TOR VERGATA UNIVERSITY

Department of Computer, Systems and Production

Geoinformation PhD Programme

XXXIII Cycle

UNIVERSITA' degli STUDI di ROMA
TOR VERGATA

UNSUPERVISED CLASSIFICATION
OF VERY HIGH RESOLUTION (VHR) OPTICAL IMAGES
AND ITS APPLICATIONS

Michele Tommaso Lazzarini

Supervisor: Prof. Fabio Del Frate

Coordinator: Prof. Domenico Solimini

SUBMITTED IN PART FULLFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT TOR VERGATA UNIVERSITY
ROME, ITALY

December 2010

Acknowledgments

Table of Contents

Abstract

Today the continuous increase in archives' size and EO (Earth Observation) sensors' variety requires new methodologies and tools for information mining and management. With this growth, new approaches for image information mining, multi-domain information management, and knowledge management and sharing (in support of information mining) are necessary. An automatic chain reduces the human assistance in the data analysis, allows to gain time on the entire image processing and can be designed to retrieve information in Near Real Time (NRT). Such concepts find an important field of application in the monitoring of urban environments, where the knowledge of changes in either natural resources or man-made structures is a valuable source of information for decision-making processes. These images opened new scenario in classification, considering the information deriving from the study of the context and pattern recognition in addition to the spectral characteristics. This information can be the used in visual interpretation, such as texture and shape, transposed to digital analysis. Besides, the availability of hyperspectral data with a spatial resolution around few meters joins the potentiality

of the VHR imagery with the use of the entire spectrum from visible to the thermal part.

This work aims at proposing an innovative methodology to extract information automatically from urban areas with Very High Resolution images (multispectral and hyperspectral) considering the application of advanced neural networks techniques. Actually, a careful literature review work led to conclude that the research in this application domain is still at a initial step.

After a pre-processing stage to retrieve reflectance values, allowing a comparison of the spectral response in all the images, an exhaustive analysis of the textural and spectral behaviour of the main land cover classes (vegetation, water, bare soil, buildings and bare soil) is performed with the objective of determining which characteristics are common in each cover class and which allow their distinction.

Successively, the accuracy of the proposed methodology has been verified in all the tested images, showing satisfactory results of at least 88 % of the overall accuracy with a significant improvement respects to a standard automatic classification technique as K-means algorithm.

Finally, the potentiality of the proposed method has been shown in the application of automatic classification to urban thermography.

Introduction

The retrieval of information in urban areas: a challenging issue

For the first time in human history, more people are living in towns and cities than in rural areas. Approximately 75 % of the world population lives in urban areas; by 2020, this will be 80 %.[1] These areas represent a small fraction of the land surface but they affect large areas due to the magnitude of the associated energy, food, water and raw material demands.

Moreover, cities are dynamic environments, where urban sprawl is re-shaping landscapes and affecting people's quality of life and the environment.

---

[1] European Environment Agency

It is easy to understand that they are one of the most complex and intricate environments and they are also among the most rapidly expanding and changing elements of the landscape (Fig. 1).

Figure 1:An urban areas in Colorado: it is possible to see changes in land cover from 1992, to 1996 and 2002 (from [9], courtesy of Digital Globe)

These preliminary considerations could explain the necessity to monitor urban areas with a certain accuracy and a high frequency.

Reliable data in populated areas is essential for urban planning: knowledge of changes in either natural resources or man-made structures is a valuable source of information for decision-making processes [1]. Conversely, lack of information contributes to problem such as ineffective urban development programs and activities, unplanned investment projects, poor functioning land markets, and disregard of the environmental impact of developments [2].

Since the late '70s, with the first Landsat missions, satellite remote sensing is increasingly being used as a timely and cost-effective source of information in a urban planning [3], [4].

During the last two decades, significant progress has been made in developing and launching satellites with instruments, in both optical/infrared and microwave regions of spectra, well suited for Earth Observation with an increasingly finer spatial, spectral and temporal resolution [5]. With the recent availability of commercial Very High Resolution (VHR) remote sensing multispectral imagery from sensors such as IKONOS and QuickBird, it is possible to identify small-scale features such as individual roads and buildings in the urban environment (roessner et. al, 2001) and have a specific potential for more detailed and accurate mapping of urban areas [6], [7]. Imagery from these sensors is an important source of timely data, which can be used for the creation and/or updating of GIS vector layer [8].

However, VHR images analysis is one of the most challenging areas for the remote sensing community due to its high spatial and spectral diversity. From the physical composition point of view, several different materials can be used for the same built-up element (for example, building roofs can be made of clay tiles, metal, concrete, glass, plastic). On the other hand, the same material can be used for different built-up elements (for example, concrete can be found in paved roads or building roofs [9]. Fig. 2 shows how the same surface (two types of roads) can be made from different material and two different surfaces (a road and a building) can be made with the same material (concrete). Misclassification occur between buildings and streets, which are caused by spectral similarities between materials covering these surfaces and the influence of shadow.

A different spectral behavior could be explained also from the effect of the atmosphere, which operates at different wavelengths.

Atmospheric correction provides estimates of the radiation emitted and reflected at the surface, and it is necessary for observing parameters that are intrinsic to the surface. accurate atmospheric correction removes the effects of changes in satellite-sun geometry and atmospheric conditions due to aerosol scattering [10]. Atmospherically corrected surface reflectance images improve the accuracy of surface type classification [11], [12] and allow to compare the spectral signatures from different images.
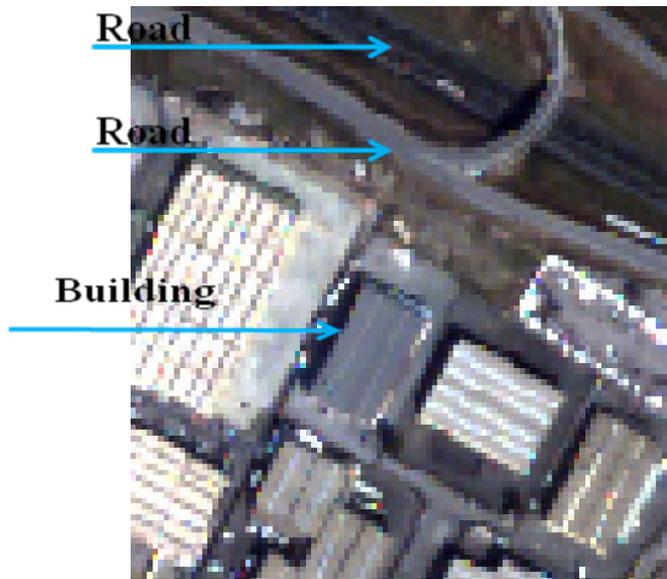
Figure 2: a detail of roads and buildings from Quickbird images (Tor Vergata area)

Therefore, the use of other information on top of spectral information for the urban land-cover classification is essential [13], [14].

This additional information can be the features used in visual interpretation, such as texture and shape (Fig. 3), transposed to digital analysis. The features used in these interpretation keys are: shape, size, pattern, tone, texture, shadows, site and association [15], [16], [17]. This additional information may help to overcome spectral similarities between specific classes [18], [19], [20].
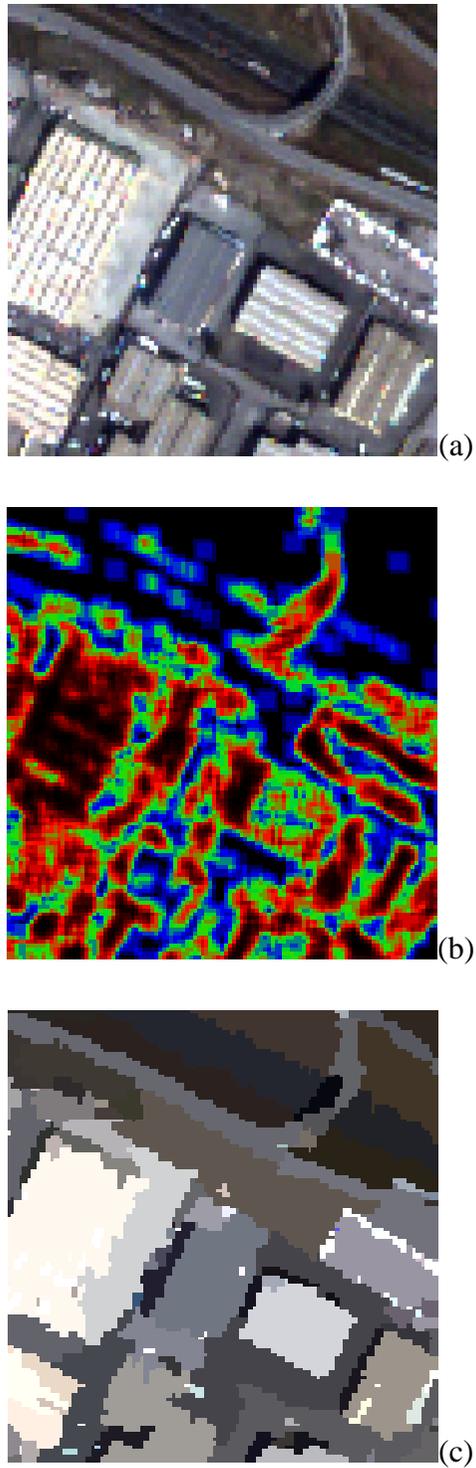
Figure 3: visualization of Quickbird image of Tor Vergata with spectral (a), texture (b) and shape (c) information

Ever since the first multispectral imagery became available from civilian remote sensing satellites in the early 1970s, considerable effort has been devoted to supervised classification of image data with the aim of producing high-quality thematic maps and establishing accurate inventories of spatial classes [21]. There have been advances under several aspects of supervised satellite image classification, developing components of the classification algorithm including the training or learning strategy and approaches to class separation based on statistical or other estimators and class separability indexes.

But today, the continuous increase in archives' size and EO sensors' variety, requires new methodologies and tools for information mining and management. The manual process performed by experts to mine information from images is currently too complex and expensive to be applied systematically on even a small subset of the acquired scenes. The issue might become even more challenging in future since more missions - including constellations - are being planned, with broader sensor variety, higher data rates and increasing complexity.

With this growth, new approaches for image information mining, multi-domain information management, and knowledge management and sharing (in support of information mining and training) are necessary, allowing the user demand for simultaneous access to multi-domain data. An automatic chain reduces the human assistance in the data analysis, allows to gain time on the entire image processing and can be designed to retrieve information in Near Real Time.

Since from the late 60's, unsupervised methods have been developed [22]: they have assumed more importance with the above mentioned reasons and now they represents one of the focal point in feature identification in urban areas [14].

The obtained classified maps of urban area can be used for several purposes, as urban planning [23], crisis management or in support to civil protection activities [24], human settlement damage [25] and urban thermography [26].

To summarize, the idea of this study is to show how to retrieve information on the land cover from VHR images (multispectral and hyperspectral), which is possible exploiting the spectral, textural and shape information from the starting images.
After a preprocessing stage, these data constitute the input for an original automatic classification tool: the proposed processing chain is able to return as output a classified image, that is a valuable source of information for decision-making processes, urban planning and NRT monitoring.

Chapter 1

Automatic classification of Very High Resolution optical images: state of the art

This chapter aims at describing the recent progress in the classification of urban areas, discussing the evolution from the first multispectral sensor from '70s until the recent Very High Resolution multispectral and hyperspectral satellites.

The advantages derived by an improved spatial resolution opened new scenarios in the amount of information to retrieve from urban areas, with a detail of a narrow road and a single building. New image information mining techniques have been developed in these last years and appreciable results derived from the proposed methodology based on neural networks classification methods.

1.1 The evolution of optical sensors

First multi-spectral Landsat sensor was launched in 1972, starting to collect information about Earth from space in several bands. Landsat satellites have taken specialized digital photographs of Earth's continents and surrounding coastal regions for over three decades, enabling people to study many aspects of our planet and to evaluate the dynamic changes caused by both natural processes and human practices[2]. Landsat sensors have a moderate spatial-resolution (30 meters): which means that it is possible to see large man-made objects such as highways and it is coarse enough for global coverage, yet detailed enough to characterize human-scale processes such as urban growth. Nowadays, optical satellite have increased their spatial resolution in order to produce Very High Resolution images.

The first SPOT mission opened the epoch of very high resolution instruments, making available new images characterized by a reduced number of spectral bands (three for visible range and one for NIR range) with a spatial resolution less than 10 meters[3]. Missions like SPOT5, Ikonos [101], Quickbird [102] and World View-II[4] are still providing multi-spectral acquisitions at few meters of spatial resolution and panchromatic data at less than one meter. New applications and products have been developed and made available for the users such as satellite cartography and land cover maps at very high resolution.

With technological improvements, another type of sensors is acquiring more and more importance: they are called hyper-spectral sensors.

A hyperspectral sensor owns several contiguous and narrow bands (Fig. 1.1), frequently reaching hundred of them. The hyperspectral measurements have demonstrated very high performances in several cases, allowing to distinguish the

---

[2] *Landsat programme home page*: http://landsat.gsfc.nasa.gov/
[3] *SPOT image homepage*: http://www.spotimage.com/
[4] *World view II homepage*: http://www.satimagingcorp.com/satellite-sensors/worldview-2.html

response of a specific material. The development of this technology for the space satellites is more complicate and very expensive in terms of payload design, maintaining and calibration. For that reasons many of these sensors are airborne (i.e. MIVIS, AHS[5], ROSIS[6]).

By the way, there are few cases of spaceborne hyperspectral sensors like Hyperion[7], developed by NASA, and CHRIS Proba-1[8], developed by a European consortium founded by ESA. The upcoming EnMap[9] mission from DLR, Prisma[10] from ASI and HySPIRI from NASA, shown an increasing interested from the Earth Observation community in the Hyperspectral sensor development.

---

[5]*AHS*:   http://bo.eufar.net/document/publi/124901161346543eefc58cf:SPIE_RS05_INTA_5978-56.pdf

[6] *ROSIS*: http://www.opairs.aero/media/download/pdf/rosis-description.pdf

[7] *Hyperion Instrument*: http://eo1.gsfc.nasa.gov/Technology/Hyperion.html

[8] *ESA Chris Proba* home page: http://earth.esa.int/missions/thirdpartymission/proba.html

[9] *EnMAP homepage*: http://www.enmap.org/

[10] *PRISMA homepage*: http://www.asi.it/en/flash_en/observing/prisma
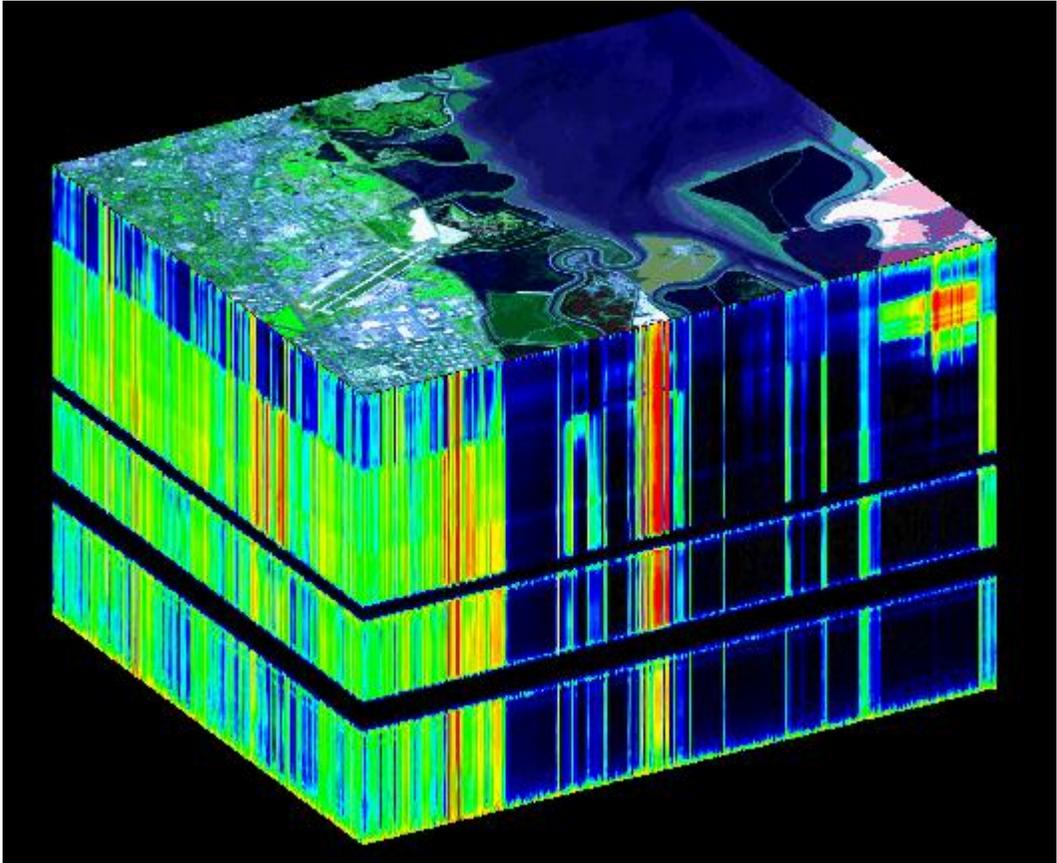
Fig. 1.1: an example of the hyperspectral cube (Moffet Field Airport –California-)

1.2 Classification methods

In order to make use of the multitude of digital data available from satellite imagery, it must be processed in a manner that is suitable for the end user. For many projects this processing includes categorizing the land into its various use functions. There are two main categories that can be used to achieve this outcome and they are called *supervised* and *unsupervised* classification techniques. In a supervised learning, the training will benefit from the assistance of the teacher and desired output is already known: the user guides the image processing software to help it decide how to classify certain features (through some ROI, Regions of interest). For an unsupervised learning rule, the image is classified without benefit of any teacher and the software does most of the processing on its own.

The state of the art of the classification techniques available is following presented with a focus on the unsupervised methods to extract information automatically from images.

1.3 Unsupervised methods

Unsupervised methods have a quite long history in remote sensing.
The most well-known techniques for unsupervised classification are the K-means [103] and ISODATA [104]: these are the fastest from the computational point of view but, they do not provide accurate results in most of the cases.
K-means is sensitive to the choice of the number of clusters, which usually is a critical issue. Different random initializations of the cluster centers result in significantly different clusters at the convergence. Thus, the algorithm is usually run many times with different initializations in an attempt to find a good solution.

Recently, in the machine learning community [105], the main difference is between methods based on:

- Learning from experimental data (examples, samples, measurements, records, patterns, or observations) by neural networks (NNs) and support vector machines (SVMs);
- Embedding existing structured human knowledge (experience, expertise, heuristics) into workable mathematics by fuzzy logic models (FLMs).

This introduction is focused on the first category, considering the unsupervised approach to the problem.

In this context, Transductive SVMs (TSVMs) [106], [107] and semisupervised SVMs (called also S^3VMs) [108] proved particularly effective in several applications and they have been adapted in Remote Sensing. In [109], TSVMs have been applied to Landsat images classification, in [110] S^3VMs have been focused to classify Hyperion data set and in [111] SVMs detected man-made objects from SPOT5 images. About NNs, their advantages will be discussed in section 1.4.1.

Another methodology is represented by the decision tree: a Decision Tree Classifier (DTC) is a hierarchical classifier that recursively partitions the classification problem in smaller sub-problems on the basis of given statistical or empirical criteria. At each node of the tree a simple decision rule with predefined threshold values is used. In [112] this method has been used to classify VHR images.

Appreciable results in automatic classification seem to derive also from the application of Self Organizing Map (SOM) developed by Kohonen [113], [114] during his long studies dedicated to this topic. Even if these particular unsupervised neural networks have been used in image analysis in the second part of the 80's,

they begin to be used in remote sensing in [115] on Landsat images clusterization. Successively, these NN's helped for SPOT image characterization [116].

Another use of SOM can be found in Laaksonen [117] where a hierarchical SOM has been utilized in Content-Based-Image Retrieval (CBIR). Due to the tree structure, the number of map units increase when moving downwards the SOM levels of the TS (Tree Structured) SOM. This concept has been extended in [118] for building detection with spectral and textural parameters in a CBIR; which means that the methodology does not provide a direct delineation of objects of interest but could point out locations in the image where there may be potentially interesting structures.

So, nowadays, there are no developed methods to classify automatically VHR image with SOM.

1.3.1 SOM

A SOM can automatically form one- or multi-dimensional maps of intrinsic features of the input data. These data are presented in mixed random order to the network which is capable of learning complicated hierarchical relations within the considered high-dimensional spaces. In remote sensing, SOMs are used to identify measurements relationships which they can re-organize in several output cluster/classes.

1.3.1.1 Net structure

A neural network consists in a series of inputs and in a n-dimensional grid of neurons. Each input is connected to all neurons of the grid: the resulting matrix of weights is used to propagate the inputs of the network to the neuron on the map.

The mathematical formulation that describes a general net is:

$$y_j = \sum_{n}^{i=1} w_{ji} x_i \qquad \text{Eq. 1.1}$$

Where w_{ji} indicates the weight of the connection between the neuron y_j and the input x_i.

An input is presented to all the neurons of the net: a winning one is generally selected select with the Euclidean Distance d:

$$d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad \text{Eq. 1.2}$$

where i = 1, … n and n is the number of the components of the input, x is the input and y is the neuron weight. It means that the neuron with the lower value d is activated.

The peculiarity of SOM is that the winning neuron has an area of influence, called *bubble*. The interaction that the winning neuron has with its neighbors is defined according to a neighborhood function λ, which normally takes value between 0 and 1.

This interaction determines a modification of the weights dependent on neighborhood function and the response of the neuron. Gauss's function can be used for this purpose:

$$\lambda_{cj} = e^{\frac{|r_c - r_j|^2}{2\sigma^2}}$$

Eq. 1.3

where r_c is the position vector of the winning neuron, r_j is the position vector of the jth neuron of the map (Fig. 1.2) and σ is named proximity parameter.
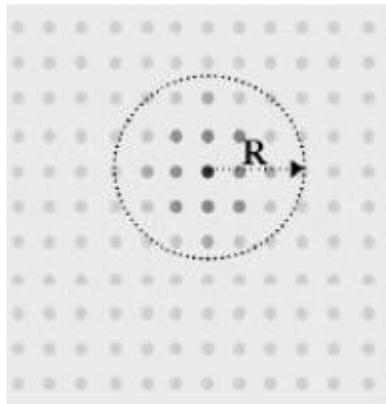
Fig. 1.2: radius of interaction $R = r_c - r_j$

When the process starts, σ parameter has a high value, thus the area of bubble is wide. During the learning phase, the bubble decreases its dimension, until a certain user-defined value.

The weights of the neurons in the bubble are, therefore, updated according to the following formula:

$$w_j(t+1) = w_j(t) + \eta \lambda_{cj}\left(x(t) - w_j(t)\right)$$

Eq. 1.4

where η is the learning rate and decreases gradually during the learning phase.

1.3.1.2 SOM Parameters Setting

Different parameters must be defined in order to synthesize a SOM. The parameters are: dimension, learning rate, neighborhood function and training cycles. In the following subsection a description of these parameters is described.

a) Dimension

A bi-dimensional configuration is usually preferred. Three-dimensional configuration does not give a substantial advantage during the elaboration of information and usually causes a complication in the management of the map itself. The number of neurons is variable and it depends on the purposes of the analysis, the input structure and the internal relationships between neurons.

b) Learning Rate

In the early cycles, η value should be 1 and then decreased until it reaches its minimum value. A possible formulation is:

$$\eta = \eta_{max} \left(\frac{\eta_{min}}{\eta_{max}} \right)^{\frac{t}{t_{max}}} \qquad\qquad \text{Eq. 1.5}$$

where $\eta_{min}=0$, $\eta_{max}=1$, t_{max} = number of cycles and t is the current cycle.

c) Neighborhood Function

This parameter defines the area of the bubble: if it is too small, the bubble does not contains an enough number of neurons inside it and the network could not be ordered after the training phase. It is possible to avoid this problem by assigning to that variable a value greater than the half of the diameter of the network. During the training cycle, it can be reduced until 0.5, which means that only the winning neuron has a modification of its weight.

$$\sigma = \sigma_{max} \left(\frac{\sigma_{min}}{\sigma_{max}} \right)^{\frac{t}{t_{max}}}$$ Eq. 1.6

d) Training Cycles

The training process is a stochastic process, so the number of cycles depends on the specific case: if it is too small, the result does not converge; on the other hand, too high value tends to incorporate many classes in a unique one

In figure 1.3, it is possible to see a Kohonen's map with two dimensions where neuron η_{kj} is connected with its neighborhood.
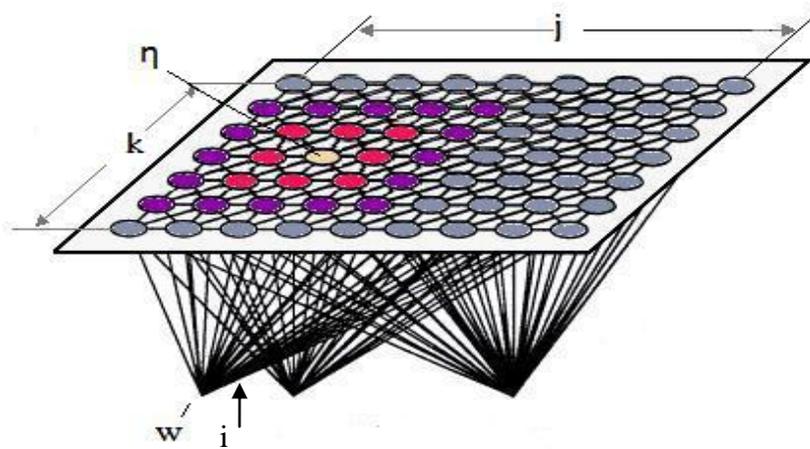
Fig. 1.3: Schematic representation of Kohonen SOM

When an input from the layer *w* activates a neuron in the η layer, the interaction that the winning neuron (selected with a user-defined rule) establishes with its neighbors is defined according to a neighborhood function, which normally takes value between 0 and 1. This interaction determines a modification of the weights dependent on the neighborhood function and the response of the neuron. The η_j neuron activity is defined by:

$$\frac{d\eta_j}{dt} = \sum_{i=1}^{n} w_{ji}\xi_i - \gamma\left(\eta_j\right) + \sum_{k \neq j}^{n} \mu_{ki}\eta_{kj} \qquad \text{Eq. 1.7}$$

Where:

η_j is the activity of jth neuron;

ξ_i is the ith component of the input and n is the total number of inputs;

w_{ji} is the weight of the connection between the jth neuron and the ith input (synapse);

γ_i is a term that considers the leaks of the process;

η_{kj} is the connection between j^{th} and k^{th} neurons.

In Kohonen map, neurons are influenced by the neighbors with a bubble model. With this model the information contained in the inputs are transferred to the synapses and their variation is defined by the following differential equation:

$$\frac{dw_{ji}}{dt} = \alpha\eta_j\,\xi_i - \beta(\eta_j)w_{ji} \qquad \text{Eq. 1.8}$$

where α controls the velocity of learning and β (η_j) the "forgetfulness". It is possible to note that the temporal evolution of synapses depends on the activity of the neurons and relative connections. Two cases may be distinguished, which depends on the presence or absence of the neuron respect to the bubble:

- Neuron inside the bubble. In this case the neuron is in the maximum of the activity ($\eta_j \approx 1$). If the other terms in the equation 1.8 are normalized in order to obtain $\alpha\eta_j \approx (\beta\eta_j)$:

$$\frac{d\mathrm{w}_{ji}}{dt} = \alpha\left(\xi_i - w_{ji}\right) \qquad\qquad \text{Eq. 1.9}$$

- Neuron outside the bubble. This neurons have a negligible activity, for this reason $\eta_j \approx 0$ and Eq. 1.8 become:

$$\frac{d\mathrm{w}_{ji}}{dt} = 0 \qquad\qquad \text{Eq. 1.10}$$

In this case, synapses of this neuron are not modified.

1.3.1.3 SOM Learning Process

The learning process of a SOM is based on three basic steps:

- Competition: every time that an input is sent to the net, only the winning neuron is activated;
- Cooperation: the winning neuron is connected to its neighbors, and defines the center of an area where the input changes the weights of the surrounding neurons;
- Upgrading of synapses: weight vectors of winner vector and of its neighbors will be updated according to learning algorithm.

The cycle that defines the learning process of the Kohonen neural networks can then schematized in this way:

1. Selection of basic parameters

2. Initialization of the weights of the map: starting values for neurons are randomly chosen from input values (avoiding the same value for different neurons)

3. Selection of a random input from the training dataset;

3. A winning neuron and its neighbors are activated;

4. Upgrade of the map weights;

5. Decreasing of σ and η parameters;

6. Restarting from step 3, with other random input from the initial dataset.

The algorithm ends when σ becomes lower of a predetermined value or when the weights on the map have a stable value or after a defined number of training cycles. When the learning process ends, the synapses are frozen: a specific neuron is assigned for each input vector x.

The most important property of Kohonen network is the "Organization" (which gives the name to the net). It means that similar neurons occupy contiguous area of the final map.

1.4 Supervised methods

Considerable effort has been devoted to classification of image data with the aim of producing high-quality thematic maps and establishing accurate inventories of spatial classes. The following techniques can be also combined between them for a multiple approach to image classification.

The main classification algorithms are considered the supervised maximum-likelihood method [119], [120], dimensional probability density function methods [121], artificial neural networks [122], [123], supervised decision trees [124], [125], discriminant analysis [126], [127], genetic algorithms [128], and spectral shape analysis [129], Spectral Angle Mapper [130], Support Vector Machine [131], [132].

Others approaches are based on "soften" classifier as fuzzy [133] and decision fusion [134].

1.4.1 Neural Networks

From the presented range of classifier, different studies have shown the effectiveness of NNs algorithms to extract information from remotely sensed imagery [135], [136]. Neural networks has the advantage of determining the input-output relationship directly from the training data with no need to seek for an explicit model of the physical mechanisms, which were often nonlinear and poorly understood [137]. Moreover, it was shown that multi-layer feed-forward networks formed a class of universal approximators, capable of approximating any real-valued continuous function provided a sufficient number of units in the hidden layer were considered. Image classification has probably been one of the most investigated fields in this context.

As far as the network topology is concerned, the feed-forward Multi-Layer Perceptron (MLP) is probably the most commonly used for remote sensing application between supervised methods. The extensive research activity confirmed that neural networks, besides representing a new and easy way of machine learning, possessed particularly interesting properties, such as the capability of capturing subtle dependencies among the data, an inherent fault tolerance due to their parallel and distributed structure, and a capability of positively merging pieces of information stemming from different sources [138]. Other neural network approach that is commonly used is related to the Radial Basin Functions [139] and the recent developed Pulse Coupled Neural Networks [140].

The main drawbacks of neural networks are referred to the conventional back-propagation learning algorithm can be stuck in a local minimum. Moreover, the choice of the NN architecture (i.e. number of hidden layers and nodes in each layer, learning rate) weight initialization and number of iterations required for training may significantly affect the learning performance.

1.5 The importance of spatial parameters in VHR images: texture and shape analysis

As it was written in the introduction, information on the pixel context (e.g., structures, patterns, objects) can be significant in Very High Resolution (VHR) image analysis [141], [142] and [143].

1.5.1 Texture

Texture is the term used to characterize the tonal or gray-level variations in an image. Texture analysis has played an increasingly important role in digital image processing and interpretation, principally motivated by the fact that it can provide supplementary information about image properties.

Many texture feature extraction methods exist. [144] identify four major categories of texture feature analysis methods:

- statistical (such as those based on the computation of the gray-level co-occurrence matrix—GLCM ([145], [146]);
- geometrical (including structural);
- model-based, such as Markov Random Fields (MRF);
- signal processing (such as Gabor filters).

It was pointed out [147] that textural features derived from GLCM are the most useful for analyzing the content of a variety of remote sensing imagery Recently, [148] demonstrated that the GLCM method has an improved discrimination ability relative to MRFs with decreasing window size. Six parameters (energy, contrast, variance, correlation, entropy and inverse different moment) are considered to be the most relevant, among the 14 originally proposed by Haralick, some of which are strongly correlated with each other [149].

Texture increased the per-pixel classification accuracy, especially in urban areas where the images are more heterogeneous [150] stated that this increase in terms of classification accuracy is dependent on the geometrical resolution of the scene. In fact, the improvement is greater for higher resolution images. With the increase of the spatial resolution of satellites such as QuickBird and Ikonos, texture features turn out to be valuable for the identification of smaller object [151].

1.5.1.1 GLCM

Haralick in 1973 and 1979 described the definition of the Grey Level Co-Occurrence Matrix and the parameters which can be extracted.

Suppose the area to be analyzed for texture is rectangular, and has N_c resolution cells in the horizontal direction, N_r resolution cells in the vertical direction, and that the gray tone appearing in each resolution cell is quantized to N_g levels.

Let $L_c = \{1,2,\ldots, N_c\}$ be the horizontal spatial domain, $L_r = \{1, 2, \ldots , N_r,\}$ be the vertical spatial domain, and $G = \{1, 2, \ldots, N_g\}$ be the set of Ng quantized gray tones. The set $L_r \times L_c$ is the set of resolution cells of the image ordered by their row-column designations. The image I can be represented as a function which assigns some gray tone in G to each resolution cell or pair of coordinates in $L_r \times L_c$; $I: L_r \times L_c \rightarrow$ G.

The gray tone co-occurrence can be specified in a matrix of relative frequencies P_{ij} with which two neighboring resolution cells separated by distance d occur on the image, one with gray tone i and the other with gray tone j. Such matrices of spatial gray tone dependence frequencies are symmetric and a function of the angular relationship between the neighboring resolution cells as well as a function of the distance between them.

The more relevant parameter used in this study in homogeneity:

$$Homogeneity = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \frac{p(i,j)}{1 + (i-j)^2}$$

Eq. 1.11

where *i, j* are the gray tones in the windows, which are also the coordinates of the co-occurrence matrix space, while *p* (*i,j*) are the normalized frequencies with which two neighboring resolution cells separated by a fixed shift occur on the image, one with gray tone *i* and the other with gray tone *j*; *N* is the dimension of the co-occurrence matrix, which has a gray value range of the original image.

1.5.2 Shape analysis: segmentation

One way to introduce this additional information is to segment the image before the classification [152]. The segmentation produces regions which are more homogeneous in themselves than with nearby regions and represent discrete objects or areas in the image. After segmentation, each image region becomes a unit analysis for which a number of features, on top of spectral features, can be measured and used during the classification [153], [154], [155]. A large variety of segmentation algorithms were developed in these last twenty years [156], [157]. Segmentation algorithms can conveniently be classified as boundary-based and region-based [156], [158], [159], [160]. Boundary-based algorithms detect object contours explicitly by using the discontinuity property; region-based algorithms locate object areas explicitly according to the similarity property [156].

The boundary approach gathers the edge detection techniques. These methods do not lead directly to a segmentation of the image because contours obtained are seldom closed; therefore, it is necessary to carry out closing edges algorithm if one wishes a complete segmentation of the image. Indeed, after contours closing, the duality contours/regions appears clearly. A region is defined as the inside of a closed line. On the other hand, the methods of the region approach lead directly to a segmentation of the image, each pixel being assigned to a single region [161].

In [152] the main segmentation algorithms have been compared:

- "optimal edge detector"[161], [162] and "watershed segmentation" [163], [164] for boundary-based algorithms;
- "multilevel thresholding technique" [165] and a "region-growing" technique for region based technique.

It emerged that the miraculous segmentation method which segments in a correct way for all types of landscape does not exist. In each of the four used methods, the choice of the parameters (i.e. thresholds) is important and has a great influence on the segmentation results.

The contour detection by "watershed" is more effective than the optimal edge detector. These methods prove to be effective for the detection of homogeneous and contrasted objects within the images [159] as in the images of urban zones where these types of objects are very common (for example, buildings). and the "Region-growing" method works well with textured images and images with not high contrast objects like in rural and forest test images.

Finally, all the objects in an image cannot be extracted with a single segmentation without over-segmentation, which means that there will be some pixels that cannot be directly assigned to a specific object. If the user decided to increase the threshold of the segmentation process to assign more pixels to the same object, other details could be lost.

Considering the results of this work, a deeper description of watershed segmentation is shown in the following paragraph.

1.5.2.1 Watershed transformation

In the "watershed segmentation", the procedure first transforms the original data into a gradient image. The resulting grey tone image is considered as a topographic surface. This surface is flooded from its minima and the merging of the waters coming from different sources is prevented, thus the image is divided into a set of watershed lines. The catchment's basin should correspond to the homogeneous regions of the image (Fig. 1.4). Before transforming the original data into a gradient image, a median filter can be applied on the image to reduce the noise. The presence of noise in an image causes an over-detection of edges by the morphological gradient. The median filter locally homogenizes the image and avoids extreme gradients, and thus disturbing contours. It is also a good means not to take into account object texture during contours detection. The image gradient can also be threshold to limit the contour sensitivity; e.g., if the threshold is 10, pixels with a gradient higher than 10 are kept, and the others are put at 0 as if there are no edges.
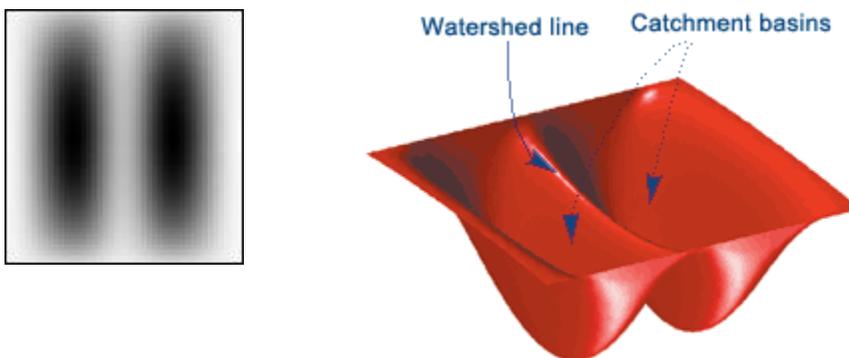
Fig.1.4: consider that the bright areas in the left image are "high" and dark areas are "low". Then it might look like the surface (right image). With surfaces, it is natural to think in terms of catchment basins and watershed lines.

Chapter 2

Used dataset

Considering the improvement of spatial and spectral resolution during these years for VHR images, it has been chosen to design the entire process on Quickbird images (very high spatial resolution and multispectral sensor) and airborne hyperspectral images (very high spatial resolution and hyperspectral sensor). In this chapter, the technical specifications of Quickbird satellite, Airborne Hyperspectral Scanner (AHS) and Multispectral Infrared and Visible Imaging Spectrometer (MIVIS) will be provided, as well as the acquired image details.

2.1 QuickBird

QuickBird (Fig. 2.1) was the first satellite in a constellation of spacecraft that Digital Globe is still developing for commercial high-resolution imagery of Earth [201]. QuickBird's global collection of panchromatic and multispectral imagery is designed to support applications ranging from map publishing to land and asset management to insurance risk assessment.

Figure 2.1: quickbird satellite

The QuickBird satellite provides images consisting of four multi-spectral (MS) channels with 2.4 m resolution and a single panchromatic (PAN) band with a 0.62 m resolution (Table 2.1). The four MS bands collect data at red, green, blue, and near-infrared wavelengths, and the data in each band is stored with an 11-bit quantization. As previously discussed, the spatial resolution plays a key role in urban monitoring related to the ability to detect fine-scale objects present in urban scenes.

Quickbird specifications	
Launch date	October 18, 2001
Altitude	450 km, 98 degree, sun-synchronous inclination
Swath	Areas of interest. - Single Area: 16.5 km x 16.5 km - Strip: 16.5 km x 115 km
Spectral resolution	- Blue: 450 - 520 nanometers - Green: 520 - 600 nanometers - Red: 630 - 690 nanometers - Near-IR: 760 - 900 nanometers - PAN: 445 - 900 nanometers
Spatial resolution	- Multi: 244 centimeter at nadir - PAN: 62 centimeter at nadir
Revisiting time	2-3 days, depending on latitude

Table 2.1: Quickbird satellite specifications

2.1.1 Quickbird image 1: Nettuno (Rome, Italy)

The first analysed image has been acquired above Nettuno area -Rome, Italy- (Fig 2.3, Table 2.2). This area was chosen for the presence of all the five main cover classes to spot. Being Nettuno a coastal urban area, it has been decide to use the automatic classification procedure to highlight water, vegetation, bare soil, buildings and roads.

