HYPERSPECTRAL MIVIS SCANNER DATA INTEGRATED INTO A GIS FOR AN INDUSTRIAL AREA

A. CAVAZZINI (*), E. MICCADEI (**), F. SILVESTRI (***) A. FERRARINI (****) N. ZACCARELLI (****)
(*) CGR (Compagnia Generale Ripreseaeree), Italy
(**) University of Chieti, Italy
(***) ISPESL, Italy
(****) CISIG (Consorzio per l'Innovazione dei Sistemi Informativi Geografici dei grandi bacini fluviali), Italy

KEY WORDS: MIVIS, hyperspectral scanner, classification, risk assessment, industrial area

ABSTRACT

ISPESL is the National Institute for the control and monitoring of risks connected with the presence of industrial plants and activities. A project between ISPESL Institute and CISIG (Consortium among National Research Council of Italy, University of Parma and Consorzio Compagnie Aereonautiche) has been developed aiming to test the integration of ground data on the principal industry in the area of the Pescara river basin and remote sensing data devived from both aerial photogrammetry and MIVIS scanner data flights. The general philosophy of the project is to integrate orthophotos (that allow a precise georeferentiation of different data sets and the possibility to precisely compare observations made in different periods) with data directly observed, *in situ* measurements of principal parameters that describe the industrial risks of the area and data recorded by the hyperspectral MIVIS scanner. The MIVIS scanner, property of National Research Council of Italy, is a 102 channels scanner covering visible and near infrared (0.43-0.83 μm), middle infrared (1.15-1.55 μm and 1.98-2.50 μm) and thermal infrared (8.21-12.70 μm) regions of the electromagnetic spectrum, providing a wealth qualitative information of the project area. The association of precisely georeferenced orthophotos, ground data and qualitative remote sensing images (MIVIS) is the state of the art technique to build GIS systems using remote sensing informations. MIVIS data in the middle and thermal infrared were found to be principally applicable for the study of the health status of vegetation, soil conditions and water turbidity. These informations are entered in a GIS for modelling the impacts resulting from geologic risks.

1. INTRODUCTION

Remote sensing and geographical information systems (GIS) have much to benefit from each other (Wilkinson, 1993). Remote sensing offers the potential for the regular provision of large spatial data sets on landscape properties required in a wide range of applications. On the other hand, the integration of remotely sensed data with other spatial properties may improve the analysis and the extraction of information from multispectral and hyperspectral data. Land cover maps are used for many purposes. In environmental engineering projects and in environmental and hydrological studies, accurate and up to date information is often required. Knowledge of changes in land cover is becoming increasingly important from both the ecological and economic point of view. Land cover information is obtained by classifying remote sensed data and the resulting thematic maps can be integrated into a GIS for later use with other data sets. The crucial aspect of this operation is the quality of the remotely sensed data. Poor quality data sets are useless for landcover mapping investigations even if ancillary data informations are added.

Starting from 1995 the MIVIS scanner, property of the National Research Council of Italy, started to fly on a CASA 212/c aircraft of Compagnia Generale Ripreseaeree, making new data sets available to the scientific community for geological, environmental and thermal analyses. This new and very powerful instrument makes possible very detailed land classifications to be integrated in GIS systems and in the connected databases.

The MIVIS scanner is a "whisk broom" scanner and, according it, has a geometrically correct scan line that, due to the movement of the aircraft, is displaced with the roll, pitch yaw, and with changes in velocity and direction. The operational flight heights of the scanner can range from 1500 meters up to 5000 meters above ground; at this height, the nadiral pixel dimension ranges from 3 meters up to 10 meters (the IFOV is 2 mrad wide). Due to these fact, the data acquisition of the scanner is affected by a lot of geometrical distortions. The cross-track scan line dimension is geometrically good in the central region, being affected laterally by a geometrical distortion; in fact, the large FOV (72°) of the scanner, makes the

