An Improved Skidmore/Turner Classifier

Abstract

A new decision rule significantly improves the classification accuracy of the nonparametric spectral classifier described by Skidmore and Turner (1988). The new rule normalizes class histograms with the mean frequency, rather than with the number of training pixels in a class, and gives a classifier which consistently outperforms the maximum-likelihood classifier. Another type of classifier, which combines a two-dimensional and a one-dimensional Skidmore/Turner (S/T) classifier, has higher classification accuracies than the S/T classifier with the improved decision rule, in two of four study areas. Tree species were classified for two study areas and land-cover classes for the other two.

Introduction

Foresters are reluctant to embrace remote sensing for operational use because of low classification accuracies (Skidmore *et al.*, 1987). This is also true for other resource management sectors. Work is progressing on improving accuracies with knowledge-based techniques (Srinivasan and Richards, 1990; Goldberg *et al.*, 1985) and textural classifiers (Franklin and Peddle, 1990; Haralick, 1979); however, Skidmore and Turner (1988) have shown that there is still room for improvement in ordinary spectral classifiers.

Parametric classifiers, such as maximum-likelihood (Richards, 1986), make implicit assumptions about the shape of data distributions (Skidmore and Turner, 1988). These assumptions permit parametric classifiers to be trained on a small training dataset, but limit the classification accuracy. Skidmore and Turner (1988) described a nonparametric classifier, the S/T classifier, which was more accurate than a maximum-likelihood classifier for determining pine tree age. The S/T classifier requires substantially sized training datasets and is effectively limited to three or fewer spectral bands; with more than three spectral bands, there is rarely enough training data to adequately define the four, or more, dimensional histograms.

This paper describes two significant improvements to the S/T classifier. The first is an improved decision rule. The resulting classifier is referred to in this paper as the "improved" S/T classifier. The second improvement is an updating procedure which enables a two-dimensional S/T classifier to be combined with a one-dimensional S/T classifier. Three spectral bands may be considered while avoiding the difficulties associated with three-dimensional feature spaces. This classifier is referred to as the "updating" S/T classifier. The "improved" S/T classifier and the "updating" S/T classifier are compared to the standard S/T classifier and the maximum-likelihood classifier in four different study areas. Tree species are classified in two of the study areas and land-cover classes in the others.

Photogrammetric Engineering & Remote Sensing,
Vol. 59, No. 5, May 1993, pp. 623–626.

0099-1112/93/5905-629$03.00/0

Skidmore/Turner Classifier

The S/T classifier uses a supervised nonparametric method. For each class (ω_c $c = 1, n$) a three-dimensional histogram ($H_c(i,j,k)$ $i,j,k = 0,255$) is calculated from training data. A value of $H_c(i,j,k)$ gives the number of training pixels in class ω_c that occur with brightness values $_i$ in the first spectral band, $_j$ in the second, and $_k$ in the third spectral band. Each histogram is normalized by the number of training pixels in its class: i.e.,

$$h_c(i,j,k) = H_c(i,j,k) \; / \; \sum_{r=0}^{255} \sum_{s=0}^{255} \sum_{t=0}^{255} H_c(r,s,t)$$

For each brightness vector (i,j,k) in a three-dimensional feature space, the class histogram with the greatest normalized frequency is found and that class (i.e., the most likely) is assigned to that position in feature space. (If all class histograms have a zero normalized frequency at that brightness vector, then the feature is set to "undefined," that is, zero). Feature space can then be used as a lookup table to rapidly classify the original bands.

In addition to the procedure outlined by Skidmore and Turner (1988), the following steps proved useful:

- Smoothing of the three-dimensional histograms with a 3 by 3 by 3 averaging box filter (McDonnell, 1981) helped create better defined histograms with fewer occurrences of singleton frequencies (i.e., $H_c(i,j,k) = 1$).
- Filling small holes of "undefined" regions in the three-dimensional feature space (caused by the sparse three-dimensional histograms) helped reduce the proportion of the image classified as "undefined." The holes were filled by using a 3 by 3 by 3 modal filter which returns the most commonly occurring non-zero class in the box. Only if all classes in the box are zero (i.e., "undefined") is a zero returned.