Image characteristics	
Image name	Nettuno (Italy)
Satellite	Quickbird
Date:	July, 22nd 2003
N° rows	5383
N° columns	4846
N° pixels	26086018

Table 2.2: Quickbird image Nettuno specifications
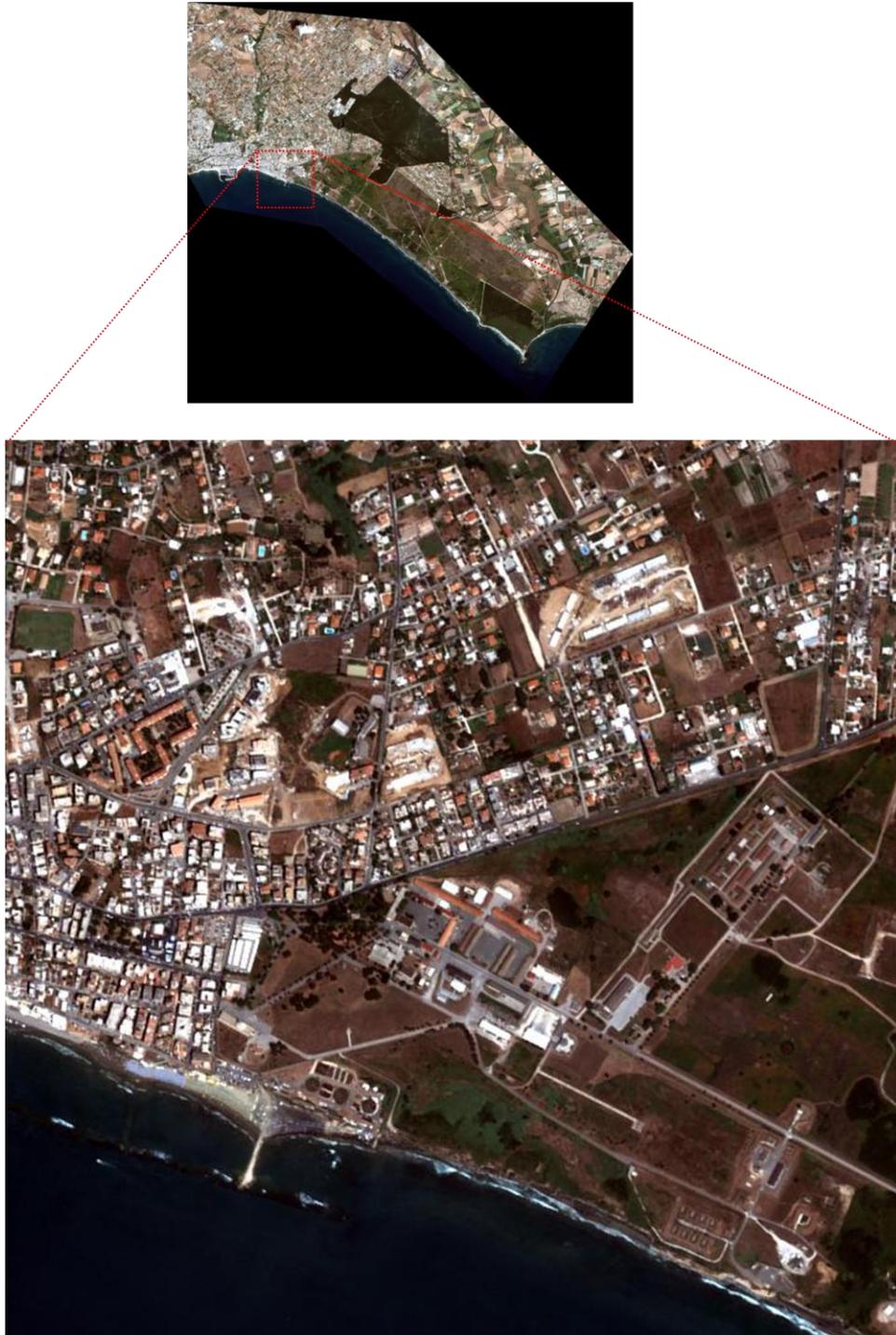
Figure 2.2: A QB image taken over Nettuno, located in Italy, South-West of Rome, on July 22, 2003.

2.1.2 Quickbird image 2: Tor Vergata area (Rome, Italy)

The second analysed image has been acquired above Tor Vergata area -Rome, Italy- (Fig 2.2, Table 2.1). This image covers the site where Tor Vergata campus is located, which means that the ground truth can be easily observed considering that this work has been developed physically in this area. Moreover, several publications on the application of Neural Networks for classification have been produced on this site, allowing to compare the accuracy of the innovative proposed methodology with the existing one [202], [203], [204], [205].

Image characteristics	
Image name	Tor Vergata area (Italy)
Satellite	Quickbird
Date:	March, 13th 2003
N° rows	2416
N° columns	1651
N° pixels	3988816

Table 2.3: Quickbird image Tor Vergata specifications

Figure 2.3: A QB image taken over the Tor Vergata University campus, located in Italy, South-East of Rome, on March 13, 2003.

2.1.3 Quickbird image 3: Denver (CO, USA)

The third analysed image (courtesy of Digital Globe) has been acquired above Denver urban area -Colorado, USA- (Fig 2.4, Table 2.3). This area was chosen to test the methodology with a different type of urbanization (more regular and dense) respects to previous Quickbird images.

Image characteristics	
Image name	Denver (USA)
Satellite	Quickbird
Date:	March, 10th 2007
N° rows	4096
N° columns	4096
N° pixels	16777216

Table 2.4: Quickbird image Denver specifications
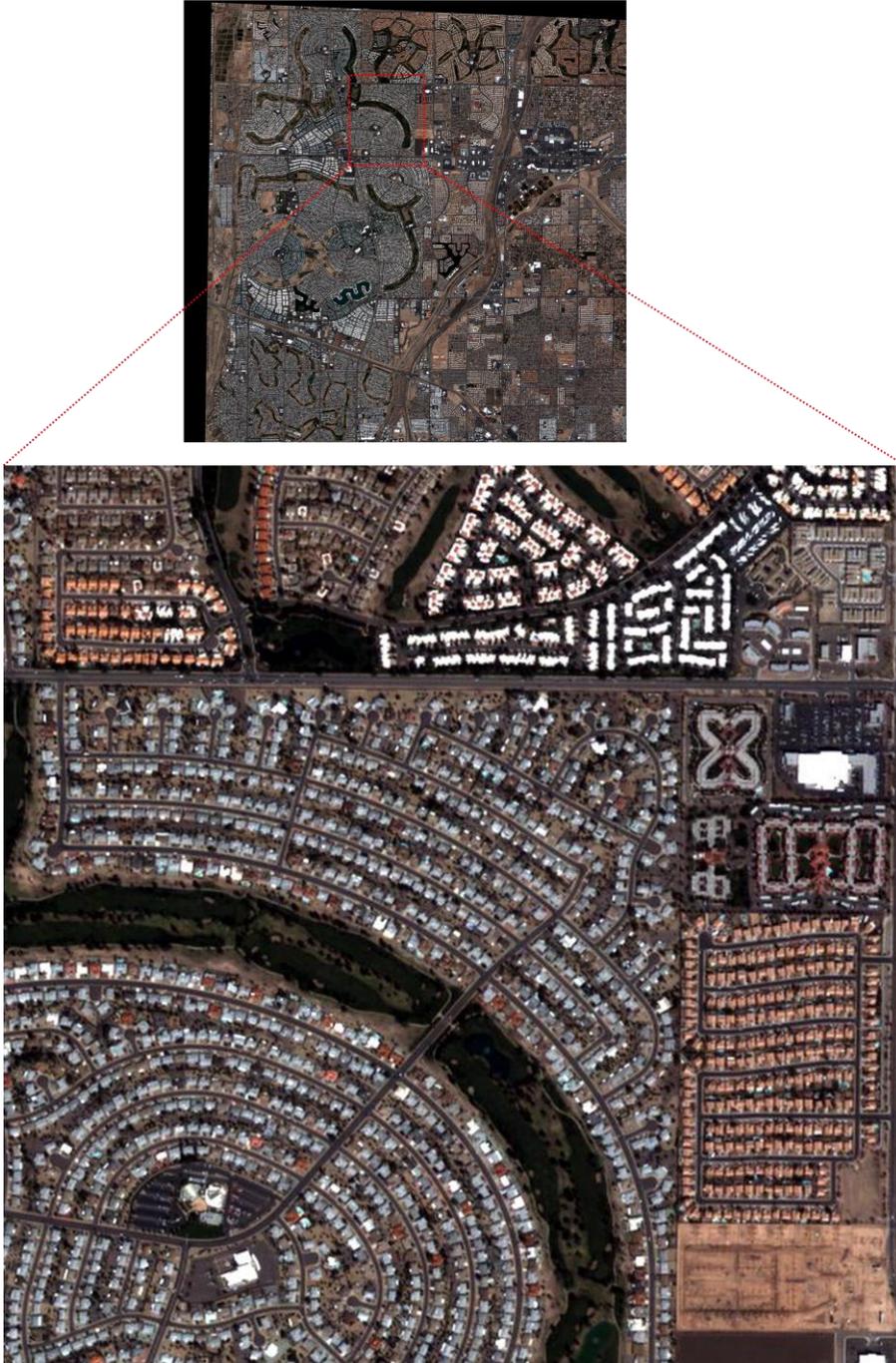
Figure 2.4: A QB image taken over Denver (Colorado, USA), on March 10, 2007 (courtesy of Digital Globe)

2.2 Hyperspectral sensors

Nowadays, hyperspectral sensors are increasing their importance in remote sensing studies. ESA Chris-Proba[11] and Hyperion[12] were the first satellite missions dedicated to this sensor. However, their spatial resolution (30 meters for Hyperion, from 18 to 30 for Chris-Proba) cannot be considered as "Very High Resolution", therefore no useful for the purposes of the study.

For this reason, it has been decided to use airborne hyperspectral sensors, which combines very high spatial resolution with several contiguous bands.

2.2.1 AHS

The Airborne Hyperspectral Scanner (AHS) is an imaging line-scanner radiometer, installed on a CASA-212 200 series aircraft owned by Spain's National Institute for Aerospace Technology (INTA)[13].

The AHS instrument (Fig. 2.5), with a whiskbroom linescanner, acquires images in 80 spectral bands covering the visible and near infrared (VNIR), short wave infrared (SWIR), mid-infrared (MIR), and thermal infrared (TIR) spectral ranges (Table 1), having a digitisation accuracy of 12 bits.

With its wide spectra, AHS has been used for several earth observation topics, as thermography [206] and agriculture [207].

---

[11] *ESA Chris Proba home page*: http://earth.esa.int/missions/thirdpartymission/proba.html

[12] *Nasa Hyperion Instrument*: http://eo1.gsfc.nasa.gov/Technology/Hyperion.html

[13] *The INTA AHS system*
http://bo.eufar.net/document/publi/124901161346543eefc58cf:SPIE_RS05_INTA_5978-56.pdf

Figure 2.5: AHS system

The image belongs to the collection of measurements carried out during the DESIREX 2008[14] (Dual-use European Security IR Experiment 2008) campaign (Table 2.4) which involved taking airborne and ground measurements with infrared sensors above the Madrid city -Spain- (Fig. 2.6) during the period June-July 2008 (desirex final report).

---

[14] Desirex campaign homepage http://www.uv.es/desirex/

AHS flight specifications			
Date	June, the 28th of 2008		
Acquisition time	11:53 am		
flight height	2497 m above sea level		
Pixel resolution	≈ 4 meters		
N° Rows	9113		
N° Columns	1473		
N° Pixels	13423449		
Spectral characteristics	Wavelength Range (μm)	Spectral Region	Bands	Band width (μm)
	0.441-1.018	VIS-NIR	1-20	0.03
	1.491-1.650	NIR	21	0.2
	2.019-2.448	NIR	22-63	0.013
	3.03-5.41	MIR	64-70	0.3
	7.95-13.17	TIR	71-80	0.4-0.5

Table 2.5: AHS Madrid image specifications

Figure 2.6: AHS image taken over Madrid (Spain), on June 28, 2008 (courtesy of Planetek)

2.2.2 MIVIS

MIVIS (Multispectral Infrared and Visible Imaging Spectrometer) is a modular hyper-spectral scanner (Fig. 2.7) composed of 4 spectrometers, which simultaneously measure the electromagnetic radiation of the Earth's surface recorded by 102 spectral bands. It covers the visible and near infrared (VNIR), short wave infrared (SWIR), mid-infrared (MIR), and thermal infrared (TIR) spectral ranges (Table 2.5), having a digitisation accuracy of 12 bits. The instrument enables advanced applications in environmental remote sensing, like Agronomy, Archaeology, Botanic, Geology, Hydrology, Oceanography, Pedology, Urban Planning, Atmospheric Sciences, and so on.

Figure 2.7: MIVIS system

The image belongs to the collection of measurements carried out during the a campaign for the project TIZIANO (Table 2.5) which involved taking airborne and ground measurements with infrared sensors above Bari city -Italy- (Fig. 2.8) during the period May 2009 [208].

MIVIS flight specifications				
Date	May, the 23rd of 2009			
Acquisition time	8.30 am			
flight height	1500 m above sea level			
Pixel resolution	$\approx$ 3 meters			
N• Rows	9251			
N• Columns	3827			
N• Pixels	35366573			
Spectral characteristics	Wavelength Range (μm)	Spectral Region	Bands	Bandwidth (μm)
	0.43-0.83	VIS-NIR	1-20	0.03
	1.15-1.55	NIR	21-28	0.2
	2.0-2.5	MIR	29-92	0.008
	8.2-12.7	TIR	92-102	0.4-0.5

Table 2.6: MIVIS Bari image specifications

Figure 2.8: MIVIS image taken over Bari (Italy), on May 23, 2009 (courtesy of Planetek)

Chapter 3

Methodology

As it has been prefaced in the first chapter, this work is based on the application of neural networks to classify automatically VHR optical images.

The following chapter will shown the developed methodology to classify the tested image (multispectral and hyperspectral). The processing chains for the two kind of remote sensing data aimed at defining some parameters (spectral, textural and shape) to highlight the common characteristics in a land cover class, allowing the discrimination between different classes.

The methodology have been applied to classify five land cover classes: vegetation, water, bare soil, buildings and roads.

3.1 First tests and the choice of a hierarchical classifier

To classify automatically an image, this work has been based on the application of Kohonen SOM. Appreciable results in automatic classification with the application of this methodology have been introduced in the first chapter; besides, encouraging results to classify VHR SAR satellite [301], [302] confirmed the potentiality of this technology on VHR images.

The first attempts were based on the idea of extracting all the land cover classes in a single step, starting from DN values. Four spectral bands from Quickbird image of Tor Vergata have been sent to a Kohonen SOM with four neurons (which should be related to vegetation, bare soil, buildings and roads). In the output image, even if vegetation was clearly distinguishable, the other classes did not correspond to the request outputs.

That was due to the spectral response of vegetation, which is completely different from the others: roughly speaking, it could be said that it is more homogeneous respects to the others. The automatic classification algorithm, which was able to discriminate the classes considering only the pixels spectral response, could not return appreciable results to distinguish *man-made classes* spectrally.

Therefore, it appeared clear that it was not possible to classify an image in a single step and it was necessary to implement a hierarchical tree methodology as the following step.

The idea of representing the data analysis process with a hierarchical tree has been under study in many pattern recognition application areas. Tree-based classifiers have represented an interesting and effective way to structure and solve complex classification problems [303], [304], [305], [306], [307]. The organization of information into a hierarchical tree allows to achieve a faster processing capability and, at times, a higher accuracy of analysis. This is mainly explained by the fact

that the nodes of the tree carry out very focused tasks, meaningless when taken individually but meaningful when taken as a whole.

Using a hierarchical tree, the first node was set to distinguish *vegetation* and *not vegetation* (with two neurons in the Kohonen SOM). The second node tried to distinguished the *non vegetation* class in the remained three classes (buildings, asphalt, and bare soil).

Even if the quality of the final classified image was improved, the mutual discrimination of the three *man made* classes was not satisfactory. This could be explained by the different spectral response in the same class and by the

Regarding the subjectivity of the classification criteria, figure 3.1 shows how the same image can be classified in different ways considering the purposes of the analysis.

If the goal is the distinction between water and land (i.e. for Land Surface Temperature retrieval, Coastal Mapping, etc.), a classification as the one in the left part of the image satisfies the request of the study.

On the other hand, if the aim of the study is the classification of sealed soils, a classification as the one shown in the center can be considered valid.

Therefore, even if the number of the classes is fixed (2 in the above example), the outputs of the classification can be different.
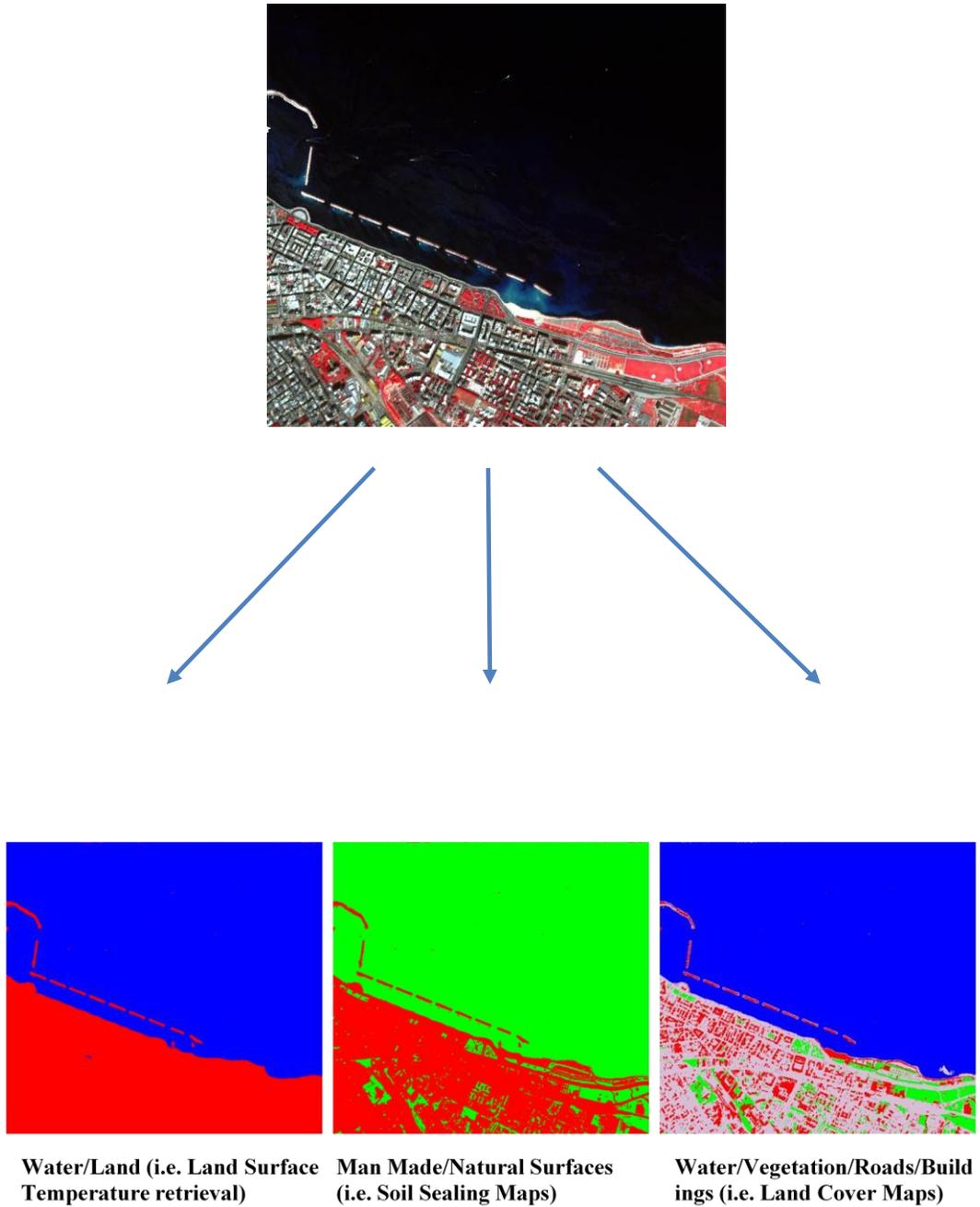
Fig.3.1: MIVIS RGB 20-12-7 (above) and different types of classification output (below)

Following these considerations, it has been decided to increase the number of neurons Kohonen map. This allowed to define an higher number of classes, with higher possibilities to match the requested output. Generally, the optimal architecture of a neural network is not known a priori for most real-world problems: in the first Kohonen map, the best results have been obtained with a map of 3 x 3 neurons, which seems the optimal set up. It means that each input signal has been compared with all the 9 neurons/weights of the map (where the initial weights of these neurons were set with random values). The winning neuron has been chosen with a minimum distance algorithm based on the Euclidean distance between input and the weight of neurons.

Some of the final clusters have to be assigned to a semantic land cover class (which depends on the aim of the study, as shown in figure 3.1): this procedure is called *labeling* procedure. The labeling procedure can be done automatically, through the comparison of the final weights of the neurons with some class values from spectral libraries [308]. The cluster which has the minimum distance with the library signature is assigned to the specific land cover.

The necessity of matching spectral signatures from the analyzed images to reference spectral signatures implied the atmospheric correction of the images. This operation allowed to convert the original DN values to radiance and, eventually, to reflectance values. The atmospheric correction will be treated in a exhaustive way in chapter four.

The labeling procedure can be done interactively from the clusterized image: in some cases, the semantics of a class [309] is too complex to be performed by a simple comparison. In other cases, the presence of two clusters in a image (as it will be shown later, red brick buildings and concrete buildings are both belonging to building class) that can be assigned to a specific class, do not allow the automatic labelling with appreciable results.

3.2 The mixed approach with Kohonen SOM and Neural Network MLP

Increasing the neurons in the Kohonen SOM, some uncertain pixels remained unclassified: some unlabelled neurons represented some mixed pixels which had to be assigned to a specific class.

This meant that the Kohonen SOM turned into a *clusterizer* from a *classifier*. To reassign these pixels, a second classification was necessary.

This latter classification is based on the Multi-Layer Perceptron MLP Neural Network, already proved as an effective classification algorithm for remote sensing images.

The mixed methodology has already been used in classification [310], [311], but the combined use of Kohonen map with MLP is original and it will be discussed in this work.

Reassuming, the complete processing chain is a mixed approach between unsupervised and supervised methods based on NNs. The concept idea is based on:

- the extraction of clusters of pixels representing land cover classes with Kohonen's SOM (unsupervised approach);

- the use of the pixels belonging to the extracted clusters to train a MLP supervised neural network (supervised approach).

In fact, the training sets (also called *Area Of Interest* or *Region Of* Interest) for the MLP have been extracted automatically from the image.

Fig. 3.2 shows the general scheme for classification, highlighting the mixed approach of unsupervised and supervised methods.

Results of the classification chains will be exhaustive discussed in chapter six.

On this general scheme of classification, other steps have been added considering the general characteristics of VHR image. To spot some land cover classes, the exclusive use of spectral signatures could be enough. Essentially, natural surfaces like vegetation, water and bare soil, can be easily extracted with their distinctive spectral signatures (chapter five).

From the physical composition point of view, several different materials can be used for the same built-up element (for example, building roofs can be made of clay tiles, metal, concrete, glass, plastic), On the other hand, the same material can be used for different built-up elements (for example, concrete can be found in paved roads or building roofs) and the information on the pixel context (e.g., structures, patterns, objects) can be significant in Very High Resolution (VHR) image analysis.
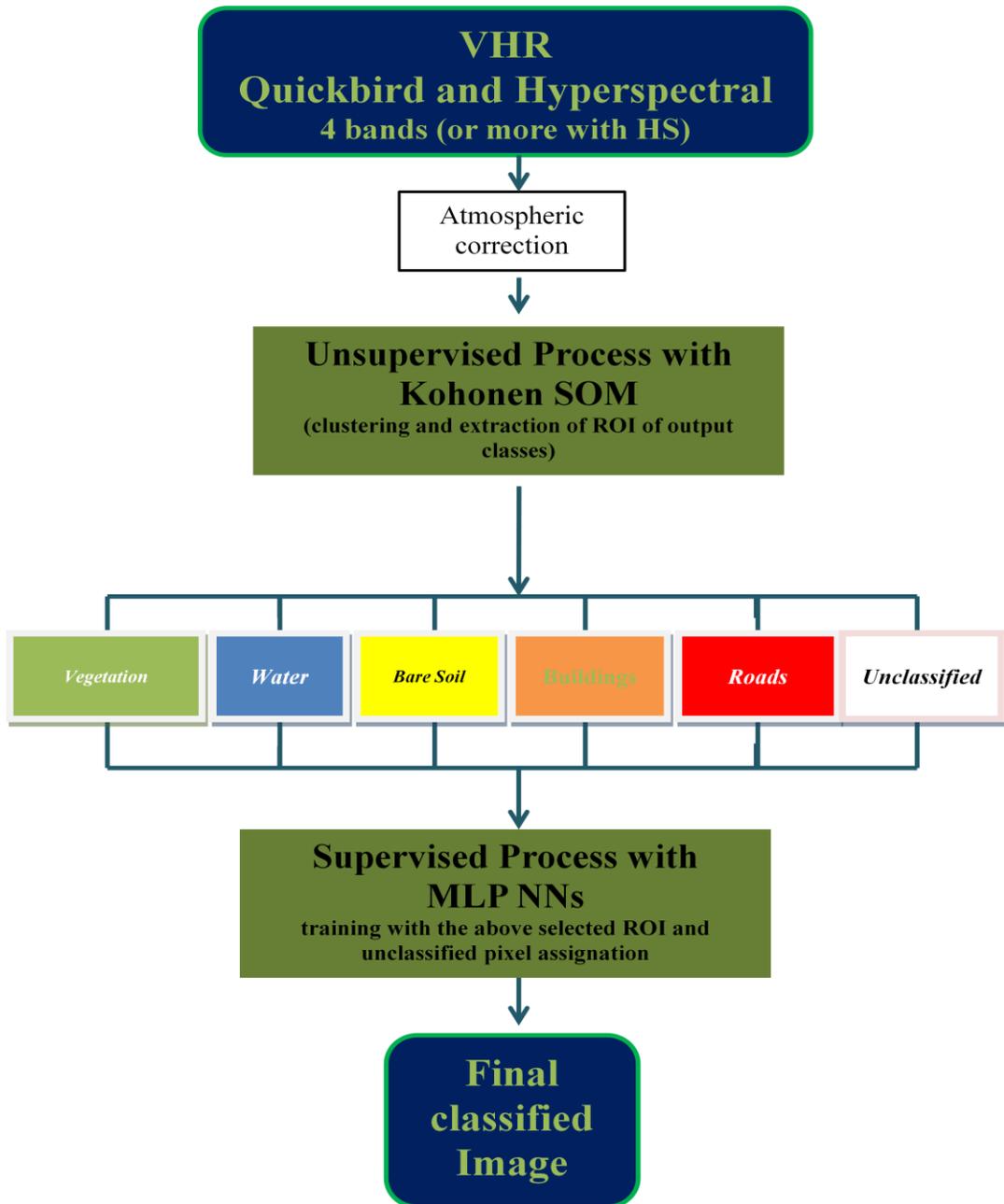
Fig 3.2: example of the mixed procedure. The input bands, after atmospheric correction, activate several neurons/clusters in the bubble (with a specific winning one) of the Kohonen SOM. After the clustering with SOM(s), neurons are labelled and they constitute the input to train the MLP Neural Network: some random pixels from these winning neurons train the supervised NN. Finally the final weights have been set and all the image is classified to define the output classes.

3.3 The pixel context: texture

Texture is the term used to characterize the tonal or gray-level variations in an image. Texture analysis has played an increasingly important role in digital image processing and interpretation, principally motivated by the fact that it can provide supplementary information about image properties.

From the available techniques to calculate textural parameters, the Grey Level Co-Occurrence Matrix from Haralick has been chosen for the previously mentioned advantages (chapter one).

A preliminary analysis have been made to test which texture parameters were more suitable to detect human features. The most relevant feature in this tests was homogeneity (Eq. 3.1), also called the "Inverse Difference Moment", as its utilization in urban feature extraction has been already proven effective [312].

$$Homogeneity = \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} \frac{p(i,j)}{1 + (i-j)^2} \qquad \text{Eq. 3.1}$$

where *i, j* are the gray tones in the windows, which are also the coordinates of the co-occurrence matrix space, while *p* (*i,j*) are the normalized frequencies with which two neighboring resolution cells separated by a fixed shift occur on the image, one with gray tone *i* and the other with gray tone *j*; *N* is the dimension of the co-occurrence matrix, which has a gray value range of the original image.

Homogeneity weights values by the inverse of the Contrast weight. To exploit a multi-scale approach, which better deals with objects having different spatial coverage in an area, homogeneity features have been computed with two different window sizes of 3×3 and 5×5 only on AHS green band. These two window sizes have been chosen considering the average building size and the spatial resolution of multi- and hyperspectral images. Indeed this latter band gave appreciable visual

results and it was preferred to avoid using a redundant number of features to feed the following supervised neural network. The first window (3×3) corresponded to nine pixels, which means a dimension on the ground of 81 m^2 (3 m $\times$ 3 m $\times$ 9), the second one (5×5) was equal to 135 m^2 (3 m $\times$ 3 m $\times$ 15). These sizes can be related to the building plant on the ground which have at least these measures: Therefore, these windows allows to spot buildings with these footprints and increasing the window size, some small buildings cannot be detected. The values of the input for this Kohonen SOM (derived from the pixel values in these spectral and textural bands) were normalized to the range of $(-1; +1)$: this step was necessary considering the different units as input for Neural Networks.

Even if a reasonable number of buildings were detected with the combined use of spectral and textural parameters, this methodology could not be always suitable to highlight some other man made features. For instance, if:

- the material of building roof and street is the same (i.e. concrete), therefore the spectral signature is similar;
- even if the spectral signature is different, the magnitude of the value difference could not be enough to allow the discrimination between asphalt and buildings (too similar);
- the texture is the same, especially for wide streets; therefore homogeneity parameter is similar.

For these reasons, it has been decided to introduce additional information segmenting the image before the classification.

3.4 *The pixel context: segmentation*

Segmentation produces regions which are more homogeneous in themselves than with nearby regions and represent discrete objects or areas in the image. After segmentation, each image region becomes a unit analysis for which a number of features, on top of spectral features, can be measured and used during the classification.

From the available techniques to segment an image, the "watershed segmentation" has been chosen for the previously mentioned advantages (chapter one).

The entire methodology consisted in the following procedure [313]:

1. an image smoothing (edge preserving) filter has been applied (the filter was a median filter applied on a kernel 3 x 3);
2. normalization of each band with Eq. 3.2

$$Data_{new} = \frac{Data_{old} - \mu_{old}}{\sigma_{old}} \qquad\qquad Eq.\,3.2$$

where *Data$_{new}$* is the rescaled pixel value, *Data$_{old}$* is the original pixel value, and μ_{old} and σ_{old} are the mean and variance value of the original dataset.

3. Principal Component Analysis *PCA* [314] was applied to get a new set of feature bands: this was the first step to reduce the original bands in a grey-level (gradient) image to apply the watershed transformation;
4. some bands from PCA are selected and added to each other, to reach the 90% of the variability of the entire dataset;
5. after the gradient map was calculated, a density function over the whole image was computed (i.e., the cumulative relative histogram): each bar of the histogram considered an interval *q$_{quant\_levels}$* described with Eq. 3.3:

$$q_{quant\_levels=v_{min}+((quat\_levels)*(V_{max}-v_{min}))/(quant\_levels)} \qquad Eq.\,3.3$$

where *quant_levels* represents the total number of bars (generally, 256) and v_{min} and v_{max} are the limits of the data distribution;

6. once the cumulative relative histogram has been calculated, it can be used along with the gradient map to calculate the gradient space scale. By choosing a relative level threshold on the cumulative relative histogram, the corresponding gradient magnitude G_T can be determined. The gradient map was then modified so that the pixels with gradient magnitude less than G_T were set to assume the value G_T. By changing the threshold from 0 to 100, a continuous gradient scale space is built. Obviously, lower is the scale level, higher is the over-segmentation (large number of objects) resulting from Watershed Transformation on the original gradient map. For detect objects like streets, a scale value of 50 is recommended;

7. after determining G_T value, the watershed transformation is performed;

8. a unique level is given to each region and the watershed lines pixels were merged to the most similar neighbor regions (with Euclidean Distance). Generally, these pixels have 0 values but, considering the purposes of the object analysis in a raster environment, these pixels had to be re-assigned to a specific object;

9. the region can be merged with the full lambda-Schedule algorithm [315]. Such algorithm iteratively merges adjacent segments based on a combination of spectral and spatial information. Merging proceeds if the algorithm finds a pair of adjacent regions, *i* and *j*, such that the merging cost $t_{i,j}$ is less than a defined threshold lambda value set as (Eq. 3.4):

$$t_{i,j} = \frac{\dfrac{|O_i| \cdot |O_j|}{|O_i| \cdot |O_j|} \cdot \|\mu_i - \mu_j\|^2}{lenght(\,\partial(O_i, O_j))} \qquad Eq.\,3.4$$

where O_i and O_j are the regions i and j of the image, $|O_i|$ and $|O_j|$ are the areas of region i *and j,* μ_i is the average value in region i, μ_j is the average value in region $j,$ $\|\mu_i - \mu_j\|^2$ is the Euclidean distance between the spectral values of regions i and $j,$ and $lenght(\partial(O_i, O_j))$ is the length of the common boundary of O_i and O_j;

10. from the final objects, it was possible to calculate several parameters describing the objects. These parameters regards the spectral, textural and shape characteristics of them.

Considering that spectral and textural parameters of the image were already been analysed with the methods described in the previous paragraph, it has been decided to focus the analysis on the shape parameters which could help, in some cases, to distinguish roads.

From these objects, a parameter like elongation [316], [317] has been calculated to highlight roads.

Elongation (Eq. 3.5) is a shape measure that indicates the ratio of the major axis of the polygon to the minor axis of the polygon. The major and minor axes are derived from an oriented bounding box containing the polygon. (i. e. elongation value for a square is 1.0, and the value for a rectangle is greater than 1.0).

$$Elongation = MaxAxis/MinAxis \hspace{3cm} Eq. 3.5$$

From several tests, an object which its elongation is greater than 3 can be assigned to roads [318] Therefore, with a simple threshold, pixels from these objects have been assigned to roads.

Considering that the output of the segmentation are objects, described as polygons in a vector layer, a following conversion to raster (with a pixel size equal to the

starting one from the original image) is necessary to insert an information layer to the MLP Neural Network.

3.5 *The final classification chain*

Once all rhe pixels related to the predefined land cover class (vegetation, bare soil, water, buildings and roads) have been selected from the clusters deriving from these Kohonen SOM, the inputs for the MLP have been set. The details of this supervised neural network will be discussed in chapter 6. The general scheme for the unsupervised clustering is shown in figure 3.3.

A deeper analysis of the two classification chains shows that roads class can be indentified from spectral signature but, if this input parameters do not give satisfactory risults in the interactive labeling, the use of segmentation to define which pixels belong to this class is preferable.
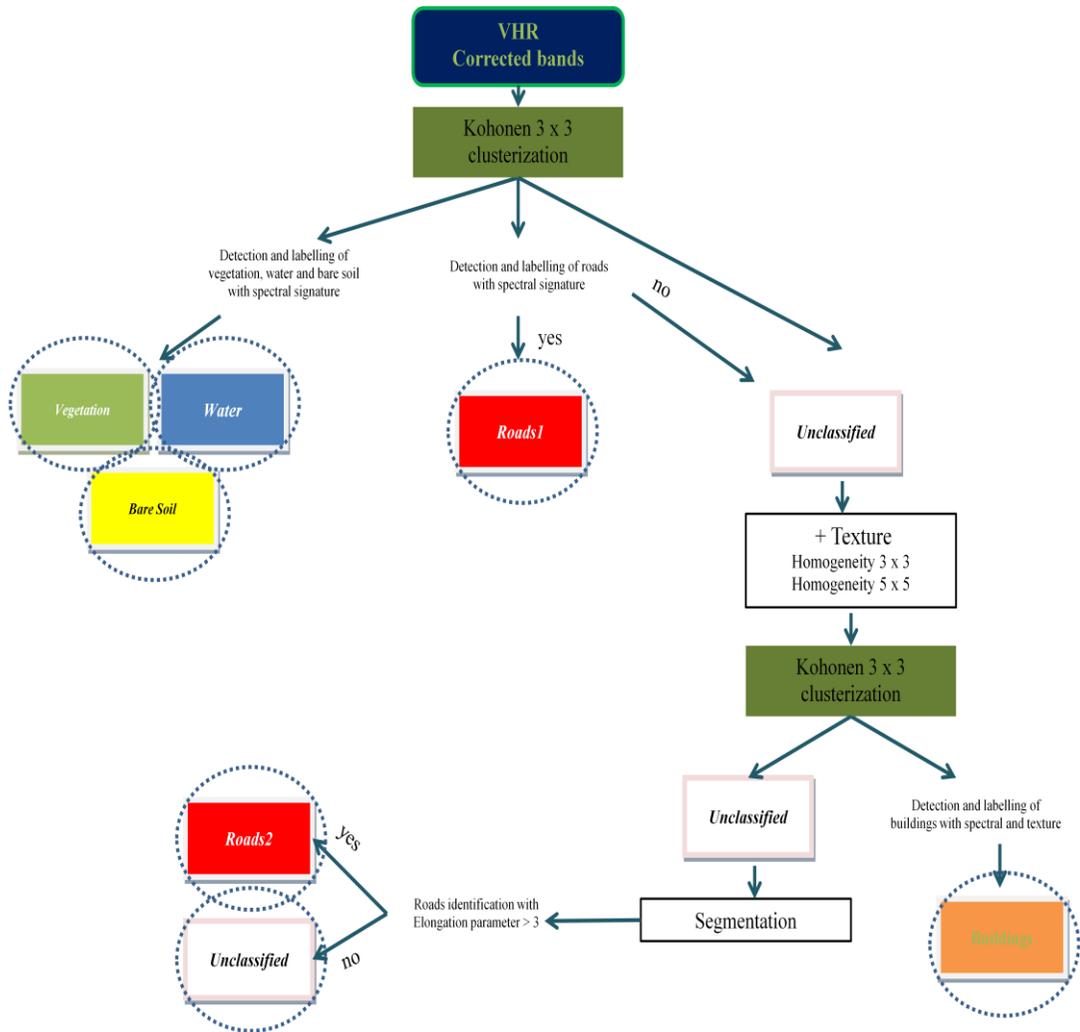
Fig. 3.3: unsupervised clustering and labelling chain

Chapter 4

Pre-processing

Before applying the developed methodology to classify hyperspectral and multispectral data, it has been decided to preprocess the images. This operation had the aim to improve the image qualities, to allow the comparison of the same surface in different images and verify if it is possible to improve the classification results modifying the input data of the Neural Networks.

The preprocessing consisted in:
- an atmospheric correction of the multispectral and hyperspectral images, to obtain reflectance values from radiance values (or Digital Numbers);
- buildings shadow removal in hyperspectral images, exploiting the thermal information.

4.1 Atmospheric correction

As it was written in the introduction, a different spectral behavior could be explained from the effect of the atmosphere, which operates at different wavelengths. Atmospheric correction provides an estimation of the radiation emitted and reflected at the surface, and it is necessary for observing parameters that are intrinsic to the surface. Accurate atmospheric correction removes the effects of changes in satellite-sun geometry and atmospheric conditions due to aerosol scattering [401]. Atmospherically corrected surface reflectance images improve the accuracy of surface type classification [402], [403] and allows to compare the spectral signatures from different images.

Atmospheric correction typically consists of the retrieval of atmospheric parameters and the solution of the Radiative Transfer (RT) equation. Generally, atmospheric parameters involves aerosol (their description, the optical depth and, if possible, the aerosol "type") and the column water amount. Successively, the solution of the RT equation for the given aerosol and column water vapor allows to transform radiance in reflectance.

A variety of methods and algorithms for atmospheric correction are available. The "empirical line method," consisting of a linear transformation derived from ground-truth spectra, remains a popular and accurate method where truth data exist. In other situations, a first-principles method is needed. ATREM, developed by Gao et al. [404] using the 5S and, later, 6S radiation transport models [405], was for many years the industry-standard algorithm. Recently, more sophisticated algorithms have been developed, focusing primarily on land imagery. These algorithms, which incorporate more accurate radiative transport models and improved methods for retrieving the atmospheric properties needed for the correction, include ATCOR [406], ACORN [407], FLAASH [408] and HATCH [409].

For this work, it has been decided to use the latest version of FLAASH (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes). FLAASH is an efficient correction code based on MODTRAN4 [410] that has been developed collaboratively by Spectral Sciences, Inc. and the Air Force Research Laboratory, with assistance from the Spectral Information Technical Applications Center (SITAC).

FLAASH is the state-of-the-art of the atmospheric correction algorithm [411]: is capable of generating accurate surface reflectance spectra from multi- and hyperspectral imagery, at least under conditions of clear to moderate aerosol/haze, low to moderate water vapor, and nadir viewing from any altitude between the ground and the top of the atmosphere.

4.1.1 The FLAASH-MODTRAN algorithm

FLAASH uses the standard equation for spectral radiance at a sensor pixel, L, in the solar wavelength range (neglecting thermal emission) from a flat Lambertian surface or its equivalent [405]. Collecting constants, the equation is reduced to the form:

$$L = \left(\frac{A\rho}{1 - \rho_e S}\right) + \left(\frac{B\rho_e}{1 - \rho_e S}\right) + L_a \qquad\qquad Eq.\,4.1$$

- **L** is the measured radiance ($mWm^{-2}sr^{-1}\mu m$)
- **ρ** is the pixel surface reflectance
- **ρ_e** is an average surface reflectance for the pixel and a surrounding region
- **S** is the spherical albedo of the atmosphere
- **L_a** is the radiance back scattered by the atmosphere ($mWm^{-2}sr^{-1}\mu m$)
- **A** and **B** are coefficients (($mWm^{-2}sr^{-1}\mu m$) that depend on atmospheric and geometric conditions but not on the surface.

Each of these variables depends on the spectral channel; the wavelength index has been omitted for simplicity. The first term in Eq. (4.1) corresponds to radiance that is reflected from the surface and travels directly into the sensor. The second term corresponds to radiance from the surface that is scattered by the atmosphere into the sensor, resulting in a spatial blending, or adjacency effect. The third term corresponds to the radiance back scattered by the atmosphere.
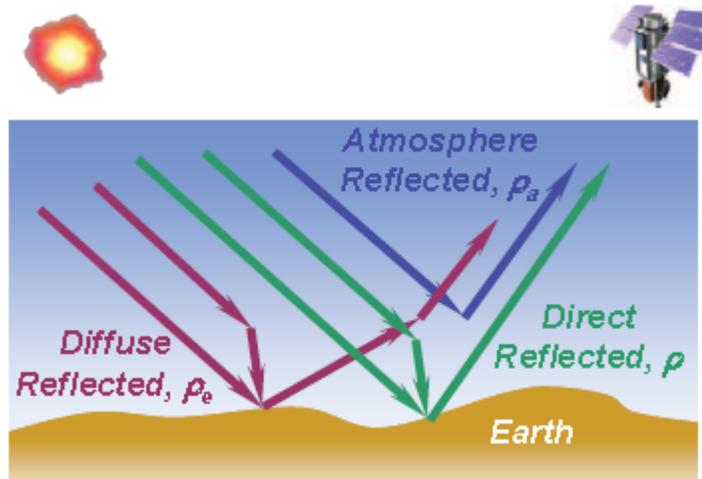
Fig. 4.1: Component of apparent reflectance [412]

In the Fig. 4.1, it is possible to observe these three components: the spectral atmospheric reflectance (in the formula 4.1 is inserted in the term L_a), the direct reflectance (which includes $T_r\uparrow$ the sensor-to-ground direct transmittances, $T_r\downarrow$ the sun to-ground direct transmittance and $\rho/(1- \rho_e\text{-}S))$ and the diffuse reflectance (which includes $T_f\uparrow$ the sensor-to-ground diffuse transmittances, $T_f\downarrow$ the sun to-ground diffuse transmittance and $\rho_e /(1- \rho_e\text{-}S))$.