geometrical distortion not insignificant. For the same reason, the pixel size on the ground is also quite different at the nadir and at the border of the same scan line. For example, for a flight height of 1500 meters the nadiral pixel size is of 3 meters while the last pixel of the scan line is 5.1 meters. The scan rate of the scanner can be varied according to the flight height from 25 scan/sec (1500 meters height), down to 6.25 scan/sec, making the scanner similar to a "push broom" sensor because of the high scan speed compared to the speed of the aircraft. The main geometrical distortions that affect the MIVIS images are due to roll, pitch yaw, variations in altitude and direction of the aircraft. The roll distortion is automatically compensated by the system itself, making the correction of the starting and ending point of registration.
The scanner has a GPS system and a gyro. It is able to record every second the data coming from these instruments together with the spectral data; the CASA 212/C aircraft on which the scanner is installed, has a complete GPS system for accuracy navigation and determination of the trajectory of the aircraft.
Both these systems are able to record data every second, that means, for low altitude flights, every 25 scan lines, and it is possible to record the position of the aircraft data set; unfortunately, it is not always possible to have a differential correction using a ground station; this aspect causes an error of an order of several meters (i.e. many pixels) in determining the right position.
Finally, terrain elevation plays the most important role in the geometric distortion of the data, both in the position of the pixels and in theirsize; a difference of 100 meters in elevation corresponds to a linear difference of 0.2 meters in pixel size, i.e. 7% for a 1500 meters flight. Different routines have been written and applied to the MIVIS data to resample the images in order to minimize all these effects.
Meanwhile, other geographic data sets have been made available to the community to be used for the adjustment of the MIVIS images. Colour orthophotos covering the whole italian country at the scale of 1:10.000 have been produced by CGR and are available for geographic co-registration of spectral data. Thanks to the use of the orthophotos, many examples of georeferenced data have been produced for different areas, with comparable results. For the purposes of this study, a second order polinomial warping algorithm has been used to coregister the MIVIS images to the orthophotos.
The advantage of the integrated use of geographical and remote sensed data is becoming apparent (Catlow et al., 1984; Van der Laan, 1988; Kenk et al., 1988). The aim is to enable feed back of remote sensing derived information to a GIS. This article briefly describes an example of the use of MIVIS instrument to define different land use classes for risk assessment in an industrial area of the Pescara river basin (Sulmona, Italy) resulting in a GIS obtained by the integration of the MIVIS data, the classification layer and the orthophotos at the scale of 1:10.000.
The study is being conducted by CISIG Consortium (Parma, Italy) for ISPESL, a National Health Research Agency that evaluates risks in industrial areas, in collaboration with the University of Chieti.

2. METHODOLOGY

A linear discriminat analysis method (Dillon et al., 1984) has been used to classify the MIVIS hyperspectral images. In previous studies and attemps (Ferrarini et al., in preparation), linear discriminant analysis has proven its suitability for classifying MIVIS images with high accuracy.
Discriminant analysis is characterized as follows: there are two types of multivariate observations - the first, called training samples, are those whose group identity (i.e., membership in a specific group) is known *a priori,* and the second type, referred to as test samples, consists of observations for which *a priori* information is not available and wish to be assigned one within the *a priori* known groups. An observation is classified into a group if the squared distance (also called the Mahalanobis distance) of the observation from the group center (mean) is the minimum. An assumption is made that covariance matrices are equal for all groups (Dillon et al., 1984; Johnson et al., 1992). There is a unique part of the squared distance formula for each group (the linear discriminant function for that group). For any observation, the group with the smallest squared distance has the largest linear discriminant function and the observation is then classified into this group. Linear discriminant analysis has the property of symmetric squared distance: the linear discriminant function of group i evaluated with the mean of group j is equal to the linear discriminant function of group j evaluated with the mean of group i.
Statistical considerations in discriminant analysis have to do with distributional assumptions concerning the observations and measures of separation among the groups. These assumptions have been tested before applying the discriminant analysis function.
Figure 1 exhibits a portion of the area that has been classified. It is an intensively developed region with residential and commercial zones, industries, cropland, herbaceous and shrub-brushland, deciduous forest and barren land. All classes of interest have been carefully selected and defined to succesfully classify MIVIS data into land-cover information. This requires the use of a classification scheme containing taxonomically correct definitions (Jensen, 1996) of classes of information. There is a fundamental difference between information classes and spectral classes (Campbell, 1987): information classes are those that human beings define; spectral classes are inherent in the remote sensed data and must be labelled by the analyst.

Nine land cover classes were distinguished: forest, grass (crops), bare soil, water, asphalt (concrete), tiles, rocks, roofs. Shadows were listed as unclassified. Because of the high degree of intermixing between residential, commercial and industrial uses, the number of anthropic landcover classes is preminent.

