and spectral/textural classifications, respectively. The overall and individual classification accuracies are illustrated in Figures 2 and 3. The inter-classification significance tests are illustrated in Figure 1, and the test results are summarized in Table 4. The tests have been divided into three groups: tests performed between the different types of spectral classifications, tests performed between the different types of spectral/textural classifications, and tests performed between these two groups.

The following observations are made from the comparisons between the spectral and the spectral/textural classifications:

(1) The overall classification accuracy of the original image increases by 9.0 percent when it is classified using spectral/textural training statistics rather than spectral training statistics. For some classes, such as the fen reed class, the increase in classification accuracy is considerably greater

TABLE 2. SPECTRAL CLASSIFICATION ACCURACIES, *khat* INDICES, AND *khat* STANDARD DEVIATIONS OF ORIGINAL AND RESAMPLED VERSIONS OF A HIGH SPATIAL RESOLUTION TEST IMAGE.

		(a) Original Image			(b) Cubic Convolution Resampled Image (Independent Statistics)			(c) Bilinear Resampled Image (Independent Statistics)			(d) Cubic Convolution Resampled Image (Dependent Statistics)			(e) Bilinear Resampled Image (Dependent Statistics)		
	Pixels[1]	%Corr[2]	*Khat*[3]	*K*-Std[4]	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std
Water	11609	83.0	0.819	0.0037	82.7	0.816	0.0037	82.8	0.817	0.0037	82.2	0.810	0.0037	81.6	0.804	0.0038
Fen Reed	30612	17.3	0.123	0.0020	18.9	0.132	0.0021	19.7	0.137	0.0021	18.0	0.131	0.0020	18.7	0.139	0.0020
Fen *Phalaris*	8570	52.2	0.489	0.0056	56.3	0.530	0.0056	61.1	0.581	0.0056	52.4	0.495	0.0056	52.8	0.502	0.0056
Wet Grass	11674	40.4	0.338	0.0048	39.6	0.332	0.0048	40.2	0.342	0.0048	44.8	0.383	0.0049	46.6	0.402	0.0050
Wet Grass (MG)	8463	36.9	0.306	0.0056	37.4	0.312	0.0056	37.3	0.313	0.0056	38.9	0.322	0.0057	41.3	0.345	0.0058
Fen Herbs	6271	30.1	0.248	0.0061	30.1	0.254	0.0060	33.9	0.294	0.0062	30.7	0.257	0.0061	31.5	0.268	0.0061
Sallow Scrub	41016	62.9	0.516	0.0028	67.3	0.569	0.0028	71.2	0.617	0.0027	66.2	0.557	0.0028	69.8	0.601	0.0027
Woodland	47617	83.6	0.758	0.0024	84.9	0.779	0.0023	85.4	0.788	0.0022	84.4	0.771	0.0023	85.1	0.783	0.0022
Overall	165832	57.1	0.476	0.0014	59.0	0.498	0.0014	60.7	0.519	0.0014	58.7	0.495	0.0014	60.2	0.513	0.0014

[1]Number of pixels defining class
[2]Percentage of correctly classified pixels
[3]*Khat* index
[4]Standard deviation of *Khat* index

than 9.0 percent. All of the classes, except the water and woodland classes, show significant increases in classification accuracy (Table 4, column Zfa). This finding confirms that the incorporation of image texture generally improves classification accuracy (e.g., Marceau *et al.*, 1990; Wang and He, 1990).

(2) The classification accuracy of the water and woodland classes does not improve under the spectral/textural classification scheme. The reasons for this are unclear but may perhaps be attributed to their textural properties, which lie at either end of the fine-to-coarse texture continuum. Consequently, these classes are likely to be texturally distinct from their immediate surroundings, and, therefore, mixed texture feature values are more likely to be introduced as the texture calculating window is moved across their boundaries. This is a fundamental limitation of all window-based texture calculating algorithms; potential solutions to this problem are discussed elsewhere (e.g., Dikshit, 1992).

(3) The spectral/textural classifications of the resampled images (performed using independent training statistics) show increases in classification accuracy similar to those observed in (1). However, the classification accuracies of the woodland class increase under the spectral/textural classification scheme. This may be because the boundaries of the woodland class are more smooth in the resampled images. In general, the increases in classification accuracy are significant (Table 4, columns Zgb, and Zhc).