The spectral atmospheric reflectance ρ_a, the sensor-to-ground direct and diffuse transmittances $T_r\uparrow$ and $T_f\uparrow$, the sun to-ground direct and diffuse transmittance $T_r\downarrow$ and $T_f\downarrow$ and the spherical albedo from the ground *S*, are all solely atmospheric parameters, independent of the ground. On the other hand, the spectral imaged-pixel reflectance ρ and the effective, spatially averaged Reflectance ρ_e are ground optical properties and thereby independent of the atmosphere.

4.1.1.1. Atmospheric correction in FLAASH

These calculations usually represent the single most computationally intensive part of the atmospheric correction. A look-up table (LUT) of A, B, S and L_a has been pre-calculated in MODTRAN and it is interpolated as needed for the specific viewing geometry, atmospheric condition, and sensor channels of interest.

The water retrieval is performed rapidly with a 2-dimensional look-up table (LUT) constructed from the MODTRAN4 outputs using a-Delaunay triangulation procedure. The water band typically used is at 1.13 µm, with the LUT for this spectral region generated on-the-fly.

For aerosol, a adjacency-corrected algorithm is implemented [413] of the 0.660 µm to 2.200 µm reflectance ratio constraint for dark land pixels (2.200 µm reflectance < ~0.1) found by [414].

Once the atmosphere is adequately characterized and the Eq. 4.1 constants are derived, calculation of the image reflectance is straightforward using a method described in [405] in which a spatially averaged radiance image is used to generate a good approximation to the averaged reflectance ρ_e. The averaging is performed using a point-spread function that describes the relative contributions to the pixel radiance from points on the ground at different distances from the direct line of sight. Presently this function is taken as a wavelength-independent radial exponential; its width is determined from a single-scattering calculation that accounts for the sensor altitude, the aerosol scale height and a nominal aerosol phase function.

4.1.2 Atmospheric correction of the used dataset with FLAASH

4.1.2.1 DN to radiance conversion

Considering that FLAASH requires data in radiance unit, Quickbird products must be converted to spectral radiance before radiometric/spectral analysis or comparision with imagery from other sensors in a radiometric/spectral manner [415].

1 - Band integrated Radiance [(W-m^{-2}-sr^{-1}]

$$L_{pixel,Band} = K(revised)L_{band} \times q_{pixel,Band} \qquad Eq.\,4.2$$

where $L_{pixel,\,Band}$ are Top-Of-Atmosphere band-integrated radiance image pixels [W-m^{-2}-sr^{-1}], K (revised)$_{band}$ is the revised absolute radiometric calibration factor [W-m^{-2}-sr^{-1}-DN^{-1}] for a given band and $q_{pixel,\,band}$ are radiometrically corrected image pixels values [DN].

2 - Band-Averaged Spectral radiance [W-m^{-2}-sr^{-1}-μm^{-1}]

$$L_{\lambda pixel,Band} = \frac{L_{pixel,Band}}{\Delta\lambda_{pixel,Band}} \qquad Eq.\,4.3$$

The Top-Of-Atmosphere band-integrated radiance has to be divided into the effective bandwidth ($\Delta\lambda$).

For hyperspectral data, this conversion was not necessary because the data had already been provided in radiance values.

4.1.2.2 Input parameters

FLAASH requires a consistent amount of input parameters to perform the atmospheric constants calculation. It is necessary to point out that many of them are quite often unknown (i.e. CO_2 mixing ratio) and, especially with Quickbird images, there are some limitations due to the restricted number of bands. With these bands, is not possible to retrieve Water Vapour with a certain precision, due to the absence of water absorption channel (1.13 μm), and aerosol, due to the absence of channel 2.200 μm to apply Kaufman-Tanre Aerosol Retrieval [414].

The required parameters are:

a) scene and sensor information
   - latitude and longitude of the scene center
   - sensor type (i.e. Quickbird)
   - FWHM (Full View at Half Maximum) and bandwidth
   - satellite altitude
   - ground elevation
   - pixel size
   - flight date
   - flight Time GMT

b) atmospheric and model settings
   - atmospheric model (between the a list of them in table 4.1)
   - water vapour retrieval option (only for Hyperspectral)
   - aerosol type, between *rural, urban, maritime,* and *tropospheric* [416]
   - aerosol retrieval option (only for Hyperspectral)
   - initial visibility between *clear sky*, *moderate haze* and *thick haze*
   - aerosol Scale Height (km)
   - CO2 Mixing Ratio (ppm)
   - adjacency correction option
   - Modtran resolution
   - Modtran multiscatter Model

All the parameters from section a) are available in the header file and from the satellite technical specification.

For atmospheric and model settings, the choice of the atmosphere model is very important especially if there is no possibility to retrieve the water vapour directly from the respective bands. If no water vapor information is available, an atmosphere should be selected according to the known or expected surface air temperature, which tends to be correlated with water vapor. For the used images, a Mid-Latitude Summer model was chosen for Tor Vergata, Nettuno and Madrid, a U.S. Standard (US) for Denver.

Model Atmosphere	Water Vapor (std atm-cm)	Water Vapor (g/cm^2)	Surface Air Temperature
Sub-Arctic Winter (SAW)	518	0.42	-16 °C
Mid-Latitude Winter (MLW)	1060	0.85	-1 °C
U.S. Standard (US)	1762	1.42	15 °C
Sub-Arctic Summer (SAS)	2589	2.08	14 °C
Mid-Latitude Summer (MLS)	3636	2.92	21 °C
Tropical (T)	5119	4.11	27 °C

Table 4.1: Column Water Vapor Amounts and Surface Temperatures for the MODTRAN Model Atmospheres

About the aerosol, the choice of model is not critical if the visibility is high (for example, greater than 40 km), and that is the condition of the analysed images (clear sky, which is the option set up in the *initial visibility* field).

The aerosol height defines the effective 1/*e* height of the aerosol vertical profile in km. Typical values are 1 to 2 km and the default value is 1.5 km. Having no information about it, it has been decided to select the default value.

For carbon dioxide (CO_2), in mixing ratio in parts per million by volume, the suggested value was 390 ppm.

With the option *adjacency correction*, the parameter ρ_e (average surface reflectance for the pixel and a surrounding region) is retrieved. If this option is not set, $\rho_e = \rho$. The command *Modtran resolution* specifies the spectral resolution of the model: for multispectral sensor, 15 cm-1 is suggested.

In the *Modtran multiscatter model*, the user could select a multiple-scattering algorithm to be used between *Isaacs* [417], *Scaled DISORT,* and *DISORT* [418]. It has been chosen a scaled DISORT with eight streams (the number of streams is related to the number of scattering directions evaluated by the model). Calculations with two or four dramatically increase computation time with little or no improvement. Increasing the number of streams for the Scaled DISORT model should not significantly increase processing time. The DISORT model provides the most accurate short-wave (less than ~ 1000 nm) corrections, however it is very computationally intensive. The Isaacs 2-stream method is fast but oversimplified. The Scaled DISORT method provides near-DISORT accuracy with almost the same speed as Isaacs. In the Scaled DISORT model, DISORT and Isaacs calculations are performed at a small number of atmospheric window wavelengths. The multiple scattering contributions in each method are identified and ratios of the DISORT and Isaacs methods are computed. This ratio is interpolated over the full wavelength range, and finally, applied as a multiple scattering scale factor in a spectral radiance calculation performed with the Isaacs method. This procedure yields near-DISORT accuracy at window wavelengths, and also improves the accuracy in absorbing regions.

Use of the DISORT multiscatter model dramatically increases FLAASH processing time, and is rarely necessary for accurate atmospheric corrections. The magnitude of multiple scattering in any scene is dependent on the amount of haze (water vapor and aerosols) that are present. Moreover, scattering preferentially affects the shorter (visible) wavelengths; longer (near infrared) wavelengths are minimally affected. The DISORT model has to be considered only when haze is very strong and critical correction of the shortest wavelengths is required.

4.1.3 Atmospheric correction results

In the next images, the results of the atmospheric correction for Quickbird images (Tor Vergata, Nettuno and Denver), and AHS image (Madrid) for vegetation (Fig. 4.2) and asphalt (Fig. 4.3) land cover are presented.
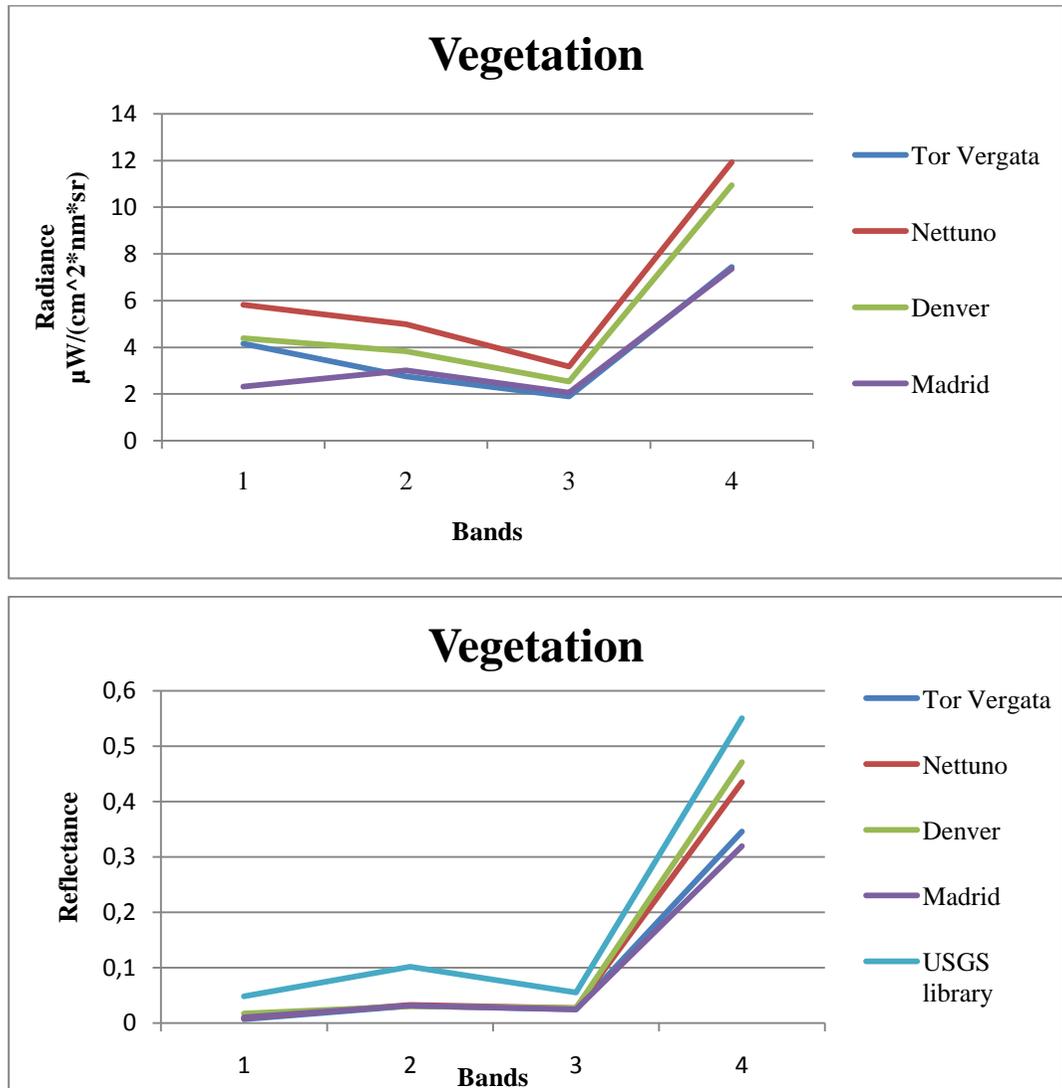
Fig. 4.2: Vegetation spectral signature in four VIS-NIR bands before (above) and after (below) atmospheric correction.

Asphalt

Asphalt
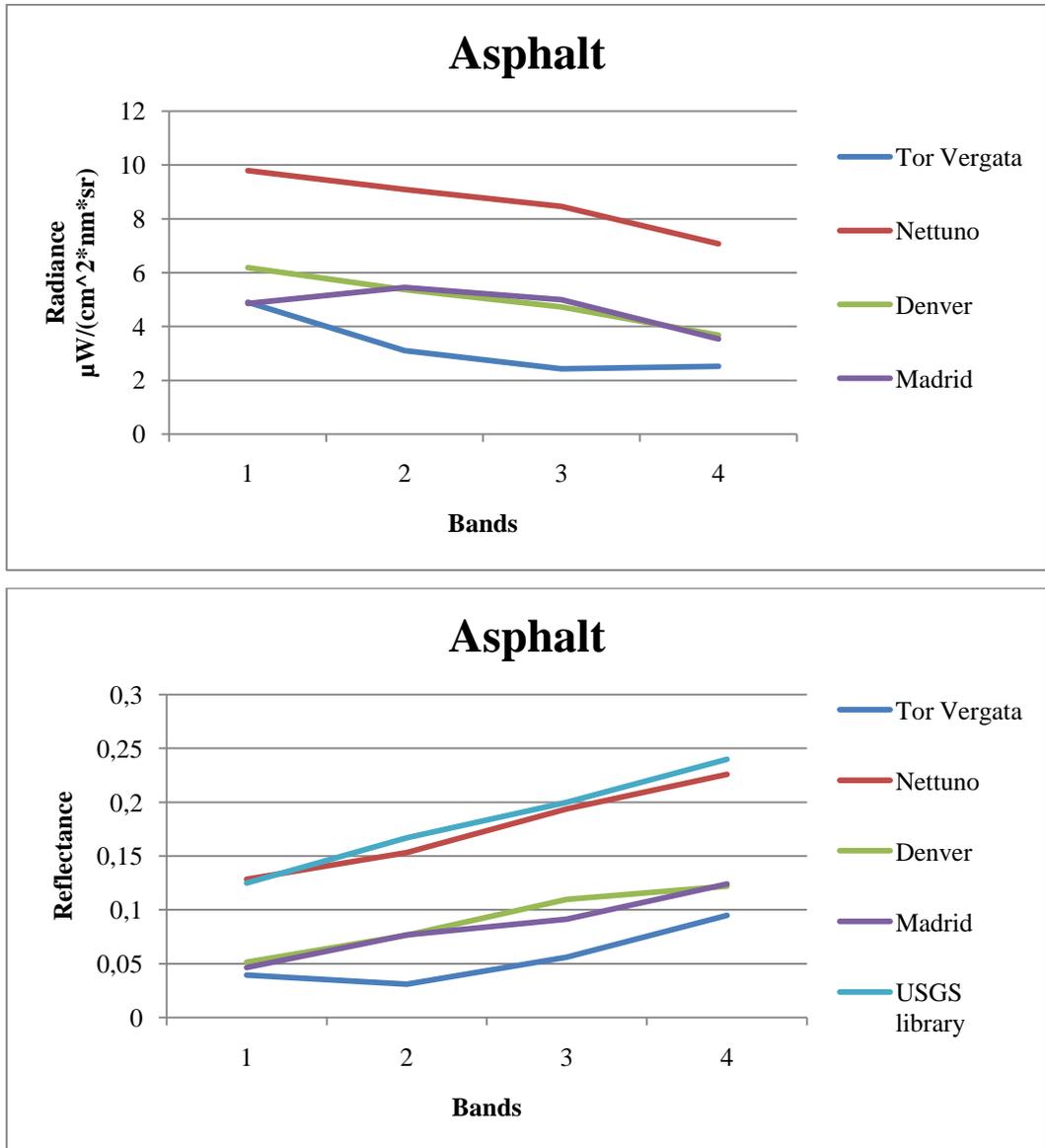
Fig. 4.3: Vegetation spectral signature in four VIS-NIR bands before (above) and after (below) atmospheric correction.

The signatures images compare the spectral response in four bands (1 = 0.42 − 0. 52 μm (Blue), 2 = 0.52 − 0.60 μm (Green), 3 = 0.63 − 0.69 μm (Red) and 4 = 0.76 − 0.89 μm (NIR)): these bands corresponds to the Quickbird bands and, consequently, the same bands have been selected from the hyperspectral image AHS in order to compare the different responses in the same interval.

Observing the radiance images, the most prominent atmospheric feature in these spectra is the consistent upward trend in blue (for the Rayleigh scattering effect). This is likely caused by atmospheric aerosol scattering, or what is often referred to as 'skylight'. An accurate atmospheric correction should compensate for the skylight to produce spectra that more truly depict surface reflectance.

In the Madrid case, this effect is not so evident like in the Quickbird images: this is due to the airborne flight which is less affected by the atmosphere Rayleigh scattering respects to the satellite sensors.

In the reflectance comparisons, standard spectral signatures from the USGS library [419] have been added. Comparing the signature from images with signatures from a spectral library allows to observe the trend of the land cover signatures with the laboratory observations, even if the type of plants and the type of asphalts could be different between images and in a single image itself.

In both of the cases, the trend has been respected: the vegetation reflectance curves now display a more characteristic shape, with a peak in the green, a chlorophyll absorption in the red, and a sharp red edge leading to higher near infrared reflectance (which varies according the plant characteristics); the asphalt signatures, where the trend was decreasing in the radiance image, now are showing an increasing trend, has it was expected from the signature from USGS library.

Other advantages of this correction will be shown in the next chapters on classification.

4.2 Building shadows removal

For remote sensing observation of urban areas, an inherent issue of high spatial resolution remote sensing images is building shadows (Fig. 4.4). A frequent problem in urban environment is measuring the real spectral property of pixels in building shadows for which shadow detection is the first step.

Fig. 4.4: buildings shadows in urban area, 1 meter resolution [420]

4.2.1 Simulated reflectance algorithm

In this paragraph, a method to reduce the effect of building shadows on urban environment, starting from the *Simulated Reflectance* algorithm developed for Airborne Thematic Mapper (ATM) images [421], is presented. The simulated reflectance can be derived from the relation between irradiance, thermal emittance, spectral reflectance and albedo based on a simplified energy conservation model of solar radiation incident on a land surface. Among the wide range of applications, the methodology has already been applied to Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) multi thermal images in which different lithologies were detected and better defined [422]. This technique highlighted sparse vegetation on shaded slopes and it permits to discriminate rocks that look uniform with a normal colour composite. Moreover, it allows to suppress topographic effect and it has the advantage of conserving surface albedo information.

However, the method carries the disadvantage of a reduction of the spatial resolution in the thermal band (90 meters).

With the increasing hyperspectral flights, where thermal bands have the same spatial resolution of the VIS-NIR bands, this technique can be very effective to detect features under building shadows and to highlight the surface that is shaded by these building in urban areas.

The solar radiation balance on the Earth can be described as:

$$B = M_a - M_e \qquad\qquad\qquad Eq.\,4.4$$

where B is the radiation balance, M_a is the absorbed energy ($M_a = \alpha E$ which depends on α, the absorptance and E, the irradiance) and M_e is the radiation emitted (thermal emission) from the land surface [423]. To satisfy conservation of energy, the following formula has to be respected:

$$\rho + \alpha = 1 \qquad\qquad\qquad Eq.\,4.5$$

where ρ is the reflectance (described by $M_r = \rho E$, the reflected solar radiation) [424]. Therefore, the Energy irradiance on a pixel is equal to:

$$E = \rho E + \alpha E = M_r + M_a = M_r + M_e + B \qquad\qquad\qquad Eq.\,4.6$$

This algorithm considers an imaging sensor system with two broad spectral bands: one is a panchromatic spectral band (visible to near-infrared) recording the reflected solar radiation M_r, which depends on the reflectance ρ) and the other one a broad thermal band recording the band thermal emission M_e. The sum image of the two bands should be equivalent to the incident irradiance on land surface E reduced by the radiation balance B, the general formula is:

$$E - B = M_r + M_e = \sum_{i=1}^{n} w_i \, \text{RefBands}_i + \sum_{j=1}^{m} w_j \text{ThermalBands}_j \qquad Eq.\,4.7$$

where the first term on the right side of equations considers a weighted sum of the *n* reflective bands of a specific sensors (it is called *simulated panchromatic* band) and the second term is generated by the weighted sum of the *m* thermal bands (called *wide thermal* band). The sum of these two terms is called *simulated irradiance band*. All the spectral bands are expressed in radiance values.

The simulated spectral reflectance of band λ is defined by:

$$\rho_{\text{sim}}(\lambda) \qquad = \frac{M_r(\lambda)}{M_r + M_e}$$

$$= \frac{\text{LRefBands}(\lambda)}{\sum_{i=1}^{n} w_i \, \text{RefBands}_i + \sum_{j=1}^{m} w_j \text{ThermalBands}_j} \qquad Eq.\,4.8$$

where LRefBands is the spectral radiance of the reflective bands refereed to a specific wavelength.

4.2.2 Application of Simulated reflectance to hyperspectral image

The above proposed methodology has been tested on the AHS image acquired above Madrid [425]. The general formulas have been adapted considering this airborne sensor characteristics:

$$E - B = M_r + M_e = \sum_{i=1}^{63} w_i \, \text{AHSRefBands}_i + \sum_{j=1}^{17} w_j \text{AHSThermalBands}_j \quad Eq.\,4.9$$

$$\rho_{\text{sim}}(\lambda) = \frac{\text{AHSRefBands}(\lambda)}{\sum_{i=1}^{63} w_i \, \text{AHSRefBands}_i + \sum_{j=1}^{17} w_j \text{AHSThermalBands}_j} \quad Eq.\,4.10$$

Likewise, the simulated emittance ε_{sim} for thermal bands (from 64 to 80) can be retrieved with the following formula:

$$\varepsilon_{\text{sim}}(\lambda) = \frac{\text{AHSThermalBands}(\lambda)}{\sum_{i=1}^{63} w_i \, \text{AHSRefBands}_i + \sum_{j=1}^{17} w_j \text{AHSThermalBands}_j} \quad Eq.\,4.11$$

Weights may be calculated using the solar radiation curve; it means that, for a specific pixel, the weight of a single band is equal to:

1) band irradiance by bandwidth (area of each column);
2) Sum of all the columns (total area);
3) Divide each column per the total area.

Solar Radiation Spectrum

UV | Visible | Infrared →

Sunlight at Top of the Atmosphere

5250°C Blackbody Spectrum

Radiation at Sea Level

H_2O

O_2

H_2O

O_3

H_2O

Absorption Bands

H_2O CO_2

H_2O

Spectral Irradiance (W/m²/nm)

Wavelength (nm)
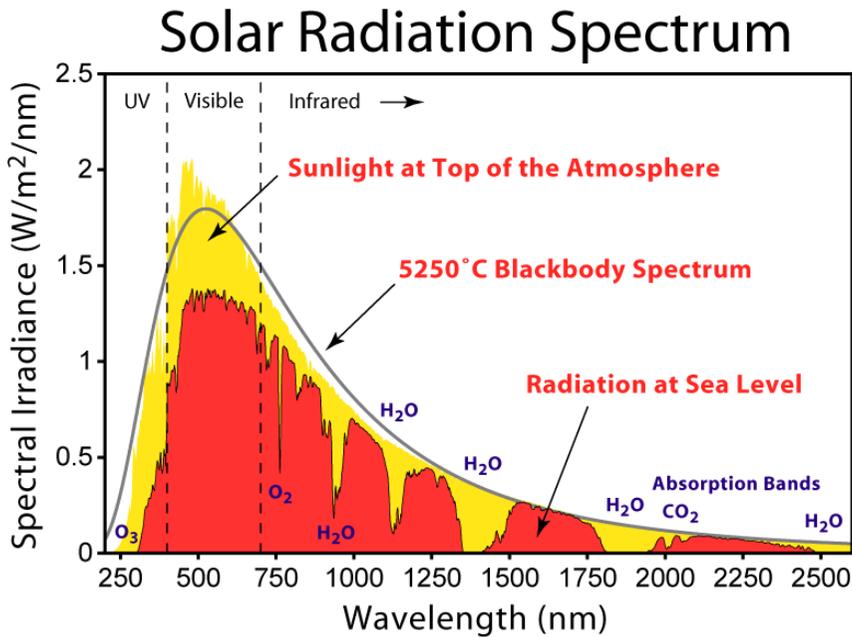
Fig. 4.5: solar radiation spectrum from 250 nm to 2500 nm

Calculating the weights, is appropriate to consider that some bands in the MIR part of the spectrum have not been used (they were affected by stripes) and the *simulated irradiance band* has been linearly stretched with a proper cut-off at both high and low ends of the histogram before calculating $\rho_{sim}(\lambda)$. With all the image bands having the same DN range after the stretch, the weights calculated from the solar radiation curve can thus be applied on all the bands.

By the way, several attempts have confirmed that *simulated panchromatic band* and *wide thermal band* can be obtained with a sum of the respective bands without considering their weight: that is probably due to the fact that the processed bands have the same bandwidth (for hyperspectral image) in the VIS-NIR part of the spectrum. The contribute of these band is the most conspicuous in the simulated panchromatic band.

4.2.3 Results

As shown in figure 4.6, the first test on an AHS testing image above Madrid showed the enhancement of the image after the simulated reflectance processing: the topographic effects of buildings have been suppressed (the entire image appears flatter) with spectral properties of various ground objects enhanced.

In the following image (4.7), where some details of the image are presented, the simulated reflectance image shows some interesting results:

- the shaded perimeter around the stadium is now visible and it can be observed that it is made of asphalt, with the same spectral properties of the lower part of the perimeter already visible before correction;
- the stand of the stadium, which was saturated and not clearly distinguishable in the left image, is now highlight due to its different response in the other bands (reflective and thermal);
- in the image below, some areas, which were shaded from buildings, show the presence of vegetation (with red) after simulated reflectance application.
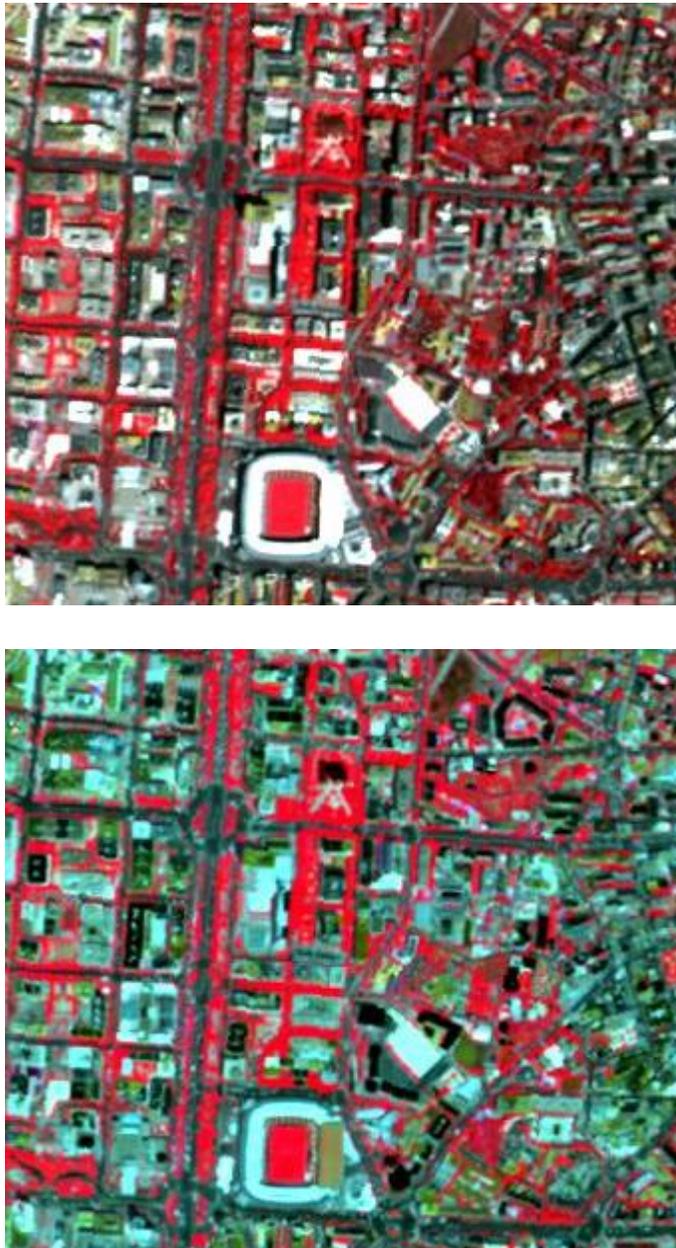
Fig. 4.6: AHS false colour composite 14-8-5 RGB obtained with the original image (left, above) and simulated reflectance image (left, below)
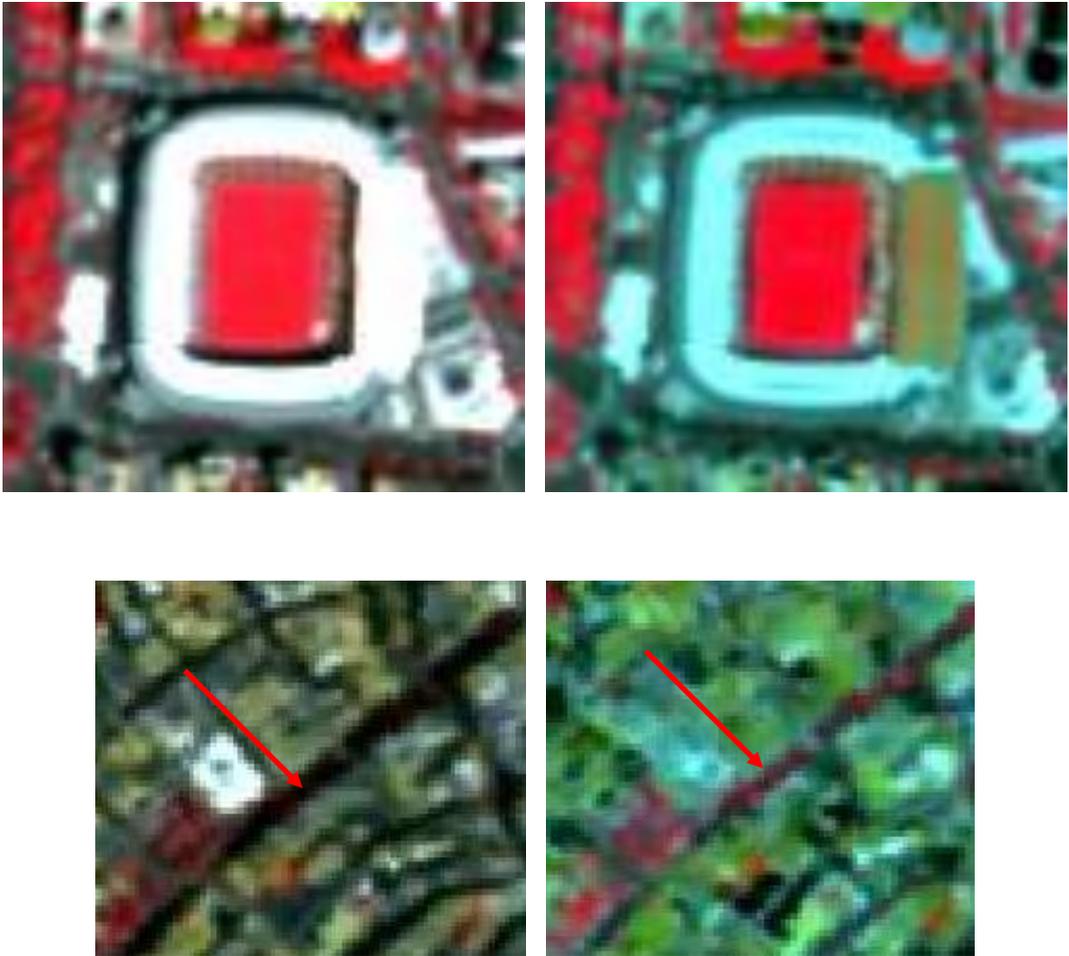
Fig. 4.7: Details of AHS false colour composite 14-8-5 RGB obtained with the original image (left) and simulated reflectance image (right). A detail of the stadium. (right, above) and shaded vegetation on the street (right, below) enhanced by the simulated reflectance is highlighted

Another useful application it can be observed to spot shadows as a class itself in a classification image. In figure 4.8, a mixed color composite with original and simulated reflectance bands allows to spot shadows (shown in red in the lower image) which can be classified in a "shadow" class with a simple threshold rule.

Fig 4.8: AHS real colour composite 8-5-2 RGB obtained with the original image (above) and colour composite (below) obtained with 8(sim)-8(real)-8(real). Using a simulated reflectance band, which has higher values of shaded pixels, allows to highlights shadows (with red).

This section can be concluded noting that the application of simulated reflectance algorithm in urban environments showed satisfactory results to remove building shadows, especially considering the use of hyperspectral sensors, where the spatial resolution of their thermal bands is the same of the other visible and near infrared bands. This methodology allowed to suppress topographic effect and it had the advantage of conserving surface albedo information. Moreover, the technique highlighted some details which cannot be observed with the only use of original spectral bands (due to the saturation of channels and the presence of shadows). The method does not require the use of external data and it is completely automatic starting from the radiance values. The initial dataset could be converted in simulated reflectance and simulated emittance bands, allowing the user to process the data with any other image processing techniques.

Chapter 5

Spectral, textural and shape characteristics of urban main land cover classes

This chapter aims at analyzing the characteristics of the land cover classes previously mentioned as objects of interest for this study: vegetation, water, bare soil, buildings and roads. These classes present some properties deriving from their electromagnetic spectrum and the pixel-context, which define each class. Some features can be easily recognized only considering their spectral signature, for some others the addition of the context of the pixel is necessary to better define the membership class.

Regarding the multispectral images, the registered radiance was assumed to be the response at the central wavelength of each band and they have been compared with reference signatures extracted from spectral libraries at the same wavelength.

5.1 Vegetation

Since early instruments of Earth Observation, such as NASA's ERTS and NOAA's AVHRR [501] acquired data in the red and near-infrared, it was natural to exploit the strong differences in plant reflectance to determine their spatial distribution in these satellite images. Leaf cells reflect solar radiation in the near infrared spectral region and absorb in the photosynthetically active radiation (PAR) spectral region, especially in red band. From these two quantities is easy to retrieve information about vegetation: many indices have been developed on this spectral relationship and the most used is called Normalized Difference Vegetation Index (NDVI).

Often NDVI is calculated using radiance, even if it is generally accepted that reflectance values give more precise results [502], [503]. By design, the NDVI itself varies between -1 and +1 [504].

Therefore, vegetation presents a distinguishable spectrum in the four main spectral bands (Blue, Green, Red and Infrared of Quickbird and Hyperspectral images). Due to this intrinsic properties, Kohonen SOM is able to distinguish this class with only spectral parameters.

The following image (Fig. 5.1) shows the vegetation reflectance values in the testing images (therefore, after the atmospheric correction from radiance values) and the reference signature from the spectral library [505] considering blue, green, red and NIR bands. It is worthwhile to observe that even without atmospheric correction, the vegetation can be easily spotted due to its different spectral characteristics respects to the other land cover classes.

On the other hand, in a multi-image context, the atmospheric correction allowed to standardize the signatures and a comparison between themselves.
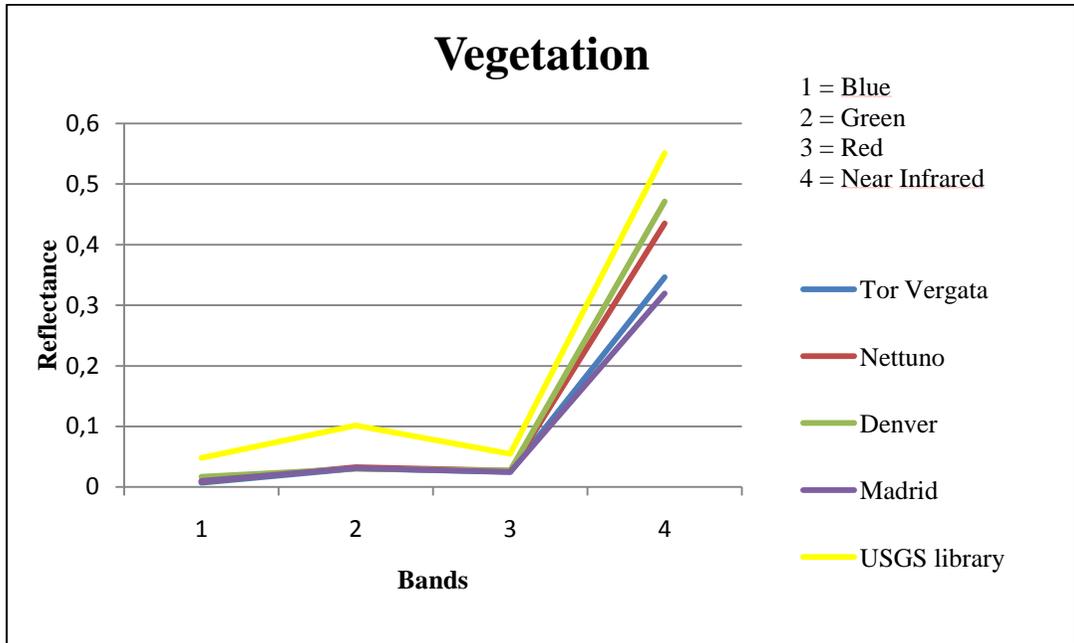
Fig. 5.1: Vegetation reflectance values (VIS-NIR bands) in Quickbird images (Tor Vergata, Nettuno and Denver), AHS image (Madrid) and spectral library.

5.2 Water

The reflectance values of water bodies (sea for Nettuno and Bari images, a small lake in Madrid image) show (Fig. 5.2) the typical spectrum of water with some local variation due to the material in the water (organic and/or inorganic material). The spectral reflectance of sea water depends on the combined effect of scattering and absorption by water itself, by the suspended particles both of organic (phytoplankton) and inorganic (sediments) nature.

The infrared intense vibrational bands make water essentially opaque at wavelengths beyond about 0.8 μm, thus reducing its reflectivity, which differs appreciably from zero only in the visible, where it exhibits a decreasing trend with wavelength. Phytoplankton modifies the trend by decreasing reflectivity at wavelength lengths below about 0.54 μm and slightly increasing it at larger wavelengths. Note that at the higher concentration of phytoplankton the spectral features of chlorophyll appear, with the evident reflectance minimum around 0.44 μm and the one just hinted about 0.66 μm. The effect of suspended inorganic sediments is different, given their large scattering section which causes the reflectance to increase with increasing concentration (this is the case of the lake in Madrid), except that at the lower wavelengths, where the chlorophyll absorption is generally present [506].
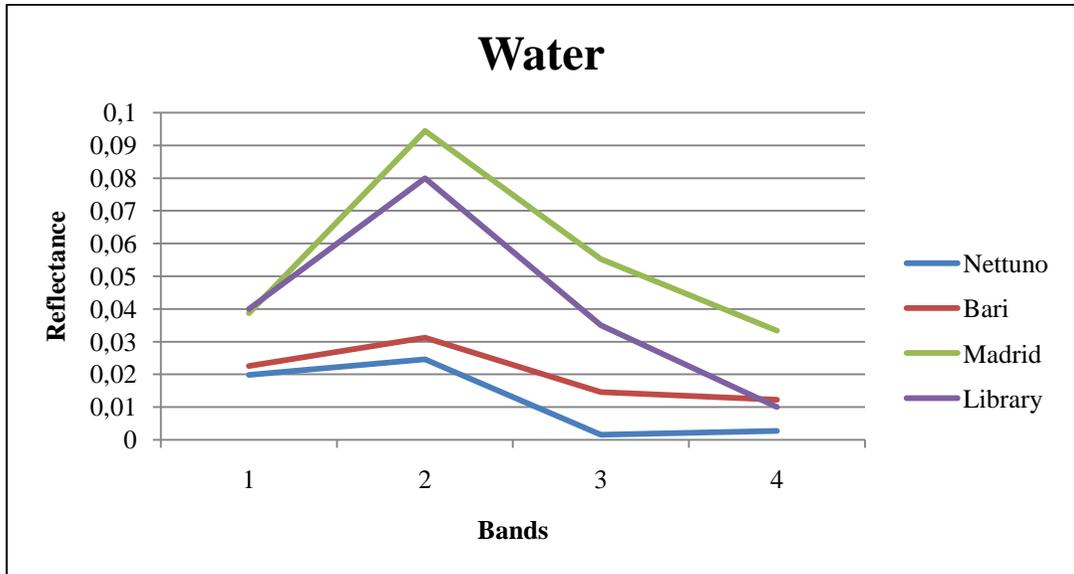
Fig. 5.2: Water reflectance values (VIS-NIR bands) in Quickbird images (Nettuno, sea), MIVIS (Bari, sea), AHS image (Madrid, lake) and spectral library [507].

5.3 Bare Soil

The interaction characteristics depend on the properties of the soil surface. The absorption mechanisms primarily involve vibrational (infrared) and electronic transitions (visible). Note that the more commonly encountered mineral substances, formed primarily by silicon, aluminum and oxygen atoms, do not present energy levels allowing transitions involving visible and near infrared photons. Instead of them, iron ions strongly absorb the shorter visible wavelengths. Organic matter, even if present in moderate quantity, has a strong influence on reflectance (high organic matter means high reflectance), given its water retentivity.

However, given the soil formation process, its reflectance may contain various spectral components of the underlying minerals: the spectra of pure rocky materials are usually less complex than those of the soil, since they depend on lattice-radiation interaction [506].