3. RESULTS

Table 1 summarizes the error matrix and figure 2 exhibits the classification image.
The classification was found to be of high quality, with 96.8 per cent of cases correctly allocated. It is also evident that much of the misclassification arose from the commission of cases from the asphalt-concrete and the roofs, being the roofs mainly composed by the same endmembers. It is interesting that the overall accuracy index and the k coefficient (Congalton and Mead, 1983) exhibit a large agreement even if the k computation incorporates the off-diagonal elements as a product of the row and column marginals. The high accuracy of the per pixel classification is indicative of land cover classification performance using MIVIS data.

Figure 1. A portion of the area that has been classified (Sulmona, Italy)

The classification image has been georeferenced using an image to map (georeferenced orthophoto) registration with a second order polinomial warping algorithm. The root mean square (RMS) error for this trasformation was 0.3 pixels (0.9 meters). As a map, a precise georeferenced orthophoto has been used.
For the purposes of this paper, spectral imagery, orthopotos and classified images have been incorporated in an ArcView GIS© project as raster layers on which vector information is displayed, this making data more significant. The methodology for the GIS implementation of MIVIS images involves four stages: 1) data extraction (MIVIS data scanning) 2) information extraction (classification algorithms and other spectrals computations) 3) data integration (use of external ground data) 4) data analysis (GIS integration, land cover and land suitability analysis, vegetation stress and water turbidity quantification, edaphic parameters mapping…).
The forest thematic layer has been isolated and a NDVI (Normalized Difference Vegetation Index) has been calculated. The association of the vegetation index with a thermal band has been useful to check and geographically locate stressed vegetation. The nearness to heavy traffic and industries was found to be the main reason for vegetation stress.
Streams and canals were isolated as a thematic layer and surface water turbidity was quantified by a blue and

Classification	Reference Data									
	shadows	*forest*	*bare soil*	*water*	*rocks*	*roofs*	*grass/crops*	*tiles*	*asphalt*	*Total*
shadows	399	1	1	0	1	0	0	3	10	415
forest	3	297	5	0	0	0	18	2	0	325
bare soil	0	1	894	0	0	2	1	2	2	902
water	0	0	0	11	0	0	0	0	1	12
rocks	0	0	1	0	114	1	0	4	2	122
roofs	0	0	1	0	2	516	0	0	22	541
grass/crops	0	5	0	0	0	0	432	0	0	437
tiles	0	0	0	0	0	0	0	499	0	499
asphalt	2	0	1	0	1	17	0	3	368	392
Total	404	304	903	11	118	536	451	513	405	3645

