Tájökológiai Lapok 6 (1–2): 119–125. (2008)

DOI: 10.56617/tl.4153

HYPERSPECTRAL DATA PROCESSING CHAIN DEVELOPMENT PERSPECTIVES FOR VEGETATION STUDIES

KÁROLY BAKOS

Telecommunication and Remote Sensing Laboratory
University of Pavia, Department of Electronics
Via Ferrata 1. 27100 PAVIA, ITALY

Keywords: Remote Sensing, Hyperspectral, Vegetation, Image processing, Digital Imagery

Abstract: In this paper the possibilities of hyperspectral data processing is investigated regarding to the application of these images in natural and ecological applications. A short overview is given of the available methods for interpretation purposes and special attention is paid on how the unique properties of hyperspectral data are affecting the choice from available methods for data processing. A broad identification of possible processing chain is discussed with the aim of developing more modular and application driven way of processing of the large data volumes. Automatic or semi-automatic procedures are proposed and key steps are to be identified that could lead to high quality mapping products by means of digital signal processing. Some experimental results are published and the broad methodology is presented that is aimed to use for identify data processing chain for vegetation mapping purposes. This work is to be continued with testing the performance at different stages of interpretation while different techniques are used. Furthermore over the HYPER-I-NET research network a document is to be supplied with the collection of results and application specific suggestions regarding to hyperspectral data processing for vegetation monitoring purposes.

Introduction

Remote sensing applications are widely used in ecological and landscape ecological studies, The great advantages of using remote sensing techniques such as the repeatability of analysis, the continuous data collection and the possibility of taking objective measures resulted greater concern about remotely sensed data and automated or semi-automated interpretation procedures for ecological studies. Presently through the intense growth of technical and technological capabilities both in terms of sensor design and interpretation techniques new possibilities are occurring in remote sensing applications for natural and ecological studies.

Such a new possibility is the application of hyperspectral sensors. Presently hyperspectral remotely sensed datasets are more and more available, therefore it is of importance to examine the possibilities of applying them. Hyperspectral image data possess a wide range of features which can make those datasets extremely useful for ecological researches. On the other hand these features make the interpretation procedure of imagery more complex hence more complicated. New algorithms and data processing chains need to be designed thus allowing users to easily exploit useful information of the image data.

Indeed hyperspectral data has such a level of complexity that not only “technical” processing has to be revised but even conceptual changes in interpretation seem to be necessary. Many studies exists that address hyperspectral data processing for different applications but there is a lack of conceptual background on coherent processing design. Most of the studies report great success in terms of interpretation results but in the same time the designed methodologies are only useful for those particular problems.

Purpose and scope

This paper leverages the results of a research project funded as support action by the European Community, entitled HYRESSA (HYperspectral REmote Sensing in Europe Specific Support Actions). This project already addressed the issues about the data processing of hyperspectral data by means of a dedicated expert workshop, hold at DLR premises in July 2006. HYRESSA findings are summarised in the final report of that workshop (HYRESSA SWOT and User Needs workshop report 2007). In the document the main findings are that the lack of standards regarding to data and data processing coupled with missing suitable processing techniques are most limiting factors of hyperspectral remote sensing applications.

This paper tries to address those problems applying an operational point of view and tries to establish conceptual design mechanisms for the application of hyperspectral remote sensing in ecological studies. This is done by reviewing remote sensing applications for natural resources and highlighting points where hyperspectral methods have particular advantages over other methodologies. While doing this special attention is paid on the aim of a coherent processing design, where data can be further processed, thus allowing users to save costs while analysing. A modular design is proposed where starting from simple tasks it is possible to reach a high level interpretation of hyperspectral imagery, while allowing the user to choose optimal processing chain. With this approach it is aimed to construct a feasible data processing mechanism which is application-driven.

According to the above-mentioned aims the document is structured as follows. First a brief introduction is given about special characteristics of hyperspectral imagery focusing only on properties that can be relevant in ecological applications. In the same section a brief introduction of present processing techniques is also given. After that, remote sensing applications for natural resources are briefly introduced in order to determine key issues and steps required in data processing chain design regarding special data characteristics. Finally, a conceptual design of a modular data processing chain is investigated over the example of vegetation monitoring application.