Soil spectra (Fig. 5.3) have a clear increasing values in all the images, with low reflectance values. The variation between images can be brought back to the different soil materials. Bare soil signature can be confused in some cases with the signature of roads (as it will be shown in figure 5.6). For this reason, bare soil is one the most difficult classes to identify and spot. By the way, if more bands are available, as in the hyperspectral images, this problem can be overtaken analyzing a wider spectrum.
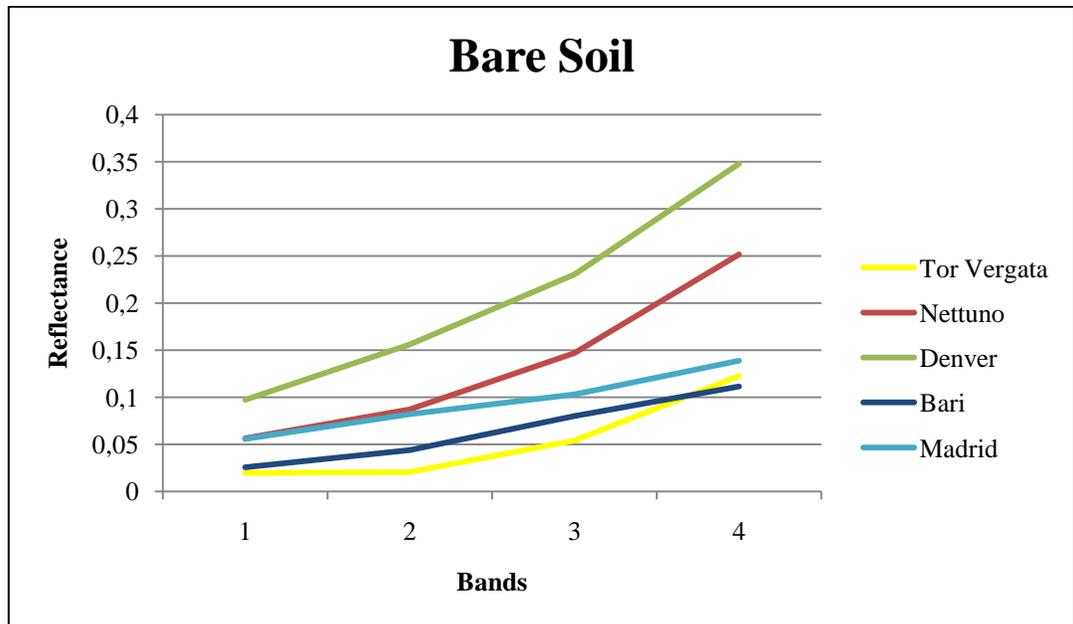
Fig. 5.3: Bare Soil reflectance values (VIS-NIR bands) in Quickbird images (Tor Vergata, Denver), MIVIS (Bari) and AHS image (Madrid).

5.4 Buildings

As it was written in the first considerations on urban areas, buildings are made of different material and it is difficult to determine a general spectral signature which allows to spot buildings with VHR images. Figure 5.4. shows the different spectral signature from buildings in radiance values. This image is also useful to demonstrate the difficulties to compare the spectral responses without the standardization that can be obtained through atmospheric correction. In chapter four it has been introduced how the use of Homogeneity could help to distinguish the buildings in some images. Homogeneity features have been computed with two different window sizes of 3×3 and 5×5 only on AHS green band (the band which gave the best appreciable results in buildings identification).

After the atmospheric correction, it is possible to compare the spectral signatures belonging to the building from different images and a spectral signature form USGS library. The spectral response appears more similar, with a increasing trend from blue to near infrared band. By the way, this trend can be observed in other man made features as well (i.e. asphalt) and a clear identification is not always possible. For instance, the red bricks of the roof of Madrid buildings (which will shown also in figure 7.2) can be associated clearly to the spectral signature from *United States Geological Survey* (USGS), but for the others, a significant difference (probably due to the different material) is still present. Therefore, the addition of textural information, especially in Very High Resolution (VHR) image analysis can be helpful. A reference building signature which combines not only spectral parameters but also textural parameters has been then considered as *standard building signature* [508]. Considering the window sizes, an high value of homogeneity can be observed in window 3 x 3 (in relation of the measure of a small building) and it tends to decrease in a window 5 x 5. This trend is respected in all the images.
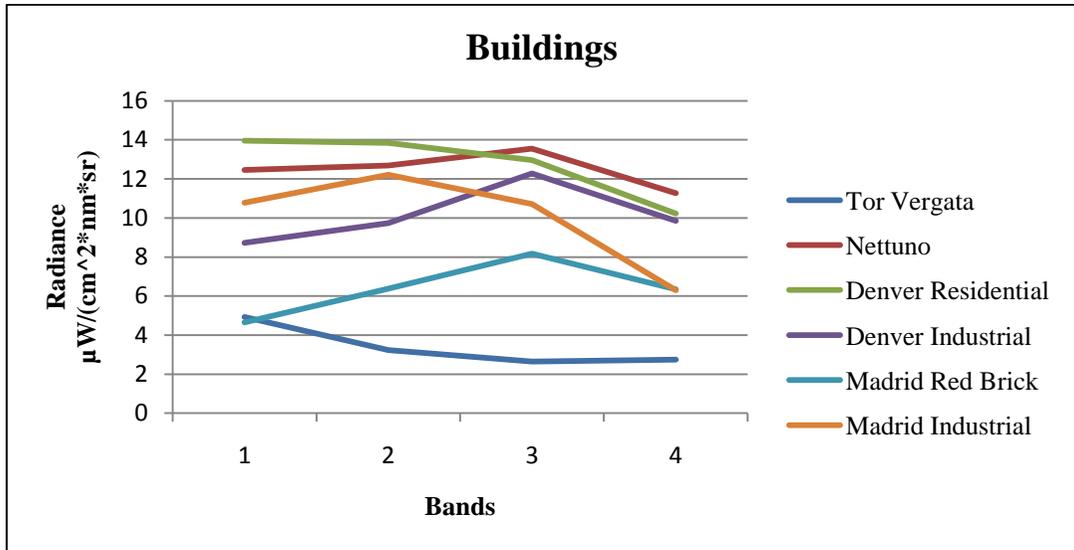
Fig. 5.4: buildings radiance values (VIS-NIR bands) in Quickbird images (Tor Vergata, Nettuno and Denver) and AHS image (Madrid)..
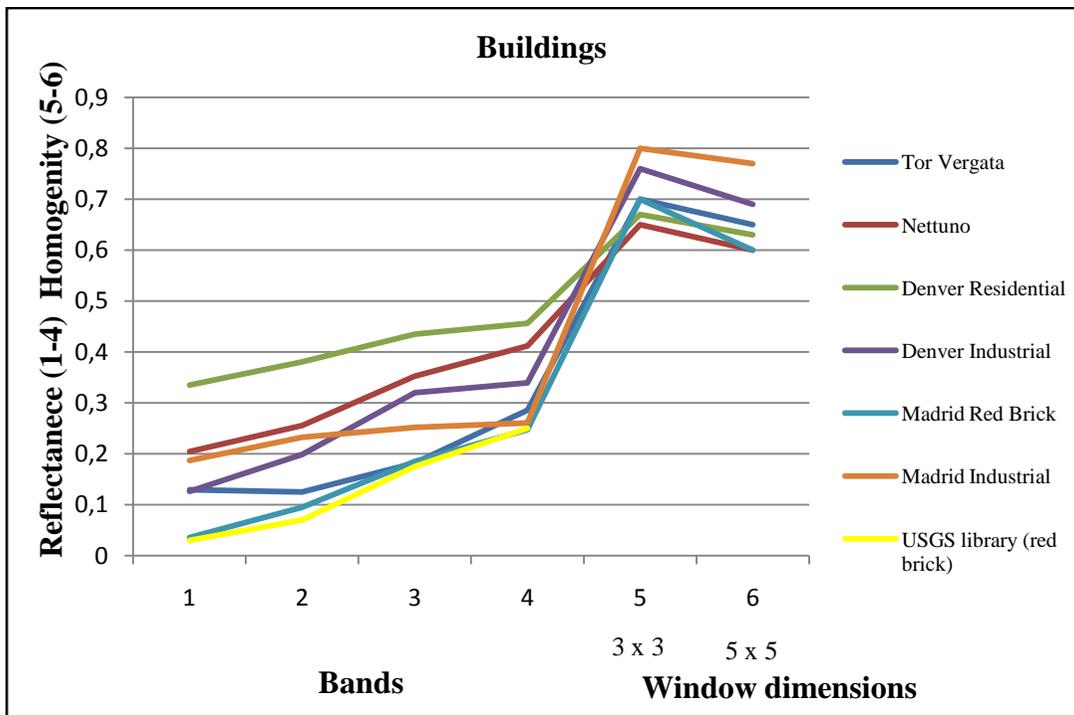
Fig. 5.5: buildings reflectance values (VIS-NIR bands) in Quickbird images (Tor Vergata, Nettuno and Denver), AHS image (Madrid) and spectral library. Hogeneity parameter from green band with two window size has been added.

5.5 Roads

Roads correspond to man-made surfaces: their identification in Very High Resolution imagery is a difficult task because of the complexity and variability within the road network. Roads can exhibit a variety of spectral responses due to the differences in age and material and vary widely in physical dimensions. Road networks in dense urban areas typically have different geometric characteristics than those in suburban and rural areas [509]. Extraction of roads directly from digital images typically may involve the use of spectral characteristics and/or the use of object detection and/or shape parameters.

As mentioned in chapter 3, in some cases roads can be confused with buildings if the material is similar (i.e. concrete) and bare soil (even if the spectral signature is different, the magnitude of the different could not be enough to allow the discrimination between them).

These considerations were confirmed after the signatures extraction from the tested image (Fig. 5.6): a general trend with increasing values from visible to infrared parts was observed. Comparing them to the standard spectral signature, the roads from the Nettuno image seemed the most similar to the library response.

In figure 5.7, spectral signatures from bare soil and asphalt are inserted in the same graph to show their similarity. In some cases, a reference spectral signature of roads for labeling can label the bare soil asphalt or vice versa, depending of the image range of values.

This limitation of the labeling technique affects the multispectral images with four bands (i.e. Quickbird): with hyperspectral images (as shown in figure 5.8) and in World View 2 satellite[15], the higher number of bands defines a more complex spectral signature where the intrinsic properties of the material are highlighted.

---

[15] *World view II homepage*: http://www.satimagingcorp.com/satellite-sensors/worldview-2.html
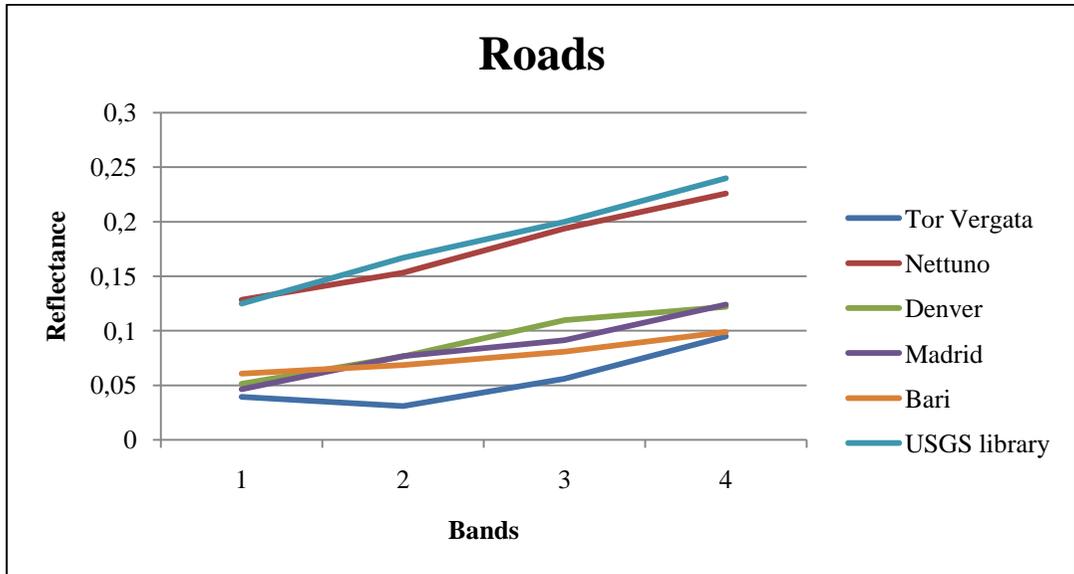
Fig. 5.6: roads reflectance values (VIS-NIR bands) in Quickbird images (Tor Vergata, Nettuno and Denver), AHS image (Madrid), MIVIS image (Bari) and spectral library (Clark).
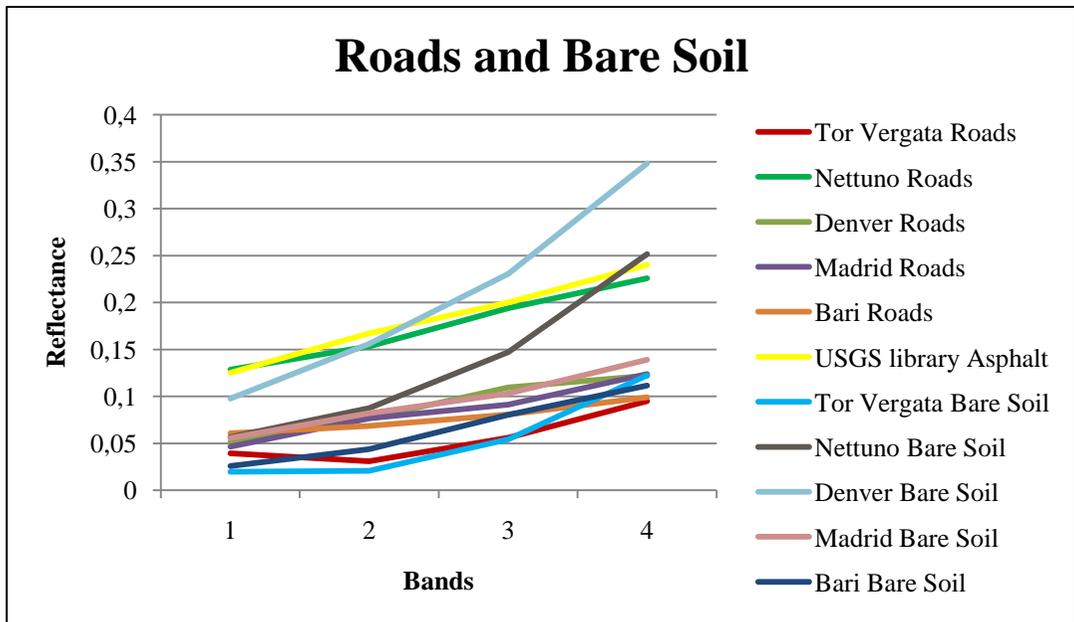
Fig. 5.7: comparison between roads signatures and bare soil signatures in all the images (Tor Vergata, Nettuno and Denver, Madrid and Bari) and asphalt from library (Clark).
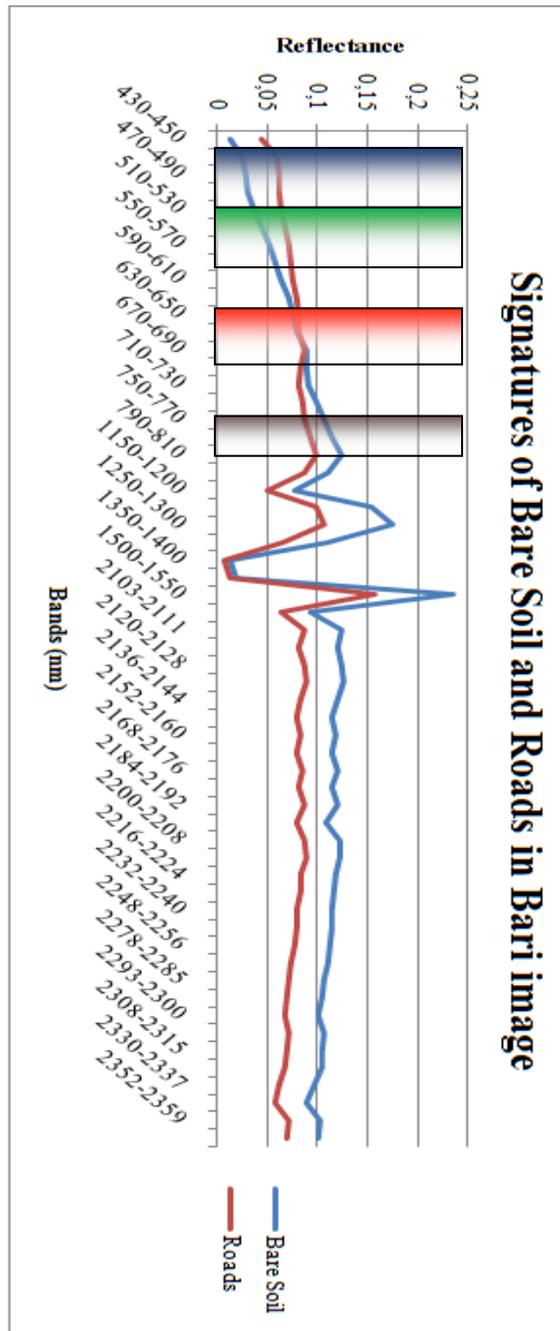
Fig. 5.8: comparison between roads and bare soil in hyperspectral image (MIVIS, Bari). The hyperspectral interval is not continuous (it has been calculated on 58 bands). The four bands available in multispectral are highlighted.

For these reasons, it has been decided to introduce additional information segmenting the image before the classification and calculating *elongation* as a shape measure that indicates the ratio of the major axis of the polygon to the minor axis of the polygon. The major and minor axes are derived from an oriented bounding box containing the polygon. (i. e. elongation value for a square is 1.0, and the value for a rectangle is greater than 1.0). In figure 5.9 a portion of the segmented image of Tor Vergata (I) is presented: after segmentation, some details of the image are lost and a generalized representation of objects allows to retrieve information about the belonging land cover class. The elongation of each polygon, setting a threshold equal to 3, could help to define roads if they have been clearly set from the segmentation process (Fig. 5.10).
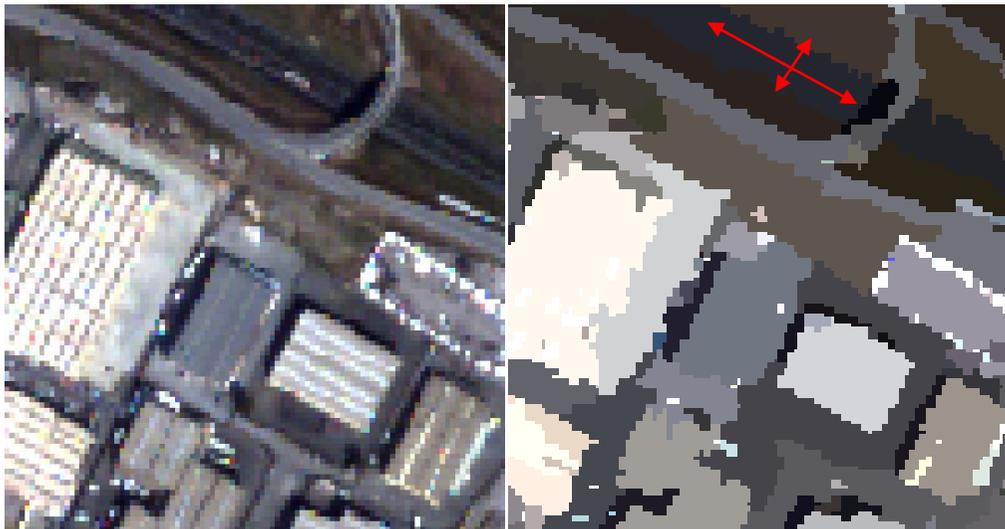
Fig. 5.9: a detail of Tor Vergata (left) and its segmentation (right). The major and the minor axis of a polygon corresponding to roads are highlighted.
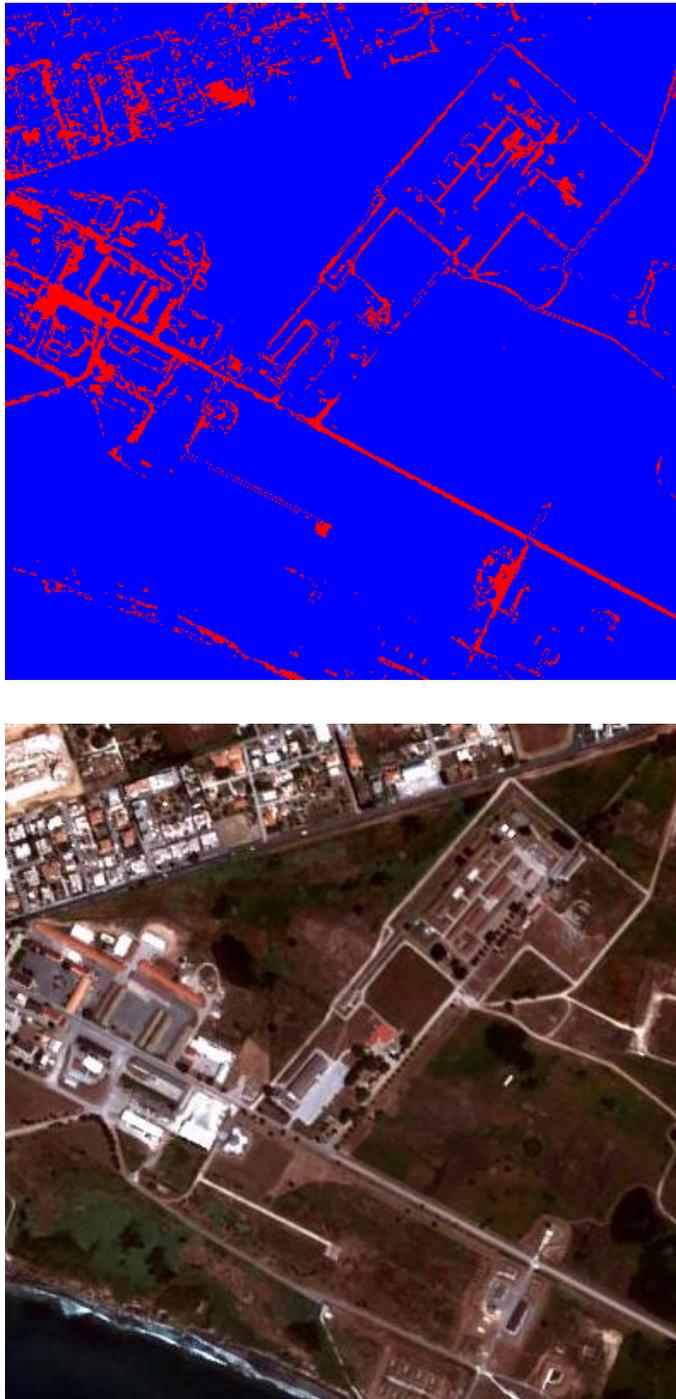
Fig. 5.10: extraction of roads from segmented image through object with elongation value greater than 3 (above) and original Nettuno image (below).

Chapter 6

Pixel based classification of VHR images: results on multispectral and hyperspectral images

In this chapter, the developed chain to classify VHR multispectral and hyperspectral images will be shown with the results on the classification accuracies in the different test images.

The classification chains will be described following the methodology described in chapter three, with a focus on the ability of the classifier to spot a specific class and the overall accuracy of the final process.

6.1 Natural features

The classes included in the natural features are vegetation, water and bare soil. Their particular spectral characteristics allow their discrimination using only the spectral signature which can be compared with the library spectra.

6.1.1 Vegetation class

As it has been written, vegetation can be easily extracted with the only use of spectral values and it can be assigned to its own semantic class through an automatic labelling.

After an automatic clusterization with Kohonen map using several SOM configurations, the mean signature of the extracted cluster have been compared with the signature from spectral library (Clark). Successively, the accuracy matrices, based on a ground truth of 1000 pixels, verified the quality of the vegetation extraction. A SOM with a different number of neurons has been trained with 50 epochs and the results have been compared.

The table 6.1 summarizes the results of vegetation extraction in Nettuno image, where the similar approach has been used to highlight the same class on the other images. The results have been compared with the classic k-means clusterization technique on the same number of classes. Considering that the process aimed at extracting and identifying vegetated areas, which will be sent to the MLP as ground truth for the vegetation class, the first parameter to check is the *percentage* of pixels in the training set which have been described as vegetation class.. If there are some other land cover classes within it (which means a value lower than 100 %), the ROI is not valid for the successive step, the supervised Neural Network.

The second parameter is the *number* of pixels which have been included in this class. Low values of this parameter means that the intrinsic variability of vegetation have not been highlighted: this lead to an higher value of the unclassified pixels. In

the following clusterization steps, these unclassified pixels may be assigned to other classes.

Conversely, a significant number of pixels already assigned at this level, means that a wide typology of vegetation cover have been detected and a lower number of pixel will be reassigned in the next steps and in the final MLP NNs.

	Clusterisation processes							
	Kohonen				**K means**			
	5 x 5	4 x 4	3 x 3	2 x 2	25	16	9	4
Accuracy in vegetation identification	100 %	100 %	100 %	85 %	100 %	97 %	93 %	100 %
Pixel identified (on 1000 pixel dataset)	681	686	727	875	556	612	693	704

Table 6.1: comparison between Kohonen and K-means classifier with different size.

Except with Kohonen SOM 2 x 2 neurons, all the pixels belonging to vegetation, are effectively vegetated (as confirmed by the ground truth): the 2 x 2 SOM, even if it spotted an higher number of vegetated pixels, has to be avoided because it stated that 15 % of the non-vegetated pixels are vegetated.

The best method from these two parameters analysis is the one with a SOM structure of 3 x 3 neurons. It can be observed that, increasing the number of neurons, the number of pixel slightly change in not significant way.

Each neuron, after a training phase, has a final weight: the values of the weights depends on the training dataset (which should be at least the 10 % of the entire dataset) and they correspond to reflectance values. Each pixel which is sent to the Kohonen is compared with these values and the neuron with a lower Euclidean distance with the input is considered the winning one.

Fig. 6.1 shows the final weights of the nine neurons of the selected SOM: the vegetation signature, due to its high response in NIR band, varies substantially from the others. The winning neuron (number 6) corresponds exactly to the one with the higher response in the NIR part of the spectrum.
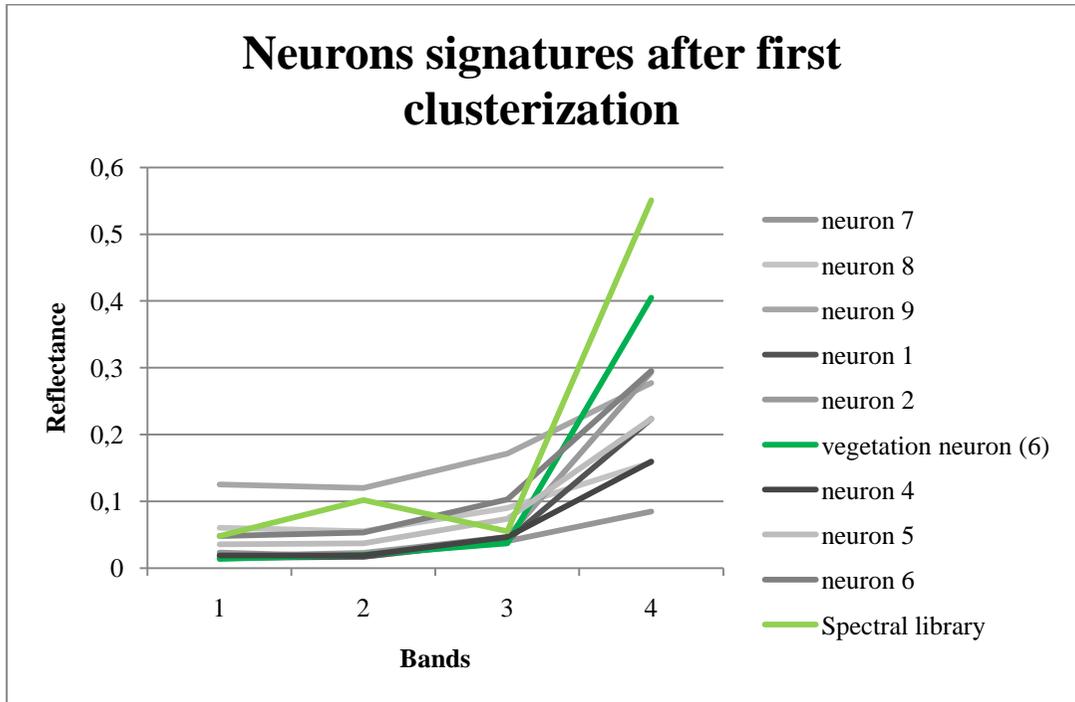
Fig. 6.1: spectral signatures of Kohonen neurons and reference signature for vegetation (Nettuno image)
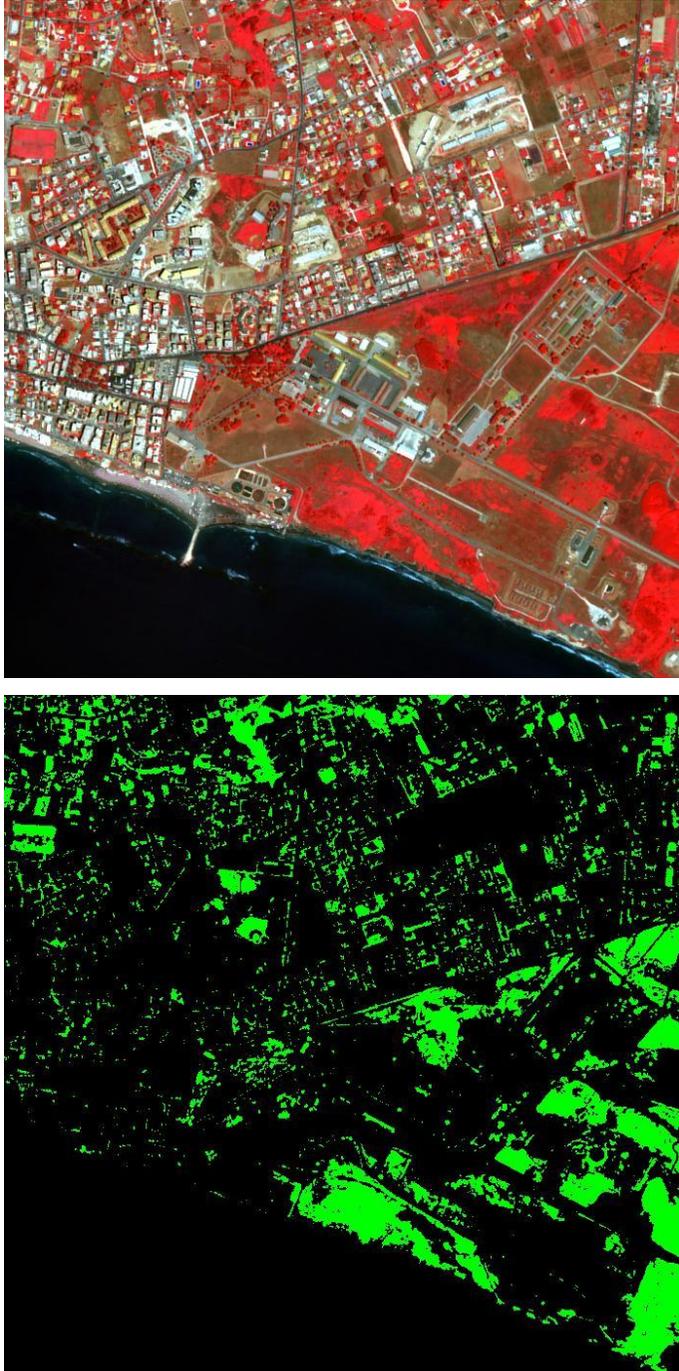
Fig. 6.2: Nettuno false color composite 432 (above) and vegetation extraction -in green- (below).

6.1.2 Water class

To test the capacity of Kohonen neural map to highlight water class, some tests have been performed on the image of Nettuno using the same configuration as previously adopted (3 x 3 neurons, 50 epochs), adding the water signature from library for automatic labelling.

As it can be observed in figure 6.3, the two classes have been spotted easily considering their typical spectral signature: moreover, the labelling is still automatic, due to difference between water spectrum and the other classes.

Fig. 6. 4 shows Nettuno image with the two identified classes (vegetation and water) highlighted.

Fig. 6.3: spectral signatures of Kohonen neurons and reference signatures for vegetation and water (Nettuno image).

Fig. 6.4: Nettuno false color composite 432 (above) and vegetation -in green- and water –in blue-extraction (below).

For Bari image from MIVIS sensor, water is the predominant part of the image: therefore, the algorithm is able to discriminate different clusters inside this class (possibly for the different bathymetry and suspended sediments).

It is worthwhile to observe the difference in clusterization, using all the available VIS-NIR spectrum from MIVIS (58 bands, excluding the bands affected by stripes), and the clusterization using only the four spectral bands homologues to Quickbird (20-12-7-2). The image 6.5 *center* use 58 bands for the classification and allowed to extract two clusters related to water (pink and sea green), but the pink is also presented in some land pixels, which means that only the cluster coloured with sea green can be considered as ground truth for the following MLP network.

The image 6.5. bottom presents three clusters related to water (pink, sea green and blue), but this cluster can be assigned only to sea class (no presence of them on land). This consideration does not necessary means that four bands allows a better classification respects to hyperspectral images but, for just macro-class distinction as land/water, the use of a limited number of bands can give satisfactory results.
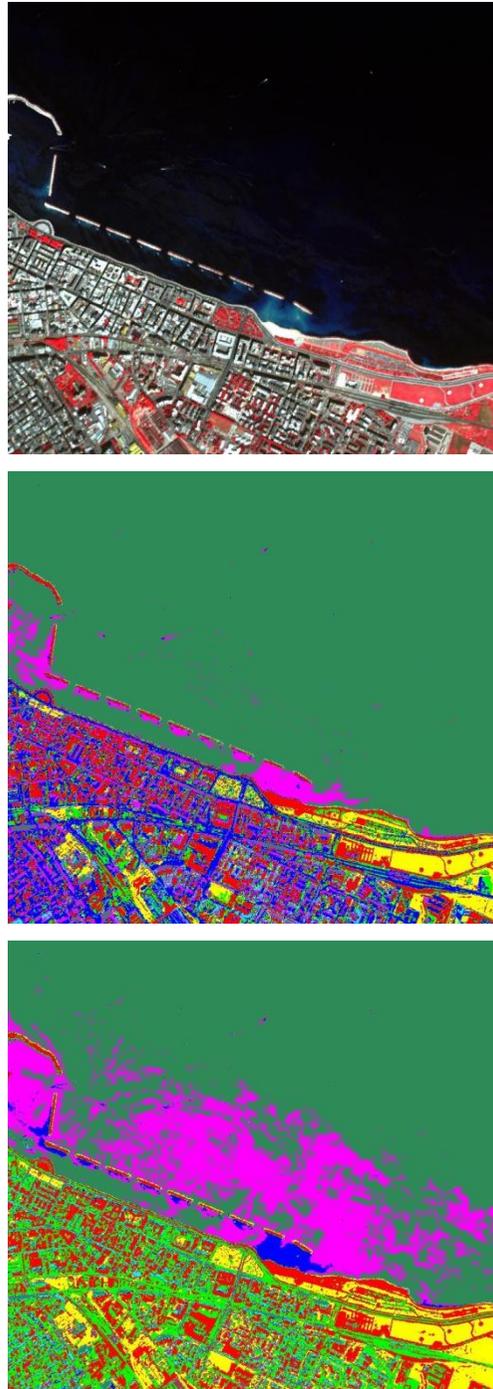
Fig. 6.5: Bari false color composite 20-12-7 (above), clusterization with 58 bands (center) and clusterization with 4 bands RGB + NIR (below).

6.1.3 Bare soil class

Adding the bare soil signature, the corresponding class has been identified in Nettuno image. As it is possible to observe from figure 6.7, the neuron values can be associated to the library spectra due to their low spectral distance. The labeling of this class has been performed in automatic way.
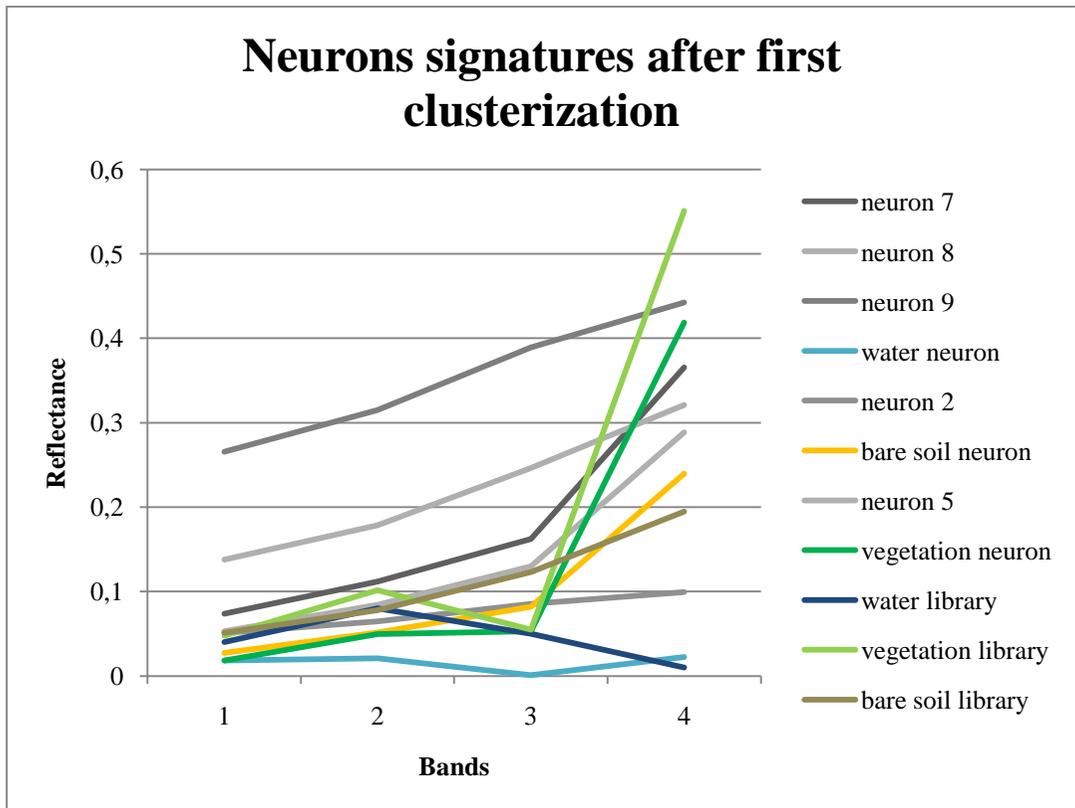
Fig. 6.6: spectral signatures of Kohonen neurons and reference signatures for vegetation, water and bare soil (Nettuno image).
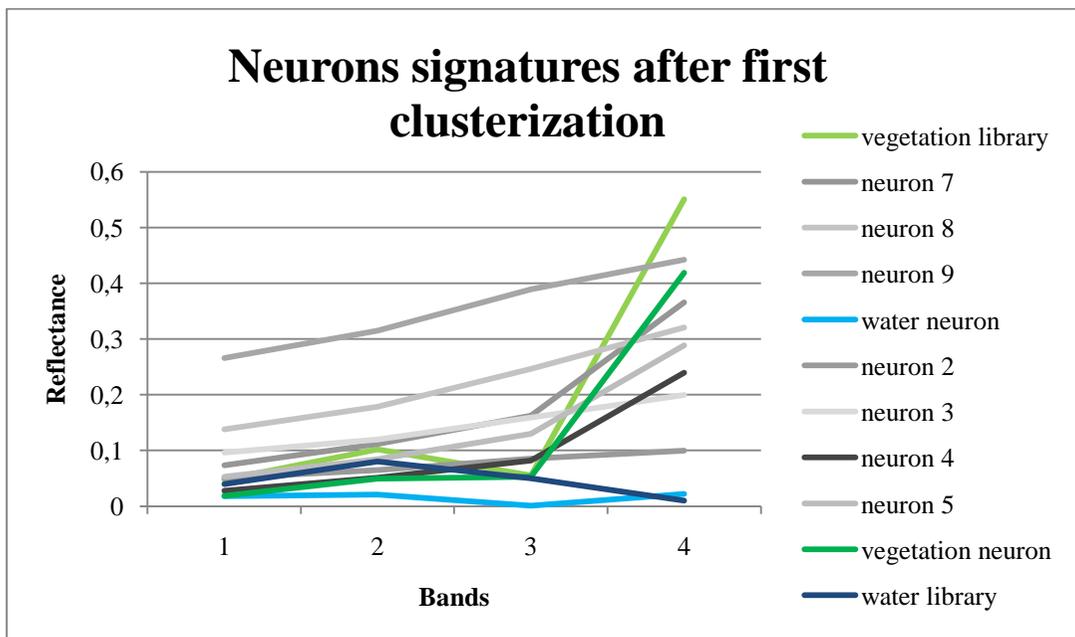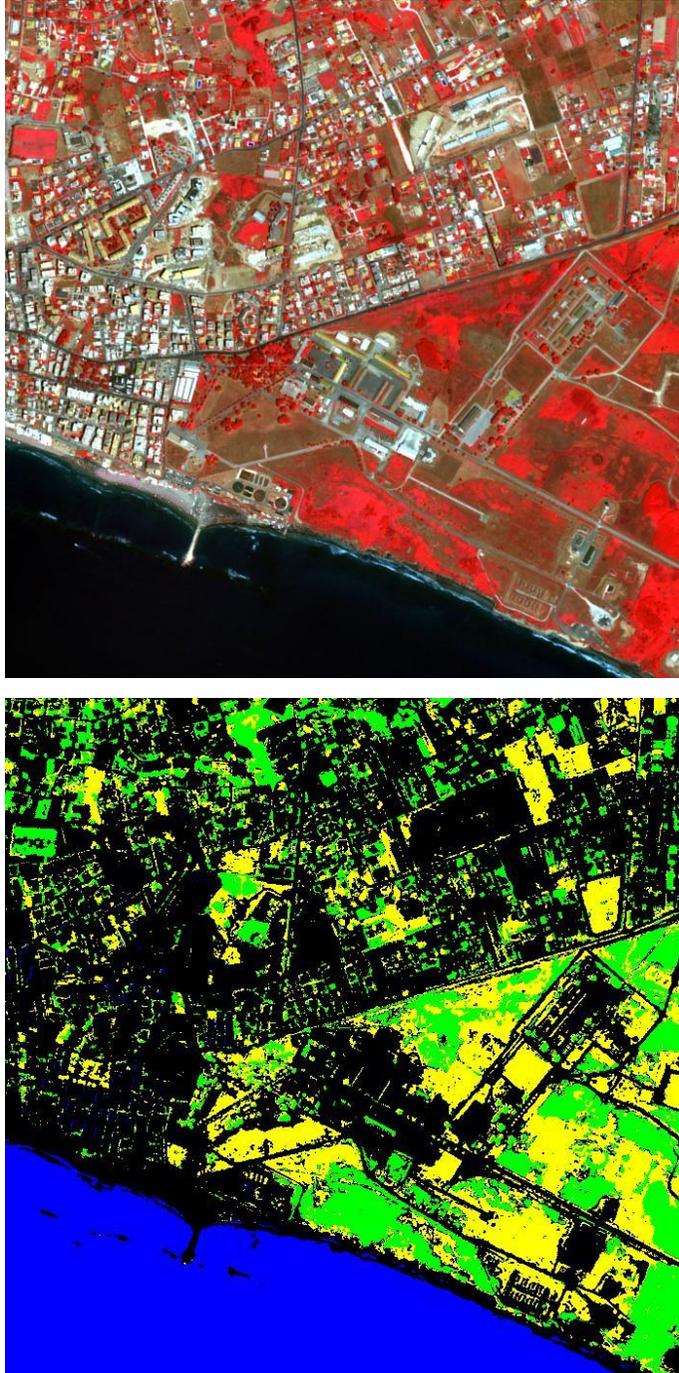
Fig. 6.7: Nettuno false color composite 432 (above) and vegetation -in green-, water –in blue- and bare soil –in yellow- extraction (below).