The following observations are made from the comparisons within the different types of spectral classification:

(4) The resampled image classifications performed using independent training statistics have higher overall classification accuracies than the original image by up to 3.6 percent. The increases in the overall classification accuracy are significant (Table 4, column Zba and Zca). Individual classification accuracies vary but are significantly higher for the texturally coarse fen reed, fen *phalaris*, sallow scrub, and woodland classes. More texturally homogeneous classes, such as the water and wet grass classes, do not give significantly different classification accuracies after resampling in this experiment.

(5) The overall classification accuracies of the bilinear and cubic convolution resampled images classified using independent training statistics are significantly different (Table 4, Column Zcb). The overall classification accuracy of the bilinear resampled image is higher than the overall classification accuracy of the cubic convolution resampled image by 1.7 percent.

(6) The overall classification accuracies of the bilinear resampled images classified using dependent and independent training statistics are significantly different (Table 4, column Zce). The difference between the overall classification accuracies is 0.5 percent. In most cases, those classes classified using independent training statistics have significantly higher classification accuracies than those classified using dependent training statistics.

(7) Comparison of the cubic convolution resampled image classifications performed using dependent and independent training statistics reveal less significantly different results than those observed in (6) (Table 4, column Zbd). The difference between the overall classification accuracies is not significant and is 0.3 percent.

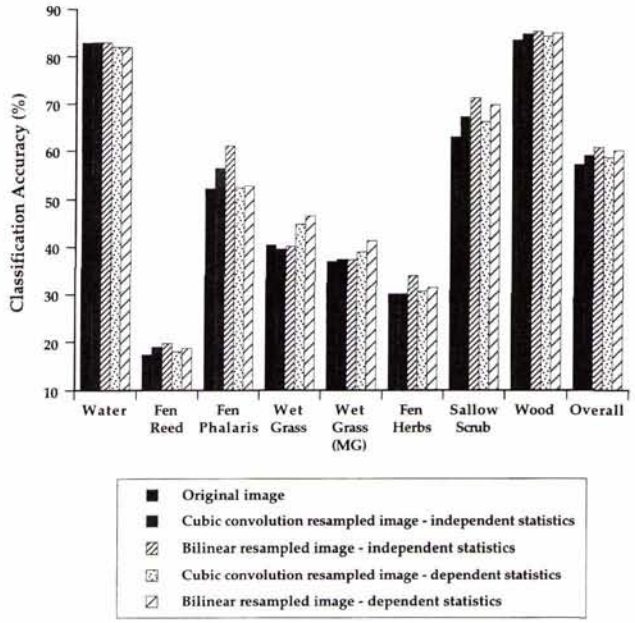

Figure 2. Spectral classification accuracies.

The following observations are made from the comparisons within the different types of spectral/textural classification:

(8) The resampled image classifications performed using independent training statistics have higher overall classification accuracies than the original image by up to 2.5 percent.

TABLE 3. SPECTRAL/TEXTURAL CLASSIFICATION ACCURACIES, *khat* INDICES, AND *khat* STANDARD DEVIATIONS OF ORIGINAL AND RESAMPLED VERSIONS OF A HIGH SPATIAL RESOLUTION TEST IMAGE.

