Research Repository UCD

Title	Classification of forestry species using singular value decomposition
Authors(s)	Danaher, Sean, Herries, Graham M., Mac Siúrtáin, Máirtín Pádraig, O'Mongain, E.
Publication date	1994-09
Publication information	Danaher, Sean, Graham M. Herries, Máirtín Pádraig Mac Siúrtáin, and E. O'Mongain. "Classification of Forestry Species Using Singular Value Decomposition" (September, 1994).
Conference details	Multispectral and Microwave Sensing of Forestry, Hydrology, and Natural Resources, Rome, Italy, September 26, 1994
Publisher	The International Society for Optical Engineering
Item record/more information	http://hdl.handle.net/10197/4033
Publisher's statement	Copyright 1995 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.
Publisher's version (DOI)	10.1117/12.200768

Downloaded 2023-04-07T10:52:29Z

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)

© Some rights reserved. For more information

Classification of forestry species using singular value decomposition.

S. Danaher, G. Herries.

Faculty of Information and Engineering Systems, Leeds Metropolitan University, Leeds, LS1 3HE.

M. MacSiúrtáin, E. O'Mongain.

University College Dublin, Belfield, Dublin 4, Ireland.

ABSTRACT

A method is defined and tested for the classification of forest species from multi-spectral data, based on Singular Value Decomposition (SVD) and key vector analysis. The SVD technique, which bears a close resemblance to multivariate statistic techniques has previously been successfully applied to the problem of signal extraction from marine data.

In this study the SVD technique is used as a classifier for forest regions, using SPOT and Landsat Thematic Mapper data. The specific region chosen is in the County Wicklow area of Ireland. This area has a large number of species, within a very small region and hence is not amenable to existing techniques. Preliminary results indicate that SVD is a fast and efficient classifier with the ability to differentiate between species such as Scots pine, Japanese larch and Sitka spruce. Classification accuracy's using this technique yielded excellent results of >99% for forest, against four background classes. The accuracy's of the individual species classification are slightly lower, but they are still high at 97-100% for the SPOT wavebands. When the Landsat TM bands 3, 4 and 5 were used on their own, accuracy's of 95-100% were achieved.

1. INTRODUCTION

Remote sensing from satellite, offers the only viable method for surveying the large-scale natural and man-made changes to forestry across the globe. Due to the recent advances in computer technology over the past 5-10 years, the data collected from these satellites can be processed extremely fast. This has enabled advanced signal processing and cognitive learning techniques to be developed and applied to such multidimensional data. Various statistical algorithms such as principal component analysis,^{1,2} maximum likelihood algorithms,^{3,4} minimum distance classifiers⁵ and Bayesian classification,⁶ have previously been applied to the classification of general land cover classes and forests from remote sensing data.^{7,8} These algorithms tend to be both ad hoc and fragmented and have many inherent problems associated with them. The accuracy of these methods is far from ideal, being typically of the order of 45-90%.^{5,6,9} An accuracy of this level is clearly unacceptable for global forest mapping, but may be sufficient for smaller areas of forest.

A new method has been developed at Leeds Metropolitan University, which is particularly suited to the extraction of the required signal (forest and individual species in this case) from spectral data sets, dominated by background (other land cover classes) and noise (instrumental and quantization), which are conditions typical of Landsat and SPOT data. This method uses Singular Value Decomposition,^{10,11,12,13} which in most cases should prove superior to the existing algorithms. SVD can be used on its own merits or as a preliminary stage in methodologies such as cluster or neural network analysis. It also allows a linear analysis technique, key vector analysis, to be used.

2. SINGULAR VALUE DECOMPOSITION

SVD represents a powerful numerical technique for the analysis of multivariate data.^{14,15} SVD can be used as a preliminary stage in most types of multivariate analysis, and can greatly increase the computational efficiency of linear techniques such as key vector analysis, and non linear techniques such as cluster analysis and neural network analysis. SVD is also an

extremely effective technique for the reduction of white noise. The inherent attributes of the SVD technique may have a considerable influence on the dataset, the more important of which may be summarised as follows.

