

Supervised Band Selection for Optimal Use of Data from Airborne Hyperspectral Sensors

Michael Riedmann and Edward J. Milton

Copyright (c) 2003 IEEE. Reprinted from IGARSS 2003 IEEE International Geoscience and Remote Sensing Symposium, Toulouse, France, CD-ROM.

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Southampton's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Supervised Band Selection for Optimal Use of Data from Airborne Hyperspectral Sensors

Michael Riedmann and Edward J. Milton

School of Geography
University of Southampton
Southampton, SO17 1BJ, UK

M.Riedmann@soton.ac.uk, E.J.Milton@soton.ac.uk

Abstract—This paper presents a practical supervised band selection procedure for airborne imaging spectrometers and Maximum Likelihood classification (MLC) as data application. The output band set is optimal in band location, width and number regarding the MLC accuracy of the classification task. The supervised algorithm is based on feature selection and requires a user-defined class set. For two given semi-natural vegetation data and class sets, the selected band sets performed superior to established vegetation band sets used in current satellite and airborne sensors, most noticeably for the first few bands. The algorithm was implemented in IDLTM/ENVITM. It may also be used for feature selection, the generation of class-discriminate colour composites, the prioritization of already existing band sets, and the determination of the intrinsic discriminant dimensionality of the data set.

Keywords—band; channel; feature; selection; hyperspectral; classification;

I. INTRODUCTION

Hyperspectral imagery allows the detection and exploitation of narrow spectral absorption or emission features of target classes of interest, leading to an improved identification and discrimination of ground targets, and characterization of their related physical or chemical properties. However, hyperspectral data have a large volume that is difficult to record, store, and transmit efficiently. For some hyperspectral sensors, e.g. the Itres Instruments Compact Airborne Spectrographic Imager (CASI-2), the data recording rate limits the amount of data to be recorded per scan. A compromise in swath width is often made for hyperspectral capability. In addition, high spectral resolution results in a reduced signal-to-noise ratio (SNR), unless coarser spatial resolution is used, which may not be in the interest of the user. At processing level, with increased number of bands to analyze, the complexity of the data and that of the classifier increase and a larger set of training samples is required in order to adequately define class statistics. As most of hyperspectral data are redundant for Earth surface targets, dimensionality reduction methods may be applied.

Band selection reduces the data volume at sensor level. A band is a wavelength interval in the electromagnetic (EM) spectrum that may correspond to the bandpass of a channel or channel assemblages. A channel is a physical CCD detector element that records signals of

specific wavelengths of the EM spectrum. In general, band selection needs to consider characteristics of

- i) the scene (target reflectance properties),
- ii) the sun (solar spectrum curve),
- iii) the atmosphere (atmospheric windows),
- iv) the sensor (spatial and spectral resolution, bandpass, point-spread function, wavelength range, SNR), and
- v) the final data application.

The first four data attributes may be taken into account by using hyperspectral data of a representative part of the scene acquired with the sensor of interest under similar solar and atmospheric conditions. For example, the CASI-2's enhanced spectral mode allows the acquisition of hyperspectral data (288 bands) at the cost of a reduced swath width (101 pixels). The band selection algorithm may then be applied to these data, before the full-swath data set (512 pixels) may be flown with the optimal band set of maximum 19 bands in CASI-2's spatial mode. Otherwise, hyperspectral scene data need to be simulated considering all four attributes.

With respect to the data application, the authors focused on Maximum Likelihood classification (MLC) for this study. Some supervised band selection methods found in the literature are based on feature selection, extracting the optimal subset of bands out of the original set according to a class-separability criterion using a sub-optimal search technique. In the context of pattern recognition, features are input dimensions into a processing algorithm and may represent image bands or transforms of bands. Search techniques aim to alleviate the computational burden of exhaustive searches which could very quickly cease to be computationally feasible. An example is the bottom-up search algorithm that constructs the final feature subset incrementally by including features from preceding subsets. Ref. [1] used a bottom-up search procedure with the probabilistic Divergence measure as single-band criterion. A band expansion stage was included immediately after each new band had been selected. Ref. [2] also used a bottom-up method, but chose and then broadened the best available band on the basis of the band set performance of the Mahalanobis distance.

In addition, feature extraction methods such as Principal Components Analysis (PCA) and Discriminant Analysis (DA) have been applied for band selection for MLC applications. They transform the original measurement space into a lower-dimensional one. Ref. [3] created the ‘discriminant power’ measure on the basis of eigenvectors and λ -values to rank bands. Ref. [4] proposed a top-down band selection method combining PCA and DA before grading the bands using eigenvalues and λ -vectors.

This paper proposes an improved supervised band selection (SBS) procedure based on feature selection that aims to find a band set that is optimized in band location, number and width with respect to MLC accuracy of the given classification task.