However, bare soil identification has been performed on different images giving different results. As it has been introduced in the study on chapter five on class characteristics, bare soil cover in VHR images presents some intrinsic difficulties for its identification:

- with the use of four spectral bands, in some cases the spectrum of bare soil can be confused with the spectrum of roads;

- the automatic labelling can be performed in not all the cases, due to the variability of the different bare soil in the same image (internal) and between a set of them (external).

Following, the example of the extraction of bare soil in Tor Vergata image is presented. With the above mentioned procedure with automatic labelling, the cluster assigned to bare soil corresponds in fact to the roads class (Fig. 6.8).

This can be explained through the analysis of spectral signatures of neurons from Kohonen map (Fig. 6.9).

In this area, bare soil signature has very low level of reflectance: as it is possible to see in the neurons comparison, these values have a higher distance respects to roads with the standard bare soil signature. Therefore, the automatic labelling assigned the pixel belonging to the wrong class to bare soil. Moreover, observing the real spectral signature of roads and bare soil in the image (extracted manually from image, not from SOM neurons), the small difference between roads value and bare soil values (Fig. 6.10), brought to a not clear distinction between them and the consequence is a mixed neuron with roads and bare soil.

In this area, the bare soil identification has been made with an interactive process after clusterisation. As shown from the classified image from the Kohonen map (6.11), one class contains pure bare soil pixel (in yellow) and another one contains mixed pixel (in orange). The cluster to consider for the next MLP is the one which contains pure bare soil pixels.
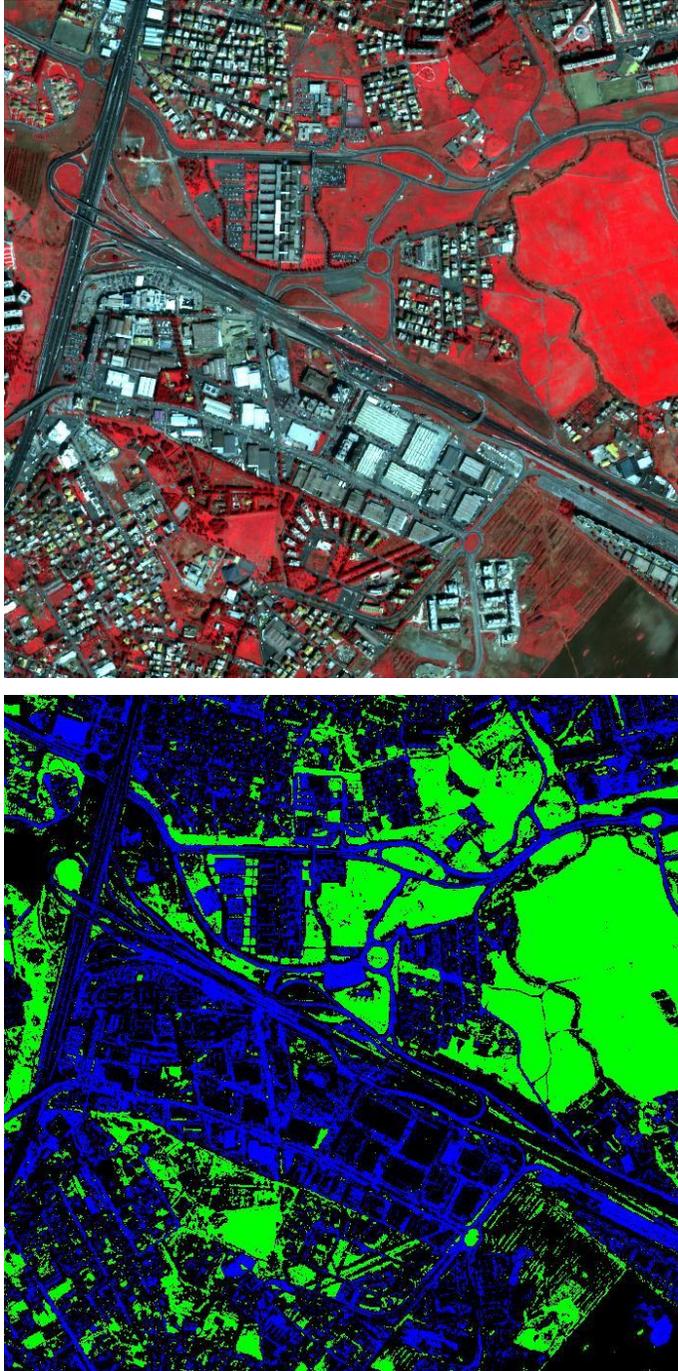
Fig. 6.8: Tor Vergata false color composite 432 (above) and vegetation -in green- and misclassification of bare soil in roads –in blue- extraction (below).
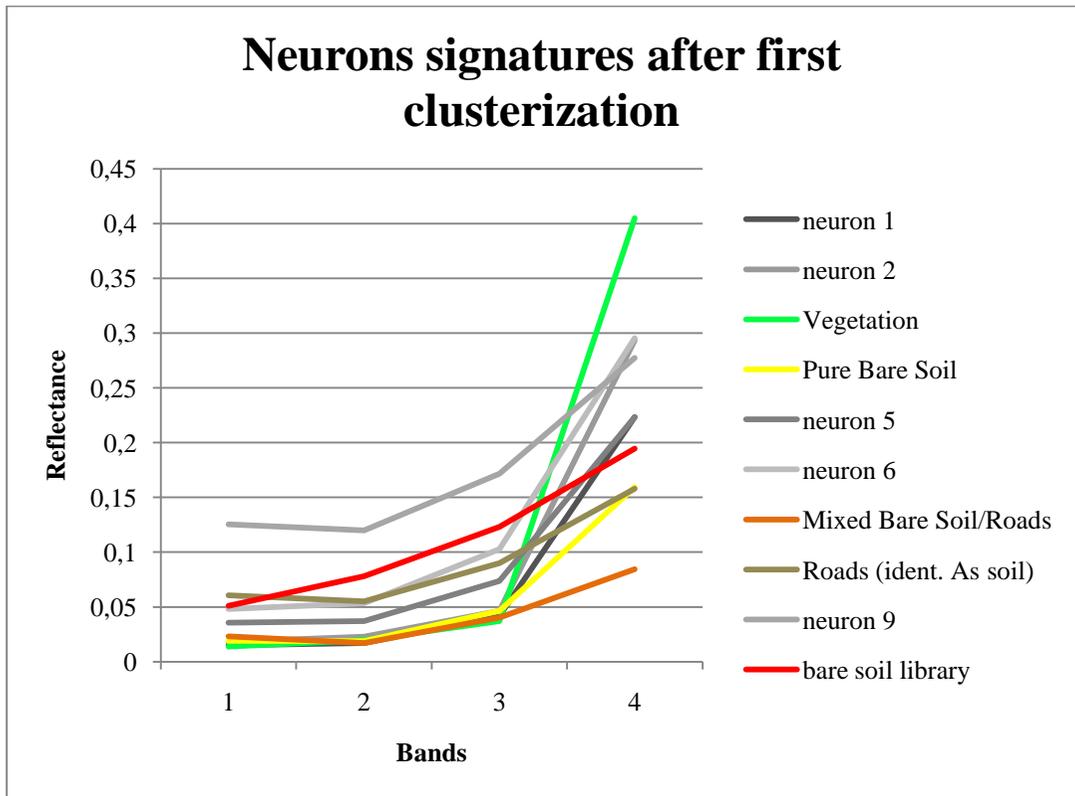
Fig. 6.9: spectral signatures of Kohonen neurons with a focus on the misclassified and mixed neurons (Tor Vergata image).
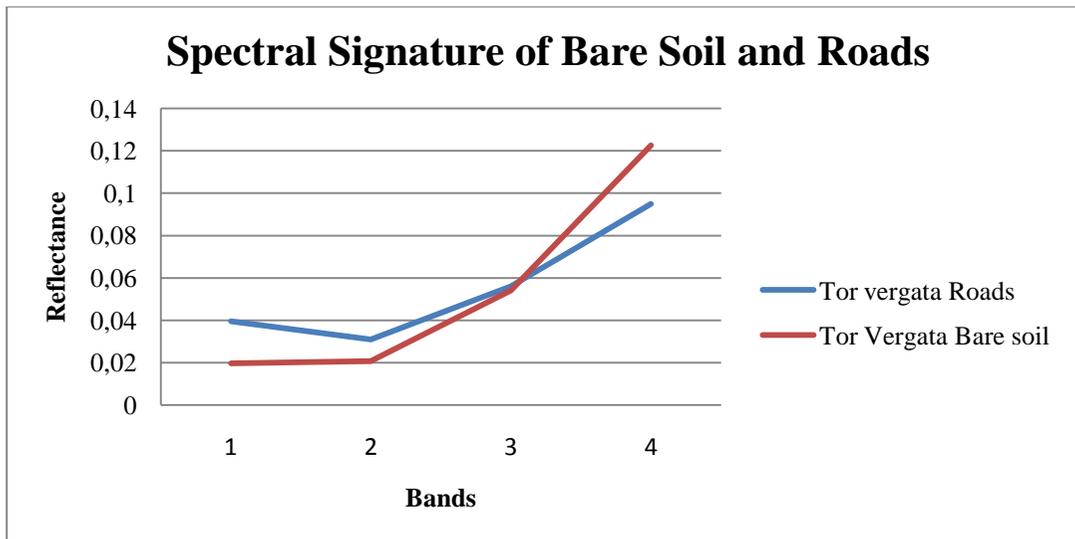
Fig. 6.10: detail of the spectral signature of roads and bare soil on Tor Vergata image.

Fig. 6.11: Tor Vergata false color composite 432 (above) and vegetation -in green-, bare soil –in yellow-, roads –in orange- and mixed class extraction –in brown-.

6.2 Man-made features

The procedure to spot the following classes very often considers not only spectral characteristics but, with regard to their internal material variability and the use of the same material for the different semantic class, required the use of textural and shape parameters.

6.2.1 Buildings class

From the physical composition point of view, several different materials can be used for the building roofs (they can be made of clay tiles, metal, concrete, glass, plastic). Therefore, different buildings types in the same image carry a different spectral signatures. As it has been shown in chapter five, homogeneity is a value which tends to be constant in buildings class in small windows size. It means that spectral signature and textural signature allows to extract buildings from images with a certain accuracy.

For the labelling procedure, the interactive procedure is preferred: the automatic algorithm cannot know how many different typologies are included in the image (i.e. industrial/concrete buildings and residential/red brick buildings). From the clusterised image, a interactive selection of the buildings clusters performed their assignation to the semantic building class.

Therefore, after the application of a mask to not include the already classified pixels (vegetation, water and bare soil), another Kohonen map has been trained with 3 x 3 neurons (even if increasing the number of neurons, the results were not substantially changed), adding the homogeneity parameters (from green band with a window 3 x 3 and 5 x 5).

The results of the cluster analysis are shown in figure 6.12: as it can be expected, the clusters signatures after the masking procedure present increasing value in the spectrum, due to the intrinsic characteristics of the remained land covers (buildings and roads).

From the interactive cluster labelling (due to the high difference between the reference signature and the one from cluster), a neuron with a homogeneity value of 0.65 has been chosen as ground truth for the MLP. The building class is also highlighted in figure 6.13.
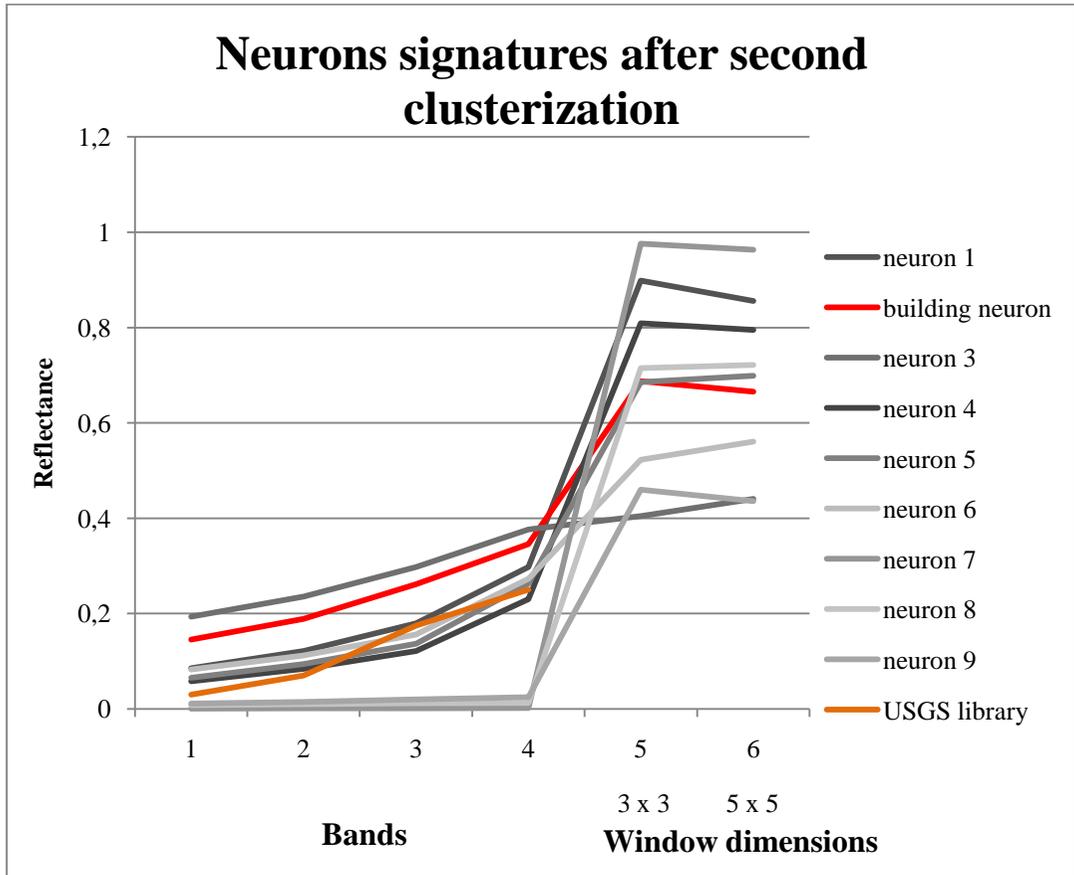
Fig. 6.12: detail of the spectral and textural signature of buildings in Nettuno image and its reference signature.
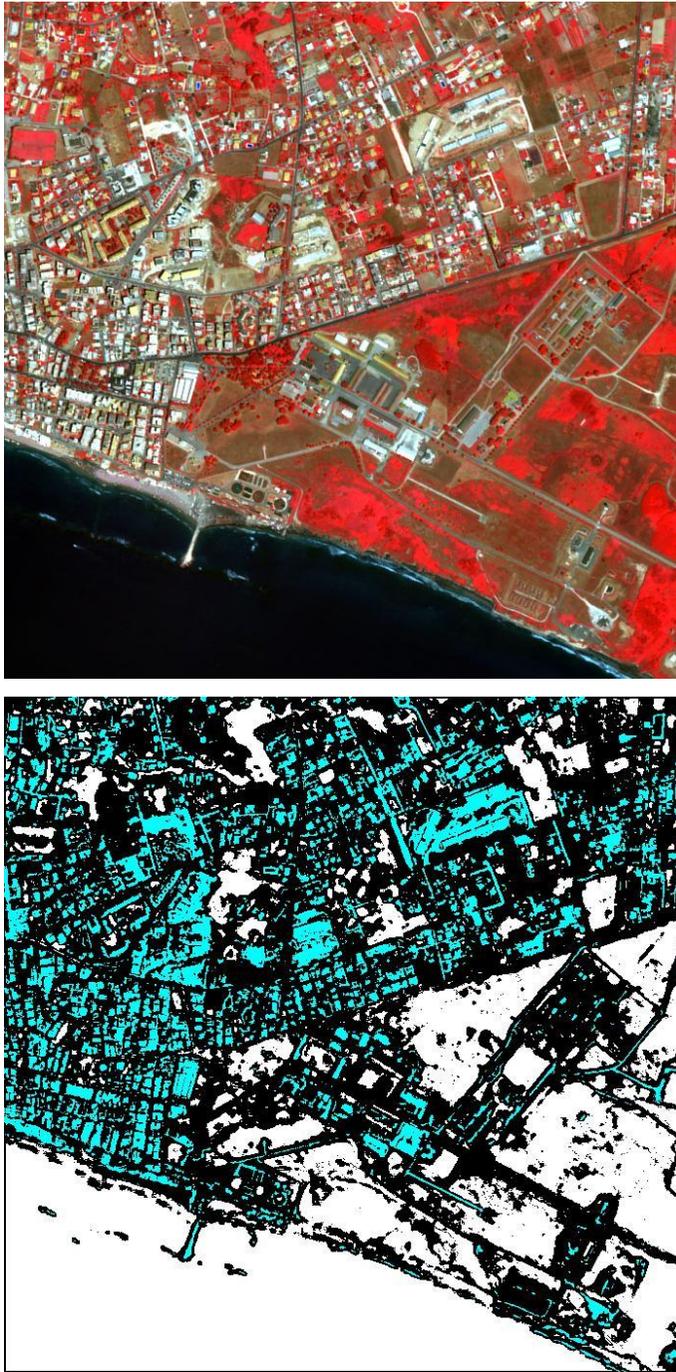
Fig. 6.13: Nettuno false color composite 432 (above) and buildings extraction -in cyan- and already classified/masked pixels –in white- (below).

In other cases, if the urban area is more complex, more cluster can be assigned to the building class (as it happened with Madrid and Denver images, where the residential areas have a different structure respects to the industrial site).

Here there is an example from the second Kohonen map for Denver which shows the two different types of buildings (Fig. 6.14 and 6.15).
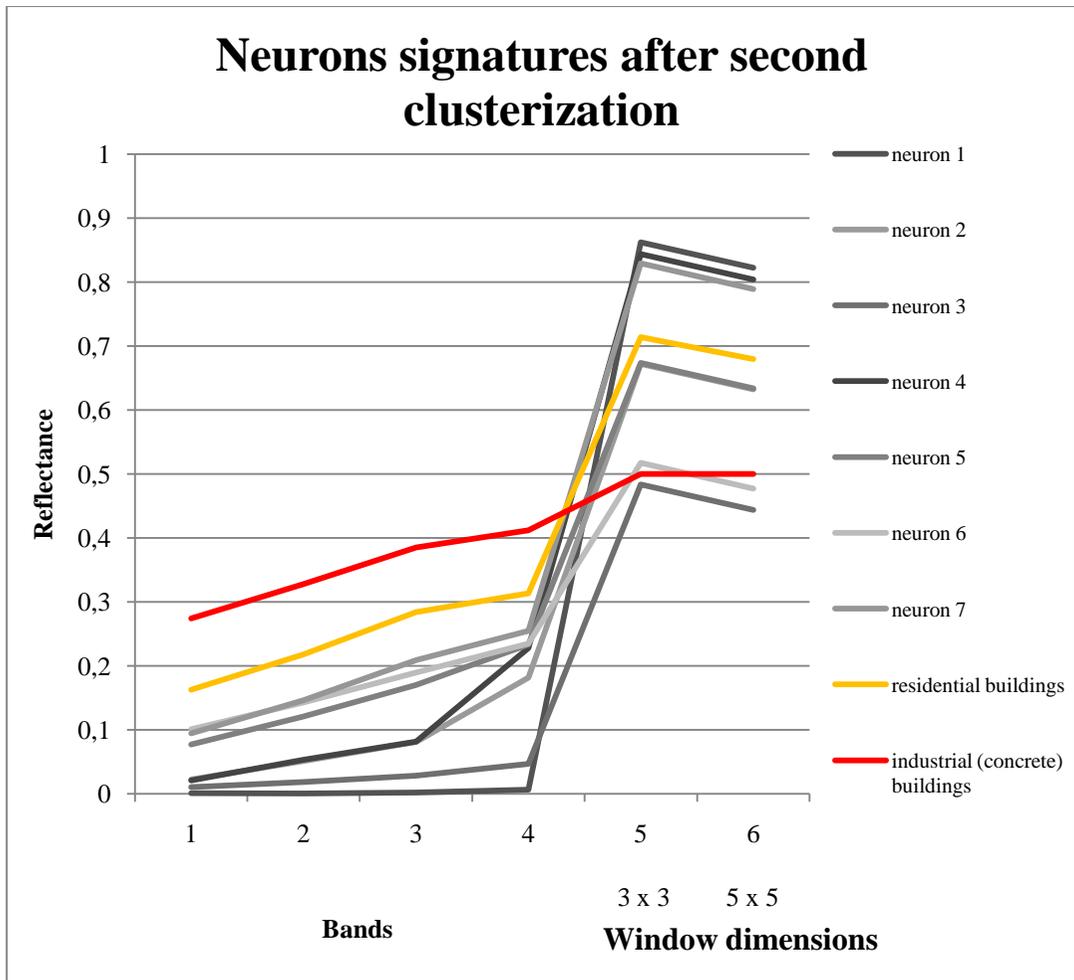
Fig. 6.14: detail of the spectral and textural signature of buildings in Denver image and its reference signature.
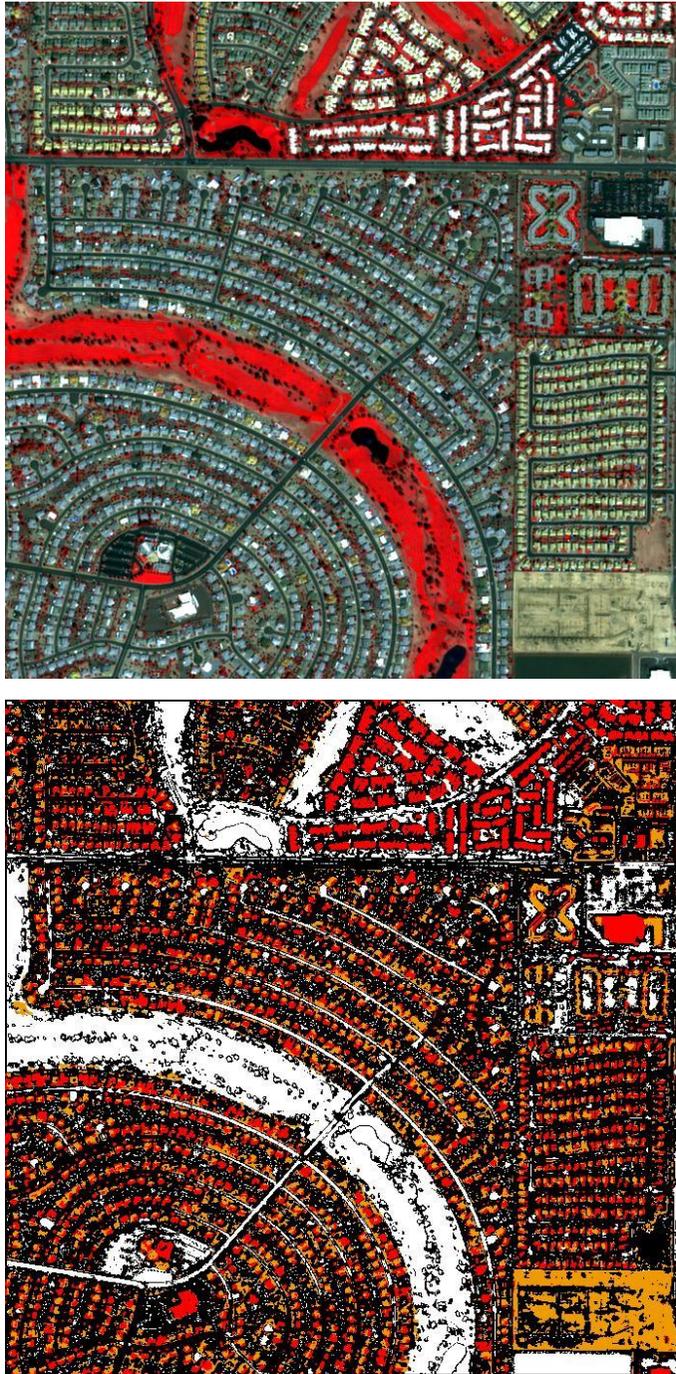
Fig. 6.15: Denver false color composite 432 (above) and residential buildings -in red-, industrial buildings –in orange- extraction and already classified/masked pixels –in white- (below).

6.2.2 Roads class

The last unsupervised step consist in the extraction of pixels related to roads.

As it was written in chapter five, roads can be spotted with the use of spectral characteristics and, it this method do not highlight the possible different type of roads, the use of shape parameters as elongation (after segmentation) can help to the discrimination).

Using only spectral features, the methodology is simpler but some misclassification errors with building roofs made of concrete and bare soil with similar spectral characteristics can get worse the accuracy of classification.

On the other hand, the segmentation process is more complicate and, as it was written in the segmentation description, there is no a miraculous method which segment all the imagery in a perfect way. What is here suggested, is to verify the classification accuracy after the first Kohonen SOM and if it is possible to assign a specific cluster to roads class. For instance, for the Denver image, the use of only spectral features allowed to obtained a final accuracy of roads extraction of 85 % (fig. 6.16).

In image 6.17, the two methods have been applied for the Nettuno image, showing that both can used to detect roads. It was astonishing that, with the only use of spectral signature, the proposed methodology was able to classify and label *automatically* (Fig. 6.18) four on five classes (except buildings, which the use of texture is often unavoidable).

In Tor Vergata image, the use of segmentation was necessary to extract roads, due to the different material which defined roads (Fig. 6.11).
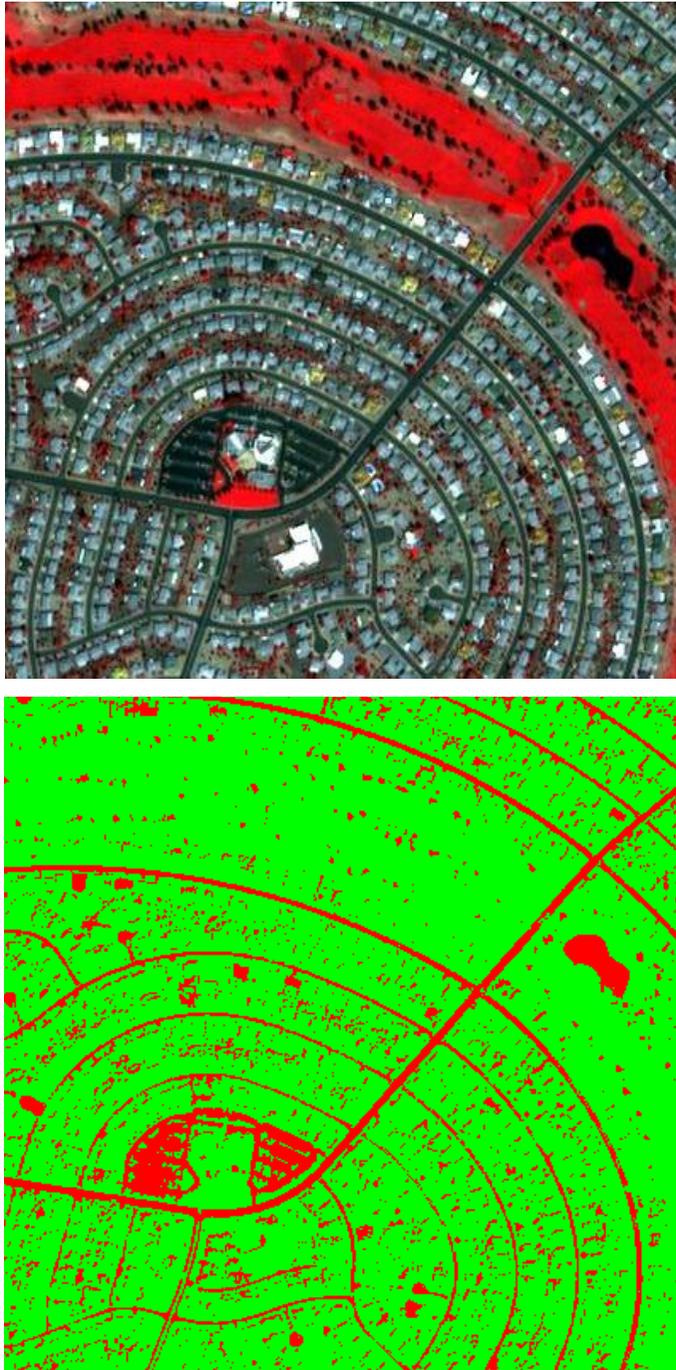
Fig. 6.16: Denver false color composite 432 (above) and roads extraction -in red- (below).
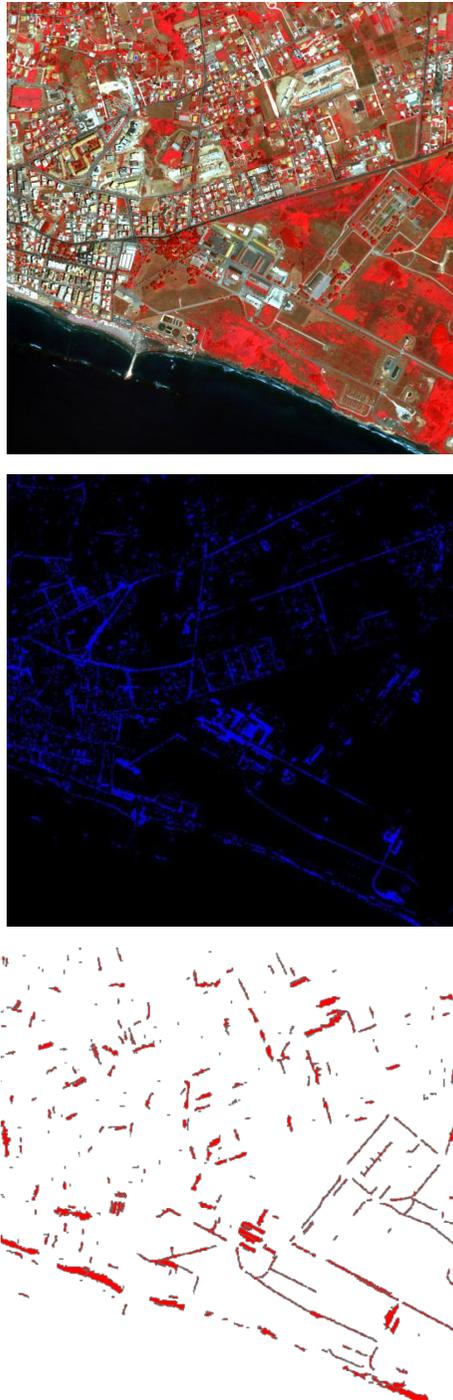
Fig. 6.17: Nettuno false color composite 432 (above), roads extraction -in blue- from spectral signature (middle) and roads extraction –in red- after segmentation through elongation (below).
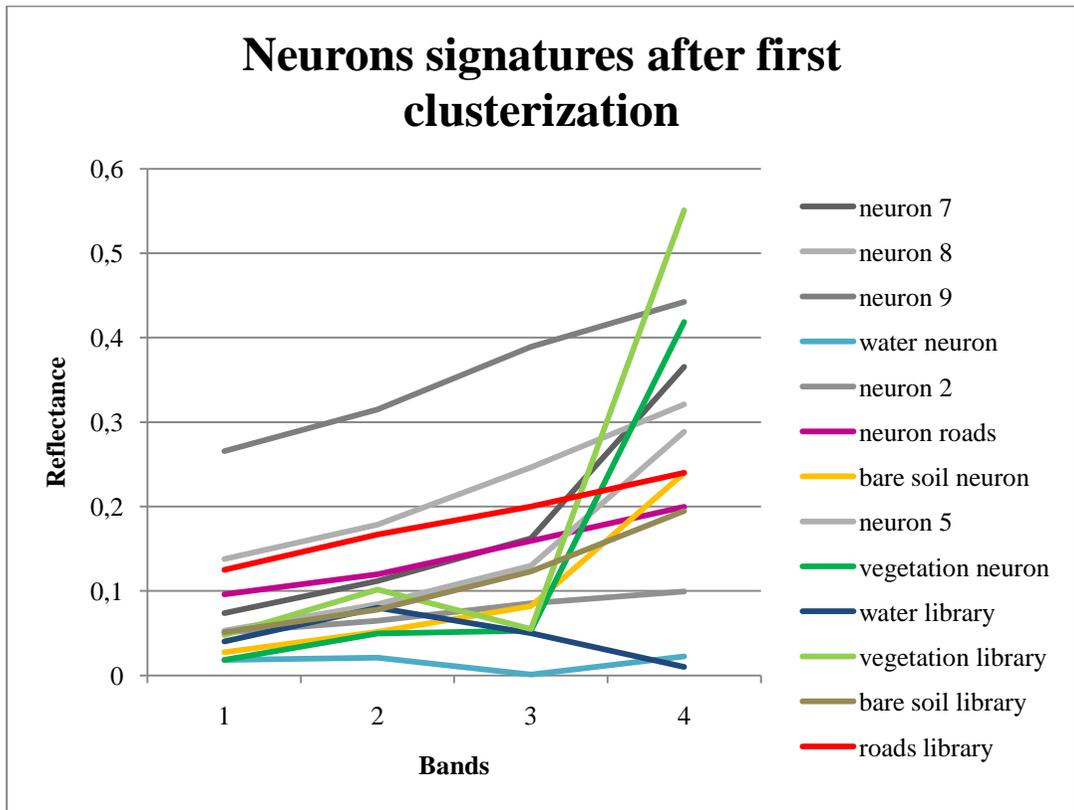
Fig. 6.18: spectral signatures of Kohonen neurons and reference signature for vegetation, water bare soil and roads (Nettuno image).

For Tor Vergata image, two different type of roads (motorway and local roads) and the misclassification with bare soil, make necessary the use of spatial information (elongation) to detect which pixels corresponds to roads (Fig. 6.19).

In particular, the segmentation contributed to extract the contribution of motorway, which joined to the roads class extract with signature, allowed to define the ROI for roads.
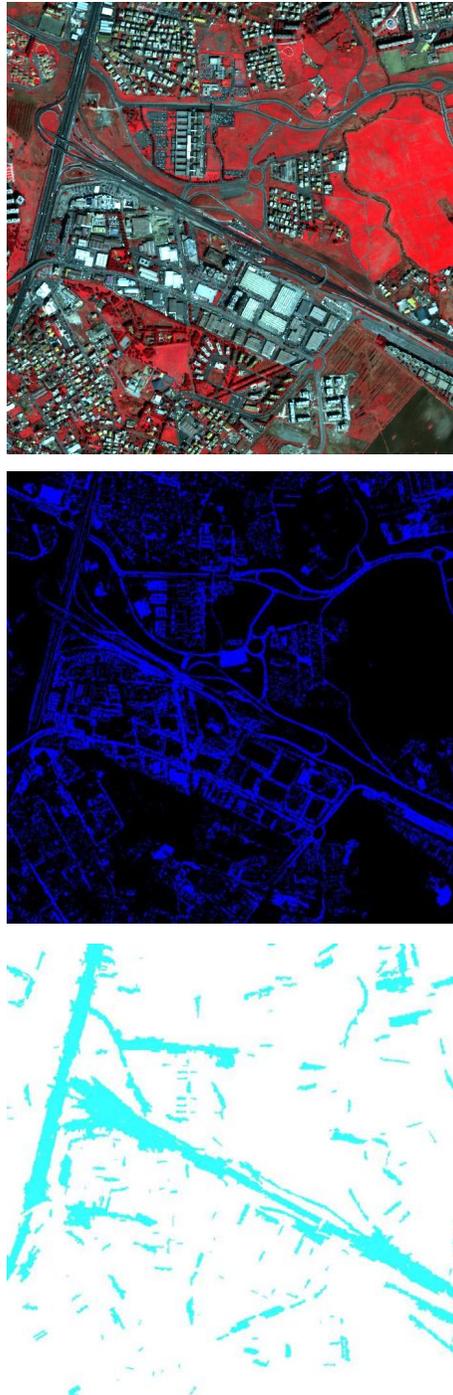
Fig. 6.19: Tor Vergata false color composite 432 (above), roads extraction -in blue- from spectral signature (middle) and roads extraction –in cyan- after segmentation through elongation (below).

6.3 MLP Neural Networks

Now that all the ROI for the predefined classes have been spotted, the pixels which belong to each class can be sent to the MLP for the final classification. This step allows to assign the uncertain pixels (those were not yet classified/clusterised) to the final land cover classes.

Some random pixels (their number depends of the dimension of the starting image but it generally corresponds to the 10 % of the entire number of pixels) for each of the five cluster (vegetation, water, bare soil, buildings and roads) have been extracted to feed the supervised network.

A network topology with 6 inputs (defined by the four spectral classes and the two homogeneity parameters), two hidden layers with sixteen neurons each one and five outputs have been chosen (Fig. 6.20): from the output weights of this neural training, the images were been classified in five classes: vegetation, water, bare soil, buildings and roads.

Regarding the hyperspectral images, two different ways can be chosen:

- the use of the corresponding four bands;
- the use of the entire available spectrum (removing the noisy and corrupted band).

The first hypothesis allows to perform an automatic labeling for the predefined class in the presence of a single cluster for a semantic class, the second one needs a further process to select the valuable bands and the labeling has to be carry out in a interactive way for all the classes (unless a reference signature for each hyperspectral sensors has been inserted to perform this activity in a automatic manner). Besides, tests on the accuracy after the insertion of all the available bands did not show an effective improvement (Lazzarini and Del Frate), at least in this

level of class discrimination (which means that deeper analysis inside each class could be helped from an higher number of bands).
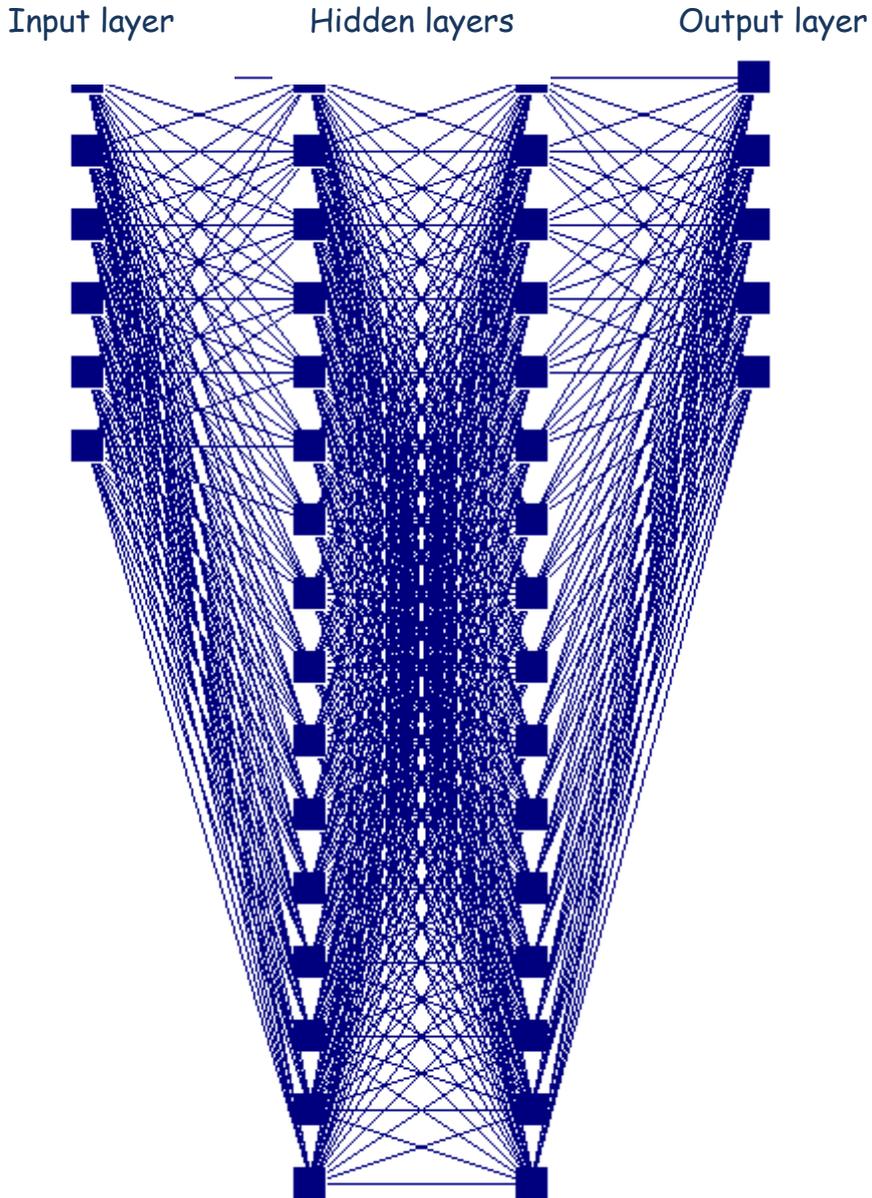
Fig. 6.20: Multi Layer Perceptron Neural Network for supervised classification with six inputs, two hidden layers with sixteen neurons each one and five outputs

6.4 Final accuracy of the tested images

Following, the final results of the classification processes are presented with a comparison with a standard unsupervised method as K-means. The final accuracy have been obtained comparing the classification output with ground truth pixels, obtaining at least 88 % in all the images, showing a increment respect to the K-means methods.