- i) **Dimensional Reduction.** Unless the parameters (bands) are completely independent or the data set is totally dominated by noise, SVD will determine a linear transformation which will convert the parameters into totally independent variables. Furthermore, it will do so in such a way that the SVD parameters (which are linear combinations of the original parameters) are chosen in decreasing order of significance. Transforming to SVD parameters is particularly important in neural network type analysis as the number of connections (determined by the number of layers) grows dramatically with the number of parameters.
- ii) **Invertibility.** A data set which has undergone SVD can be regenerated by simple matrix multiplication.
- iii) **Data Compression.** The data set can be regenerated with high accuracy, by ignoring parameters (dimensions) which make minimal contribution.
- iv) **Noise Reduction.** No data set is totally noise free. If the parameterized data is considered to form an n dimensional space, where n is the number of parameters, white noise will spread uniformly over the space, having no preferred direction. An orthogonal transformation, such as SVD, is equivalent to a rotation/reflection in n -space. Dimensions which are deliberately excluded will have exactly the same noise content as those which are included (within statistical fluctuations). The signal to noise level is improved by discarding the higher dimensions. Using SVD for data compression reduces the noise level of the data.
- v) **Orthogonality.** If the data is subsequently used for a cluster type analysis the fact that the SVD parameters are totally independent allows separation of events to be defined (if the 2-norm is used) as

$$r = \sqrt{\sum_{i=1}^{i=k} (w_i - \langle w_i \rangle)^2}$$

where r is the normalised distance from the centroid of a cluster to every point, k is the number of included SVD dimensions and w_i are the individual SVD parameters.

The standard cluster analysis technique uses the equivalent 2-norm definition for a non orthogonal basis set

$$r = \sqrt{\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} A_{ij} (b_i - \langle b_i \rangle) (b_j - \langle b_j \rangle)}$$

where r is the normalised distance from the centroid of a cluster to every point, b_i are the individual band parameters and A is the inverse of the variance-covariance matrix.

- vi) **Efficiency.** SVD is an extremely efficient and robust technique. For a data set of 10000 forest pixels and 10000 background pixels and using seven bands, a SVD takes less than one minute on a 486 PC.
- vii) **Key Vector analysis.** SVD allows a technique called key vector analysis to be used. Whereas key vector analysis is not expected to be as effective as other more sophisticated techniques such as cluster analysis or neural network analysis it has a number of advantages:

- 1) The key vector is unique.
- 2) The key vector can be determined rapidly using a set of Monte-Carlo simulations or a subset of the data.
- 3) Once the key vector has been obtained a score value can be determined on the original data set with only n floating point operations per event, where n is the number of bands used. It is therefore far more computationally efficient than most of the other techniques used.
- 4) Due to the fact that the key vector is highly constrained, the likelihood of spurious results is significantly reduced.

Before proceeding to a formal definition of SVD it is worthwhile stating that it bears a very close relationship to Characteristic Vector analysis.¹⁶ The difference is in the elegance of the formalism. If $\mathbf{O}$ is a $m \times n$ data matrix the SVD of $\mathbf{O}$ is defined by:

$$\mathbf{O} = \mathbf{W}\mathbf{L}\mathbf{V}$$

The matrices $\mathbf{W}$, $\mathbf{L}$ and $\mathbf{V}$ can be defined as follows: $\mathbf{V}$ is a $n \times n$ matrix containing the unit eigenvectors of $\mathbf{O}^T\mathbf{O}$, (the variance/covariance matrix), on each row. These are sorted in descending significance. If the matrix $\mathbf{O}$ has had its mean subtracted (column by column) the matrix $\mathbf{V}$ will be a matrix of the characteristic vectors as defined in multivariate statistics. $\mathbf{L}$ is a $m \times n$ matrix of the form:

$$\mathbf{L} = \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}$$

where $\mathbf{D}$ is diagonal. The diagonal elements of $\mathbf{D}$ are the square roots of the eigenvalues of $\mathbf{O}^T\mathbf{O}$ sorted in descending order.

The matrix $\mathbf{W}$ is a $m \times m$ containing the unit eigenvectors of $\mathbf{O}\mathbf{O}^T$ on each column. It also may be considered as a matrix containing the weights or scalar multipliers of the characteristic vectors $\mathbf{V}$ in the data matrix $\mathbf{O}$.

There is also an 'economy sized' SVD where $\mathbf{W}$ is $m \times n$, $\mathbf{L}$ is $n \times n$ and $\mathbf{V}$ is $n \times n$. For remote sensing data, m (the number of pixels) can be very large and hence a $m \times n$ matrix would be too large to be manageable. Whereas all the eigenvectors will generally be needed to define $\mathbf{O}$ exactly, in many cases a sufficiently good approximation to $\mathbf{O}$ can be achieved by taking only a few eigenvectors. Then:

$$\mathbf{O} \approx \mathbf{W}\mathbf{L}\mathbf{V}$$

with $\mathbf{W}$ being $m \times k$, $\mathbf{L}$ being $k \times k$ and $\mathbf{V}$ being $k \times n$; taking the first k columns of $\mathbf{W}$, the first k rows and columns of $\mathbf{L}$ and the first k rows of $\mathbf{V}$.