Hyperspectral imagery

Hyperspectral sensors (often referred to as Imaging spectrometers) are instruments that acquire images in many, very narrow, contiguous spectral bands throughout visible, near-IR, mid-IR and thermal-IR portions of the spectrum (LILLESAND 2004). These systems typically collect 200 or more bands of data, which enables the construction of an effectively contiguous reflectance spectrum of every pixel in the scene. These systems are able to discriminate among features of earth surface that have diagnostic absorption and reflection characteristic over narrow wavelength intervals that are not present while relatively coarse bandwidths of conventional multispectral scanners are used and for a given geographic extent the data can be viewed as a cube, having two dimensions corresponding to spatial position and one that represents wavelength. The comparison and the data cube concept are illustrated in Figure 1.

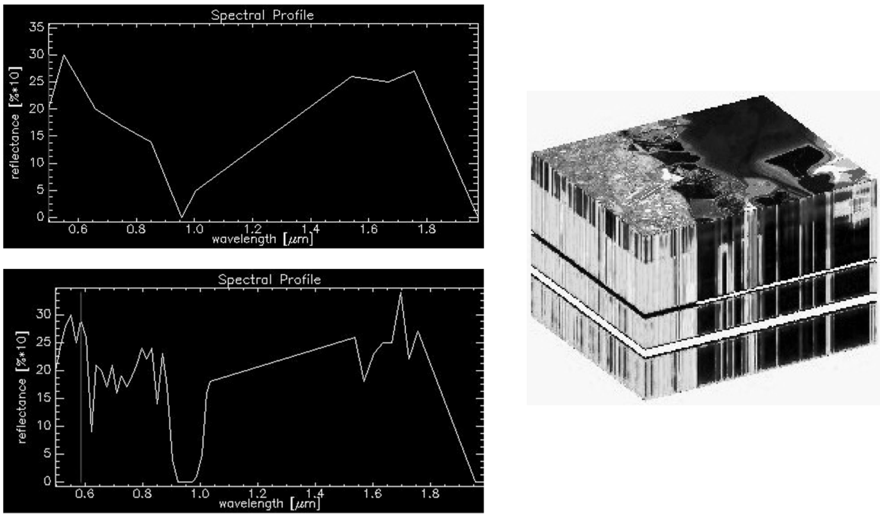

Figure 1. The comparison of pixel brightness values in multispectral and hyperspectral datasets (a) and the illustration of hyperspectral data cube concept (b)
1. ábra Multispektrális és hiperspektrális adatok összehasonlítása tárolt adatok tekintetében (a) és a hiperspektrális adathalmazok reprezentálása (b)

On one hand the property of having large number of consistent descriptive data from each pixel is an advantage as it allows more comprehensive comparison and differentiation among those pixels. On the other hand from a processing point of view this property makes particularly difficult the data processing and analysis. The difficulties can be originated from the high dimensionality of data as it increases computational time and costs while requires sophisticated methods for processing. Another problem is that statistical methods based on training data and its probability distribution, although used very efficiently in multispectral data interpretation, have strong limitations in hyperspectral processing. This is because of the high number of training data needed due to the high number of wavebands often referred to as Huges phenomenon (SWAIN and DAWIS 1978).

According to these, standard data processing requires further steps before classification could be done. A regular approach consists of feature extraction and feature selection stage prior to classification. Feature extraction and selection are techniques used to reduce the size of the datasets and making the classification procedure easier. For feature extraction rule based expert systems (SRINIVASA 1991, SRINIVASA and RICHARDS 1993), source separation (spectral mixture analysis) (CHANG 2007) or different transformations can be used. These transformations include the Principal Component Analysis (PCA) (LANDGREBE 2003), independent component analysis (HYAVARINEN et al. 2001) projection pursuit (LANDGREBE 2003) and Minimum noise fraction (MNF). Discriminant Analysis Feature Extraction (DAFE) (LANDGREBE 2003) and Decision Boundary Feature Extraction (DBFE) (LEE and LANDGREBE 1997) are also available for transformation purposes. If spatial context is to be considered during image processing (QUATTROCHI and GOODCHILD 1997, RAMSTEIN and RAFFY 1989; SERRA 1989; WOODCOCK et al. 1988) some methods such as co-occurrence texture analysis (HARALICK et al. 1973) and the semivariogram analysis (MATHERON 1997) or even mathematical morphology (SOILLE 2003, BENEDIKTSSON et al. 2003) can be used. For feature selection purpose usually indices are used such as he Euclidean Distance-, the Mahalanobis Distance-, the Transformed Divergence-, the Jeffries-Matsushita (J-M) Distance-, or the Bhattacharyya Distance and the Histogram Distance Separability Indices.