As general considerations, classes like vegetation and water can be generally spotted and labeled in automatic way. Bare soil can be highlighted with Kohonen map: the labelling is often interactive, due to with its variability in composition. The addition of homogeneity allowed to distinguish buildings: this class can be made of different buildings clusters, therefore the labeling depends on the number of buildings typology. Roads can be spotted with spectral characteristics and, if this parameter is not enough, the use of an additional segmentation step could help in its retrieve.

6.4.1 Nettuno (I)

For this image, the process has been performed in a almost totally automatic way: it means that the clusters extraction and the labelling of vegetation, water, bare soil and asphalt have been done autonomously (6.21). The extraction of building have been done with a second Kohonen SOM adding the homogeneity parameters. After the unsupervised step, 59 % of the entire pixels have been defined as ROIs.

A comparison with five ROIs from the respective classes of 2000 pixels each one has been performed.

The final accuracy, as shown in table 6.1a, was around 90.8 % with an increment of 13 % respect to the K-means method (table 6.1b): vegetation and water have been perfectly detected, the accuracy decreased with some misclassification of roads and bare soil classes, and the detection of some roads made of gravel (confused with buildings material).

Overall accuracy 90.80 % (SOM+MLP)					
	vegetation	**bare soil**	**buildings**	**roads**	**water**	**Total**
vegetation	100	0	0	0	0	20
bare soil	0	78.5	0	0	0	15.7
buildings	0	9.5	97.5	22	0	25.8
roads	0	12	2.5	78	0	18.5
water	0	0	0	0	100	20
Total	100	100	100	100	100	100

Table 6.1a: Nettuno accuracy matrix with the proposed method

Overall accuracy 77.5 % (k-means)					
	vegetation	**bare soil**	**buildings**	**roads**	**water**	**Total**
vegetation	100.00	0	0	0	0	20.00
bare soil	0	70.50	2.50	37.00	0	22.00
buildings	0	0	60.00	6.00	0	13.20
roads	0	29.50	37.50	57.00	0	24.80
water	0	0	0	0	100.00	20.00
Total	100	100	100	100	100	100

Table 6.1b: Nettuno accuracy matrix with K-means method
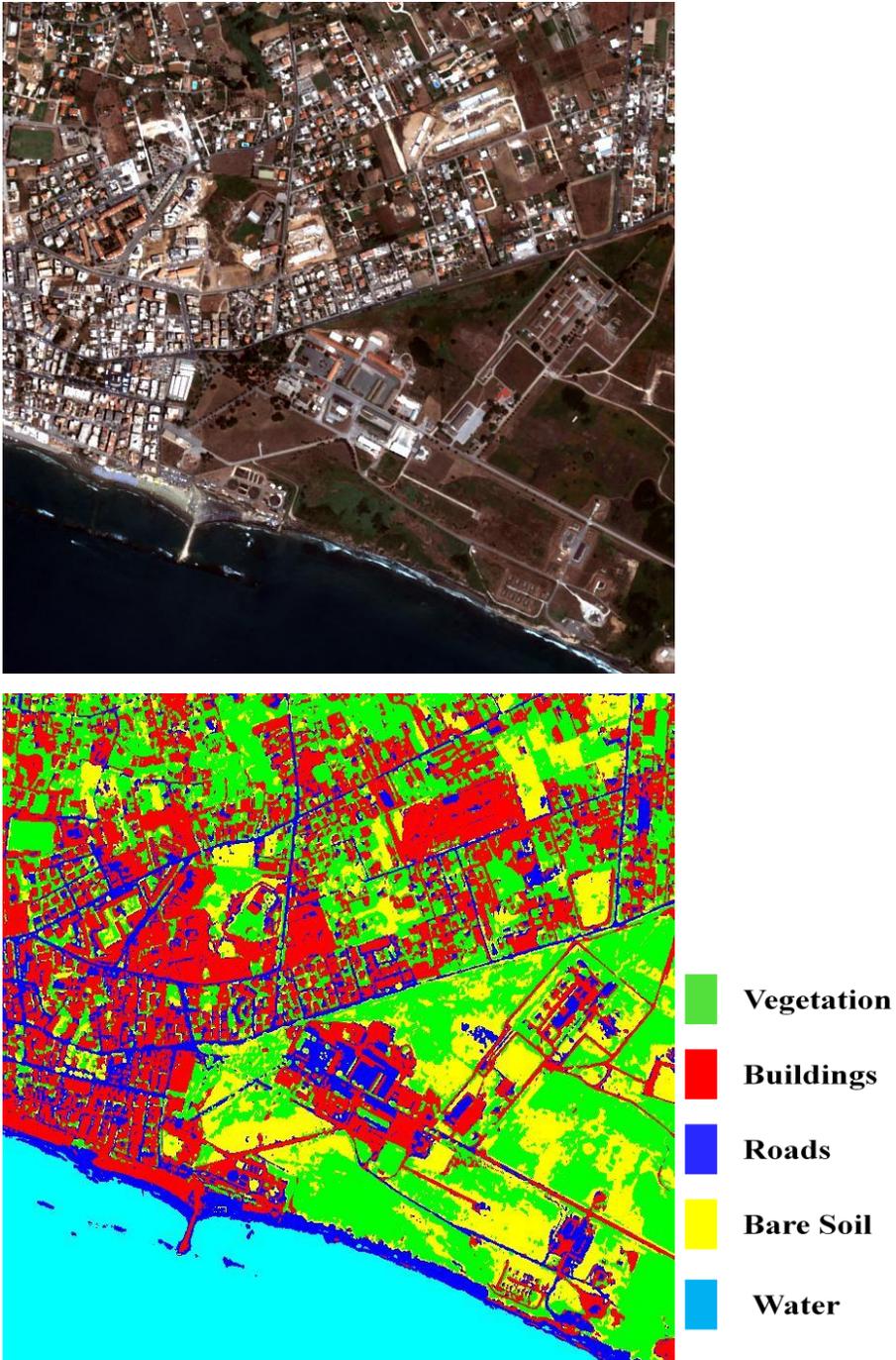
Fig. 6.21: Nettuno false color composite 432 (above) and final classified map

6.4.2 Tor Vergata (I)

This image is one of the most complicated from its structure and the lack of an urbanization plan. Vegetation have been spotted and labeled automatically; bare soil has an intrinsic variability and, in some pixels, a similar response to roads (as previously discussed); two types of roads have been recognized, one with spectral signature, the other one with the help of elongation parameter. The extraction of building have been done with a second Kohonen SOM adding the homogeneity parameter. After the unsupervised step, 54 % of the entire pixels have been defined as ROIs.

A comparison with four ROIs from the respective classes of 2000 pixels each one has been performed.

The final accuracy can be considered satisfactory with the proposed method (table 6.2a) with a value of 88.125 %, conversely a very poor accuracy has been obtained with the standard method: the highest accuracy have been obtained with buildings, it tends than to decreases with the other classes but the values were over 84 % in all the classes.

Overall accuracy 88.125 % (SOM+MLP)					
	vegetation	**bare soil**	**buildings**	**roads**	**Total**
vegetation	84	0	4	0	22
bare soil	16	88	0.5	18.8	28.88
buildings	0	5	95	3.5	25.88
roads	0	11	0.5	85.5	23.25
Total	100	100	100	100	100

Table 6.2a: Tor Vergata accuracy matrix with the proposed method

Overall accuracy 50.225% (K-means)					
	vegetation	**bare soil**	**buildings**	**roads**	**Total**
vegetation	40.40	0	64.40	0.10	26.23
bare soil	10.50	66.50	1.30	17.90	24.05
buildings	34.70	6.20	16.40	4.40	15.43
roads	14.40	27.30	17.90	77.60	34.30
Total	100	100	100	100	100

Table 6.2b: Tor Vergata accuracy matrix with k-means
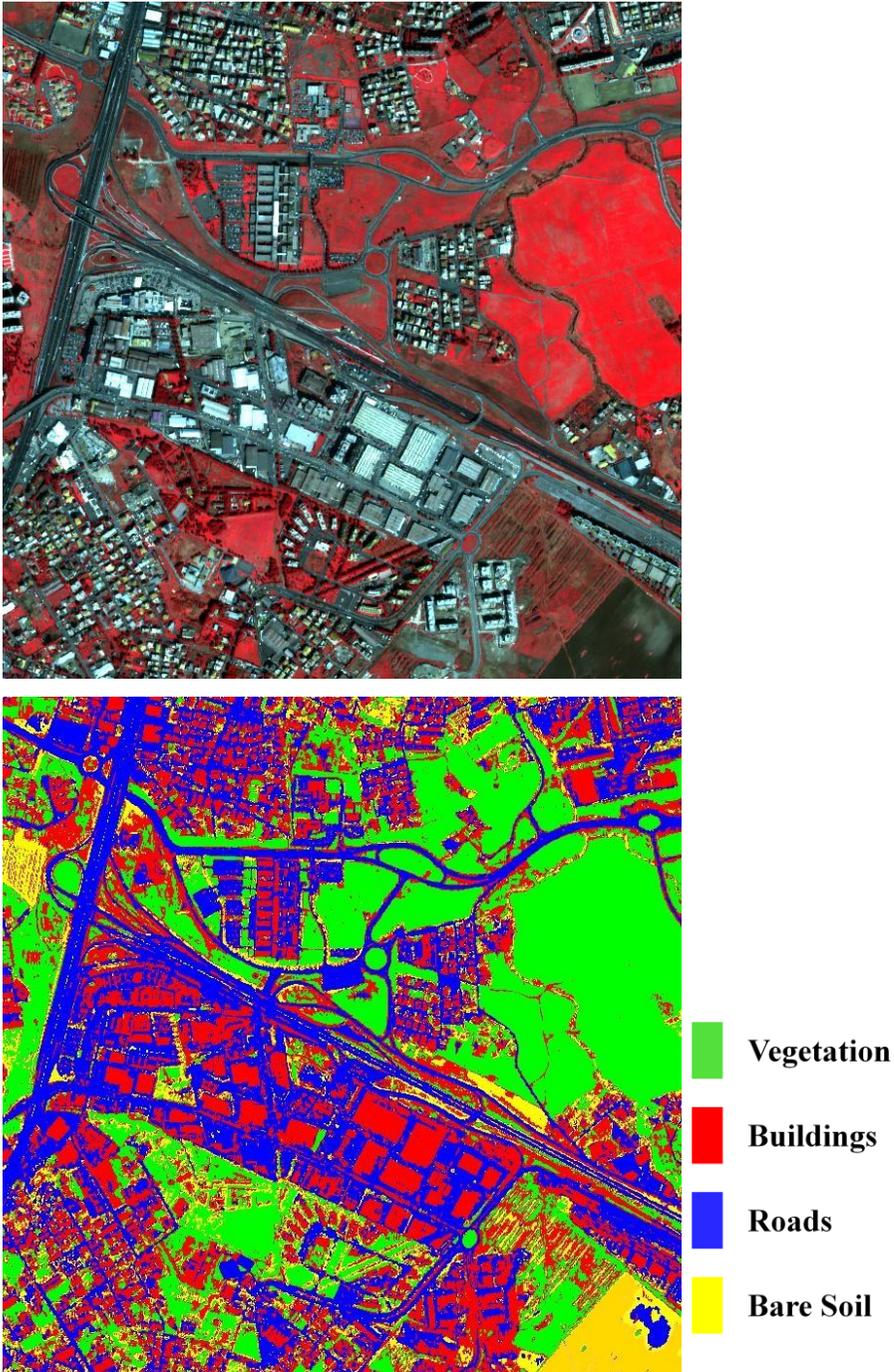
Fig. 6.22: Tor Vergata false color composite 432 (above) and final classified map

6.4.3 Denver (US)

The final accuracy of Denver was the higher of the entire tested dataset: the homogeneous urban structure allowed to obtained the higher value in buildings recognition and road extraction (Fig. 6.23). Water and vegetation have labeled automatically, roads and buildings (two clusters) interactively.

After the unsupervised step, 65 % of the entire pixels have been defined as ROIs.

A comparison with four ROIs from the respective classes of 2000 pixels for vegetation, buildings and roads and 500 pixels for water (due to its limited presence in the image) has been performed.

As general observation, all the classes have an accuracy higher of 90 % (Table 6.3a). With k-means, the accuracy value decreased drastically due to the fact that water class has not be spotted at all but it has been confused with roads (table 6.3b).

With the proposed method, bare soil class has not be detected considering the dense urban area which not contains this kind of cover. In the lower right part of the image there is a constructive site: even its conceptual classification in bare soil or buildings is difficult to determine. Vegetation has some misclassification errors due to its shadows: in some cases, the spectral responses of tree shadowed has been confused with asphalt response and water response.

Overall accuracy 95.55 %				
	water	**vegetation**	**buildings**	**roads**	**Total**
water	91.4	0	0	0	22.85
vegetation	0	93	0	0	23.25
buildings	0	0	99.4	1.6	25.25
roads	8.6	7	0.6	98.4	28.65
Total	100	100	100	100	100

Table 6.3a: Denver accuracy matrix with the proposed method

Overall accuracy 62.05 %				
	water	**vegetation**	**buildings**	**roads**	**Total**
water	0	0	43.00	0	10.75
vegetation	0	94.60	0	3	23.65
buildings	0	0.80	56.60	0	15.10
roads	100	4.60	0.4	97	50.50
Total	100	100	100	100	100

Table 6.3b: Denver accuracy matrix with k-means
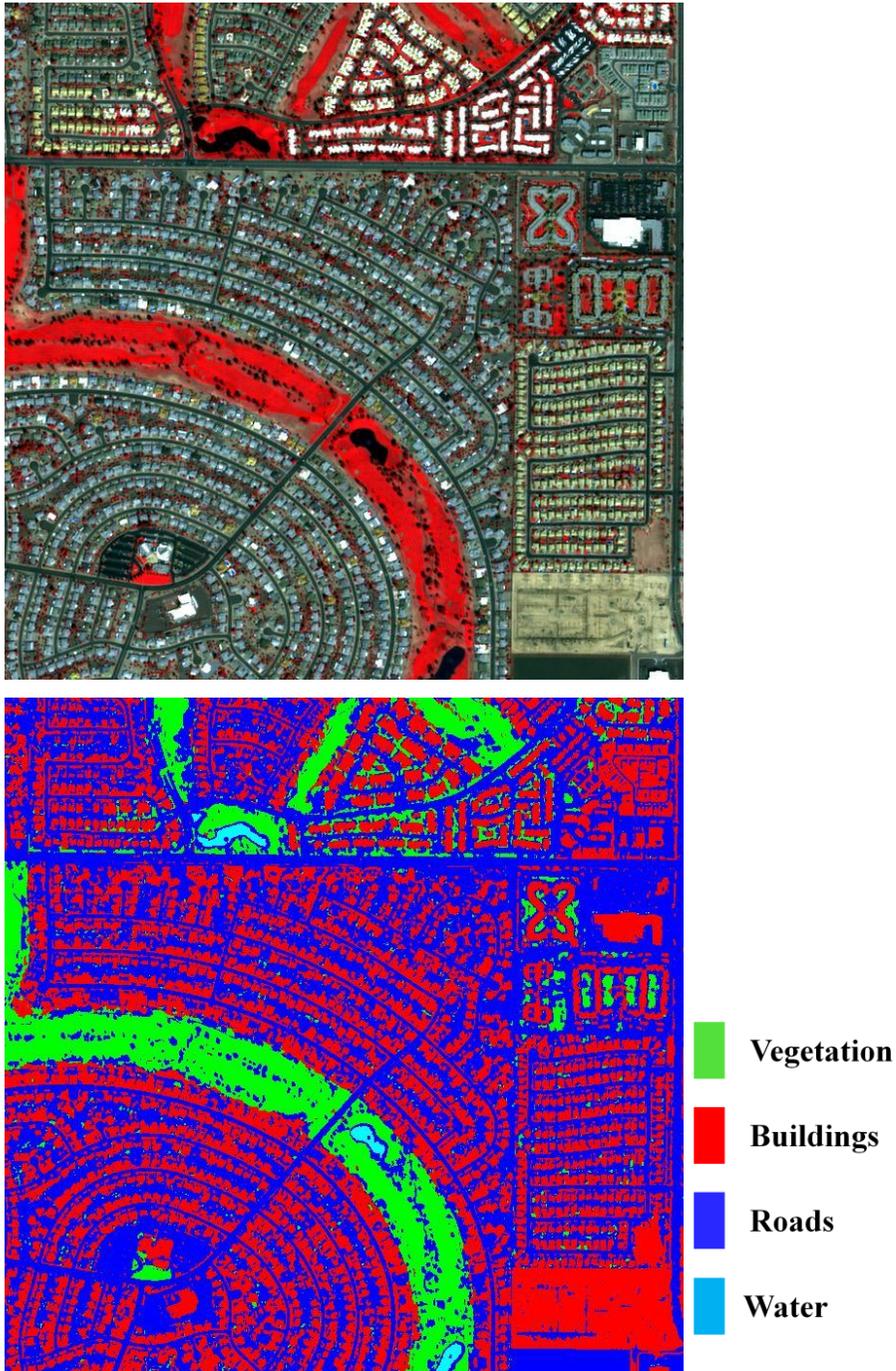
Fig. 6.23: Denver false color composite 432 (above) and final classified map

6.4.4 Bari (I)

For Bari image, vegetation has been extracted and labeled automatically. Water have been detected automatically and labeled interactively for the presence of different clusters belonging to the same class. Roads have been highlighted with the only use of spectral features, due to the presence of very narrow streets, most of the time with the presence of vegetation, not allowing the use of segmentation. A single cluster of buildings have been extracted with the Kohonen second SOM and the use of homogeneity parameter (6.24).

After the unsupervised step, 75 % of the entire pixels have been defined as ROIs.

A comparison with four ROIs from the respective classes of 2000 pixels each one has been performed.

Following the accuracy results, using the four bands corresponding to Quickbird (Table 6.4a) and fifty-eight bands on the entire VIS+NIR spectrum (Table 6.5b). Both of the final results were verified with the same ground truth data.

The accuracy with the two input data are very similar (with a slight higher value with the use of hyperspectral data). A significant difference can be viewed in the detection of buildings, higher with the use of multispectral input and roads, conversely higher with the use of hyperspectral input.

Also with this image, the results from a simple K-means classifier have been significantly improved (Table 6.4b and 6.5b).

Overall accuracy 92.15 % (SOM + MLP)				
	vegetation	**water**	**buildings**	**roads**	**Total**
vegetation	97.20	0	0	0	24.30
water	0	99.20	0	0	24.80
buildings	0	0.20	82.60	10.40	23.30
roads	2.80	0.60	17.40	89.60	27.60
Total	100	100	100	100	100

Table 6.4a: Bari accuracy matrix obtained with the RGB+NIR bands (SOM + MLP)

Overall Accuracy 70.35 % (k-means)				
	vegetation	**water**	**buildings**	**roads**	**total**
vegetation	99.40	0.30	27.00	0.90	49.16
water	0	90.12	0.10	70.00	22.81
building	0	0	72.40	9.70	20.51
roads	0.60	9.58	0.50	19.40	7.52
Total	100	100	100	100	100

Table 6.4b: Bari accuracy matrix obtained with k-means

Overall Accuracy 93 % (SOM + MLP)					
	vegetation	water	buildings	roads	total
vegetation	98.20	0.00	0.00	0.00	24.55
water	0.00	99.40	0.00	0.00	24.85
building	1.20	0.00	74.80	0.40	19.10
roads	0.60	0.60	25.20	99.60	31.50
Total	100.00	100.00	100.00	100.00	100.00

Table 6.5a: Bari accuracy matrix obtained with the 58 available bands (MLP + SOM)

Overall accuracy 69.60 % (k-means)					
	vegetation	water	buildings	roads	Total
vegetation	93.01	0	26.9	73.80	24.91
water	0	96.71	0	2.80	48.41
buildings	0	0	72.10	6.90	19.74
roads	6.99	3.29	1	16.50	6.94
Total	100	100	100	100	100

Table 6.5b: Bari accuracy matrix obtained with the 58 available bands (k-means)
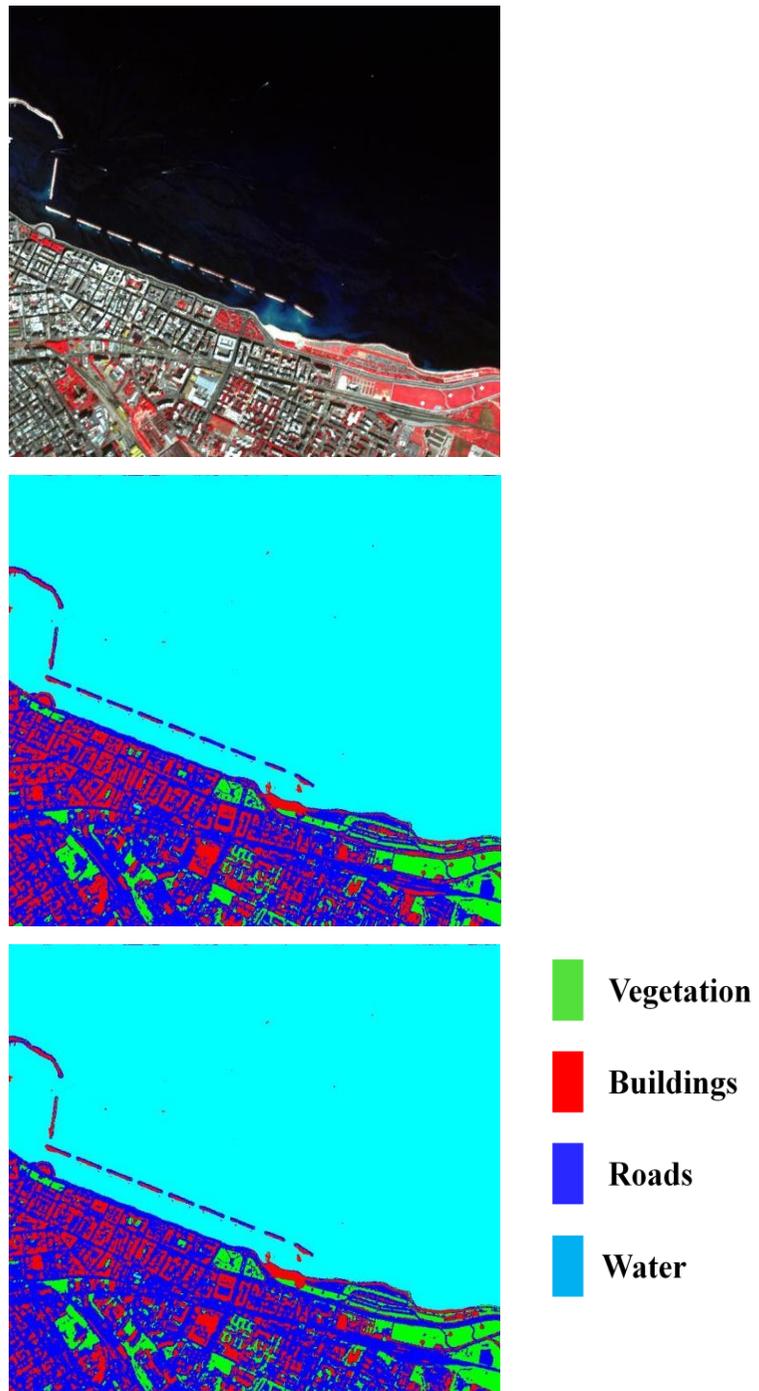
Fig. 6.24: Bari false color composite 20-12-7 (above) and final classified map with the use of the four RGB+NIR band and the entire available spectrum

6.4.4 Madrid (E)

The AHS image from Madrid has been classified using the analogous four bands from Quickbird (14-8-5-2) in four classes (6.25). The methodology considers the extraction and labeling of vegetation automatically, the extraction of bare soil and roads with the use of spectral information (interactive labelling) and the extraction of buildings with the second Kohonen map and homogeneity (two clusters of residential/red brick and industrial/concrete). After the unsupervised step, 64 % of the entire pixels have been defined as ROIs.

A comparison with four ROIs from the respective classes of 2000 pixels each one has been performed. The accuracy value was above 88 % (table 6.6a) with some misclassification errors belonging to buildings and bare soil.

Overall accuracy 89.62% (MLP + SOM)				
	vegetation	**bare soil**	**buildings**	**roads**	**Total**
vegetation	99.22	0	0	0	24.80
bare soil	0.40	94.60	17.40	10.93	30.83
buildings	0.29	4.33	78.33	2.73	21.42
roads	0.09	1.07	4.27	86.33	22.93
Total	100	100	100	100	100

Table 6.6a: Madrid accuracy matrix with the proposed method

Overall accuracy 55.9 % (k-means)				
	vegetation	**bare soil**	**buildings**	**roads**	**Total**
vegetation	99.07	41.07	1.87	11.00	38.25
bare soil	0.13	0	34.00	0	8.50
buildings	0	0.20	39.40	3.87	10.90
roads	0.80	58.73	24.73	85.13	42.35
Total	100	100	100	100	100

Table 6.6b: Madrid accuracy matrix with k-means

Fig. 6.25: Madrid false color composite 14-8-5 (above) and final classified map

Chapter 7

Application of automatic classification in urban thermography

In this application, it has been designed a novel and automatic scheme for the estimation of the temperature associated to buildings: this approach relies on a cascade of neural networks (NN) and retrieval procedures and can represent an important technique to improve cities energy balance models. The entire process has been tested on an Airborne Hyperspectral Scanner (AHS) image acquired above Madrid in summer 2008 and the results assessed by comparison with the state of the art on LST retrieval [701].

7.1 Urban Heat Island and Land Surface Temperature

It has been written that one of most important application of automatic processing techniques is represented by the environmental parameters retrieval in Near Real Time (NRT): in the monitoring of urban environments, where the knowledge of changes in either natural resources or man-made structures is a valuable source of information for decision-making processes [702], it may be necessary to obtain data in the fastest possible way. Of particular interest in this context is the occurrence of Urban Heat Islands (UHI) phenomenon and the retrieval of Land Surface Temperature.

Regarding UHIs, high densely built-up areas trap the heat causing city centers be up to 10º C warmer than surrounding rural areas and this is often manifested particularly at night [703], [704]. For energy balance and urban planning aiming at reducing the number of casualties when temperatures soar, the UHI information can be also updated on monthly basis. However, the retrieval of temperature should be in NRT to support civil protection in monitoring health impact over the weakest population with a daily frequency[16]. Remotely sensed observations of UHI are possible using satellite and aircraft platforms and they could open new avenues for their analysis [705].

In 2008, ESA has launched a set of activities to aid decision and policy makers in mitigating the effects of UHIs through appropriate alert systems and, in terms of reducing risk, through improved urban planning. Within this framework, ESA is funding the Urban Heat Island and Urban Thermography (UHI) project under the DUE - Data User Element - program. The main purpose of this project is developing, validating and demonstrating a set of services that assimilate satellite remote sensing observations - mainly Thermal InfraRed (TIR) channels - with urban weather stations measurements (temperature, relative humidity and wind

---

[16] [16]http://dup.esrin.esa.int/Files/News/summary_DESIRE.pdf, *final draft of "DESIRE workshop"*

speed) into urban meteorological and climate modelling to help decision and policy makers in better preventing the impact of Urban Heat Island (UHI) during heat waves, through appropriate alert systems, and in reducing the risk, through dedicated urban land planning. A second objective of the project is to measure the energy balance of the cities for a better response to the energy efficiency policies. The third and last goal of the project is to study the optimum mission requirements for a dedicated TIR satellite sensor with higher spatial resolution and revisiting time for an adequate provision of Land Surface Temperature retrievals in the metropolitan areas of European cities.

The project, started on 1st November 2008, is analyzing the UHI trends over 10 European cities (Athens, Bari, Brussels, Budapest, Lisbon, London, Madrid, Paris, Seville, Thessaloniki) over the last 10 years, using a multi-sensor approach.

In Urban Thermography studies, Land Surface Temperature is one of the key parameters which can be retrieved from thermal infrared (TIR) remotely sensed data. Most of the energy detected by the sensor in this spectral region is directly emitted by the land surface: hence, except for solar irradiance components, most of the fluxes at the surface/atmosphere interface can only be parameterized by means of surface temperature.

Several techniques have been proposed to retrieved LST from thermal infrared data: these methods are generally classified in single-channel methods (which used only one thermal band [706], two-channel or split-window methods (which use a combination between two thermal bands [707]), two-angle methods (which use one thermal band and two view angles [708]), multispectral method (which use several thermal bands, like TES method [709] and hybrid methods (which use a combination of other methodologies [710]).

7.2 How classification techniques can help urban thermography

Image automatic classification represents an important issue in Land Surface Temperature and Surface Energy Budget (SEB) retrieval (Fig. 7.1). First of all, a specific class adds information to a temperature map, considering the surface where thermal differences are measured: thermal anomalies could be determined by different material with different emissivity properties (i. e,. metal and concrete).

Moreover, classification map contributes to improve the management of the urban heat island phenomenon through other surface parameters products (i.e. roughness map). [711] studied the UHI effect with formula 7.1 to estimate z_0:

$$z_0 = \exp\left\{\frac{1}{n}\left(\sum_{i=1}^{n}\ln z_{0i}\right)\right\}$$

Eq. 7.1

where z_{0i} is the roughness length for the corresponding surface coverage type (defined in a look-up table) and n is the number of pixels for a selected window Therefore, to apply this method, a land cover map is necessary.

Roughness map, as well as LST map, constitute some of the input in the city energy balance model: with other input parameters, they help to obtain energy efficiency maps, using a model of surface energy budget estimation [712].

In these energy budget models, classification maps bind the parameters of the Energy balance model for the specific reference surface (i.e., building heat transfer model parameters to be apply on building cover class).
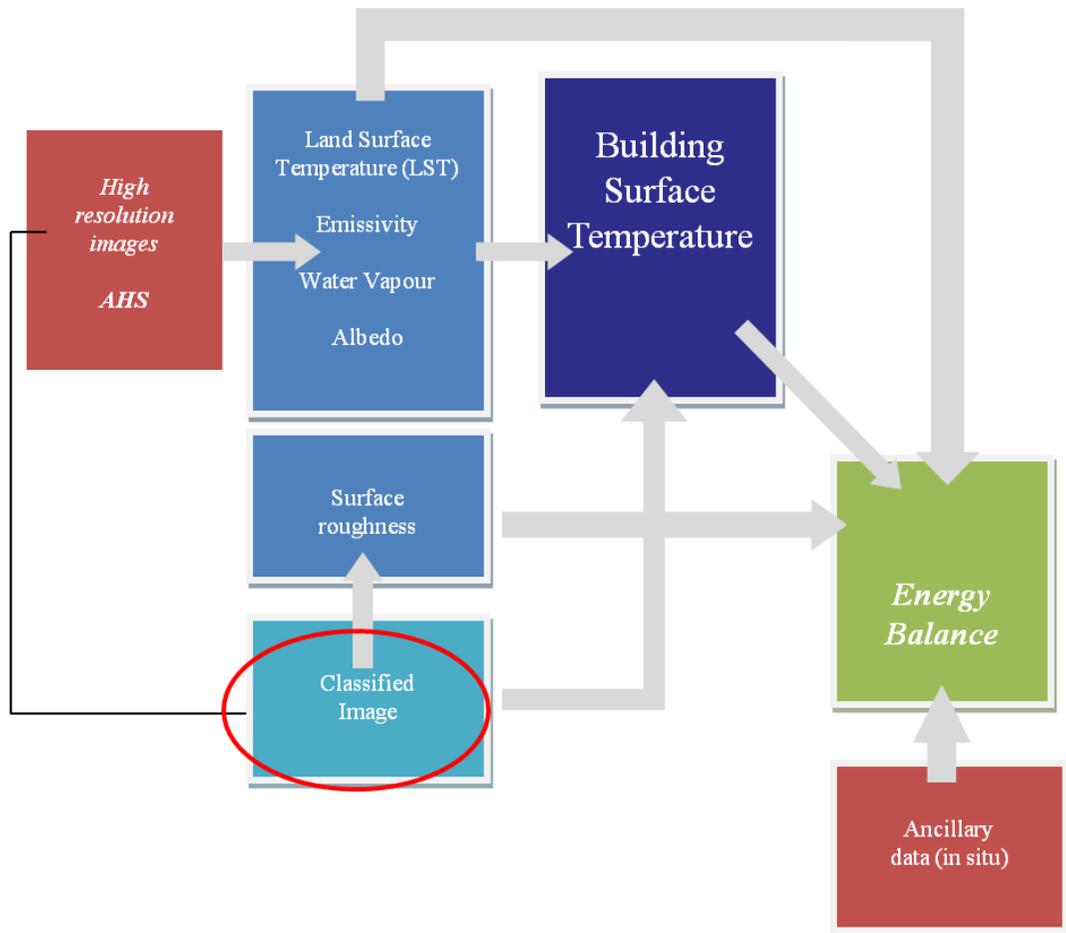
Figure 7.1: role of classification in urban thermography

7.3 Building extraction

Buildings have been extracted with the mixed approach described in chapters three and five: Kohonen's map has been used to clusterize the AHS image and to extract the building signatures considering both spectral and textural characteristics (CLUSTERING procedure). Classes "Buildings" has been recognized comparing their value to *standard building signature*s (LABELING procedure). The extracted building pixels have been used as training set of a supervised neural network performing the ultimate building detection. The standard building signature includes AHS bands 2 (blue, 0.4836 μm), 5 (green, 0.5717 μm), 8 (red, 0.6586 μm) and 14 (NIR, 0.8334 μm) and the Grey Level Co-Occurrence Matrix (GLCM) parameters homogeneity (also called the "Inverse Difference Moment") as its utilization in urban feature extraction has been already proven effective [713]. To exploit a multi-scale approach, which better deals with objects having different spatial coverage in an area, homogeneity features have been computed with two different window sizes of 3×3 and 5 x 5 only on AHS green band. Indeed this latter band gave appreciable visual results and it was preferred to avoid using a redundant number of features to feed the following supervised neural network. The dataset that constituted the input for the unsupervised learning was formed by 6 bands: blue, green, red, near infrared, homogeneity of band 2 (from a window 3 x 3) and homogeneity-bis of band 2 (from a window 5 x 5). The values of the input vector were normalized to the range of $(-1; +1)$: this step was necessary considering that different measure units (radiance and homogeneity) were used. The extracted building pixels have been then used as training set of a supervised neural network performing the ultimate building detection. The mixed approach has been made to improve the classification accuracy observed using only unsupervised methodologies.

The optimal NN architecture have been obtained with a map of 5 x 5 neurons, which seems the optimal architecture to detect buildings. It means that each input signal has been compared with all the 25 neurons/weights of the map (where the initial weights of these neurons were set with random values). The winning neuron has been chosen with a minimum distance algorithm based on the Euclidean distance between input and neuron. After 100 epochs, the unsupervised training was completed and the final weights of the 25 clusters were fixed: between these 25 clusters, two of them could be distinguished as buildings (Fig. 7.2).

Fig 7.2: building typology highlighted by the unsupervised neural networks. A (left) represents red brick roof building, B (right) represents industrial (concrete) building

6000 pixels from each clusters have been randomly chosen and they have been sent to the supervised neural networks. Once the training sets have been defined, they have been joint together as input to the supervised neural network. A very simple network topology has been chosen, with six inputs, a single hidden layer with eight neurons and two outputs: after the training of the new network, the entire image has been classified in two classes: buildings and no buildings (fig. 7.3 and fig. 7.4). To verify the accuracy of the building extraction method, 2500 pixels for each class have been chosen to represent the ground truth: the overall accuracy of the classification reached a satisfactory level, with a value of 83.7 % (Table 7.1). In

table 7.2 the improvement with the used method, adding different texture windows and the supervised step, is shown. The method picked up some pixels which are not covered by buildings (in figure 7.4, right) it could be observed some misclassification error between roads and bare soil).
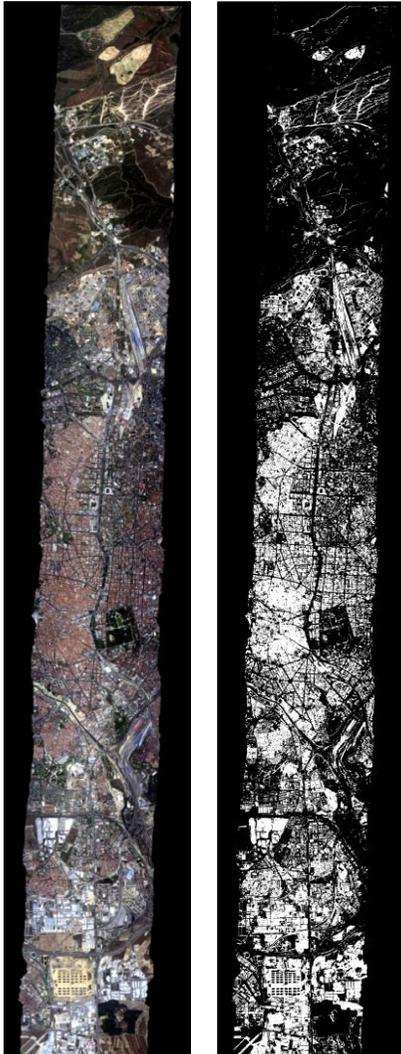
Figure 7.3: AHS swath above Madrid. RGB 8-5-4 real colour composite (left), extracted building mask (center) and Building Temperature map (right, legend figure 3).

Classification Accuracy 83.7 %			
Class	Buildings	No buildings	Total
Buildings	**86.96**	19.48	53.22
No buildings	13.04	**80.52**	46.78
	100.00	100.00	100.00

Table 7.1: classification accuracy in building extraction

Classification Method	Accuracy in building detection
4 Spectral bands -Unsupervised-	60. 2 %
4 Spectral bands + Homogenity 3 x 3 -Unsupervised-	70.65 %
4 Spectral bands, Homogenity 3 x 3 and 5 x 5 -Unsupervised-	74. 825 %
4 Spectral bands, Homogenity 3 x 3 and 5 x 5 -Unsupervised + Supervised-	**83.7 %**

Table 7.2: improvment in classification accuracy with the proposed method
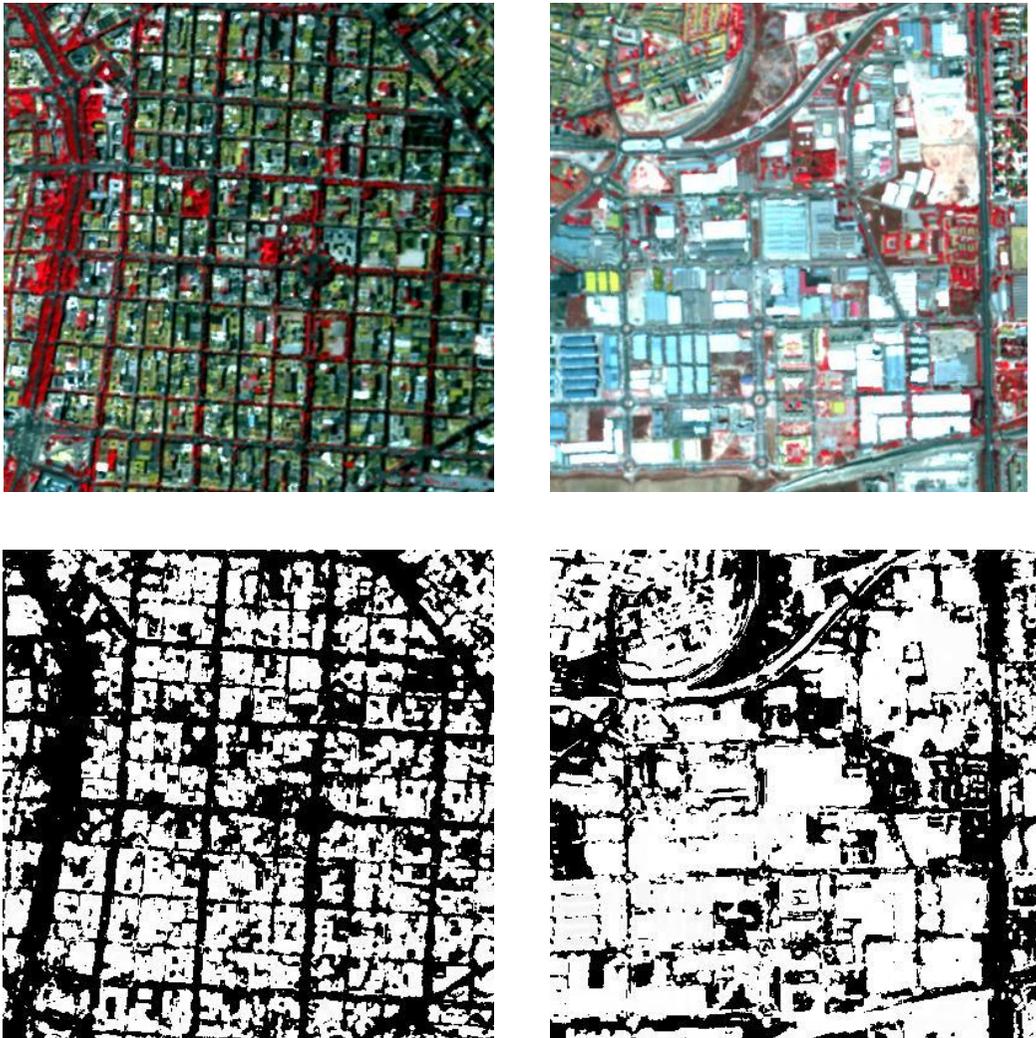
Fig 7.4: details of RGB 14-8-5 standard false colour composite (up) and the final classification (below), where buildings are highlighted (in white).