3. SVD APPLIED TO THE FORESTRY DATA

The first step in a two stage process involves finding that key vector which is most efficient at separating a mixed forest population and background population. (The background will consist of all the non forest regions in the image). Due to the high variability of background classes, the technique may have to be applied a number of times; filtering the pixels classified as forest. The second stage of the operation is to consider the desired species as one population and the other types of forest as a background class. The data is again filtered to separate the desired species. Each pixel is considered to be a vector in n dimensional space where n is the number of bands.

$$\vec{o} = (h_1, h_2, h_3, \dots, h_n)$$

The simulation matrix $\mathbf{O}$ was generated by combining a set of forest and background classes.

$$\mathbf{O} = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \dots & h_{1n} \\ h_{21} & h_{22} & h_{23} & \dots & h_{2n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ h_{m1} & h_{m2} & h_{m3} & \dots & h_{mn} \end{pmatrix}$$

In SVD we wish to find an optimal single parameter (which will be a linear combination of the original pixel parameters) which will separate the two populations. The basis of this separation is a key vector which is produced in the SVD process. Without SVD it would only be possible by trial and error to determine this optimal combination. A best estimator would be of the form

$$s = \sum_{i=1}^{i=m} a_i h_i$$

Where s is the score value, a_i are the weights associated with the key vector and b_i are the parameters. Ideally all the pixels containing forest would have a score of say +1 and the background pixels another score of say 0. Unfortunately no combination of parameters will produce such a perfect separation. In practice the score values for both the pixel types will be spread around the two means. These score values are often plotted as histograms (Fig. 2), using the technique known as histogram separation. The goodness of an estimator may be defined by how closely it maps the two image types onto the respective values of 0 and 1. As the estimator improves, the width of the standard deviations, (or variances), of the two histograms will decrease. It is natural therefore to define the best estimator as the one which minimises the standard deviations of the two score values as much as possible. The optimal estimator is the one which will map the respective pixel types onto 1 and 0, subject to the minimisation of

$$e = \frac{1}{n_f} \sum_{i=1}^{i=n_f} (s_{f_i} - 1)^2 + \frac{1}{n_b} \sum_{i=1}^{i=n_b} (s_{b_i})^2$$

Where n_f is the number of forest pixels and n_b is the number of background pixels and s_f and s_b are the scores corresponding to particular events. This is often referred to as minimisation in the least squares sense. This optimal choice of parameter weights ($a_1, a_2, \dots, a_n$), can be considered as a vector in n dimensional space. This direction defines the key vector.¹⁰ It is the direction in multidimensional space which provides optimal separation of the two populations.

As a simple introduction to singular value decomposition it is instructive to look at a model two dimensional example consisting of two independent populations of events, each characterised by parameters 1 and 2. One of the populations has a circularly symmetric distribution while the other is elliptical, Fig. 1. In this example it is assumed that there is no correlation between the distributions in the x and y directions (parameters 1 and 2). The first characteristic vector lies along the direction v_1 , the direction along which the major variation in the dataset is observable. The second characteristic vector v_2 , is orthogonal to the first. The key vector is defined by the direction along which, in the least squares sense, the two populations are optimally separated in the two dimensional space. This separation is evident from the distribution of the histograms along the key vector direction.

The key vector is obtained using singular value decomposition. We note that the definition given is equivalent to finding a vector k such that

$$\mathbf{O}k = a$$

Where a is a column vector of size $1 \times m$. The n th element of a is one if the corresponding row in $\mathbf{O}$ contains forest parameters and 0 otherwise. That is, the scalar product of the multi-dimensional vector associated with each event and the key vector, will be 1 if the pixel contains forest and 0 if the pixel contains background. Therefore

$$\mathbf{WLV}k = a$$

or

$$k = \mathbf{V}^{-1} \mathbf{L}^{-1} \mathbf{W}^{-1} a$$

Note as both $\mathbf{V}$ and $\mathbf{W}$ are orthogonal these matrices can be inverted by taking their transpose. Also as $\mathbf{L}$ is diagonal, $\mathbf{L}^{-1}$ can be obtained by taking the reciprocal of the diagonal terms. Therefore once the SVD has been calculated there is little computational overhead in obtaining the key vector. In practice it is usual to constrain the number of included dimensions to those in which the signal to noise ratio is reasonable. This will reduce the accuracy of the solution, but render it more robust with respect to noise and spurious variations in a small number of pixels.