Overall Accuracy = 0,968 K coefficient = 0,963

Table 1. Error matrix

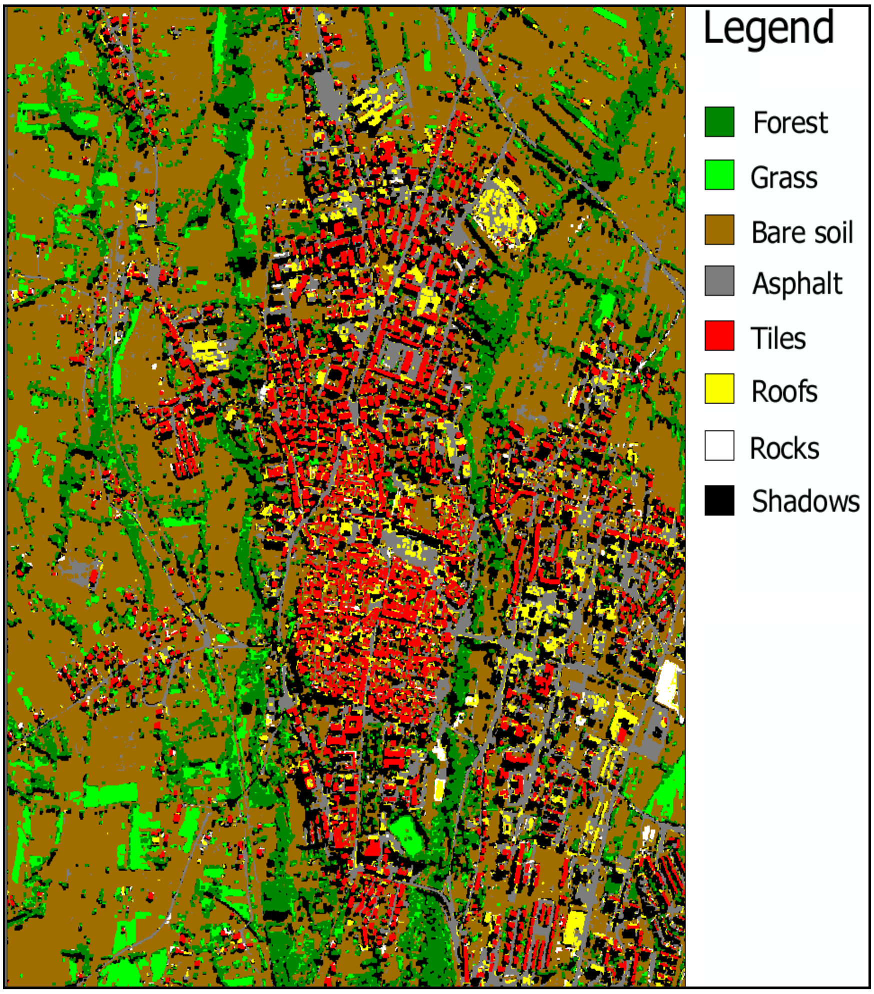


Figure 2. The classified image

thermal wavelenghts analysis. Ground data were useful for calculating a regression between blue wavelengths brightness values and surface water turbidity (figure 3). Vegetation stress and water turbidity data were transferred into ArcView GIS© both with orthophotos and classification image thematic layers.