II. SUPERVISED BAND SELECTION ALGORITHM

The algorithm was developed using hyperspectral HyMAP and CASI data sets acquired over the New Forest area in Hampshire (UK) in June 2000, and over the River Severn Estuary in October 1997, respectively. For both data sets information classes were defined. A flowchart of the SBS algorithm is presented in Fig. 1.

A. Criterion Function

The optimal criterion for band selection for MLC as application would be the MLC accuracy, which is, however, computationally inefficient. Possible surrogates include class-separability measures, which quantify the distances between class distributions in feature space. Probabilistic separability measures are preferred to heuristic ones as they exploit differences in the class covariance matrix as well as in the class means. Saturated measures, such as Transformed Divergence (TD) and Jeffries-Matusita are more justified than unsaturated ones (Divergence and Bhattacharyya, respectively) as the entire distance range will be evaluated in this algorithm. The authors chose the TD measure as criterion, as it gave the highest correlation with the MLC accuracy for the given data sets. The ‘maximum average’ multi-class form of the criterion was favoured to the ‘maximum minimum’ type, as the latter may introduce a bias by focussing only on the class pair hardest to separate. The band achieving the highest TD band set performance when added to the set was selected for the final set.

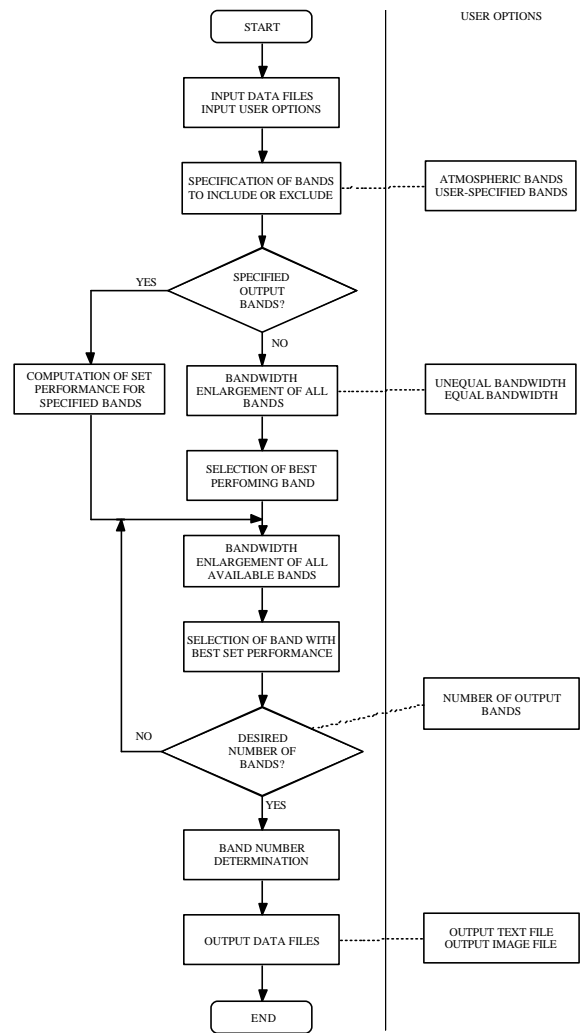
B. Search Method

It can be shown that an exhaustive search for a typical band selection task will result in an infeasible computational task using the surrogate TD measure. The ‘bottom-up’ forward selection was preferred to the ‘top-down’ equivalent, as the number of features to be selected for the final set is much smaller than the number of original bands in the hyperspectral data set. Monte Carlo search methods were avoided as they produce inconsistent results.

Figure 1. Flowchart of the supervised band selection algorithm.

C. Bandwidth Increase

Two neighbouring bands may be merged, i.e. their signals summed, to improve band set performance. Contrary to [1] and [2], bandwidth increase in this algorithm takes place before a new band is added to the set, as bandwidth increase alters the performance of the band set. In addition, adjacent narrow bands of the final band set



were not merged by default, as they may, for example, exploit the top and bottom of a narrow spectral feature.

This subroutine starts by comparing the mergers of a band with both of its neighbours in terms of overall set performance, and the merger leading to the better criterion performance is carried out. The process is repeated until both left and right mergers decrease the overall set accuracy. A band may not be increased further in width if the smallest of all band means in the set falls below a user-defined percentage (lower limit 20%) of its mean, to assure that all bands achieve a sufficient SNR. Expanded bands that do not reach an adequate signal level are not considered for the final band set.

D. Band Number Determination

This subroutine aims to find the number of bands that are necessary to achieve sufficient classification accuracy, say 95% accuracy of the full band set. The authors defined the proportion of maximum achievable Transformed Divergence (PMATD) as approximation of the proportion of maximum achievable MLC accuracy. It is calculated by dividing the TD value of a band subset by the TD value achieved with the entire band set. The PMATD is based on the linear relationship between the TD and the MLC accuracy and assumes that the maximum TD value is reached when using the entire band set. The optimal number of bands equals to the dimension of the smallest band set, which achieves a PMATD of at least 95%.