Fig. 1. Example of the SVD technique applied to a two-dimensional model, * Population1, o Population 2.

The optimal key vector $\mathbf{k}$ is now applied to the simulation matrix $\mathbf{O}$ in order to compute a vector $\mathbf{s}$ of m score values, one for each pixel.

$$\mathbf{s} = \mathbf{O}\mathbf{k}$$

A threshold value can now be applied, typically with a value of about 0.5. The pixels with a score above this value are considered to be forest and those below background. More formally, assuming that the score values are normally distributed for both the forestry and background data it can be shown that the optimal value at which to locate the threshold (that which will give the greatest classification accuracy) is given by

$$\frac{N_f}{\sigma_f^2} \exp \left(\frac{(s - \mu_f)^2}{2 \sigma_f^2} \right) = \frac{N_b}{\sigma_b^2} \exp \left(\frac{(s - \mu_b)^2}{2 \sigma_b^2} \right)$$

Where N_f and N_b are the number of forest and background pixels respectively, μ_f and μ_b the means (ideally 1 and 0) and σ_f and σ_b the standard deviations.

4. METHODOLOGY

4.1 Acquisition and processing of satellite data

The satellite data used in this study comprises of the following bands with dates:

SPOT P, XS1, XS2, XS3.	28-04-93
Landsat TM bands 5, 4 and 3.	02-05-90

As can be seen the dates for the data are approximately 3 years apart, but the images were taken at the same time of year in mid-spring, when forest is easiest to classify. The images were geometrically corrected using the nearest neighbour method and were subsequently corrected for the Irish national grid. SPOT XS1-XS3 and Landsat TM bands, were then resampled to 10m resolution by pure scaling and a 3x3 median filter was then applied.

The location of the selected area is in the County Wicklow area of Ireland at Lat 53° N, Lon 6° 13' W and 320000m E, 197000m N on the Irish national grid. The forest in this area is very mixed and comprises of approximately 20 different species, including Corsican pine, Oak, Ash, Grand fir, Lodgepole pine. The mean area of the various stands is approximately 20000m².

4.2 SVD applied to this data

Due to the large variation of land cover classes and species within the forests, a multiple key vector technique was used to filter the data in stages. An area of mixed forest comprising pure Sitka spruce, pure Scots pine, mixed Sitka Spruce, Scots pine and mixed Nobel fir, was selected and extracted for the analysis of forest and background classes. Four spectrally identifiable background classes were chosen from the background and data sets extracted for these classes. The data was split in two, with one half used for training and the other used for testing the results of the SVD process and key vector analysis.

The first 3 dimensions of the SVD were used to generate all the key vectors, as this was found to be the optimum value. Thresholding of the key vector for each background class is then applied by calculating the optimum threshold level, from the training data. The SVD process developed for this analysis plots the distribution of the classified data about the ideal score values of 0 (background ie. non-forest) and 1 (signal ie. general forest) and calculates an estimated accuracy based on the two distributions as can be seen in Fig. 2.

The key vector generated from the training data was then applied to the test data to verify the integrity of the key vector, and the accuracy of the result was then calculated and the distribution of score values plotted Fig. 2. To verify that the result was statistically accurate a theoretical classification accuracy based on the test data was calculated, so that small sample sizes did not invalidate the results. This process was repeated for each background class. The results are tabulated in table 1.

Areas of pure species were then selected to enable the classification of 3 independent species, Sitka spruce, Scots pine and Japanese larch. Each species was then characterised using SVD against the each of the other two species in turn, eg. Sitka spruce became the signal and Japanese larch then Scots pine, became the background. This resulted in 6 key vectors for the species being produced, each with their own theoretical and actual classification accuracy's. The results are tabulated in table 2. A typical key vector using the SPOT & Landsat bands, for Sitka spruce as signal and Scots pine as background is:

$$K = \begin{bmatrix} +0.0195 \\ -0.0135 \\ +0.0212 \\ +0.0169 \\ -0.0018 \\ -0.0005 \\ +0.0139 \end{bmatrix}$$

To classify individual species in an image the forest/non-forest key vectors must first be applied to filter out the background classes from the image, retaining forest areas. The species key vectors can then be applied to the forest data to classify the species in the forest.