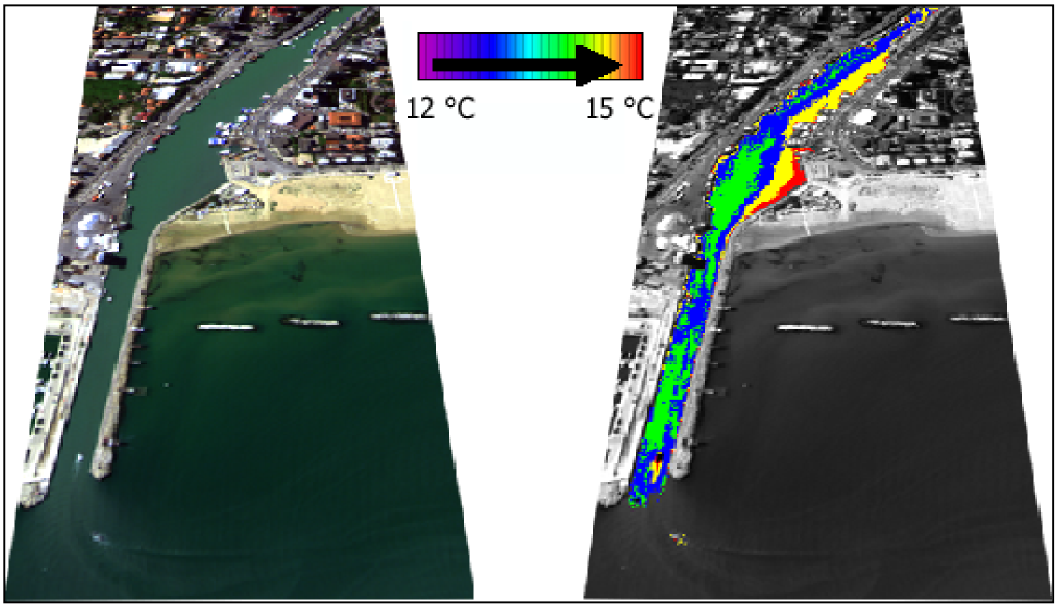


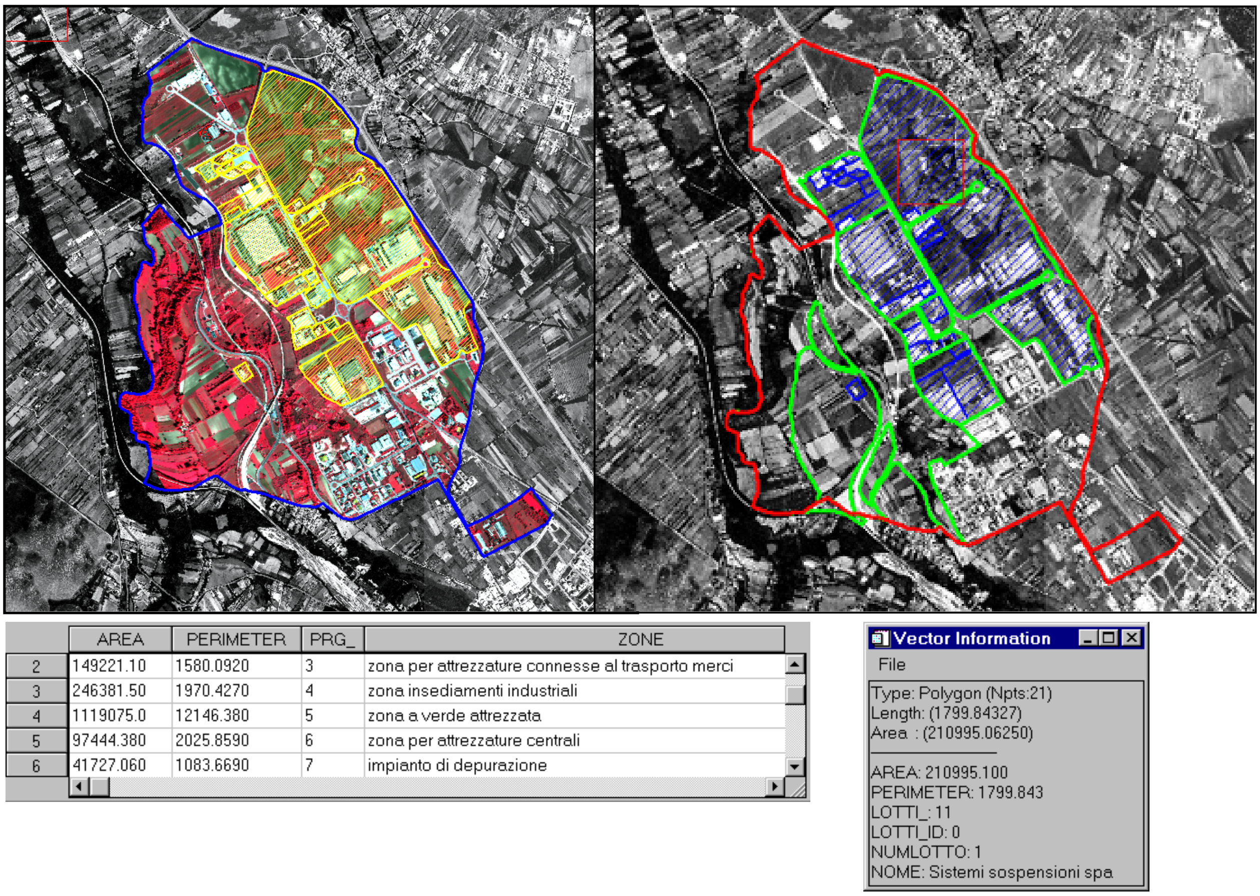
Figure 3. A thermal infrared classification of the Pescara stream (Sulmona, Italy)

Again, all relevant boundaries from the classified image were digitized and stored as a polygon layer in ArcView. Data base creation was also required for the perfect integration of MIVIS data. A complete combination of soil, vegetation, physical and chemical data was used. The geology and hydrography overlays digitized for the study area were interfaced with the slope angle and elevation data from the DEM. Thematic overlays were superimposed and studied as composite images. For the purposes of this research, a very detailed focus on the industrial area was implemented (figures 4 and 5) by the registration of several quantitative and qualitative information about the industrial installations. As a result, composites of thematic overlays on the MIVIS classified images were useful in discerning the spatial pattern of known phenomena. Also, such tecniques were found useful to examine phenomena and specific attributes not appearing on the orthophotos but readily apparent on the GIS images.
Further applications will be computed. Geologic mapping of the area is an image treatment that is possible through edge-detection and edge-enhancement filtering of suitable MIVIS wavelenghts. The spatial pattern of geologic structures can be analyzed and various rock and mineral types highlighted.

4. CONCLUSIONS

Geographical Information Systems can be used for processing spatial data to assess the risks of environmental contamination. Their use depend upon the amount and quality of the available data, data processing procedures and models for the calculation of health risk.
In this paper, MIVIS data have been used to assess the utility of hyperspectral imagery in detecting environmental anomalies such as stressed vegetation, water pollution and soil contaminated areas. A variety of data were involved: MIVIS images, map data (elevation and georeferenced orthophotos) and field reports. The MIVIS utility to thematic layers development has resulted in a soundly based use of GIS to support monitoring of contaminated areas by allowing calculation of pollution contours and environmental risk assessment (not shown here).
Procedures for environmental sanitation are still hampered by financial and methodological problems. The potential economic and scientific gain of using MIVIS hyperspectral images can be recognized given the quality of the data and the high geometric resolution achieved by the sensor. MIVIS sensor offers facilities to accurately classify the landscape, get advanced water and pedogenetic information, study the spatial variability of spectral informations and locate stressed vegetation. The integration of MIVIS data with GIS support is a crucial element. The objective of using this source of

information is a system constructed to handle environmental data, to support risk assessment and to derive sensible conclusions from the hyperspectral data.