III. ALGORITHM EVALUATION

A. Reliability

The reliability of the band selection algorithm was judged by inquiring the appropriateness of its assumptions. Class samples are assumed to be normally distributed which is usually appropriate for a sufficient number of class training samples. In addition, the suitability of the linear relationship between the MLC accuracy and the Transformed Divergence was shown for the given data sets, but needs to be verified for other class and data sets. The sub-optimality of the search procedure used in the SBS cannot be quantified, but is thought to be small for the first few bands of the final band set as the latter are only weakly correlated. The band number determination routine assumes that the maximum achievable MLC accuracy is produced by the entire band set, which is accurate if a sufficient number of training samples are provided to avoid Hughes phenomenon.

B. Effectiveness

The effectiveness of the algorithm was judged by comparing, in terms of the MLC accuracy, the SBS optimal band set with the one generated with an unsupervised correlation-based band selection (UBS) algorithm presented in [5],[6], and vegetation band sets of current satellite and airborne band sets (see Fig. 2). Both optimal band sets achieved similar or higher MLC accuracy than the simulated band sets with only the first 2 or 3 bands for both data sets. In addition, the accuracy of optimal band sets was compared to the one of randomly and equally spaced band sets. The optimal sets performed superior only for the first three dimensions, indicating that band selection is most effective for the firstly selected bands for the given class and data sets.

IV. CONCLUSION

This paper has presented the first results from an improved supervised band selection approach to optimizing data collection for Maximum Likelihood classification using programmable airborne hyperspectral sensors. The evaluation showed that for the given data sets an optimal performance can be expected particularly for the firstly selected bands. The algorithm was implemented in IDLTM/ENVITM. It may also be used for feature selection, the generation of class-discriminate colour composites, the prioritization of already existing band sets, and the determination of the intrinsic discriminant dimensionality of the data set.

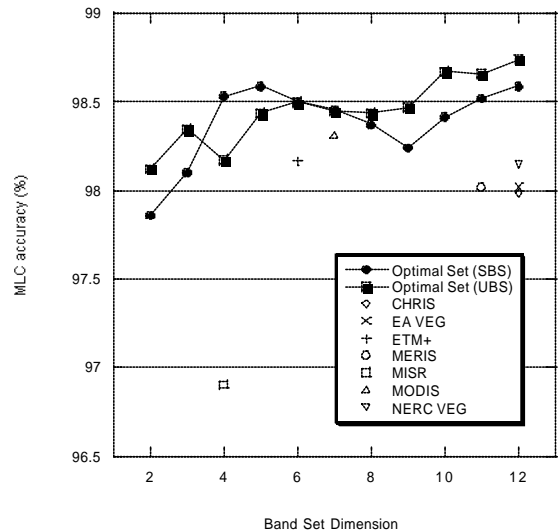


Figure 2. MLC accuracies of the optimal SBS and UBS band sets compared with those of satellite and airborne band sets for the New Forest data set.

ACKNOWLEDGMENT

The HyMAP data were collected for the BNSC/NERC SAR/Hyperspectral Airborne Campaign (SHAC). The CASI data were acquired by the Environment Agency, UK. The authors thank Karen Anderson for her advice in the training site selection for the New Forest data.

REFERENCES

- [1] G.M. Petrie, P.G. Heasler, and T.A. Warner, "Optimal band selection strategies for hyperspectral data sets," *IEEE IGARSS'98 Sensing and Managing the Environment*, pp. 1582-1584, 1998.
- [2] P.J. Withagen, E. den Breejen, E.M. Franken, A.N. de Jong, and H. Winkel, "Band selection from a hyperspectral data-cube for a real-time multispectral 3CCD camera," *SPIE Aerosense, Algorithms for Multi-, Hyper-, and Ultraspectral Imagery VII*, S.S. Shen and M.R. Descour, Eds., pp. 84-93, 2001.
- [3] T.-M. Tu, C.H. Chen, J.L. Wu., and C.I. Chang, "A fast two-stage classification method for high-dimensional remote sensing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, pp. 182-191, 1998.
- [4] F. Csillag, L. Pasztor, and L.L. Biehl, "Spectral band selection for the characterization of salinity status of soils," *Remote Sensing of Environment*, vol. 43, pp. 231-242, 1993.
- [5] M. Riedmann and E.J. Milton, "Development and evaluation of an in-flight band selection methodology for optimal use of data from airborne hyperspectral sensors," *Annals of Geophysics*, in press.
- [6] M. Riedmann, *Development and evaluation of a band selection methodology for optimal use of data from airborne hyperspectral sensors*, PhD Thesis, School of Geography, University of Southampton, UK, 2003, unpublished.