After feature extraction and selection procedure is done the data has to be classified. For classification statistical methods such as maximum likelihood, minimum distance, parallelepiped algorithms, geometric approaches such as Spectral Angle Mapper (SAM) (KRUSE et al. 1993) or Support Vector Machines (SVM) (GUALTIERI et al. 1998) or even non-parametric or multi-stage approaches including Neural Networks (NN) and Decision Tree Classifiers (DTCs) can be used (RICHARDS 2006). As mentioned before, due to the huge dimensionality of datasets usage of statistical approaches is limited to cases where feature extraction and selection can reduce the data into an appropriate volume (FUKUNAGA 1990).

As can be seen above several way exist for the interpretation, hence considering the application it is easier to define the needed processing chain that may be suitable for the given application.

Remote sensing for natural resources

Remote sensing tools are often used in natural studies. Not only quantitative parameters can be measured but also qualitative assessment of the resources can be carried out by means of interpretation techniques (RICHARDS 2005). Monitoring of vegetation, soil, water or even atmosphere is possible and furthermore throughout hyperspectral data interpretation we can acquire more accurate and detailed information for analytical purposes than ever before (MARINO et al. 2000). Quantitative analysis of image data is mostly referring to the process of producing thematic maps based on the image data. In this procedure there are particular advantages of hyperspectral imagery, allowing more accurate distinction among different land cover classes present on a scene. These advantages are very useful in vegetation mapping. Throughout qualitative analysis of hyperspectral images more detailed analysis of image data can be carried out therefore more accurate modelling of underlying processes is possible. For example by applying spectral mixture analysis it is possible to identify the underlying materials for each pixel on the scene that the particular reflectance signature consists of, therefore very detailed information on surface materials can be derived (CHANG 2007). This technique is also useful for estimating parameters of the land covers such as moisture content of soils, eutrophication level of surface waters or different parameters of vegetative species (GONG et al. 1992). As an example, it is possible to estimate and model canopy chlorophyll content of vegetation or the accurate estimation regarding to vegetation condition of different species can also be done. However there are many experiments with really promising results in the above areas but usually those experiments are driven by a unique processing technique. This highlights the need of designing processing chains that are based on the desired application and allows the beneficial and efficient usage of hyperspectral remotely sensed data.

Vegetation mapping using hyperspectral imagery

With the above aims extensive research is carried out at the University of Pavia within the Hyper-I-Net project to define processing chain for vegetation analysis. Although the results will be published in a different article some interim results are presented here. Possible processing techniques at each stage of the image interpretation are tested and performance is assessed. The aim of the process is to identify feasible and modular way of data processing thus allowing the user to consider more choices for interpretation. Besides this, accuracy assessment of the thematic products is investigated too, and a consistent way to assess classification errors is desired to be defined.

For the initial experiments a 1025×1025 pixel 3m spatial resolution DAIS sensor image was used covering a mountainous vegetated area in the Tatra Mountains. The vegetation map of the area was made available by the University of Warsaw. For the experiments a set of classifications was carried out with applying different data processing chain. For feature reduction Principal Components forward rotation, DBFE and transformed divergence based feature selection was done and then the images were processed with different classification methods using different sets of training data with different spatial distribution and resolution. For classification mainly maximum likelihood as statistical and SAM as geometric approaches were used but it is aimed to further extend the number of studied algorithms. In terms of overall accuracy, the best result could be obtained by using a maximum likelihood algorithm on 10 spectral bands selected by transformed divergence feature selection procedure, but for particular land cover types I.e. SAM applied on the first 3 bands of PCA forward rotated data enabled higher user and mapping accuracy. The figures below are outcomes of the experiments showing the mapping accuracy values while applying different input, classification and accuracy assessment techniques for a cover type (*Vaccinietum myrtilli*) that was difficult to identify.