	AREA	PERIMETER	PRG_	ZONE
2	149221.10	1580.0920	3	zona per attrezzature connesse al trasporto merci
3	246381.50	1970.4270	4	zona insediamenti industriali
4	1119075.0	12146.380	5	zona a verde attrezzata
5	97444.380	2025.8590	6	zona per attrezzature centrali
6	41727.060	1083.6690	7	impianto di depurazione

Figure 4. Registration of quantitative and qualitative informations about the industrial installations (Sulmona, Italy) as raster and vector layers into the ArcView GIS©

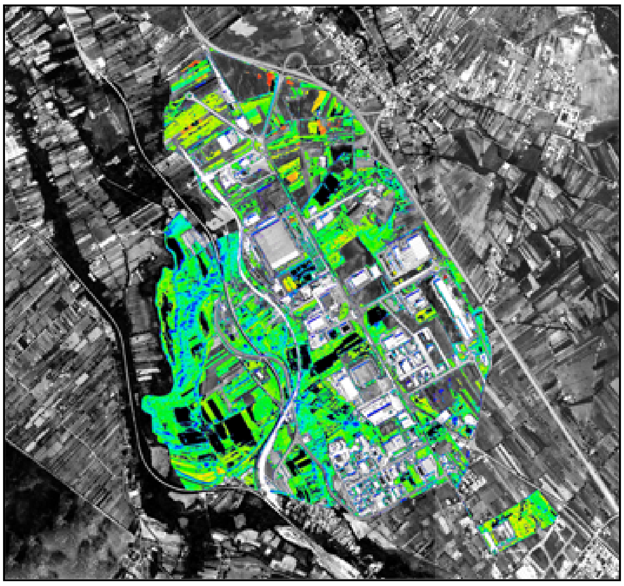


Figure 5. Thermal anomalies (red pixels) of bare soil in the industrial area of Sulmona

The results that have been achieved are promising enough to propose MIVIS hyperspectral sensor as a tool for supporting and integrating geographical information systems for risk assessment in an industrial area. When MIVIS images and GIS are combined, a powerful tool is created. Spectral imagery provides critical input to GIS thematic layers, maintaining the currency and accuracy of the data and serving as a base for the data layers. GIS, on the other hand, provides the capability to have additional information incorporated and provides further information data layers that cannot easily be extracted from the spectral imagery only.

ACKNOWLEDGEMENTS

We are pleased to acknowledge the comments and suggestions of dr. Riccardo De Paulis and dr. Ennio Giusti and to confirm that none of them is responsible for any errors that still remain.

REFERENCES

Campbell J., 1987. Introduction to Remote Sensing, New York, Guilford Press, 551 p.

Catlow D.R., Parsell R.J., Wyatt B.K., 1984. The integrated use of digital cartographic data and remotely sensed imagery. Proceedings Integrated Approaches in Remote Sensing Guildford, UK, ESA SP., 214, pp. 41-45.

Congalton R.G. and Mead R.A., 1983. A quantitative method to test for consistency and correctness in photointerpretation. Photogrammetric Engineering and Remote Sensing, 49(1): 69-74.

Dillon W. and Goldstein M., 1984. Multivariate Analysis, Methods and Applications. John Wiley & Sons.

Kenk E., Sondheim M. and Yee B., 1988. Methods for improving accuracy of thematic mapper ground cover classification. Canadian. Journal of Remote Sensing 14, 1:17-31.

Ferrarini A., Zaccarelli N., Confronto tra algoritmi parametrici di classificazione per dati iperspettrali MIVIS (in preparation).

Johnson R. and Wichern D., 1992. Applied Multivariate Statistical Methods, Third Edition, Prentice Hall.

Jensen J.R., 1996. Introductory Digital Image Processing: A Remote Sensing Prospective. Prentice Hall Series in Geographic Information Science, Keith C. Clarke Series Editor.

Van der Lann F.B., 1988. Improvement of classification results from a satellite image using context information from a topographic map. Proceedings 16th Intern. Congress ISPRS, Kyoto, Japan, Vol.27, Part B8: VII, pp. 98-108.

Wilkinson G., 1993. Neural Networks: a new analysis tool for GIS and remote sensing, GIS Europe, 2, 38-40.