		(f) Original Image			(g) Cubic Convolution Resampled Image (Independent Statistics)			(h) Bilinear Resampled Image (Independent Statistics)			(i) Cubic Convolution Resampled Image (Dependent Statistics)			(j) Bilinear Resampled Image (Dependent Statistics)		
	Pixels[1]	%Corr[2]	*Khat*[3]	*K*-Std[4]	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std	%Corr	*Khat*	*K*-Std
Water	11609	82.2	0.811	0.0037	81.2	0.800	0.0038	81.6	0.804	0.0038	84.1	0.831	0.0036	84.6	0.836	0.0035
Fen Reed	30612	43.8	0.353	0.0029	50.1	0.418	0.0030	51.1	0.429	0.0030	43.1	0.349	0.0029	25.6	0.188	0.0024
Fen *Phalaris*	8570	69.8	0.656	0.0056	66.6	0.629	0.0056	68.7	0.653	0.0055	39.2	0.318	0.0057	11.0	−0.046	0.0039
Wet Grass	11674	53.1	0.505	0.0047	48.7	0.458	0.0047	45.0	0.421	0.0047	21.9	0.179	0.0038	8.3	0.025	0.0026
Wet Grass (MG)	8463	47.9	0.464	0.0054	43.5	0.419	0.0054	42.3	0.406	0.0054	53.1	0.504	0.0056	35.1	0.315	0.0053
Fen Herbs	6271	31.2	0.288	0.0059	29.9	0.274	0.0058	31.6	0.290	0.0059	39.3	0.333	0.0066	38.1	0.273	0.0071
Sallow Scrub	41016	74.4	0.649	0.0027	73.0	0.636	0.0027	72.1	0.629	0.0027	64.4	0.539	0.0028	27.4	0.177	0.0021
Woodland	47617	79.8	0.710	0.0025	85.5	0.786	0.0023	89.1	0.835	0.0021	75.7	0.660	0.0025	77.0	0.659	0.0026
Overall	165832	66.1	0.579	0.0014	67.8	0.598	0.0014	68.6	0.608	0.0014	59.3	0.503	0.0014	43.9	0.330	0.0013

[1]Number of pixels defining class
[2]Percentage of correctly classified pixels
[3]*Khat* index
[4]Standard deviation of *Khat* index

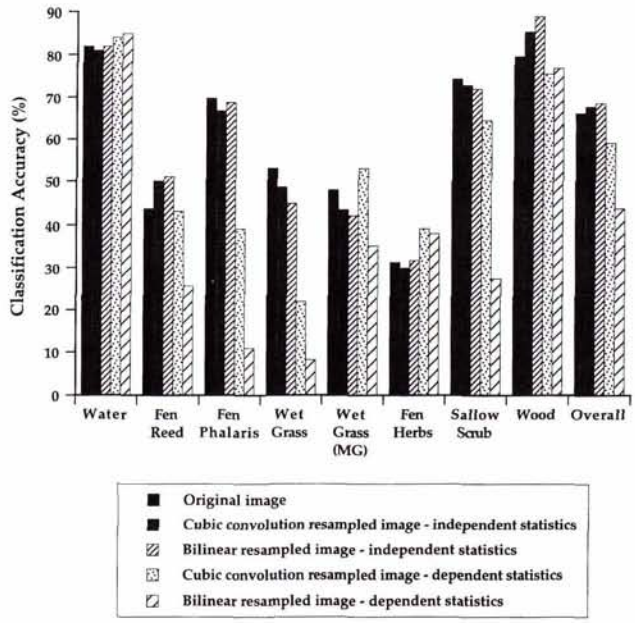

Figure 3. Spectral/textural classification accuracies.

The increases in the overall classification accuracy are significant (Table 4, Column Zgf and Zhf), but there is no clear pattern among the different classes.

(9) The overall classification accuracies of the bilinear and cubic convolution resampled images classified using independent training statistics are significantly different (Table 4, column Zhg). The overall classification accuracy of the bilinear resampled image is higher than the overall classification accuracy of the cubic convolution resampled image by 0.8 percent.

(10) The overall classification accuracies of the bilinear resampled images classified using dependent and independent training statistics are significantly different (Table 4, columns Zhj). The classification performed using independent training statistics has an overall classification accuracy 24.7 percent higher than the classification performed using dependent training statistics.

(11) The overall classification accuracies of the cubic convolution resampled images classified using dependent and independent training statistics are significantly different (Table 4, columns Zgi). The classification performed using independent training statistics has an overall classification 8.5 percent higher than the classification performed using dependent training statistics.

Conclusions and Discussion

The impact of bilinear and cubic convolution resampling upon the supervized maximum-likelihood classification of a high spatial resolution test image has been investigated. The classifications were performed using spectral class training data and combined spectral/textural class training data. Examination of the average transformed divergence values of the training data indicated increased overall class separabilities when the data were collected from resampled versions of the test image. This was supported by examination of the overall classification accuracies of the resampled images which were significantly higher than the overall classification accuracies of the original test image. The overall spectral classification accuracies increased from 57.1 percent (original image) to a maximum of 60.7 percent (bilinear resampled image). The overall spectral/textural classification accuracies increased from 66.1 percent (original image) to a maximum of 68.6 percent (bilinear resampled image). The bilinear resampled images had significantly higher overall classification accuracies than the cubic-convolution resampled images by 1.7 percent and 0.8 percent when spectral and spectral/textural training statistics were used, respectively. These findings may be explained by consideration of the interaction between the local smoothing properties of the resampling algorithms and the grey-level structure of the image.