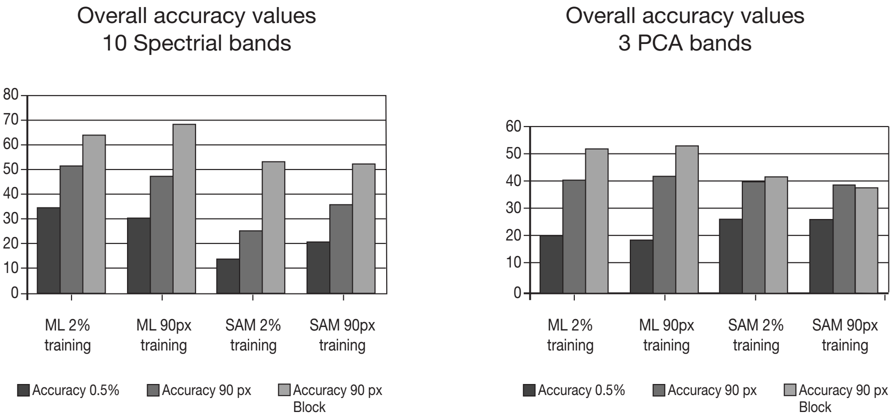

Figure 2 Accuracy values measured for *Vaccinietum myrtilli* classifying two set of inputs under different conditions. The group of bars represents the way the classifier was trained and the colours of bars represent how the accuracy was assessed
2 ábra Mért pontossági érékek (*Vaccinietum myrtilli*) két különböző bemenetei adathalmaz és pontossági mérés esetén. A z oszlopcsoportok az osztályozás tanítási részét reprezentálják az oszlopok színe pedig igazodik a pontosság mérésének módjához.

Through this example it can be seen that maximum likelihood algorithm performs better when spectral information is being used but the difference is lower when transformed dataset is used as input for the classifier. It is also important to mention that PCA bands gave more balanced result, which can be important if there are other classes to be mapped too. This could help to make an optimal trade-off between costs and quality of interpretation.

As was highlighted above it is important to address the issue of hyperspectral data processing for vegetation studies. By examining the performance of different processing chain it is possible to design a modular processing chain that allows to reach targeted quality level while limiting processing costs. By selecting applications for which the processing methodology is designed helps users to apply hyperspectral data more efficiently than the application is suited to the available processing techniques. The research aim of the ongoing experiments to find a feasible methodology of applying hyperspectral remotely sensed images in ecological and natural studies can be reached and application driven processing chains can be designed. It is also clear that more experiments are needed to find optimal solution for feature extraction, -selection and classification of imagery. Other methodologies such as support vector machines (SVM), Neural networks (NN) or decision tree classifiers should be investigated to serve more comprehensive comparison of possible solutions. The re-usability of intermediate results at each stage of the processing should be considered too thus allowing the set up a really modular design.

Acknowledgements

This work is carried out at the University of Pavia as a part of HYPER-I-NET project. I would like to thank to Csaba Centeri for his endless patience, to Dániel Kristóf for his help when it was needed and to Paolo Gamba for the guidance and advices for preparing this article. I would also like to thank for my family and friends.

References

BENEDIKTSSON J. A., PESARESI M., ARNASON K. 2003: Classification and Feature Extraction of Remote Sensing Images from Urban Area based on Morphological Approaches, IEEE Trans. On Geoscience and Remote Sensing, 41: 1940–1949.

CHANG C. I. 2007: Hyperspectral data exploitation. Theory and applications. John Willey and Sons Inc.

FUKUNAGA K. 1990: Introduction to Statistical Pattern Recognition. Academic Press, Second Edition

GONG P., HOWARTH P. J. 1992: Land-use classification of SPOT HRV data using a cover-frequency method. International Journal of Remote Sensing, 13: 1459–1471.

GUALITIERI J. A., CROMP R. F. 1998: Support vector machines for hyperspectral remote sensing classification. In: Merisko, R. J. (ed.), 1998: Proc. SPIE–27th AIPR Work-shop Advances in Computer Assisted Recognition, 3584: 221–232.

HARALICK R. M., SHANMUGAM K., DINSTEIN I. 1973: Texture features for image classification. IEEE Trans. Systems Man Cybernet. 3: 610–621.

HYVARINEN A., KARHUNEN J., OJA, E. 2001: Independent Component Analysis, John Wiley and Sons, New York,

HYRESSA SWOT and User Needs workshop report, accessed on line at http://www.hyressa.net/documents/, April 2007.

KRUSE F. A., LEFKOFF A. B., BOARDMAN J. B., HEIDEBRECHT K. B., SHAPIRO A. T., BARLOON P. J., GOETZ A. F. H. 1993: The Spectral Image Processing System (SIPS) – Interactive Visualization and Analysis of Imaging spectrometer Data. Remote Sensing of the Environment, 44: 145–163.

LANDGREBE D. A. 2003: Signal Theory Methods in Multispectral Remote Sensing. John Wiley and Sons, Hoboken, New Jersey

LEE C., LANDGREBE D. A. 1997: Decision Boundary Feature Extraction for Neural Networks, IEEE Trans. on Geoscience and Remote Sensing, 8: 75–83.