5. RESULTS

5.1 Analysis of forest v. non-forest classification

The above process was applied to the general forest data against each individual background class. This was extended to see if the bands chosen affected the classification. Three sets of key vectors were generated, one set for each of the following bands: SPOT P XS1-XS3, Landsat TM and SPOT P XS1-XS3 with the Landsat TM bands. The results are tabulated in table 1, with corresponding histogram plots for the SPOT bands shown in Fig. 2.

Table 1. Results of SVD and Key vector analysis on test data.

Background Class	SPOT P, XS1-XS3 bands only			Landsat TM bands 3, 4, 5 only			Landsat TM and SPOT bands 3		
	1	2	3	4	5	6	7	8	9
	Acc %	T.Acc %	Thresh	Acc %	T.Acc %	Thresh	Acc %	T.Acc %	Thresh
1	100	100	0.6177	100	100	0.491	100	100	0.6178
2	100	100	0.4535	100	100	0.6358	100	100	0.704
3	100	100	0.7638	100	100	0.3832	100	100	0.7638
4	100	100	0.7581	100	99.854	0.5589	100	100	0.6967

Where: Acc denotes classification accuracy for test data, T.Acc denotes Theoretical classification accuracy and Thresh is the threshold level computed from training data.

As can be seen from table 1, and Fig. 2, the SVD process is extremely good at separating forest from non-forest with accuracy's in excess of 99%. The theoretical accuracy's based on the distribution of the signal and background data sets are also excellent. These results are probably over-optimistic, due to the fact that both the forest and background data were chosen from un-mixed pixels, in the centre of the cover class.

Figure 2. Plots showing distribution of score values. a,c,e,g are results of training for each background class. b,d,f,h show results of key vectors applied to test data sets. a,b-background1. c,d- background2. e,f- background3. g,h- background4.

5.2 Analysis of species classification

Species classification has always been inherently difficult, due to the small variations in reflectance at different wavelengths, but as shown SVD is excellent in separating these variations. This will be possible using SVD, providing that no other species change in the same way across the spectrum. Six key vectors were produced from training data for each chosen species against every other. This is excessive, as 3 key vectors one for each species, could be generated without the need for every variation, but as can be seen in table 2, the results do vary between corresponding key vectors. Further filtering would enhance classification's over large areas.

Table 2. Results of SVD and Key vector analysis on species test data.

Species Class eg. Spruce v Pine	SPOT P, XS1-XS3 bands only 1			Landsat TM bands 3, 4, 5 only 2			Landsat TM and SPOT bands 3		
	Acc %	T.Acc %	Thresh	Acc %	T.Acc %	Thresh	Acc %	T.Acc %	Thresh
SS v SP	98.81	98.16	0.4237	97.63	82.81	0.5725	94.47	98.64	0.4035
SS v JL	100	99.95	0.6639	98.9	99.40	0.5707	81.87	91.19	0.6349
SP v SS	98.81	97.96	0.5742	95.26	75.2	0.3959	95.26	98.39	0.5939
SP v JL	97.78	100	0.7425	99.26	92.40	0.5438	77.04	91.75	0.6398
JL v SP	100	100	0.2538	100	99.9	0.4853	85.93	98.75	0.2618
JL v SS	100	100	0.3348	100	99.99	0.475	87.91	97.8	0.2795

Where: Acc denotes classification accuracy for test data, T.Acc denotes Theoretical classification accuracy and Thresh is the threshold level computed from training data, SS=Sitka spruce, SP=Scots pine, JL=Japanese larch.

As can be seen in table 2 and Fig. 3, the species classification accuracy varies between 0-4% for the (1) SPOT and (2) Landsat TM, and there is even greater variation of 4.7-23% for (3) Spot & Landsat. The SPOT bands on their own are far better at classifying species than the Landsat or a combination of SPOT & Landsat, yielding an overall classification accuracy of 99.2%. The resolution of the SPOT data complements the area of forest chosen for this study, because of the large variation over small areas.

3a

3b

Figure 3. Plots for species classification. a,c,e show the distribution of score values for the training data. b,d,f show the distribution of test data after the key vector has been applied. The threshold levels and accuracy's are also displayed. a,b- SS/SP. c,d- SS/JL. e,f- JL/SP.