These steps also proved useful with the "improved" S/T classifier.

Improved Skidmore/Turner Classifier

The standard S/T classifier normalizes each class histogram by dividing by the number of training pixels in the class. An alternative is to normalize by dividing by the mean frequency, excluding zero frequencies in the mean: i.e.,

$$\begin{aligned} h_c'(i,j,k) &= H_c(i,j,k) \; / \; \Big(\sum_{r=0}^{255} \sum_{s=0}^{255} \sum_{t=0}^{255} H_c(r,s,t)/N_c \Big) \\ &= N_c \cdot h_c(i,j,k) \end{aligned}$$

where N_c is the number of non-zero values of $H_c(i,j,k)$.

Assuming we have one spectral dimension, then multiplying by N_c is effectively the same as multiplying by the standard deviation of the histogram, because N_c is proportional to it. It can be seen that the decision rule using this normalization is analogous to a Mahalanobis type classifier (Richards, 1986) where the class with the smallest z statistic (the deviation from the mean divided by the standard deviation) is chosen. The resulting decision boundaries will there-

John R. Dymond
Landcare Research New Zealand, Ltd, Private Bag 11052, Palmerston North, New Zealand

fore give approximately equal *class proportions* of misclassified pixels.

The S/T classifier with this new decision rule is referred to as the "improved" S/T classifier. Table 1 gives a one-dimensional numerical example of how the normalization is done for the case of two classes.

Updating Skidmore/Turner Classifier

The weakness of the S/T classifier is that in three dimensions, or more, the histograms become sparsely populated and a high proportion of histogram frequencies occur as singleton frequencies. In two dimensions this sparcity is very much reduced. An "updating" classifier is therefore proposed in which a two-dimensional "improved" S/T classifier is updated by a one-dimensional "improved" S/T classifier defined on the third spectral band.

The "updating" S/T classifier records the three most likely classes from the two-dimensional histograms, and their associated normalized frequencies (as normalized in the "improved" S/T classifier), for each pixel in the spatial domain. The three normalized frequencies ($h'_c(i,j)$ $c = \alpha, \beta, \gamma$) are regarded as empirical probabilities and are updated (i.e., multiplied) by the normalized frequency ($f'_c(k)$ $c = \alpha, \beta, \gamma$) from the one-dimensional histogram of the same class. The decision rule chooses the class out of ω_c $c = \alpha, \beta, \gamma$ with maximum $h'_c(i.j).f_c(k)$. This classifier loses some of the structure in the co-joint three-dimensional histograms, but reduces the problem of data sparcity in feature space.

Data

Images from four separate study areas were chosen to test the different classifiers under a range of radiance conditions.

Wanganui Hill Country

The Wanganui image is a color infrared aerial photograph scanned by a large flatbed scanner through blue, green, and red filters. The photograph was taken with a large format camera at an altitude of 7600 m, with a lens of focal length 152 mm, to give a film scale of 1:50,000. Elevation of the sun was 46 degrees. The scanning aperture of 100 μm gave a ground pixel size of 5 m which was subsequently re-sampled to 20 m. This gave an image of size 512 by 512 pixels.

The image contains mainly pastoral hill country with some production forestry. Surface elevations range from 10 to 336 m and 20 percent of slopes exceed 30 degrees. The main land-cover classes are pasture, pine forest, scrub/regenerating native forest, burnt vegetation, and water.

The training and test ground datasets for land-cover class were created by a combination of photointerpretation and fieldwork. The ease of creating ground datasets, due to the presence of large homogeneous areas of land cover, permitted the luxury of having a very large test dataset: 70 polygons of average size 1400 pixels. However, the size of the training dataset was kept to a more normal size: 34 polygons of average size 150 pixels.

TABLE 1. EXAMPLE OF HOW TO NORMALIZE HISTOGRAMS FOR THE "IMPROVED" S/T CLASSIFIER. THERE ARE TWO POSSIBLE CLASSES.