LILLESAND T. M., KIEFER R. W. 2004: Remote Sensing and Image Interpretation. Wiley & Sons, Fifth Edition.

MATHERON G. 1997: Principles of geostatistics. Econ. Geol. 58: 1246–1266.

MARINO C. M., PANIGADA C., BUSETTO L., GALLI A., BOSCHETTI M. 2000: Environmental applications of airborne hyperspectral remote sensing: asbestos concrete sheeting identification and mapping, Proc. of the 14th International Conference and Workshops on Applied Geologic Remote Sensing

QUATTROCHI D. A., GOODCHILD, M. F. (eds.) 1997: Scale in remote sensing and GIS. Boca Raton: CRC Lewis Publishers

RAMSTEIN G., RAFFY, M. 1989: Analysis of the structure of radiometric remotely-sensed images. International Journal of Remote Sensing 10: 1049–1073

RICHARDS J. A. 2005: Analysis of remotely sensed data: the formative decades and the future. IEEE Trans. Geosci. Remote Sensing, 43: 422-432.

RICHARDS J. A., XIUPING, J. 2006: Remote Sensing Digital Image Analysis: an introduction. Springer Verlag, Berlin (4th edition)

SERRA J. P. 1989: Image analysis and mathematica morphology. International Journal of Remote Sensing 10: 1049–1073.

SOILLE P. 2003: Morphological Image Analysis: Principles and Applications. Springer Verlag, Berlin (2nd edition)

SRINIVASAN A. 1991: An artificial intelligence approach to the analysis of multiple information sources in remote sensing. Ph.D. thesis, Univ. New South Wales, School of Elect. Eng., Kensington, Australia

SRINIVASAN A., RICHARDS J. A. 1993: Analysis of GIS spatial data using knowledge-based methods. Int. J. Geograph. Inf. Syst., 7: 479–500.

SWAIN P. H., DAWIS S. M. (eds) 1978: Remote Sensing. The Quantitative Approach. New York McGraw Hill

WOODCOCK C. E., STRAHLER A. H., JUPP D. L. B. 1988: The use of semivariogram in remote sensing and simulated images: Real digital images. Remote Sensing Environments 25: 349–379.

HIPERSPEKTRÁLIS ADATFELDOLGOZÁSI LÁNC MEGALAPOZÁSA NÖVÉNYTANI KUTATÁSOKHOZ

KAROLY BAKOS

Telecommunication and Remote Sensing Laboratory
University of Pavia, Department of Electronics
Via Ferrata 1. 27100 PAVIA, ITALY

Kulcsszavak: távérzékelés, hiperspektrális, vegetáció, képfeldolgozás, digitális leképezés

Jelen cikk a hiperspektrális légi- és űrfelvételek feldolgozásával foglalkozik, különös tekintettek ezen felvételek alkalmazási lehetőségeit vizsgálja természetvédelmi és ökológiai kutatásokban. Röviden bemutatásra kerülnek azon módszerek, melyek segítségével egy hiperspektrális adatfeldolgozó lánc állítható össze. A vizsgálódás célja egy olyan módszertan és adatfeldolgozó lánc megalkotása, mely általánosan alkalmazható vegetációs térképezési célokra hiperspektális felvételek segítségével. Jelenleg számos példa áll rendelkezésre specifikus módszerekre, amelyekben hiperspektrális adatok osztályozásával különböző paramétereket nyernek ki ökológiai alkalmazások számára, de ezen módszerek alkalmazhatósága általában csak az adott ökológiai alkalmazásra korlátozódik. A jelen cikkben megfogalmazott célkitűzések és bemutatott részeredmények a paviai egyetemen a HYPER-I-NET program keretében zajló kutatás részét képezik, melyben a hiperspektrális adatok alkalmazási lehetőségeit és tulajdonságait vizsgáljuk általános vegetációs alkalmazások számára. A cikk célja a kutatási munka bemutatása és tudományos megalapozása, valamint az, hogy áttekintést nyújtson a további kutatások irányáról. A program várható eredménye egy olyan moduláris és alkalmazás-centrikus adatfeldolgozó lánc és annak dokumentációja, mely nagyban segítheti a hiperspektrális távérzékelési technológiák nyújtotta lehetőségek kihasználását természetvédelmi és ökológiai tervezési és monitorozási folyamatok során.