7.4 LST retrieval

In this work, LST has been obtained considering a split window (SW) technique: the basis of the technique is that the radiance attenuation for atmospheric absorption is proportional to the radiance difference of simultaneous measurements at two different wavelengths, each subject to different amounts of atmospheric absorption. The SW technique uses two thermal bands typically located in the atmospheric window between 10 and 12 μm. In [714] a general formula to retrieve LST for AHS high flights has been defined and coefficients have been readapted in [715]. It is based on a combination between AHS bands 75 (10.07 μm) and 79 (12.35 μm):

$$T_s = T_{75} + 0.723\ (T_{75\text{-}}\ T_{79}) + 0.04275\ (T_{75\text{-}}\ T_{79})^2 \text{-} 0.08463 +$$

$$(45.49\text{-}5.17W)\ (1\text{-}\ \varepsilon) + (\text{-}60.81 + 16.93W)\ \Delta\varepsilon \qquad\qquad \text{Eq. 7.2}$$

where T_{75} and T_{79} are the at-sensor brightness temperatures at the thermal bands (in Kelvin), ε is the mean emissivity of two bands (ε_{75} and ε_{79}), $\Delta\varepsilon$ is the emissivity difference ε_{75}-ε_{79} and W is the total atmospheric water vapour content (in g cm^{-2}). From a research in literature, it seemed that nowadays there are no developed methods to calculate LST from image data as AHS without specific atmospheric correction and the use of ancillary data (i.e. radiosonde). On the other hand, the entire technique presented in this paper aims at increasing the level of automatism in extracting building temperature considering only the processing of the image data. In fact, one of the purposes of the study is to investigate on the price to be paid in terms of error on the sought final quantities for increasing the level of

automatism. In next paragraphs we explain how we tried to estimate both emissivity and water vapour content directly from the used AHS image.

However, first of all, georeferencing and a conversion from radiance L_i (m$Wm^{-2}sr^{-1}\mu$) to temperature T_i (K) for thermal bands has been applied with Planck's formula considering the equation:

$$T_i = \frac{1}{(a+b*\ln(L_i))}$$ Eq. 7.3

where a and b are optimized coefficients (Table 7.3) provided by INTA for AHS flights.

a coefficients	b coefficients	Error
8,62823E-03	-5,78661E-04	band 71 max. error [K] = 0.03
8,94013E-03	-6,10497E-04	band 72 max. error [K] = 0.03
9,24115E-03	-6,42080E-04	band 73 max. error [K] = 0.04
9,52452E-03	-6,72616E-04	band 74 max. error [K] = 0.05
9,79788E-03	-7,02837E-04	band 75 max. error [K] = 0.05
1,00895E-02	-7,35967E-04	band 76 max. error [K] = 0.06
1,04274E-02	-7,75533E-04	band 77 max. error [K] = 0.08
1,07362E-02	-8,12848E-04	band 78 max. error [K] = 0.09
1,10336E-02	-8,49946E-04	band 79 max. error [K] = 0.10
1,13322E-02	-8,88480E-04	band 80 max. error [K] = 0.12

Table 7.3: applied coefficients to convert radiance to temperature (INTA)

7.4.1 Emissivity

Land Surface emissivity (ε) is a proportionality factor that scales blackbody radiance (Planck's law) to predict emitted radiance, and it is the efficiency of transmitting thermal energy across the surface into the atmosphere. An automatic way to calculate this parameter is with NDVI[THM] *(NDVI thresholds method)* [716], whose assessment by comparison with other more complex methods has been already carried out in [715]. It has to be observed that these latter complex methods rely on TES (*Temperature Emissivity Separation*) and TISI (*Temperature Independent Spectral Indices*) algorithms which, assuming an atmospheric correction of the data by means of radiative transfer models such as MODTRAN [17], perform better results but do not permit a high level of automation.

NDVI[THM] uses certain NDVI values (thresholds) to distinguish between soil pixels (NDVI < NDVI$_s$) and pixels of full vegetation (NDVI > NDVI$_v$). For those pixels composed of soil and vegetation (mixed pixels, NDVI$_s$ ≤ NDVI ≤ NDVI$_v$). The general formula to calculate NDVI is:

$$NDVI = \frac{(rNIR - rRED)}{(rNIR + rRED)} \qquad \text{Eq. 7.4}$$

where rNIR is the radiance of band 14, centered at 0.862 and rRED is the radiance of band 8, centered at 0.689 μm.

The NDVI[THM] considers the following simplified conditions:

$$\epsilon_\lambda = \begin{cases} \epsilon_{v\lambda} \\ \epsilon_{s\lambda} + (\epsilon_{v\lambda} - \epsilon_{s\lambda})P_V \\ a + b(\rho) \end{cases} \quad \text{if} \quad NDVI = \begin{cases} NDVI > NDVI_v \\ NDVI_s \leq NDVI \leq NDVI_v \\ NDVI < NDVI_s \end{cases} \quad \text{Eq. 7.5}$$

where ε_v and ε_s are, respectively, the soil and vegetation emissivities, ρ is the reflectivity of red band and PV is the proportion of vegetation (also referred to as fractional vegetation cover, FVC).

$$PV = \left[\frac{NDVI - NDVI_S}{NDVI_v - NDVI_S}\right]^2 \qquad\qquad \text{Eq. 7.6}$$

Typical constant values of $\varepsilon_v = 0.99$, $\varepsilon_s = 0.97$, $NDVI_v = 0.5$ and $NDVI_s = 0.2$ [27vecchia ref] have been proposed to apply the method in global conditions.

Therefore, there are three different possibilities:

1) $NDVI > 0.5$ $\qquad\qquad \varepsilon_{75} = \varepsilon_{79} = 0.99$ $\qquad\qquad$ Eq. 7.7

2) $0.2 < NDVI < 0.5$

$$\begin{bmatrix} \varepsilon_{71} \\ \varepsilon_{72} \\ \varepsilon_{73} \\ \varepsilon_{74} \\ \varepsilon_{75} \\ \varepsilon_{76} \\ \varepsilon_{77} \\ \varepsilon_{78} \\ \varepsilon_{79} \\ \varepsilon_{80} \end{bmatrix} = \begin{bmatrix} 0.945 \\ 0.967 \\ 0.971 \\ 0.969 \\ 0.974 \\ 0.979 \\ 0.980 \\ 0.981 \\ 0.985 \\ 0.985 \end{bmatrix} + \begin{bmatrix} 0.045 \\ 0.023 \\ 0.019 \\ 0.021 \\ 0.016 \\ 0.011 \\ 0.010 \\ 0.009 \\ 0.005 \\ 0.005 \end{bmatrix} P_V \qquad \text{Eq. 7.8}$$

3) $NDVI < 0.2$

In the case of AHS with bands from 1 to 20 located in the VNIR, emissivities for pixels with $NDVI < NDVI_s$ could be estimated as:

$$\epsilon_\lambda = a_\lambda + \sum_{i=1}^{20} b_{\lambda i}\rho_i \qquad\qquad \text{Eq. 7.9}$$

where ε_s is the soil emissivity and ε_v is the vegetation emissivity (assumed to be both a constant value and PV is the fractional vegetation cover estimated from the NDVI.

7.4.2 Water Vapour

Knowledge of the total atmospheric water vapour content is necessary to improve the precision of the estimates of land surface temperature obtained from satellite data by means of split window algorithms. Radiosonde data represent one way to obtain atmospheric water vapour content: however, this is not always possible, also considering that radiosonde observations are not carried out in a systematic way. Considering the fact that there are no methods to calculate water vapour from AHS data without using radiation transfer models, a methodology has been proposed based on the split window technique [718], [719] and starting from the following expression:

$$W = a\left(T_i - T_j\right) + b \qquad\qquad\qquad \text{Eq. 7.10}$$

where W is the atmospheric water vapour content in g cm^{-2} and T_i T_j are the at-sensor brightness temperatures measured in two different thermal bands (in K). With a regression technique, it was possible to recalculate the coefficients *a* and *b* for application to AHS flights. The AHS sensors has ten thermal bands (71 to 80, with effective wavelengths of 8.190, 8.660, 9.170, 9.600, 10.080, 10.560, 11.160, 11.720, 12.320 and 12.890 µm), so different combinations can be used (excluding band 80, which was affected by stripes). With the radiosoundings launched from Barajas airport, the regression have been performed with eight ground truth data synchronized with AHS flights from 25th June (Day and Night), 26th June (N), 28th June (D-N), 1st July (D), 2nd July N and 4th July (D). In Table 7.4, the obtained

correlations considering all the possible coefficient combinations are shown: the highest correlation coefficient (0.963) corresponds to 74-76 bands, with an *a* coefficients of 1.5073 and *b* coefficients of 2.7465. The water vapour map on the test image has been then extracted using these coefficients: the obtained mean value was 0.1 g/cm^2 higher than the ground truth data.

	71	72	73	74	75	76	77	78	79
71		0.4479	0.4706	0.5514	0.5645	0.6205	0.6547	0.7493	0.0307
72			0.5874	0.7514	0.8557	0.8985	0.2220	0.0831	0.3326
73				0.0004	0.8376	0.4802	0.0581	0.2029	0.3633
74					0.6176	0.9637	0.1051	0.2765	0.4134
75						0.1051	0.3689	0.3863	0.4457
76							0.4850	0.4496	0.4757
77								0.7943	0.4757
78									0.4987
79									

Table 7.4: correlation values of water vapour retrieval for each band combination

Now, all the terms in eq. (2) have been calculated and the LST map can be yielded.

7.5 Final results

Afterwards, the building mask and the LST map have been merged to display the Building Surface Temperature map (Fig. 7.4). From a first qualitative analysis we distinguished a colder area in the north part of the swath (which is located at a higher altitude) and individuate the areas of the city characterized by higher temperatures values, such as an industrial site in the southern part. A more quantitative assessment was performed considering the AHS LST values. Due to an insufficient numbers of specific ground truth in correspondence of buildings during the AHS flight, the assessment was carried out relying on the TES method for the emissivity and on the radiosoundings measurements for the water vapor estimation [715], as a benchmark. We remind that this latter method is considered as the "state of the art" about LST retrieval. The statistics of the results are shown in figure 7.5. On a set of 80 buildings randomly selected over the image mask, a RMSE (Root Mean Square Error) value of 1.59 K has been obtained. It can be observed a general good correlation (0.99) between estimated values with the proposed method and values obtained with radiosoundings and TES method.
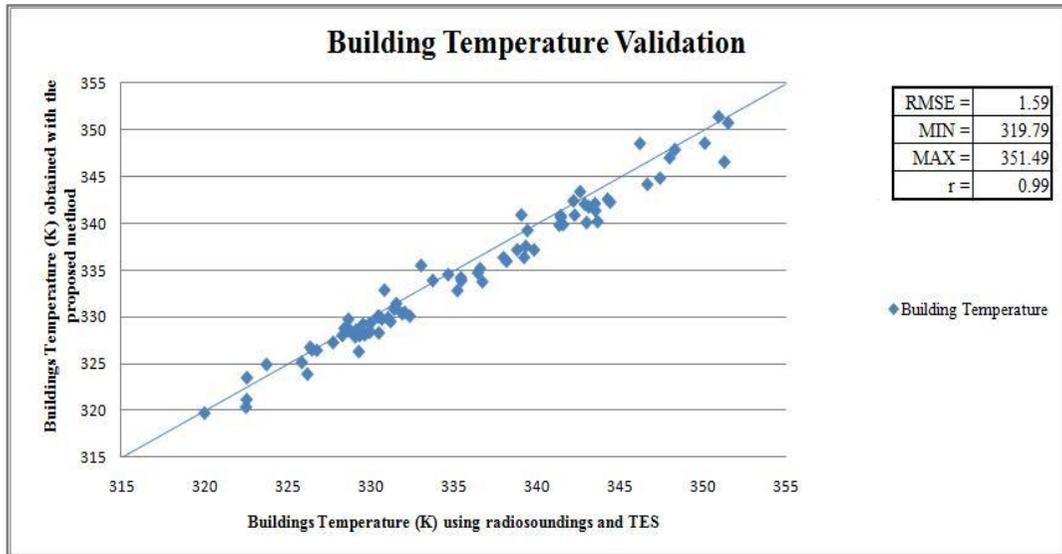
Figure 7.5: Validation of Building Temperature values obtained using radiosoundings and emissivity from TES algorithm versus Building Temperature values obtained using Split Window for water vapour and NDVI[THM]

Temperature (K)

🟥	> 330
🟧	325 – 330
🟧	320 – 325
🟨	315 – 320
🟨	310 – 315
🟩	305 – 310
🟩	300 – 305
🟩	< 300
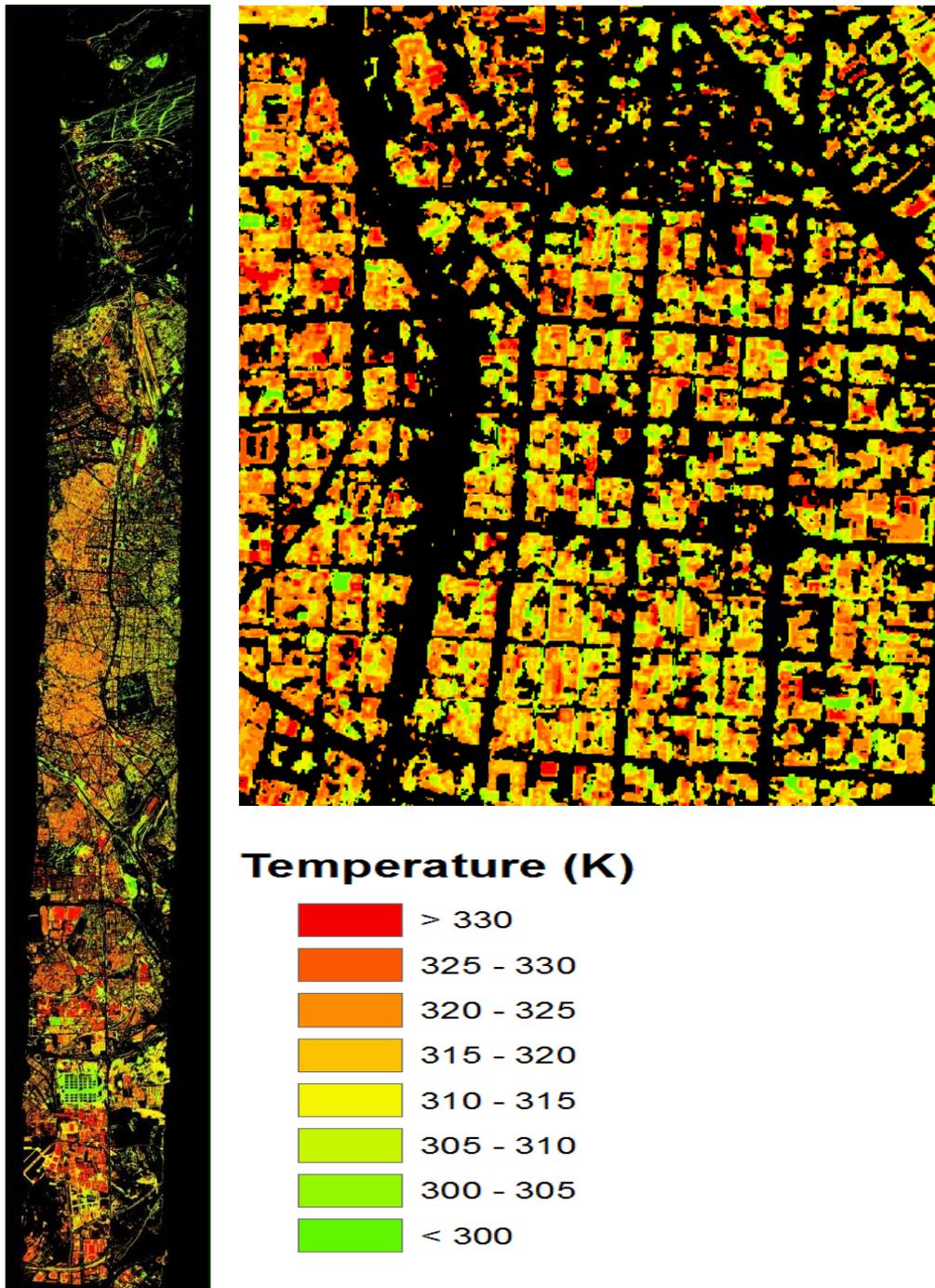
Fig . 7.4: Building Temperature map (left) and a detail of it (right)

Conclusions

In this work, an innovative processing chain for the automatic classification of urban area with Very High Resolution (VHR) images is presented. The complexity of urban environment and the key role of the context parameters in VHR images requires new automatic approaches for land cover detection (chapter 1).

The presented methodology has been tested on two types of VHR imagery: satellite multispectral as Quickbird, with four bands in visible (RGB) and near infrared band; airborne hyperspectral as AHS and MIVIS, with more than seventy bands from visible to thermal part of the spectrum. Both imagery have a comparable spatial resolution (chapter 2).

The idea was to create an original scheme for the automatic production of classified maps, stemming from the fusion of different and last generation remote sensing techniques such as automatic image classification with neural networks using spectral, textural and shape characteristics of the main land cover classes (chapter 3).

The features to extract from the VHR imagery corresponded to vegetation, water, bare soil, buildings and roads. Before performing the classification step, a pre-processing stage has been considered to standardise the data (chapter 4). The atmospheric correction, through the retrieval of reflectance values with information from the scene, the sensors, atmospheric and model settings, allowed to compare the response of the main land cover classes in different images and correct the influence of the atmosphere on the surfaces spectra. Besides, the exploitation of the thermal information in hyperspectral images let to remove the building shadows in urban area, enhancing the image quality through the application of simulated reflectance algorithm in urban environments. This showed satisfactory results in the removal of building shadows, especially considering the use of hyperspectral sensors, where the spatial resolution of the thermal bands is the same of the other visible and near infrared bands. This methodology allowed to suppress topographic effect and it had the advantage of conserving surface albedo information.

Following, a deeper analysis of the land cover classes characteristics in all the images allowed to explain their sensitivity to the classification process (chapter 5). This step showed that natural classes (vegetation and water) has a distinctive spectra: on the other hand, man-made has generally a wider variability and they also need additional parameters from their context for their identification.

The classification has been achieved through a mixed approach by combining Kohonen's SOMs and a MLP network, based on:

a) the extraction of clusters of pixels representing the land cover classes with a Kohonen's SOM (unsupervised approach);

b) a labelling step, to assign the spotted cluster to a semantic land cover class;

c) the use of the pixels belonging to the labeled clusters to train a MLP supervised neural network (supervised approach).

In fact, the training sets for the MLP have been extracted automatically from the image. A first Kohonen SOM was applied to extract vegetation, water, bare soil and, in some cases, roads. After masking the classified pixels, a second Kohonen SOM has been successively applied to spot buildings (exploiting the homogeneity parameter) and roads (if they were not detected before). Finally, the pixels from each clusters have been randomly chosen and they have been sent to the supervised neural networks.

We observed that classes like vegetation and water can be generally spotted and labeled in automatic way (chapter 6). Bare soils can be highlighted with Kohonen map: the labelling is often interactive, due to its variability in composition. The addition of homogeneity allowed to distinguish buildings: this class can be made of different buildings clusters, therefore the labeling depends on the number of buildings typology. Roads can be spotted with spectral characteristics and, if this parameter is not enough, the use of an additional segmentation step could help in its retrieval.

For the extraction of these macro-classes, the use of the entire hyperspectral spectrum instead of the only four analogous bands of Quickbird seems to slightly improve the results of few percentages points.

To complete this study, an innovative application of the proposed automatic classification methodology to urban thermography has been proposed (chapter 7). It was based on the fusion of automatic buildings detection and Land Surface Temperature (LST) retrieval from hyperspectral data. The final result is an original scheme for the production of building temperatures maps over urban areas. The method has been applied to an image taken over the city of Madrid and the obtained results shows a RMSE of 1.59 K of LST . Considering that an increased use of hyperspectral imagery for UHI management can be foreseen for the next

years, the developed technique may contribute to improve the characterization of the building heat transfer processes in the more general context of Energy Balance Models.

As final remarks, this study shows the capacity of the Kohonen SOM to cluster VHR images in a predefined number of features which can be assigned to a semantic class.

This labelling is more efficient for natural surfaces. Therefore, output image of water/land maps, vegetation/sealed areas maps can be performed often autonomously. In more complex environments, especially with human structure, the automatic labeling cannot be applied in all the different cases and more clusters can be assigned to the same semantic class. By the way, the use of contextual parameter (as texture and shape) helps in this discrimination.

Concluding, from the present work and the literature research, even if some very satisfactory results have been obtained, it can be hypothesized that a universal and objective classification criteria which is able to classify automatically each kind of image in the desired output classes has not been found yet.

These results stem from an objective evaluation that a classification criteria is always subjective in its strict sense. It has been shown how the same image can be classified in several ways, even if the number of the output class is the same (Fig. 3.1). *Therefore, the classification criteria and its output depend on the purposes of the analysis and the intrinsic composition of the image itself.*

It is worthwhile to observe that human assistance will be reduced to a quick view of the clusterization output, with a fast re-label of the clusters if the automatic results are not satisfactory.

In my personal opinion, now the efforts in this field have to be directed to the study of the input of the classifier, considering that the classification accuracy reached a satisfactory value. A deeper study of the common and the distinctive parameters

(i.e. topological relationship) in each cover classes could help in the direction of building more powerful classification chains and move the entire process toward a more automatic processor.

Bibliography

[1]   J. R. Jensen and D. C. Cowen, "Remote sensing of urban/suburban infrastructure and socio-economic attributes", *Photogramm. Eng. Remote Sens.*, vol. 65, no. 5, pp. 611–622, May 1999.

[2]   G. Trianni, *Techniques for fusion of remotely sensed data over urban environments*, Ph.D. thesis, Pavia University, Pavia, Italy, 2007.

[3]   S. W. Wharton, "A Spectral-Knowledge-Based Approach for Urban Land-Cover Discrmination", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, no. 3, pp. 272 – 282, 1987.

[4]   S. Khorram, J.A. Brockhaus, and H. M. Cheshire, "Comparson of Landsat MSS and TM Data for Urban Land-Use Classification", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, no. 2, pp. 238- 243, 1987.

[5]   ESA, "Annual report", *Tech. Rep*., European Space Agency, 2006 [on-line].

[6]   Herold, M. Gardner and D.A. Roberts, "The spectral dimension in urban land cover mapping from high-resolution optical remote sensing data", *In Proceedings of the 3rd Symposium on Remote Sensing of Urban Area*, June 2002, Istanbul, Turkey, 2002.

[7]   Herold, M. Gardner and D.A. Roberts, "Spectral resolution requirement for mapping urban areas*", IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1907–1919, 2003.

[8]   J.P. Donnay, M.J. Barnsley and P. Longley, *Remote Sensing and urban Analysis*, Taylor and Francis, pp. 3-18, .2001.

[9]   F. Pacifici, *Novel Neural Network-based algorithms for urban classification and change detection from satellite imagery*, Ph.D. thesis, Tor Vergata University, Rome, Italy, 2010.

[10] P. M. Teillet, "An algorithm for the radiometric and atmospheric correction of AVHRR data in the solar reflective," *Remote Sens. Environ.,* vol. 41, pp. 185–195, 1992.

[11] R. S. Fraser, O. P. Bahethi, and A. H. Al-Abbas, "The effect of the atmosphere on the classification of satellite observations to identify surface features," *Remote Sens. Environ.,* vol. 6, pp. 229–249, 1977.

[12] Y. J. Kaufman, "The atmospheric effect on the separability of field classes measured from satellites," *Remote Sens. Environ.,* vol. 18, pp. 21–34, 1985.

[13] P. Carleer and E. Wolff, "Urban land covermulti-level region-based classification of VHR data by selecting relevant features", *International Journal of Remote Sensing*, vol. 27, pp. 1035−1051, 2006.

[14] Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, and S. D'Elia, "Information mining in remote sensing image archives: system concepts", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2923–2936, Dec. 2003.

[15] J.E. Estes, E.J. Hajic and L.R. Tinney, "Fundamentals of image analysis: Analysis of visible and thermal infrared data", *Manual of Remote Sensing*, 2nd ed., R.N. Colwell (Ed.), pp. 987–1124, 1983.

[16] T.M. Lillesand and R.W. Kiefer, *Remote Sensing and Image Interpretation*, John Wiley & Sons, Inc, New York, 1994.

[17] J. B. Cambell, *Introduction to Remote Sensing*, Taylor & Francis, London, 1996.

[18] Guindon, "Combining diverse spectral, spatial and contextual attributes in segment-based image classification", *Paper presented at the ASPRS Annual Conference*, 22–26, May, Washington DC, 2000.

[19]   A.K. Shackelford and C.H. Davis, "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1920–1932, 2003.

[20]   Thomas, C. Hendrix and R.G. Congalton, "A comparison of urban mapping methods using high-resolution digital imagery" *Photogrammetric Engineering and Remote Sensing*, vol. 69, pp. 963–972, 2003.

[21]   G.G. Wilkinson, "Results and implications of a study of _fteen years of satellite image classification experiments, *IEEE Trans. Geosci. Remote Sens.*, 43, pp. 433-440, 2005.

[22]   J.B. MacQueen, "Some Methods for classification and analysis of multivariate Observations", *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability",* Berkeley, University of California press, 281-297, 1967.

[23]   K. Shackelford and C. H. Davis, "A Hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 9, pp. 1920−1932, 2003.

[24]   E. Pagot, and M. Pesaresi, "M. Systematic study of the urban postconflict change classification performance using spectral and structural features in a support vector machine", *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, vol. 2, pp. 120−128, 2008.

[25]   Chini, N. Pierdicca, and W. J. Emery, "Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake", *IEEE Transaction on Geosciences and Remote Sensing*, vol. 47, pp. 145–152, 2009.

[26]   M. Lazzarini, F. Del Frate and G. Ceriola, "Automatic Generation of Building Temperature Maps from Hyperspectral Data", IEEE Geoscience and Remote Sensing Letters, vol. 8, n. 2, March 2011.

[101] Digital Globe, *Quickbird imagery products: Product guide*, Tech Rep., Digital Globe, Longmont, CO, USA, 2008.

[102] Geoeye, Ikonos Imagery product guide, Geoeye, Dulles, VA, USA, 2006.

[103] J. MacQueen, "Some methods for classification and analysis of multivariate observations", *In Proc. Fifth Berkeley Symposium on Mathematics, Statistics and Probability*, pp. 281–296, 1967.

[104] D. Hall and G. Ball, "Isodata: a novel method of data analysis and pattern classification", *Technical report*, Stanford Research Institute, 1965.

[105] V. Kecman, Learning and soft computing: *Support Vector Machines, Neural Networks and Fuzzy Logic Models*, The MIT press, Cambridge, Massachusetts, 2001.

[106] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[107] V. N. Vapnik, *The Nature of Statistical Learning Theory*, 2nd ed. Berlin, Germany: Springer-Verlag, 1999.

[108] K. P. Bennett and A. Demiriz, "Semi-supervised support vector machines," in *Proc. Adv. Neural Inform. Process. Syst.*, 1998, vol. 10, pp. 368–374.

[109] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for semisupervised classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 11, pp. 3363–3373, Nov. 2006.

[110] M. Chi and L. Bruzzone, "Semisupervised classification of hyperspectral images by SVMs optimized in the primal," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1870–1880, Jun. 2007.

[111] J. Inglada, "Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features", *ISPRS Journal of Photogrammetry and Remote Sensing,* vol.62, pp. 236-248, 2007.

[112] A. Baraldi, T. Wassenaar, and S. Kay, "Operational Performance of an Automatic Preliminary Spectral Rule-Based Decision-Tree Classifier of Spaceborne Very High

Resolution Optical Images", *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3482–3502, Sept. 2010.

[113] T. Kohonen, *Self Organization and Associative memory*, Springer-Verlag, Berlin, 1984.

[114] T. Kohonen, *Self-Organizing Maps*. 3rd ed., Berlin, Springer, 2001.

[115] A. Baraldi and F. Parmiggiani, "A Neural Network for Unsupervised Categorization of Multivalued Input Patterns: An Application to Satellite Image Clustering", *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 305–316, March 1995.

[116] N. Abou Dagga, M. F Courel, F. Badran and B. Anselm, B. "The evolution study of the natural hazard zones by the Kohonen self-organising map classification", *International Archives of Photogrammetry and Remote Sensing*, vol. 35, pp. 503-508, 2004.

[117] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja, "Self-organizing maps as a relevance feedback technique in content-based image retrieval," *Pattern Analysis & Applications*, vol. 4, no. 2+3, pp. 140–152, June 2001.

[118] M. Molinier, J. Laaksonen, and T. Hame, "A Self-Organizing Map Framework for Detection of Man-Made Structures and Changes in Satellite Imagery", *proocedings on IEEE Geoscience and Remote Sensing Symposium*, 1-4 August, Denver (CO), USA, 2006.

[119] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, "Neural network approaches versus statistical methods in classification of multisource remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 540–552, Jul. 1990.

[120] B. G. Frizzelle and A. Moody, "Mapping continuous distributions of land cover: A comparison of maximum-likelihood estimation and artificial neural networks," *Photogramm. Eng. Remote Sens.*, vol. 67, no. 6, pp. 693–705, 2001.

[121] H. Cetin, T. A. Warner, and D. W. Levandowski, "Data classification, visualization, and enhancement using n-dimensional probability density functions (nPDF): AVIRIS, TIMS, TM and geophysical applications," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 12, pp. 1755–1764, 1993.

[122] P. Dreyer, "Classification of land cover using optimized neural nets on SPOT data," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 5, pp. 617–621, 1993.

[123] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, "Use of neural networks for automatic classification from high-resolution imagery", *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 800–809, Apr. 2007.

[124] M. Hansen, R. Dubayah, and R. Defries, "Classification trees: An alternative to traditional land cover classifiers," *Int. J. Remote Sens.*, vol. 17, no. 5, pp. 1075–1081, 1996.

[125] A. S. Kumar and K. L. Majumder, "Information fusion in tree classifiers," *Int. J. Remote Sens.*, vol. 22, no. 5, pp. 861–869, 2001.

[126] S. E. Franklin, "Discrimination of subalpine forest species and canopy density using digital CASI, spot PLA, and Landsat TM data," *Photogramm. Eng. Remote Sens.*, vol. 60, no. 10, pp. 1233–1241, 1994.

[127] P. J. Hardin, "Parametric and nearest-neighbor methods for hybrid classification: A comparison of pixel assignment accuracy," *Photogramm. Eng. Remote Sens.*, vol. 60, no. 12, pp. 1439–1448, 1994.

[128] B. C. K. Tso and P. M. Mather, "Classification of multisource remote sensing imagery using a genetic algorithm and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1255–1260, May 1999.

[129] M. J. Carlotto, "Spectral shape classification of Landsat thematic mapper imagery," *Photogramm. Eng. Remote Sens.*, vol. 64, no. 9, pp. 905–913, 1998.

[130] F. A. Kruse, A.B. Lefkoff, J. B. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz,, "The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging spectrometer Data", *Remote Sensing of the Environment*, vol. 44, pp. 145-163, 1993.

[131] L. Bruzzone and L. Carlin, "A multilevel context-based system for classification of very high resolution images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 2587–2600, 2006.

[132] B. Schölkopf and A. Smola, *Learning With Kernels—Support Vector Machines, Regularization,* Optimization and Beyond. Cambridge, MA: MIT Press, 2001.

[133] J. C. Seong and E. L. Usery, "Fuzzy image classification for continental scale multitemporal NDVI series images using invariant pixels and an image stratification method," *Photogramm. Eng. Remote Sens.*, vol. 67, no. 3, pp. 287–294, 2001.

[134] J. A. Benediktsson and I. Kanellopoulos, "Classification of multisource and hyperspectral data based on decision fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1367–1377, May 1999.

[135] Bishof, W. Schneider, and A. J. Pinz, "Multispectral classification of Landsat-images using neural networks", *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp. 482–490, May 1992.

[136] C. Bishop, *Neural Networks for Pattern Recognition*. New York: Oxford Univ. Press, 1995.

[137] K. Hornik, M. Stinchcombe and H. White, " Multilayer feedforward networks are universal approximators", Neural Networks, vol.2, no. 5, pp. 359-366. 1989.

[138] F. Pacifici, *Novel Neural Network-based algorithms for urban classification and change detection from satellite imagery*, Ph.D. thesis, Tor Vergata University, Rome, Italy, 2010.

[139] L. Bruzzone and D. Fernández-Prieto, "A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 1179–1184, Mar. 1999.

[140] F. Pacifici and F. Del Frate, "Automatic Change Detection in Very High Resolution Images with Pulse-Coupled Neural Networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 1, pp 58-62. 2010.

[141] P. Gong, D. J. Marceau and P.J. Howart, "A comparison of spatial feature extraction algorithms for land-use classification with spot hrv data", *Remote Sensing of Environment*, vol. 40, pp. 137-151, 1992.

[142] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, and S. D'Elia, "Information mining in remote sensing image archives: system concepts", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2923–2936, Dec. 2003.

[143] A. P. Carleer and E. Wolff, "Urban land covermulti-level region-based classification of VHR data by selecting relevant features", *International Journal of Remote Sensing*, vol. 27, pp. 1035−1051, 2006.

[144] M. Tuceryan and A.K. Jain, "Texture Analysis," *Handbook Pattern Recognition and Computer Vision*, C.H. Chen, L.F. Pau, and P.S.P. Wang, eds., Singapore: World Scientific, pp. 235-276, 1993.

[145] R. Haralick, K. Shangmugam, and L. Dinstein, "Textural Features for Image Classification", *IEEE Trans. Systems, Man, and Cybernetics*, vol. 3, pp. 610-621, 1973.

[146] R. Haralick, "Statistical and structural approaches to texture", *Proceeding of the IEEE,* vol. 67, 786−804, 1979.

[147] K. S. Shanmugan, V. Narayanan, V. S. Frost, J. A. Stiles, and J. C. Holtzman, "Textural features for radar image analysis", *IEEE Trans. Geosci. Remote Sensing*, vol. 19, pp. 153−156, 1981.

[148] D. A. Clausi and B. Yue, "Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 1, pp. 215–228, 2004.

[149] R. Cossu, *Segmentation by means of textural analysis*. Pixel, 1, 21-24,1988.

[150] D. Chen, D. A. Stow and P. Gong, "Examining the effect of spatial resolution and texture window size on classification accuracy: an urban environment case" *International Journal of Remote Sensing*, vol. 25, pp. 2177-2192, 2004.

[151] F. Pacifici, M. Chini, and W. J. Emery, "A neural network approach using multi-scale textural metrics from very-high resolution panchromatic imagery for urban land-use classification", *Remote Sensing of Environment*, vol. 113, pp. 1276-1292, June 2009.

[152] A.P. Carleer, O. Debeir, and E. Wolff, "Assessment of Very High Spatial Resolution Satellite Image Segmentations", Photogrammetric Engineering and Remote Sensing, vol. 71, 1285-1294, 2005.

[153] M. Herold, X. Liu, and K.C. Clarke, "Spatial metrics and image texture for mapping urban land use", Photogrammetric *Engineering & Remote Sensing*, vol. 69, pp. 991–1001, 2003a.

[154] M. Herold, M.E. Gardner, and D.A. Roberts, "Spectral resolution requirement for mapping urban areas", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1907–1919, 2003.

[155] N. Thomas, C. Hendrix, and R.G. Congalton, "A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering & Remote Sensing*, vol. 69, pp. 963–972, 2003.

[156] Y.J. Zhang, "Evaluation and comparison of different segmentation Algorithms", *Pattern Recognition Letters*, vol. 18, pp 963–974, 1997.

[157] R.M. Haralick, and L.G. Shapiro, "Survey: Image segmentation Techniques", *Computer Vision, Graphics, and Image Understanding*, Vol. 29, pp. 100–132, 1985.

[158] B. Nicolin and R. Gabler, "A knowledge-based system for the analysis of aerial images", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 25, pp. 317–329, 1987.

[159] L.L.F. Janssen and M. Molenaar, "Terrain objects, their dynamics and their monitoring by the integration of GIS and remote sensing", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 33, pp. 749–758, 1995.

[160] B. Guindon, "Computer-based aerial image understanding: A review and assessment of its application to planimetric information extraction from very high resolution satellite images", *Canadian Journal of Remote Sensing*, vol. 23, pp. 38–47, 1997.

[161] J. Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8, pp. 679–698, 1986.

[162] J.P. Cocquerez and S Philipp, *Analyse d'images: filtrage et segmentation*, Masson, Paris, pp. 457. 1995.

[163] L. Vincent and P. Soille, "Watershed in digital spaces: An efficient algorithm based on immersion simulations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, pp. 583–598, 1991.

[164] O. Debeir, *Segmentation Supervisée d'Images*, Ph.D., Faculté des Sciences Appliquées, Université Libre de Bruxelles, 2001.

[165] F. Deravi and S.K. Pal, "Grey level thresholding using secondorder Statistics", *Pattern Recognition Letters*, vol. 1, pp. 417–422, 1983.

[201] Digital Globe, *Quickbird imagery products: Product guide*, Tech Rep., Digital Globe, Longmont, CO, USA, 2008.

[202] C. Solimini, *High Resolution Urban Monitoring by Neural Network and Transform Algorithms*, Ph.D. thesis, Tor Vergata University, Rome, Italy, 2007.

[203] F. Pacifici, *Novel Neural Network-based algorithms for urban classification and change detection from satellite imagery*, Ph.D. thesis, Tor Vergata University, Rome, Italy, 2010.

[204] F. Pacifici, F. Del Frate, C. Solimini and W.J. Emery, "An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery", *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2940–2952, September. 2007.

[205] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, "Use of neural networks for automatic classification from high-resolution imagery", *IEEE Trans. Geosci. Remote Sens*., vol. 45, no. 4, pp. 800–809, Apr. 2007.

[206] J. A. Sobrino, J. C Jiménez-Muñoz, P. J. Zarco-Tejada, G. Sepulcre-Cantó, and E. De Miguel, "Land Surface Temperature derived from Airborne Hyperspectral Scanner Thermal Infrared data", *Remote Sensing of Environment,* 102, 99-115, 2006.

[207] ESA and DLR*, AGRISAR (Agricultural Bio-/Geophysical Retrievals from Frequent Repeat SAR and Optical Imaging) Final Report,* 14 January 2008.

[208] R.Polieri, V. Barbieri, R. Balice and D. Lorusso, "Sistema di Monitoraggio Quali-Quantitativo dei Corpi Idrici Sotterranei della Puglia", *proceedings on 13° Conferenza Nazionale ASITA*, 1-4 December 2009, Bari (Italy), Bari, 2009.

[301] C. Putignano "PolSOM and TexSOM in Polarimetric SAR Classification", *PhD thesis*, Tor Vergata University of Rome, 2009.

[302] C. Pratola, M. Del Greco, F. Del Frate, G. Schiavon, D. Solimini, "Towards fully automatic generation of land cover maps from polarimetric and metric resolution SAR data", proceedings on IGARSS 2010, Honolulu, Hawaii (USA), 25-30 July, 2010.

[303] P. H. Swain and H. Hauska, "The decision tree classifier: Design and potential," *IEEE Trans. Geosci. Electron.*, vol. GE-15, pp. 142–147, 1977.

[304] B. Kim and D. A. Landgrebe, "Hierarchical classifier design in high-dimensional, numerous class cases," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 518–528, July 1991.

[305] J. T. Morgan, A. Henneguelle, M. M. Crawford, J. Ghosh, and A. Neuenschwander, "Adaptive feature spaces for land cover classification with limited ground truth data," in *Proc. 3rd Int. Workshop on Multiple Classifier Systems—MCS 2002*, Cagliari, Italy, June 2002,

[306] M. Datcu, F. Melgani, A. Piardi, and S. B. Serpico, "Multisource data classification with dependence trees," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 609–617, Mar. 2002

[307] F. Melgadi and L. Bruzzone, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines", *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 8, pp. 1778-1790, 2004.

[308] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, S.J. Sutley, *USGS digital spectral library splib06a*, U.S. Geological Survey, Digital Data Series 231, 2007.

[309] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, and S. D'Elia, "Information mining in remote sensing image archives: system concepts", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 12, pp. 2923–2936, Dec. 2003.

[310] D. Chanchal and B. Chakraborty, "Acoustic Characterization of Seafloor Sediment Employing a Hybrid Method of Neural Network Architecture and Fuzzy Algorithm", *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 743-747, October 2009.

[311] C. Pratola, *Progettazione di un algoritmo per la classificazione completamente automatic d'immagini satellitari*, MSc Thesis, Tor Vergata University, Rome, Italy, 2008.

[312] F. Pacifici, M. Chini, and W. J. Emery, "A neural network approach using multi-scale textural metrics from very-high resolution panchromatic imagery for urban land-use classification", *Remote Sensing of Environment*, vol. 113, pp. 1276-1292, June 2009.

[313] X. Jin, Segmentation-based image processing system, *US Patent 20090123070*, pp. 34, 2009.

[314] K. Pearson, "On lines and planes of closest fit to systems of points in space", *Philosophical Magazine* 2, vol. 6, pp. 559-572, 1901.