Bilinear and cubic-convolution resamplers can be expected to smooth the image and therefore reduce the image noise and the spectral variability of the image classes. Consequently, the separability of spectral class training data can be expected to increase after resampling, giving improvements in spectrally based classification accuracies. This effect will be counteracted, however, by smoothing of the boundaries between spectrally distinct classes. This will increase the likelihood of mixed boundary pixels and, therefore, increase the likelihood of their misclassification.

It is likely that the cubic-convolution resampler does not give as large improvements in classification accuracy as the bilinear resampler because it produces less smooth imagery that more closely resembles the original image. The degree of resampling-induced image smoothing will also be dependent upon the local grey-level structure of the image. The bilinear and cubic-convolution resampling algorithms interpolate grey-level values from 2- by 2- and 4- by 4-pixel neighborhoods, respectively. Consequently, resampling-induced smoothing will be most apparent in those parts of the image which have grey-level variations occurring at the same spa-

TABLE 4. RESULTS OF PAIRWISE INTER-CLASSIFICATION SIGNIFICANCE TESTS (ILLUSTRATED IN FIGURE 1).

	Within Spectral Classifications					Within Spectral/Textural Classifications					Between Spectral and Spectral/Textural Classifications		
	Zca[1]	Zba	Zcb	Zce	Zbd	Zhf	Zgf	Zhg	Zhj	Zgi	Zfa	Zhc	Zgb
Water	−0.32*	−0.61*	0.29*	2.45	0.98*	−1.35*	−2.04	0.70*	−6.17	−5.90	−1.45*	−2.47	−2.88
Fen Reed	4.88	3.34	1.54*	−0.97*	0.27*	18.02	15.47	2.53	62.82	16.54	65.73	79.48	78.49
Fen *Phalaris*	11.65	5.28	6.35	10.10	4.56	−0.42*	−3.46	3.06	104.07	38.98	21.19	9.25	12.44
Wet Grass	0.46*	−0.97*	1.43*	−8.71	−7.36	−12.67	−7.01	−5.62	74.34	46.10	24.62	11.85	18.76
Wet Grass (MG)	0.92*	0.75*	0.17*	−3.90	−1.30*	−7.62	−5.87	−1.75*	12.14	−10.90	20.32	12.03	13.87
Fen Herbs	5.26	0.63*	4.65	2.99	−0.35*	0.23*	−1.76*	1.98	1.91*	−6.72	4.72	−0.44*	2.38
Sallow Scrub	25.86	13.43	12.39	4.30	3.08	−5.13	−3.44	−1.68*	132.86	24.99	33.96	3.13	17.14
Woodland	9.31	6.41	2.89	1.59*	2.49	39.13	22.78	16.18	52.78	37.04	−14.26	15.48	2.05
Overall	22.02	11.47	10.54	2.92	1.67*	14.55	9.30	5.23	145.00	47.69	52.21	45.65	50.55

[1]Zca is test statistic for significance analysis between classifications c and a
*Denotes *insignificant* difference at the 95% confidence level

tial scale as these pixel neighborhoods. Texturally fine cover types are characterized by localized grey-level variation and are, therefore, more likely to fall into this category than more coarse cover types characterized by grey-level variations occurring over larger spatial scales. This may explain why the spectral classifications of the texturally smooth classes are similar in the original and the resampled images, whereas the spectral classifications of the less texturally smooth classes are observed to be significantly improved after resampling. These explanations cannot be used reliably to explain the spectral/textural classification results because the interaction of the texture measure with the grey-level structure of the image is not easily understood.

The impact of using training statistics collected from the original test image to classify resampled versions of the test image was also examined. In these experiments, the overall spectral/textural classification accuracies of the resampled images were found to be significantly poorer when they were performed using training statistics collected from the original image. This phenomenon was not observed when the images were classified using only spectral training statistics. This indicates that, when resampling is necessary prior to classification, texturally based classification procedures may use training statistics collected from the resampled and not the original image.