6. CONCLUSIONS

As can be seen in section 5, the results are very good and are extremely promising for the classification of this type of remote sensing data. Although these results are for a best case scenario it is anticipated that this method will be still valid for general forest and species classification. It is estimated that with mixed pixels that the classification accuracy will be between 85-100% depending on the number of filter stages (using key vectors) applied to the data and the variation of the land cover classes.

In forested areas such as the one chosen for this study, where there are a large number of different species, key vectors for each species would have to be generated. This technique should be able to classify mixed pixels in an image with the aid of a rule or intelligent decision system such as Fuzzy logic or Neural networks.

Future work will concentrate on building a knowledge base of species key vectors and developing an intelligent classification system that can classify mixed pixels and other forest characteristics. This will extensively involve the simulation of data to increase the number of pixels available for training and verification.

7. ACKNOWLEDGEMENTS

The SPOT and Landsat TM data utilised in this study was provided through a collaboration between the Forest Institute of Remote Sensing Technology, University College Dublin and Leeds Metropolitan University. Funding to perform the statistical analysis and image processing was provided by the Faculty of Information and Engineering Systems, Leeds Metropolitan University.

8. REFERENCES

1. Richards, J.A., "Thematic mapping from multitemporal image data using the principal components transformation". *Remote Sensing of Environment*, vol. 7, pp. 1653-1667, 1984.
2. Conese, C., Maracchi, G., Miglietta, F., Maselli, F., and Sacco, V.M., "Forest classification by principal component analyses of TM data". *International Journal of Remote Sensing*, vol. 9, no. 10 & 11, pp. 1597-1612, 1988.
3. Mather, P.M., "A computationally-efficient maximum-likelihood classifier employing prior probabilities for remotely-sensed data". *International Journal of Remote Sensing*, vol. 6, no. 2, pp. 369-376, 1985.
4. Bolstad, P.V., and Lillesand, T.M., "Rapid Maximum Likelihood Classification". *Photogrammetric Engineering & Remote Sensing*, vol. 57, no. 1, pp. 67-74, 1991.
5. Wilson, J.D., "A comparison of procedures for classifying remotely-sensed data using simulated data sets". *International Journal of Remote Sensing*, vol. 13, no. 2, pp. 365-386, 1992.
6. Benediktsson, J.A., Swain, P.H., and Ersoy, O.K., "Neural network approaches versus statistical methods in classification of multisource remote sensing data". *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 540-552, 1990.
7. Brockhaus, J.A., and Khorram, S., "A comparison of SPOT and Landsat-TM data for use in conducting inventories of forest resources". *International Journal of Remote Sensing*, vol. 13, no. 16, pp. 3035-3043, 1992.
8. De Wulf, R.R., Goosens, R.E., De Roover, B.P., and Borry, F.C., "Extraction of forest stand parameters from panchromatic and multispectral SPOT-1 data". *International Journal of Remote Sensing*, vol. 11, no. 9, pp. 1571-1588, 1990.
9. Benediktsson, J.A., Swain, P.H., and Ersoy, O.K., "Conjugate-gradient neural networks in classification of multisource and very-high-dimensional remote sensing data". *International Journal of Remote Sensing*, vol. 14, no. 15, pp. 2883-2903, 1993.
10. Danaher, S., and O'Mongain, E., "Singular Value Decomposition in multispectral radiometry". *International Journal of Remote Sensing*, vol. 13, no. 9, pp. 1771-1777, 1992.
11. Danaher, S., O'Mongain, E., and Walsh, J., "A new cross-correlation algorithm and the detection of rhodamine-B dye in sea water". *International Journal of Remote Sensing*, vol. 13, no. 9, pp. 1743-1755, 1992.
12. Danaher, S., Fegan, D.J., and Hagan, J., "Application of singular value decomposition (SVD) in high energy gamma-ray astronomy". *Astroparticle Physics*, vol. 1, pp. 357-368, 1993.
13. Herries, G., "Characterisation of forestry species using satellite images". Unpublished-Internal Project Report, Faculty of Information and Engineering Systems 34/93, Leeds Metropolitan University.
14. Press, W.H., et al., Numerical Recipes - The Art of Scientific Computing. Cambridge Univ. Press, Cambridge, 1986.
15. Golub, G.H., and Van Loan, C.F., Matrix Computations. Johns Hopkins Univ. Press, Baltimore, MD, 1983.
16. Simonds, J.L., "Applications of characteristic vector analysis to photographic and optical response data". *Journal of Optical Society of America*, vol. 53, pp. 968-975, 1963.