Class 1								
brightness	13	14	15	16	17	18		$N_1 = 6$
H_1	1	3	3	3	1	1		$\Sigma_{s=0}^{255} H_1(s) = 12$
$h_1 = H_1/12$	1/12	3/12	3/12	3/12	1/12	1/12		
$h'_1 = 6.h_1$	1/2	3/2	3/2	3/2	1/2	1/2		
Class 2								
brightness					17	18	19	$N_2 = 3$
H_2					4	6	5	$\Sigma_{s=0}^{255} H_2(s) = 15$
$h_2 = H_2/15$					4/15	6/15	5/15	
$h'_2 = 3.h_2$					4/5	6/5	5/5	
decision	1	1	1	1	2	2	2	

Saltwater Ecological Area

The Saltwater image is a 730 by 730 pixel extract from a SPOT multispectral scene. Elevation of the sun was 23 degrees.

The image contains native forest, a saltwater lagoon, and some swamps. There are three forestry types of interest: kahikatea (*Dacrycarpus dacrydioides*)-dominated forest, kamahi (*Weinmannia racemosa*)-dominated forest, and rimu (*Dacrydium cupressinum*)-dominated forest. Surface elevations generally increase away from the coast to a maximum of 200 m.

Ground data for the three forestry types came from a generalized vegetation map published by Norton and Leathwick (1990). These data were visually checked against the SPOT image. It was difficult obtaining enough reliable data this way, so all ground data were amalgamated to create a combined training and test dataset – 19 polygons of average size 1100 pixels. This procedure is not generally recommended because assessed classification accuracies can be biased (Thomas and Allcock, 1984). However, the accuracies can be reasonably used to compare the relative performance of classifiers to classify the training data.

Kaingaroa Forest

The Kaingaroa image is a 749- by 796-pixel extract of a SPOT multispectral scene. Elevation of the sun was 28 degrees.

The image contains several hundred production forestry woodlots, primarily of *Pinus radiata* and Douglas fir (*Pseudotsuga menziesii*) trees. Tree age varies between zero and 65 years. The terrain is mainly flat with drainage channels dissecting the area.

Ground data for the two tree types were extracted from a comprehensive set of woodland maps and forest inventories supplied by New Zealand Timberlands. Only trees over ten years old were included in the training and test datasets in order to avoid areas with incomplete canopy closure. The training dataset consisted of 17 polygons of average size 700 pixels while the test dataset consisted of 14 polygons of average size 750 pixels.

Foxton Coast

The Foxton coast image is a 1024- by 512-pixel extract of a SPOT multispectral scene. Elevation of the sun was 53 degrees.

The image shows a sandy coast which exhibits an inland gradation into protection and production forestry, and finally into pastoral farming. The main land-cover classes are sand, trees, lakes, poor pasture, good pasture, sea, and urban. The terrain is flat.

A combination of photointerpretation and field work created the training and test ground datasets. The training dataset consisted of 89 polygons of average size 200 pixels, and the test dataset consisted of 256 polygons of average size 140 pixels.

Assessment of Classification Accuracy

The confusion matrix is the standard summary of classification accuracy (Hay, 1979). However, when comparing the performance of different classifiers, it is convenient to summarize a confusion matrix with one parameter. The overall classification accuracy is often used, but it does not really take into account the classification accuracy of classes which occupy small proportions of the image area. It is therefore convenient to define a "summary" classification accuracy to be the average of the overall classification accuracy, the average true class accuracy, and the average map class accuracy. The overall classification accuracy is the sum of diagonal terms in the confusion matrix (Hay, 1979). The true class accuracy is the class average of the percentage correct given a true class; and similarly for the average map class accuracy. These parameters derive from confusion matrices which have been adjusted for differential sampling rates of test data within classes (Card, 1982). Table 2 gives an example of how the "summary" classification accuracy is calculated.

Results

Figure 1 shows the "summary" classification accuracies for the three types of S/T classifier, and for a maximum-likelihood classifier. (For completeness the other classification accuracy parameters are given in Table 3.) All classified images have had a 3 by 3 modal filter (Booth and Oldfield, 1989) passed over them before determination of classification accuracy. This consistently improved accuracies. The three-dimensional S/T classifiers used 3 by 3 by 3 boxes for histogram smoothing and hole filling, while the two-dimensional S/T classifiers used 3 by 3 boxes.