The results described in this paper are empirical; therefore, it cannot be certain that they will be applicable to other images and resampling and texture calculating algorithms. However, they indicate that remotely sensed images may be classified after they have been resampled without causing a reduction in the classification accuracy. In some cases, the bilinear and cubic-convolution resampling techniques may improve the classification accuracies because of their local smoothing effect. However, the benefits of image smoothing prior to classification will be subject to counteracting trends of spatial inter-class and intra-class variability (Townshend and Justice, 1981). Consequently, any classification improvements associated with resampling-induced smoothing will be dependent upon the grey-level structure of the image and will not necessarily be always improved, particularly when images containing many spectrally distinct class boundaries are used.

The motivation of this study has been the perceived requirement for automated image classification procedures. The findings described in this paper indicate that such procedures may be implemented using georeferenced (i.e., resampled) imagery without a reduction in the classification accuracy. This may be particularly important because recent trends in image classification have been towards the incorporation of contextual information and knowledge-based rules (e.g., Mason *et al.*, 1988; Argialas and Harlow, 1990; Ton *et al.*, 1991; Kontoes *et al.*, 1993) using techniques which are amenable to automation in this manner.

Acknowledgments

The authors acknowledge the U.K. Natural Environment Research Council (NERC) for provision of the ATM imagery and Cambridge University Geography Department for provision of image processing facilities. The ground reference map was produced and provided by Robin Fuller of the NERC Institute of Terrestrial Ecology, Monkswood, U.K.

References

Argialas, D.P., and C.A. Harlow, 1990. Computational image interpretation model: an overview and perspective, *Photogrammetric Engineering & Remote Sensing*, 56(6):871–886.

Atkinson, P., 1988. The interrelationship between resampling method and information extraction technique, *Proceedings of IGARSS' 88 Symposium*, Edinburgh, 13-16 September, pp. 521–527.

Baker, J.R., G.W. Marks, and E.M. Mikhail, 1975. Analysis of digital multispectral scanner (MSS) data, *Bildmessung und Luftbildwesen*, 43:22–27.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland, 1975. *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, Massachusetts.

Callison, R.D., P. Blake, and J.M. Anderson, 1987. The quantitative use of airborne thematic mapper thermal infrared data, *International Journal of Remote Sensing*, 8(1):113–126.

Castleman, K.R., 1979. *Digital Image Processing*, Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Congalton, R.G., and R.A. Mead, 1986. A review of three discrete multivariate analysis techniques used in assessing the accuracy of remotely sensed data from error matrices, *IEEE Transactions on Geoscience and Remote Sensing*, 24(1):169–174.

Conners, R.W., and C. Harlow, 1980. A theoretical comparison of texture algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(3):204–222.

Curran, P.J., and M.I. Pedley, 1989. Land cover classification with an airborne multispectral scanner, *Proceedings of the NERC Workshop on Airborne Remote Sensing*, Institute of Fresh Water Ecology, Windermere, U.K., pp. 179–193.

Cushnie, J.L., 1987. The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies, *International Journal of Remote Sensing*, 8(1): 15–29.

Dikshit, O., 1992. *The Classification of Texture in Remotely-Sensed Environmental Imagery*, Ph.D. thesis, University of Cambridge, Cambridge, U.K.

Etheridge, J., and C. Nelson, 1979. Some effects of nearest neighbour, bilinear interpolation, and cubic convolution resampling on Landsat data, *Machine Processing of Remotely Sensed Data Symposium*, Purdue University, Indiana, p. 84.

Ferneyhough, D.G., and C.W. Niblack, 1977. *Resampling Study*, IBM Final Report NASA contract NAS5-21865 for Goddard Space Flight Center, Greenbelt, Maryland.

Foley, J., A. Van Dam, S. Feiner, and J. Hughes, 1990. *Computer Graphics: Principles and Practice*, Addison-Wesley Publishing Company, Inc., Reading, Massachusetts.

Fuller, R.M., N.J. Brown, and M.D. Mountford, 1986. Taking stock of changing broadland, Part I: Air photo-interpretation and digital cartography, *Journal of Biogeography*, 13:313–326.

Gool, L.V., P. Dewaele, and A. Oosterlinck, 1985. Texture analysis anno 1983, *Computer Vision, Graphics and Image Processing*, 29:336–357.

Haralick, R., 1979. Statistical and structural approaches to texture, *Proceedings of IEEE*, 67(5):786–804.