In all four images the "improved" S/T classifier performed much better than the standard S/T classifier. The "updating" S/T classifier performed better than the "improved" S/T classifier in the Wanganui and Kaingaroa images, but not in the Saltwater and Foxton images. In all images the maximum-likelihood was outperformed by either the "improved" or "updating" S/T classifier.

The maximum-likelihood performed much better than the S/T classifier in all images. This result contradicts that of Skidmore and Turner (1988) where the S/T classifier outperformed the maximum-likelihood in classifying pine tree age. An explanation could be that pine tree age is a special case where the S/T classifier is particularly suitable. The performance of classifiers on pine tree age could not be tested on the Kaingaroa image because spectral signatures are affected by regular pruning.

Discussion

The reason for the relatively poor performance of the "updating" S/T classifier on the Saltwater image is uncertain. This highlights the fact that the perfomance of any classifier, relative to others, depends on the statistical nature of the training and test data. Therefore, a selection of different classifiers should be tried in order to optimize classification accuracy.

The "updating" S/T classifier is potentially a good classifier for more than three spectral dimensions. Rather than updating with one-dimensional frequencies, the updating could be done with two-dimensional frequencies. Further updating could be done to increase dimensions beyond four if necessary. In this way a many dimensional nonparametric classifier could be generated. Another useful feature of the "updating" classifier is that because it only uses two-dimensional co-joint histograms, it does not need as much training data as the standard three-dimensional S/T classifier.

Despite the reasonably high "summary" classification accuracies of the Wanganui and Saltwater classified images (i.e., in the high eighties), they are still far from satisfactory. The classified Wanganui image underestimates scrub in the bottom-left, top-left, and bottom-right corners; overestimates

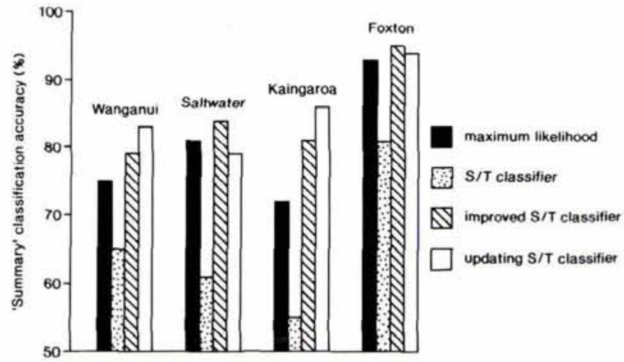

Figure 1. "Summary" classification accuracies for the three types of S/T classifier and a maximum-likelihood classifier.

TABLE 2. EXAMPLE OF HOW TO CALCULATE "SUMMARY" CLASSIFICATION ACCURACY FROM A CONFUSION MATRIX. OVERALL CLASSIFICATION ACCURACY = 0.118 + 0.340 + 0.315 = 0.77. AVERAGE TRUE CLASS ACCURACY = (0.88 + 0.72 + 0.80)/3 = 0.80. AVERAGE MAP CLASS ACCURACY = (0.64 + 0.94 + 0.69)/3 = 0.75. "SUMMARY" CLASSIFICATION ACCURACY = (0.77 + 0.80 + 0.75)/3 = 0.77.

		MAP kahikatea	MAP kamahi	CLASS rimu	probability correct given true class
TRUE CLASS	kahikatea	0.118	0.000	0.016	0.88
	kamahi	0.012	0.340	0.123	0.72
	rimu	0.055	0.021	0.315	0.80
probability correct given map class		0.64	0.94	0.69	

TABLE 3. CLASSIFICATION ACCURACY PARAMETERS OF DIFFERENT CLASSIFIERS

	Overall classification accuracy (%)	Average true class accuracy (%)	Average map class accuracy (%)
WANGANUI IMAGE			
maximum-likelihood	82	70	73
S/T	70	61	65
improved S/T	82	72	83
updating S/T	88	80	80
SALTWATER IMAGE			
maximum-likelihood	79	81	85
S/T	62	65	56
improved S/T	84	84	84
updating S/T	78	78	81
KAINGAROA IMAGE			
maximum-likelihood	79	85	53
S/T	77	50	38
improved S/T	82	81	81
updating S/T	87	85	86
FOXTON IMAGE			
maximum-likelihood	95	90	98
S/T	95	74	74
improved S/T	97	94	93
updating S/T	95	93	93

forest on shaded slopes; and assigns parts of the river channel to burnt vegetation. These problems are caused by a hotspot (Hugli and Frei, 1983) to the bottom-left of the image, lens fall off, and topographic effects (Holben and Justice, 1980). The Saltwater image overestimates Kahikatea-dominated forest on shaded slopes. The high "summary" classification accuracies of these unsatisfactory classifications suggests that classification accuracies in the high nineties, like that of the Foxton image, are required. Also suggested is the need to consider the spatial distribution of classification errors in addition to the purely statistical summary of the confusion matrix. It is humbling to note that the high classification accuracy of the Foxton image was achieved only under ideal conditions of high sun elevation, flat land, and very generalized land-cover classes (Anderson *et al.*, 1976).

Acknowledgments
I thank R. H. Wilde and H. H. Araghi for providing the Foxton ground datasets; L. Brown for the Saltwater ground dataset; New Zealand Timberlands for provision of ground data for the Kaingaroa image; and Dr. C. Trotter for making helpful comments on an earlier draft of this paper.

References

Anderson, J. R., E. E. Hardy, J. T. Roach, and R.E. Witmer, 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. U.S. Geological Survey Professional Paper 964.

Booth, D. J., and R. B. Oldfield, 1989. A comparison of classification algorithms in terms of speed and accuracy after the application of a post-classification modal filter. *Int. J. Remote Sensing*, Vol. 10, No. 7, pp. 1271-1276.

Card, D. H., 1982. Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering & Remote Sensing*, Vol. 48, No. 3, pp. 431-439.

Franklin, S. E., and D. R. Peddle, 1990. Classification of SPOT HRV imagery and texture features. *Int. J. Remote Sensing*, Vol. 11, No. 3, pp. 551-556.

Goldberg, M., D. G. Goodenough, M. Alvo, and G. M. Karam, 1985. A hierarchical expert system for updating forestry maps with Landsat data. *Proc. IEEE*, Vol. 73, No. 6, pp. 1054-1063.

Haralick, R. M., 1979. Statistical and structural approaches to texture. *Proc. IEEE*, Vol. 67, No. 5, pp. 786-804.

Hay, A. M., 1979. Sampling designs to test land-use map accuracy. *Photogrammetric Engineering & Remote Sensing*, Vol. 45, No. 4, pp. 529-533.

Holben, B. N., and C. O. Justice, 1980. The topographic effect on spectral response from nadir-pointing sensors. *Photogrammetric Engineering & Remote Sensing*, Vol. 46, No. 9, pp. 1191-1200.

Hugli, H., and W. Frei, 1983. Understanding anisotropic reflectance in mountainous terrain. *Photogrammetric Engineering & Remote Sensing*, Vol. 49, No. 5, pp. 671-683.

McDonnell, M. J., 1981. Box filtering techniques. *Optronics J.*, Vol. 17, pp. 65-70.

Norton, D. A., and J. R. Leathwick, 1990. The lowland vegetation pattern, South Westland, New Zealand: 1. Saltwater Ecological Area. *New Zealand J. Botany*, Vol. 28, pp. 41-51.

Richards, J. A., 1986. *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, New York.

Skidmore, A. K., G. B. Wood, and K. R. Shepherd, 1987. Remotely sensed digital data in forestry: a review. *Aust. For.*, Vol. 50, No. 1, pp. 40-53.

Skidmore, A. K., and B. J. Turner, 1988. Forest mapping accuracies are improved using a supervised nonparametric classifier with SPOT data. *Photogrammetric Engineering & Remote Sensing*, Vol. 54, No. 10, pp. 1415-1421.

Srinivasan, A., and J. A. Richards, 1990. Knowledge-based techniques for multisource classification. *Int. J. Remote Sensing*, Vol. 11, No. 3, pp. 505-525.

Thomas, I. L., and G. M. K. Allcock, 1984. Determining the confidence level for a classification. *Photogrammetric Engineering & Remote Sensing*, Vol. 50, No. 10, pp. 1491-1496.

(Received 29 May 1991; accepted 17 June 1992; revised 3 September 1992)