Haralick, R.M., K. Shanmugan, and I. Dinstein, 1983. Textural features for image classification, *I.E.E.E. Transactions on Systems, Man and Cybernetics*, 3:610–621.

Jensen, J.R., and M.E. Hodgson, 1987. Interrelationships between spatial resolution and per-pixel classifiers for extracting information classes, Part I: The urban environment, *ASPRS-ACSM Annual Convention*, Baltimore, The American Society for Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, 1:121–129.

Keys, R.G., 1981. Cubic convolution interpolation for digital image processing, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):1153–1160.

Kontoes, C., G.C. Wilkinson, A. Burril, S. Goffredo, and J. Megier, 1993. An experimental system for the integration of GIS data in knowledge-based image anaylsis for remote sensing of agriculture, *International Journal of Geographic Information Systems*, 7(3):247–262.

Logan, T.L., and A.H. Strahler, 1979. The errors associated with digital image resampling of Landsat forest imagery for multidate registration, *Proceedings of the Eighth Annual Remote Sensing of Earth Resources Conference*, Tullahoma, pp. 163–181.

Marceau, D.J., P.J. Howarth, J.M. Dubois, and D.J. Gratton, 1990. Evaluation of the grey level co-occurrence matrix method for land-cover classification using SPOT imagery, *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):513–519.

Mason, D.C., D.G. Corr, A. Cross, D.C. Hogg, D.H. Lawrence, M. Petrou, and A.M. Tailor, 1988. The use of digital map data in the segmentation and classification of remotely sensed images, *International Journal of Geographic Information Systems*, 2(3):195–215.

Mather, P.M., 1987. *Computer Processing of Remotely Sensed Images: An Introduction*, John Wiley and Sons, Chichester.

Morris, K.P., and M.J. Barnsley, 1989. An assessment of various procedures for the radiometric correction of the sensor view angle effect, *Proceedings of the 15th Annual Conference of the Remote Sensing Society*, University of Bristol, 13-15 September, pp. 155–164.

Park, S.K., and R.A. Schowengerdt, 1983. Image reconstruction by parametric cubic convolution, *Computer Vision, Graphics and Image Processing*, 23:258–272.

Roy, D.P., and O. Dikshit, 1994. Investigation of image resampling effects upon the textural information content of a high spatial resolution remotely sensed image, *International Journal of Remote Sensing*, 15(5):1123–1130.

Rifman, S.S., and D.M. McKinnon, 1974. *Evaluation of Digital Correction Techniques for ERTS Images*, TRW Report 20634-6003-TV-02.

Sali, E., and H. Wolfson, 1992. Texture classification in aerial photographs and satellite data, *International Journal of Remote Sensing*, 13(18):3395–3408.

Shlien, S., 1979. Geometric correction, registration, and resampling of Landsat imagery, *Canadian Journal of Remote Sensing*, 5(1): 74–89.

Smith, J.L, and B. Kovalick, 1985. A comparison of the effects of resampling before and after classification on the accuracy of a Landsat derived cover type map, *Proceedings of the International Conference of the Remote Sensing Society and the Centre for Earth Resources Mangement*, University of London, 9-12 Spetember, pp. 391–399.

Swain, P.H., and S.M. Davis, (editors), 1978. *Remote Sensing: The Quantitative Approach*, McGraw-Hill, New York.

Thomas, I.L., V.M. Benning, and N.P. Ching, 1987. *Classification of Remotely Sensed Images*, Hilger, Bristol.

Ton, J., J. Sticklen, and A. Jain, 1991. Knowledge-based segmentation of Landsat images, *IEEE Transactions on Geoscience and Remote Sensing*, 29(2):222–231.

Townshend, J., and C. Justice, 1981. Information extraction from remotely sensed data. A user view, *International Journal of Remote Sensing*, 2(4):313–329.

Wang, L., and D. He, 1990. A new statistical approach for texture analysis, *Photogrammetric Engineering & Remote Sensing*, 56(1): 61–66.

Wharton, S.W., 1989. Knowledge-based spectral classification of remotely sensed image data, *Theory and Applications of Optical Remote Sensing* (G. Asrar, editor), John Wiley and Sons, Inc., New York.

(Received 2 June 1994; accepted 29 November 1994; revised 7 February 1995)