[315] D. J. Robinson, N. J. Redding and D. J. Crisp, "Implementation of a fast algorithm for segmenting SAR imagery", *Scientific and Technical Report*, Australia: Defense Science and Technology Organization, 2002.

[316] D. H. Douglas and T. K.. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature", *Cartographica*, 10(2), pp. 112-122, 1973.

[317] J. C. Russ, *The Image Processing Handbook*. Fourth Edition. Boca Raton, FL: CRC Press, 2002.

[318] M. Lazzarini, F. Del Frate, "Features extraction from hyperspectral images: an approach based on spectral, textural and spatial information applied to urban environments", *proceedings on Hyperspectral 2010 Workshop*, ESA-ESRIN, Frascati, (I), March 17- 19, 2010.

[401] P. M. Teillet, "An algorithm for the radiometric and atmospheric correction of AVHRR data in the solar reflective," *Remote Sens. Environ.,* vol. 41, pp. 185–195, 1992.

[402] R. S. Fraser, O. P. Bahethi, and A. H. Al-Abbas, "The effect of the atmosphere on the classification of satellite observations to identify surface features," *Remote Sens. Environ.,* vol. 6, pp. 229–249, 1977.

[403] Y. J. Kaufman, "The atmospheric effect on the separability of field classes measured from satellites," *Remote Sens. Environ.,* vol. 18, pp. 21–34, 1985.

[404] B.-C. Gao, K.B. Heidebrecht and A.F.H. Goetz, *Atmosphere Removal Program (ATREM) Version 2.0 Users Guide,* Center for the Study of Earth from Space/CIRES, University of Colorado, Boulder, Colorado, 26, 1996.

[405] E.F. Vermote, D. Tanre, J.L. Deuze, M. Herman and J.J. Morcrette, "Second Simulation of the Satellite Signal in the Solar Spectrum (6S)", *6S User Guide Version 6.0*, NASAGSFC, Greenbelt, Maryland, 134, 1994.

[406] R. Richter, "Correction of atmospheric and topographic effects for high spatial resolution satellite imagery", *Int. J. Remote Sensing*, vol. 18, pp. 1099-1111, 1997.

[407] Z. Qu, B. Kindel and A. F. H. Goetz, "Aerosol Retrieval and Spectral Calibration of AVIRIS with HATCH," *2001 AVIRIS Geoscience Workshop,* Jet Propulsion Laboratory, Pasadena, CA, 2001.

[408] A. Berk, S.M. Adler-Golden, A.J. Ratkowski, G.W. Felde, G.P. Anderson, M.L. Hoke, T. Cooley, J.H. Chetwynd, J.A. Gardner, M.W. Matthew, L.S. Bernstein, P.K. Acharya, D. Miller, P. Lewis, "Exploiting MODTRAN radiation transport for atmospheric correction: The FLAASH algorithm", *Proceedings of the Fifth International Conference on Information Fusio*n, vol.2, no., pp. 798- 803, 2002.

[409] R. Green, *Atmospheric Correction Now (ACORN)*, developed by ImSpec LLC, available from Analytical Imaging and Geophysics LLC, 2001.

[410] A. Berk, G. P. Anderson, P. K. Acharya, M. L. Hoke, J. H. Chetwynd, L. S. Bernstein, E. P. Shettle, M. W. Matthew, and S. M. Adler-Golden, "MODTRAN4 Version 3 Revision 1 Users´s Manual*",* Bedford, MA: Air Force Res. Lab., Hanscom AFB, 2003.

[411] M. W. Matthew, S. M. Adler-Golden, A. Berk, G. Felde, G. P. Anderson, D. Gorodetzky, S. Paswaters, M. Shippert, "Atmospheric Correction of Spectral Imagery: Evaluation of the FLAASH Algorithm with AVIRIS Data," *Proceedings on 31st Applied Imagery Pattern Recognition Workshop*, pp.157, 2002.

[412] A. Berk, P. Acharya, G. Anderson and B. Gossage, "Recent developments in the modtran atmospheric model and implications for hyperspectral compensation", *proceedings on IGARSS 2009*, Cape Town, South Africa, 12-17 July, 2009.

[413] L.S. Rothman, C.P. Rinsland, A. Goldman, S.T. Massie, D.P. Edwards, J.-M. Flaud, A. Perrin, C. Camy-Peyret, V. Dana, J.-Y. Mandin, J. Schroeder, A. McCann, R.R. Gamache, R.B. Wattson, K. Yoshino, K.V. Chance, K.W. Jucks, L.R. Brown, V. Nemtchinov and P. Varanasi, "The HITRAN Molecular Spectroscopic Database and HAWKS (HITRAN Atmospheric Workstation): 1996 Edition," J. Quant. Spectrosc. Radiat. Transfer, Vol. 60, pp. 665-710, 1998

[414] Y. J. Kaufman, D. Tanre, L. A. Remer, E. F. Vermote, A. Chu, and B. N. Holben, "Operational Remote Sensing of Tropospheric Aerosol over Land from EOS Moderate Imaging Spectroradiometer," *J. Geophys. Res*., vol. 102, 1997.

[415] Digital Globe, *Quickbird imagery products: Product guide*, Tech Rep., Digital Globe, Longmont, CO, USA, 2008.

[416] L.W. Abreu and G.P. Anderson, *The MODTRAN 2/3 Report and LOWTRAN 7 Model*, Prepared by Ontar Corporation for PL/GPOS, 1996.

[417] R. G Isaacs, W. C. Wang, R. D. Worsham, and S. Goldenberg, "*Multiple Scattering LOWTRAN and FASCODE Models"*, *Applied Optics*, Vol. 26, pp. 1272-1281, 1987.

[418] K. Stamnes, S.-C. Tsay, W. Wiscombe, and K. Jayaweera, "Numerically Stable Algorithm for Discrete-Ordinate-Method Radiative Transfer in Multiple Scattering and Emitting Layered Media*"*, *Applied Optics*, Vol. 27, pp. 2502-2509, 1988.

[419] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, S.J. Sutley, *USGS digital spectral library splib06a*, U.S. Geological Survey, Digital Data Series 231, 2007.

[420] S. Watanabe, K. Miyajima, and N. Mukawa, "Detecting changes of buildings from aerial images using shadow and shading model", *Proceedings of IEEE International Conference on Pattern Recognition*, 20-20 Aug., vol.2, pp- 1408 – 1412, 1998.

[421] J. G. Liu, J. McM. Moore and J. D. Haigh, "Simulated reflectance technique for ATM image enhancement", *International Journal of Remote Sensing*, vol. 18, issue 2, pp. 243-254, 1997.

[422] M. Lazzarini, "Spectral enhancement of alteration associated with epithermal gold deposits in region III, Northern Chile", *MSc in Remote Sensing Dissertation,* University College London and Imperial College London, Academic Year 2006-2007.

[423] N. Robinson, "*Solar Radiation".* Amsterdam, London and New York: Elsevier Publishing Company. pp. 196-221, 1966.

[424] C. Elachi, "*Introduction to the Physics and Techniques of Remote Sensing*". New York: John Wiley & Sons., pp. 114-141, 1997.

[425] M. Lazzarini, J.G. Liu, F. Del Frate, "Hyperspectral image enhancement using thermal bands: a methodology to remove building shadows", *proceedings on International Geoscience and Remote Sensing Symposium 2010*, Honolulu, Hawaii (USA), 25-30 July, 2010.

[501] P. J. Sellers "Canopy reflectance, photosynthesis, and transpiration", *International Journal of Remote Sensing*, Vol. 6, pp. 1335-1372, 1985.

[502] Y. J. Kaufman and D. Tanre "Atmospheric resistant vegetation index (ARVI) for EOS-MODIS" IEEE. *Transactions on Geoscience and Remote Sensing*, vol. 30, pp. 261-270, 1992.

[503] A. R. Huete and H. Liu "An error and sensitivity analysis of the atmospheric- and soil-correcting variants of the NDVI for the MODIS-EOS", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, pp. 897-905, 1994.

[504] C.Perry, J. Lautenschlager and L. Lautenschlager, "Functional equivalence of spectral vegetation indices", *Journal Remote Sensing of Environment*, vol. 14, pp. 170-182, 1984.

[505] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, S.J. Sutley, *USGS digital spectral library splib06a*, U.S. Geological Survey, Digital Data Series 231, 2007.

[506] D. Solimini, "Understanding Earth Observation: electromagnetic foundations of remote Sensing", *advanced course in the "Geoinformation"* PhD programme, Tor Vergata University, Roma, 2010.

[507] J.T. Houghton, A.H. Cook and H. Charnock, *The study of the Ocean and the land surface from Satellites,* The royal Society, London, 1983.

[508] M. Lazzarini and F. Del Frate, "Features extraction from hyperspectral images: an approach based on spectral, textural and spatial information applied to urban environments", *proceedings on Hyperspectral 2010 Workshop, ESA-ESRIN*, Frascati, (I), March 17- 19, 2010.

[701] M. Lazzarini, F. Del Frate and G. Ceriola, "Automatic Generation of Building Temperature Maps from Hyperspectral Data", *IEEE Geoscience and Remote Sensing Letters,* in press, October 2010.

[702] J. R. Jensen and D. C. Cowen, "Remote sensing of urban/suburban infrastructure and socio-economic attributes," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 5, pp. 611–622, May 1999.

[703] S. Kato and Y. Yamaguchi, "Estimation of storage heat flux in an urban area using ASTER data", *Remote Sensing of Environment*, vol 110, pp. 1– 17, 2007.

[704] J.A. Voogt and T.R. Oke, "Thermal remote sensing of urban climates", *Remote Sensing of Environment*, Vol. 86, pp. 370 – 384, 2003.

[705] H. Tran, D. Uchihama, S. Ochi, and Y. Yasuoka, "Assessment with satellite data of the urban heat island effects in Asian mega cities", *International Journal of Applied Earth Observation and Geoinformation*, Vol. 8, pp. 34-48, 2006.

[706] J. C. Jiménez-Muñoz and J. A. Sobrino, "A generalized single-channel method for retrieving land surface temperature from remote sensing data*," J. Geophys. Res.,* vol. 108, no. D22, p. 4688, Nov. 2003.

[707] L. M. McMillin, "Estimation of sea surface temperatures from two infrared window measurements with different absorption", *J. Geophys. Res*., vol. 80, no. 36, pp. 5113–5117, 1975.

[708] R. Ferri, N. Pierdicca, S. Talice, "Mapping sea surface temperature from aircraft using a multi-angle technique: an experiment over the Orbetello lagoon", *International Journal of Remote Sensing,* Volume 21, Issue 16, pages 3003 – 3024, November 2000.

[709] A. R. Gillespie, T. Matsunaga, S. Rokugawa and S. J. Hook, "A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 4, pp. 1113-1126, July 1998.

[710] L. F. Peres, J. A. Sobrino, R. Libonati, J. C. Jiménez-Muñoz, C. C. Dacamara, M. Romaguera, "Validation of a temperature emissivity separation hybrid method from airborne hyperspectral scanner data and ground measurements in the SEN2FLEX field campaign", *International Journal of Remote Sensing*, Vol. 29, No. 24, pp. 7251 – 7268, December 2008.

[711] S. Kato, and Y. Yamaguchi, "Analysis of urban heat-island effect using ASTER and ETM+ Data: Separation of anthropogenic heat discharge and natural heat radiation from sensible heat flux", *Remote Sensing of Environment*, Vol. 99, pp. 44 – 54, 2005.

[712] V. Masson, "A physically-based scheme for the urban energy balance in atmospheric models", *Bound.-Layer Meteor.*, Vol. 94, No. 3, pp. 357-397, 2000.

[713] M. Lazzarini and F. Del Frate, "Features extraction from hyperspectral images: an approach based on spectral, textural and spatial information applied to urban environments", *proceedings on Hyperspectral 2010 Workshop,* ESA-ESRIN, Frascati, (I), March 17- 19, 2010.

[714] J. A. Sobrino, J. C Jiménez-Muñoz, P. J. Zarco-Tejada, G. Sepulcre-Cantó, and E. De Miguel, "Land Surface Temperature derived from Airborne Hyperspectral Scanner Thermal Infrared data", *Remote Sensing of Environment,* 102, 99-115, 2006.

[715] J. A. Sobrino, et al, "Desirex final report", *ESA Earth Observation Campaign*, March 2009.

[716] J. A. Sobrino, J. C. Jiménez-Muñoz, G. Sòria, M. Romaguera, L. Guanter, J. Moreno, A. Plaza, and P. Martínez, "Land Surface Emissivity Retrieval From Different VNIR and TIR Sensors", *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 46, No. 2, pp. 316-327, February 2008.

[717] A. Berk, G. P. Anderson, P. K. Acharya, M. L. Hoke, J. H. Chetwynd, L. S. Bernstein, E. P. Shettle, M. W. Matthew, and S. M. Adler-Golden, *MODTRAN4 Version 3 Revision 1 Users´s Manual,* Bedford, MA: Air Force Res. Lab., Hanscom AFB, 2003.

[718] K.P. Czajkowski, S.N. Goward, D. Shirey, and A. Walz, "Thermal remote sensing of near-surface water vapor", *Remote Sensing of Environment*, Vol. 79, pp. 253–265, 2002.

[719] J. C. Jiménez-Muñoz and J. A Sobrino, "Atmospheric water vapour content retrieval from visible and thermal data in the framework of the DAISEX campaigns"**,** *International Journal of Remote Sensing*, Vol. 26, No. 15, pp. 3163-3180, 2005.

List of acronyms and abbreviations

6S	Second Simulation of a Satelllite Signal in the Solar Spectrum		DLR	Deutsches zentrum für Luft und Raumfahrt
(A)ATSR	(Advanced) Along-Track Scanning Radiometer		DN	Digital Number
			DTC	Decision Tree Classifier
ACORN	Atmospheric CORrection Now		DUE	Data User Elements
AHS	Airborne Hyperspectral Scanner		EnMAP	Environmental Mapping
			EO	Earth Observation
AOI	Area Of Interest		ERTS1	Earth resources Yechnology Satellite 1
ASI	Agenzia Spaziale Italiana		ESA	European Space Agency
ASITA	Associazioni Scientifiche per le Informazioni Territoriali e Ambientali		FLAASH	Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer		FVC	Fractional Vegetation Cover
			FWHM	Full View at Half Maximum
			GLCM	Grey Level Co-occurence Matrix
ATCOR	ATmospheric CORrection		GMT	Greenwich Mean Time
ATM	Airborne Thematic Mapper		HATCH	High accuracy Atmospheric Correction for Hyperspectral Data
ATREM	ATmospheric REMoval			
AVHRR	Advanced Very High Resolution Radiometer			
			HyspIRI	HYperSPectral InfraRed Imager
CBIR	Content-Based-Image Retrieval		IGARSS	International Geoscience And Remote Sensing Symposium
CHRIS	Compact High Resolution Images Spectrometer		INTA	National Institute for Aerospace Technology
DASVM	Domain Adaptation Support Vector Machine		IR	InfraRed
DESIREX	Dual-use European Security IR Experiment		KIM	Knowledge-driven Information Mining
DISORT	DIScrete Ordinate Radiative Tranfer		Landsat	Land Satellite
			LST	Land Surface Temperature

LUT	Look-Up Table	RBF	Radial Basin Function
MERIS	Medium Resolution Imaging Spectrometer	RGB	Red Green Blue
MIR	Mid InfraRed	RMSE	Root Mean Square Error
MIVIS	Multispectral Infrared and Visible Imaging Spectrometer	ROI	Region Of Interest
MLP	Multi Layer Perceptron	ROSIS	Reflective Optics System Imaging Spectrometer
MODIS	MODerate resolution Imaging Spectroradiometer	RT	Radiative Tranfer
MODTRAN	MODerate resolution atmospheric TRANsmission	SAM	Spectral Angle Mapper
MRF	Markov Random Fields	SAR	Synthetic Aperture Radar
MS	Multi Spectral	SEB	Surface Energy Budget
NASA	National Aeronautics and Space Administration	SITAC	Spectral Information Technical Applications Center
NDVI	Normalized Difference Vegetation Index	SOM	Self Organizing Map
NDVI^{THM}	NDVI THresholds Method	SPOT	Système Pour l'Observation de la Terre
NIR	Near InfraRed	SVM	Support Vector Machine
NN	Neural Networks	SW	Split Window
NOAA	National Oceanic and Atmospheric Administration	SWIR	Short Wave InfraRed
NRT	Near Real Time	TES	Temperature Emissivity Separation
PAN	Panchromatic	TIR	Thermal InfraRed
PAR	Photosynthetically Active Radiation	TISI	Temperature Independent Spectral Indices
PCA	Principal Component Analysis	TOA	Top Of the Atmosphere
PRISMA	PRecursore IperSpettrale della Missione Applicativa	TS	Tree Structured
PROBA	PRoject for On-Board Autonomy	UHI	Urban Heat Island
PV	Proportion of Vegetation	USGS	United States Geological Survey
		VHR	Very High Resolution
		VIS	Visible

Curriculum Vitae

Michele Lazzarini started to work in Remote Sensing and GIS in 2002, with a thesis (cum laude) regarding the human impact on vegetation.

These scientific interests have been in-depth studied with a First level Master in "Cartography and GIS" at Università degli Studi di Trieste (Italy) in 2004 and with a Master of Science course in "Remote Sensing" at University College London and Imperial College London (UK) in 2007.

Since 2007 he is a PhD candidate in Geoinformation at Tor Vergata University of Rome and in 2010 he worked six months to support the activities of the Research and Service Support Section at ESA/ESRIN.

In these years, he has been involved in Global Monitoring Environment and Security (GMES) and Data User Element (DUE-ESA) projects. He has obtained a good command of the main image processing software and IDL, having a wide view about the application of information technologies to Earth Observation.

His main research activities regard the processing of multispectral and hyperspectral images, from the pre-processing stage until the information extraction for end-users. His goal is extending the use of Remote sensing to "real life" applications, increasing its catchment area to other users.

Conferences

M. Lazzarini and F. Del Frate, "Features extraction from hyperspectral images: an approach based on spectral, textural and spatial information applied to urban environments", *proceedings on Hyperspectral 2010 Workshop*, ESA-ESRIN, Frascati, (I), March 17- 19, 2010.

M. Lazzarini, F. Del Frate, G. Ceriola. "Automatic buildings extraction from hyperspectral data and its application to urban thermography". Image Information Mining ESA-EUSC 2009, EUSC, Torrejon air base - Madrid (Spain), November 3- 5, 2009.

F. Del Frate, M. Lazzarini, G. Licciardi, C. Putignano, C. Pratola, "Una procedura completamente automatica per la classificazione di immagini satellitari", XIV Riunione Annuale CeTeM / V Workshop AIT / XIV Giornata MECSA sull'Ingegneria delle Microonde, Roma, October 23-24, 2008.

F. Pacifici, E. Angiuli, M. Lazzarini, G. Licciardi, "Automatic Classification of SAR imagery with Neural Networks**,** *2008 IEEE GOLD Remote Sensing Conf.*, ESA/ESRIN, Frascati, Italy, May 2008.

Publications

M. Lazzarini, F. Del Frate and G. Ceriola, "Automatic generation of Building Temperature Maps from Hyperspectral data", *IEEE Geoscience and Remote Sensing Letters*, vol. 8, n. 2, March 2011 (online available).

M. Lazzarini, J.G. Liu and F. Del Frate, "Hyperspectral image enhancement using thermal bands: a methodology to remove building shadows", *proceedings on IEEE International Geoscience and Remote Sensing Symposium 2010*, Honolulu, Hawaii (USA), 25-30 July, 2010.

C. Putignano, D. Solimini, A. Burini, F. Del Frate, G. Schiavon, G. Licciardi, M. Lazzarini, P. Manunta and F. De Biasi, "TerraSAR-X imaging for unsupervised land cover classification and fire mapping", *proceedings on IEEE International Geoscience & Remote Sensing Symposium 2008*, Boston, Massachusetts, U.S.A., July 6-11, 2008.

P. Bevilacqua and M. Lazzarini, "Environmental monitoring: geostatistics applied to the Lagoon of Grado and Marano (italy) characterisation", *Siti Contaminati*, 2/2004, pp. 36-48, 2004.

A. Comuzzi and M. Lazzarini, "Preliminary studies of the correlation between the concentration of pollutants and their hyperspectral response: a methodological approach", *Siti Contaminati*, 2/2004, pp. 50-62, 2004.

Automatic Generation of Building Temperature Maps From Hyperspectral Data

Michele Lazzarini, *Student Member, IEEE*, Fabio Del Frate, *Senior Member, IEEE*, and Giulio Ceriola

Abstract—In this letter, a method to automatically retrieve building surface temperature maps using hyperspectral imagery is presented. The approach can be conceptually described by considering two different problems. The first consists in the design of an automatic procedure for the extraction of building surfaces from the hyperspectral image. Such an issue has been addressed using both unsupervised and supervised neural networks. The second problem deals with the retrieval of land surface temperature from the same image. The final step is the merging of the temperature map with the building mask. It is worthwhile to observe that the proposed approach aims at retrieving the temperature values by reducing the manual editing and the use of ancillary data to a minimum level. The obtained results show an accuracy in the building identification of 83.7% and a root-mean-square error (rmse) in the temperature retrieval of 1.59 K. The importance of this methodology has to be considered within the studies on *urban heat islands*, which is becoming an important issue in urban management politics.

Index Terms—Automatic classification, hyperspectral data, Kohonen self-organizing map (SOM), land surface temperature (LST), neural networks (NNs), urban heat island (UHI).

I. INTRODUCTION

TODAY, the continuous increase in the archives' size and the Earth observation sensors' variety requires new methodologies and tools for information mining and management. With this growth, new approaches for image information mining, multidomain information management, and knowledge management and sharing (in support of information mining) are necessary. An automatic chain, reducing the human assistance in the data analysis, allows to gain time on the entire image processing and can be designed to retrieve information in near real rime (NRT). Such concepts find an important field of application in the monitoring of urban environments, where the knowledge of the changes in either natural resources or man-made structures is a valuable source of information for decision-making processes [1]. Of particular interest in this context is the occurrence and detection of an urban heat island (UHI) phenomenon, which has grown in magnitude over the last decade. High densely built-up areas trap the heat, causing the city centers to be warmer, up to 10 °C, than the surrounding rural areas, and this is often manifested particularly at night [2], [3]. For the energy balance and urban planning aiming at reducing the number of casualties when temperatures soar,

the UHI information can be also updated on a monthly basis. However, the retrieval of temperature should be in NRT to support the civil protection in monitoring the health impact over the weakest population with a daily frequency.[1] The remotely sensed observations of the UHI are possible using satellite and aircraft platforms, and they could open new avenues for their analysis [4].

In this letter, we aim at the design of a novel and automatic scheme for the estimation of the temperature associated to buildings: This approach relies on a cascade of neural networks (NNs) and retrieval procedures and can represent an important technique to improve cities' energy balance models. The entire process has been tested on an airborne hyperspectral scanner (AHS) image acquired above Madrid in summer 2008, and the results were assessed by comparison with the state of the art on the land surface temperature (LST) retrieval.

II. SUPERVISED AND UNSUPERVISED NNs

Different studies have shown the effectiveness of NN algorithms in extracting information from remotely sensed imagery [5]. A common distinction is made, considering the type of their learning [6]: In a supervised learning network, the training will benefit from the assistance of the teacher, and the desired output is already known; for an unsupervised learning rule, the network is trained without the benefit of any teacher. These networks learn to adapt based on the experiences collected through the previous training patterns, and they are able to re-organize themselves. Hence, an unsupervised NN has no target output, and they find them only by considering the original data structure. Among the supervised methods, the multilayer perceptron (MLP) seems to be the most widely used topology for classifying satellite images [7]. In satellite image classification, the input dimension is often associated to the number of spectral bands, and the output dimension is the number of predefined classes. However, the choice of input bands is not always strictly related to the spectral bands: Band ratios, as well as other features derived by band combinations, or even textural bands could be used to feed the NN instead or in addition to spectral bands [8]. Considering the unsupervised methods, appreciable results in such research fields seem to derive from the application of an NN developed by Kohonen [9], named the *self-organizing map* (SOM). A SOM can automatically form 1-D or multidimensional maps of the intrinsic features of the input data. These data are presented in mixed random order to the network which is capable of learning complicated hierarchical relations within the considered high-dimensional
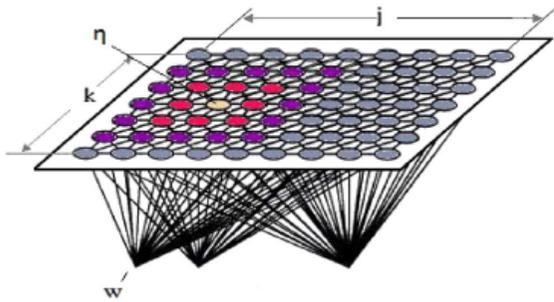
Fig. 1. Schematic representation of Kohonen SOM.

spaces. In remote sensing, the SOMs are used to identify the measurement relationships which they can reorganize in several output clusters/classes [10]. In Fig. 1, it is possible to see a Kohonen map with two dimensions where the neuron η_{kj} is connected with its neighborhood. When an input from the layer w activates a neuron in the η layer, the interaction that the winning neuron (selected with a user-defined rule) establishes with its neighbors is defined according to a neighborhood function, which normally takes a value between zero and one. This interaction determines the modification of the weights dependent on the neighborhood function and the response of the neuron. The η_j neuron activity is defined by

$$\frac{d\eta_j}{dt} = \sum_{i=1}^{n} w_{ji}\xi_i - \gamma(\eta_j) + \sum_{k \neq j}^{n} \mu_{kj}\eta_{kj} \qquad (1)$$

where η_j is the activity of the jth neuron, ξ_i is the ith component of the input, n is the total number of inputs, w_{ji} is the weight of the connection between the jth neuron and the ith input (synapse), γ_i is a term that considers the leaks of the process, and η_{kj} is the connection between the jth and kth neurons.

III. APPLICATION TO BUILDING DETECTION

The AHS is an imaging line-scanner radiometer, installed on a CASA-212 200 series aircraft owned by Spain's National Institute for Aerospace Technology (INTA). The AHS instrument has 80 spectral bands covering the visible and near infrared, short-wave infrared, midinfrared, and thermal infrared (TIR) spectral ranges. The imagery analyzed in this letter belongs to the collection of measurements carried out during the *Dual-use European Security IR Experiment 2008* campaign and has been taken over the city of Madrid (Spain). The testing image (see left of Fig. 2) has been acquired on June 28, 2008 (see Table I).

To obtain a detailed map of the buildings from the image, it has been decided to use a mixed approach between the unsupervised and supervised methods based on the NN. The concept idea is based on the following: 1) the extraction of clusters of the pixels representing buildings using Kohonen's SOM (unsupervised approach) and 2) the use of the pixels belonging to the extracted clusters to train an MLP-supervised NN (supervised approach). In fact, the training sets for the MLP have been extracted automatically from the image. To define the building signatures, both spectral and textural characteristics have been computed. In fact, particularly in very high resolution (VHR) image analysis [11], the information on the
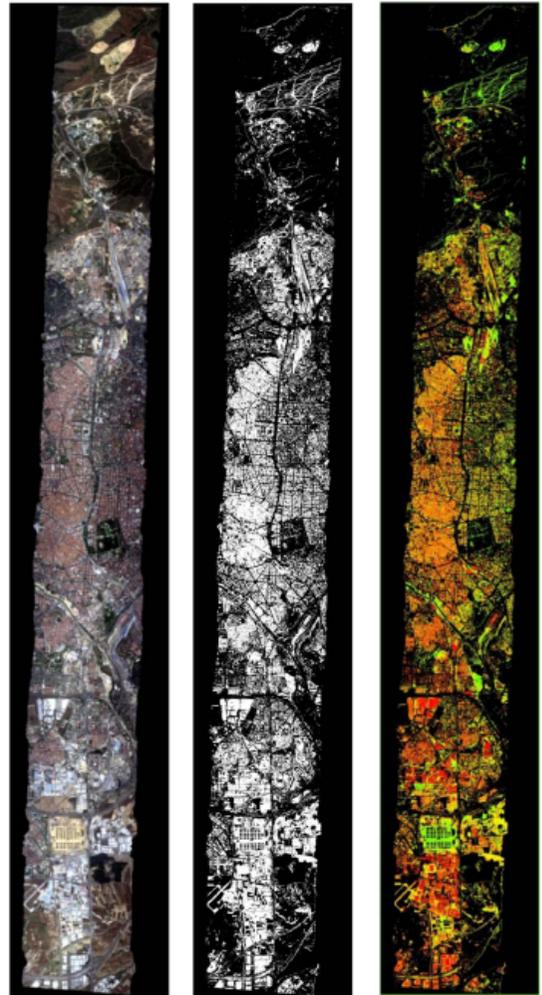

Fig. 2. AHS swath above Madrid. (Left) RGB 8-5-4 real color composite, (center) extracted building mask, and (right) building temperature map (see the legend in Fig. 3).

TABLE I
AHS IMAGE CHARACTERISTICS

Date	June, the 28th of 2008	
Acquisition time	11:53 am	
flight height	2497 m above sea level	
Pixel resolution	≈ 4 meters	
Spectral characteristics	Wavelength Range (µm)	Bands
	0.441-1.018	1-20
	1.491-1.650	21
	2.019-2.448	22-63
	3.03-5.41	64-70
	7.95-13.17	71-80

pixel context (e.g., structures, patterns, and objects) can be significant. A reference building signature which combines not only the spectral parameters but also the textural parameters has then been considered as the *standard building signature* [12]. In order to analyze the potential of the approach also for VHR multispectral images such as IKONOS or QuickBird,

the standard building signature only includes the AHS bands 2 (blue, 0.4836 μm), 5 (green, 0.5717 μm), 8 (red, 0.6586 μm), and 14 (near infrared, 0.8334 μm). As far as the textural parameters are concerned, they have been obtained from the grey level cooccurrence matrix (GLCM). The GLCM describes the frequency of one gray tone appearing in a specified spatial linear relationship with another gray tone, within the area under investigation [13]. From the range of texture parameters which can be calculated from the GLCM, the use of homogeneity (also called the "inverse difference moment") has been chosen as its utilization in urban feature extraction has already been proven effective [8]. Homogeneity weights values by the inverse of the contrast weight. To exploit a multiscale approach, which better deals with objects having different spatial coverages in an area, homogeneity features have been computed with only two different window sizes of 3 × 3 and 5 × 5 on the AHS green band. Indeed, this latter band gave appreciable visual results, and we preferred to avoid using a redundant number of features to feed the following supervised NN. An analysis of the signature over different types of imagery (AHS and QuickBird) and Mediterranean locations (mainly, the Madrid and Rome areas have been considered) revealed a general good stability, particularly as far as the homogeneity parameters are concerned, while the spectral content, as expected, showed some conditioning with the particular building roof material. Therefore, the data set that constituted the input for the unsupervised learning was formed by six bands: the blue, green, red, NIR, and homogeneity of band 2 (from a 3 × 3 window) and the homogeneity-bis of band 2 (from a 5 × 5 window). The values of the input vector were normalized to the range of $(-1, +1)$: This step was necessary, considering that different measure units (radiance and homogeneity) were used. The extracted building pixels have then been used as the training set of a supervised NN performing the ultimate building detection. The mixed approach has been made to improve the classification accuracy observed using only unsupervised methodologies.

A. Network Topologies

Generally, the optimal architecture of an NN is not known *a priori* for most real-world problems: In this case, the best results (confirmed also by the tests on QuickBird images) have been obtained with a map of 5 × 5 neurons, which seems to be the optimal architecture to detect buildings. It means that each input signal has been compared with all the 25 neurons/weights of the map (where the initial weights of these neurons were set with random values). The winning neuron has been chosen with a minimum distance algorithm based on the Euclidean distance between the input and the neuron. After 100 epochs, the unsupervised training was completed, and the final weights of the 25 clusters were fixed: Between these 25 clusters, two of them could be distinguished as buildings. 6000 pixels from each clusters have been randomly chosen, and they have been sent to the supervised NNs. Once the training sets have been defined, they have been joined together as the input to the supervised NN. A very simple network topology has been chosen, with six inputs, a single hidden layer with eight neurons, and two outputs: After the training of the new network, the entire image has been classified into two classes: buildings and no buildings (see center of Fig. 2).

IV. LST Retrieval

In all environmental studies, the LST is one of the key parameters which can be retrieved from TIR remotely sensed data. Most of the energy detected by the sensor in this spectral region is directly emitted by the land surface: Hence, except for the solar irradiance components, most of the fluxes at the surface/atmosphere interface can only be parameterized by means of the surface temperature. Several techniques have been proposed to retrieve the LST from thermal infrared data. In this letter, the LST has been obtained considering a split window (SW) technique [14]: The basis of the technique is that the radiance attenuation for atmospheric absorption is proportional to the radiance difference of simultaneous measurements at two different wavelengths, each subject to different amounts of atmospheric absorption. The SW technique uses two thermal bands typically located in the atmospheric window between 10 and 12 μm. In [15], a general formula to retrieve the LST for AHS high flights has been defined, and the coefficients have been readapted in [16]. It is based on a combination between the AHS bands 75 (10.07 μm) and 79 (12.35 μm)

$$T_s = T_{75} + 0.723(T_{75} - T_{79}) + 0.04275(T_{75} - T_{79})^2 - 0.08463$$
$$+ (45.49 - 5.17W)(1 - \varepsilon) + (-60.81 + 16.93W)\Delta\varepsilon \quad (2)$$

where T_{75} and T_{79} are the at-sensor brightness temperatures at the thermal bands (in kelvin), ε is the mean emissivity of two bands (ε_{75} and ε_{79}), $\Delta\varepsilon$ is the emissivity difference $\varepsilon_{75} - \varepsilon_{79}$, and W is the total atmospheric water vapor content (in grams per square centimeter). From research in the literature, it seemed that, nowadays, there are no developed methods, such as the use of an AHS without specific atmospheric correction and without the use of ancillary data (i.e., a radiosonde), to calculate the LST from the image data. On the other hand, the entire technique presented in this letter aims at increasing the level of automatism in extracting the building temperature considering only the processing of the image data. In fact, one of the purposes of this letter is to investigate on the price to be paid in terms of error on the sought final quantities for increasing the level of automatism. In the next paragraphs, we explain how we tried to estimate both the emissivity and water vapor content directly from the used AHS image. However, first of all, georeferencing and a conversion from radiance L_i (in milliwatts per steradian square meter micrometer) to temperature T_i (in kelvin) for thermal bands have been applied with Planck's formula considering the following:

$$T_i = \frac{1}{(a + b * \ln(L_i))} \quad (3)$$

where a and b are optimized coefficients (see Table II) provided by the INTA for AHS flights.

A. Emissivity

The land surface emissivity (ε) is a proportionality factor that scales the blackbody radiance (Planck's law) to predict the emitted radiance, and it is the efficiency of the transmitting thermal energy across the surface into the atmosphere. An automatic way to calculate this parameter is with the *normalized difference vegetation index (NDVI) threshold method* $NDVI^{THM}$ [17], whose assessment by comparison with other

TABLE II
APPLIED COEFFICIENTS TO CONVERT
RADIANCE TO TEMPERATURE (INTA)

a coefficients	b coefficients	Error
8,62823E-03	-5,78661E-04	band 71 max. error [K] = 0.03
8,94013E-03	-6,10497E-04	band 72 max. error [K] = 0.03
9,24115E-03	-6,42080E-04	band 73 max. error [K] = 0.04
9,52452E-03	-6,72616E-04	band 74 max. error [K] = 0.05
9,79788E-03	-7,02837E-04	band 75 max. error [K] = 0.05
1,00895E-02	-7,35967E-04	band 76 max. error [K] = 0.06
1,04274E-02	-7,75533E-04	band 77 max. error [K] = 0.08
1,07362E-02	-8,12848E-04	band 78 max. error [K] = 0.09
1,10336E-02	-8,49946E-04	band 79 max. error [K] = 0.10
1,13322E-02	-8,88480E-04	band 80 max. error [K] = 0.12

more complex methods has already been carried out in [16]. It has to be observed that these latter complex methods rely on *temperature emissivity separation* (TES) and *temperature-independent spectral index* algorithms which, assuming an atmospheric correction of the data by means of radiative transfer models such as MODTRAN [18], perform better results but do not permit a high level of automation.

B. Water Vapor

The knowledge of the total atmospheric water vapor content is necessary to improve the precision of the estimates of the LST obtained from satellite data by means of split-window algorithms. Radiosonde data represent one way to obtain the atmospheric water vapor content: However, this is not always possible, also considering that the radiosonde observations are not carried out in a systematic way. Considering the fact that there are no methods to calculate water vapor from AHS data without using radiation transfer models, a methodology has been proposed based on the split-window technique [19], [20] and starting from the following expression:

$$W = a(T_i - T_j) + b \qquad (4)$$

where W is the atmospheric water vapor content in grams per square centimeter and T_i and T_j are the at-sensor brightness temperatures measured in two different thermal bands (in kelvin). With a regression technique, it was possible to recalculate the coefficients a and b for application to the AHS flights. The AHS sensors has ten thermal bands (71 to 80, with effective wavelengths of 8.190, 8.660, 9.170, 9.600, 10.080, 10.560, 11.160, 11.720, 12.320, and 12.890 μm), so different combinations can be used (excluding band 80, which was affected by stripes). With the radiosoundings launched from the Barajas Airport, the regression has been performed with eight ground-truth data synchronized with the AHS flights from the 25th of June [day (D) and night (N)], 26th of June (N), 28th of June (D–N), 1st of July (D), 2nd of July (N), and 4th of July (D). In Table III, the obtained correlations considering all the possible coefficient combinations are shown: The highest correlation coefficient (0.963) corresponds to the 74–76 bands, with an a coefficient of 1.5073 and a b coefficient of 2.7465. The water vapor map on the test image has then been extracted using these coefficients: The obtained mean value was 0.1 g/cm^2 higher than the ground-truth data.

Now, all the terms in (2) have been calculated, and the LST map can be yielded.

TABLE III
CORRELATION VALUES OF WATER VAPOR
RETRIEVAL FOR EACH BAND COMBINATION

	71	72	73	74	75	76	77	78	79
71		0.447	0.470	0.551	0.564	0.620	0.654	0.749	0.030
72			0.587	0.751	0.855	0.898	0.222	0.083	0.332
73				0.000	0.837	0.480	0.058	0.202	0.363
74					0.617	0.963	0.105	0.276	0.413
75						0.105	0.368	0.386	0.445
76							0.485	0.449	0.475
77								0.794	0.475
78									0.498
79									

TABLE IV
IMPROVEMENT IN CLASSIFICATION ACCURACY
WITH THE PROPOSED METHOD

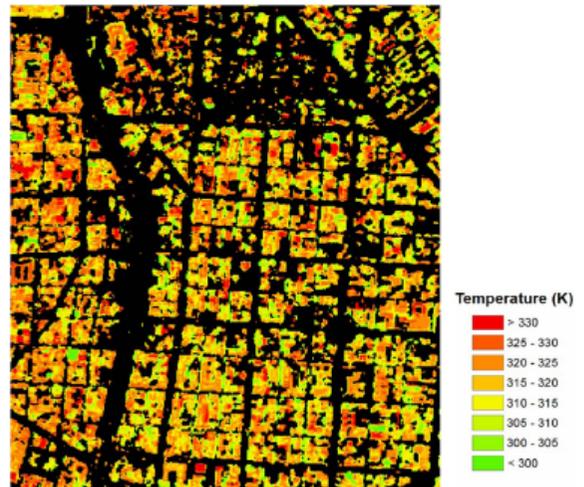| Classification Method	Accuracy in building detection
4 Spectral bands -Unsupervised-	60. 2 %
4 Spectral bands + Homogenity 3 x 3 -Unsupervised-	70.65 %
4 Spectral bands, Homogenity 3 x 3 and 5 x 5 -Unsupervised-	74. 825 %
4 Spectral bands, Homogenity 3 x 3 and 5 x 5 -Unsupervised + Supervised-	**83.7 %**

Fig. 3.   Detail of building temperature map.

V. RESULTS

To verify the accuracy of the building extraction method, 2500 pixels for each class have been chosen to represent the ground truth: The overall accuracy of the classification reached a satisfactory level, with a value of 83.7%. In Table IV, the improvement with the used method, adding different texture windows and the supervised step, is shown. The method picked up some pixels which are not covered by buildings; in particular, there are some misclassification errors with the roads and bare soil. Afterward, the building mask and the LST map have been merged to display the building surface temperature map (see right of Fig. 2 and a detail in Fig. 3). From the first qualitative analysis, we distinguished a colder area in the north part of the swath (which is located at a higher altitude) and individuated the areas of the city characterized by higher temperature values,
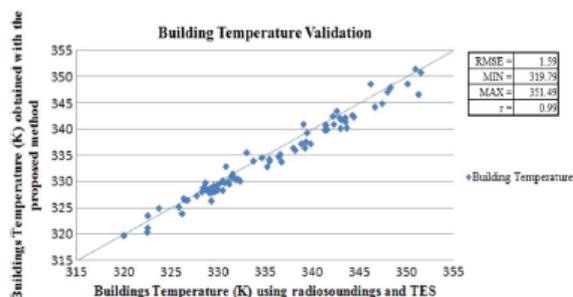
Fig. 4. Validation of building temperature values obtained using radiosoundings and emissivity from TES algorithm versus building temperature values obtained using split window for water vapor and $NDVI^{THM}$.

such as an industrial site in the southern part. A more quantitative assessment was performed considering the AHS LST values. Due to an insufficient number of specific ground truths in the correspondence of buildings during the AHS flight, the assessment was carried out, relying on the TES method for the emissivity and on the radiosounding measurements for the water vapor estimation [15], as a benchmark. We remind that this latter method is considered as the "state of the art" of the LST retrieval. The statistics of the results is shown in Fig. 4. On a set of 80 buildings randomly selected over the image mask, a root-mean-square error (rmse) value of 1.59 K has been obtained. A generally good correlation (0.99) between the estimated values obtained with the proposed method and the values obtained with the radiosoundings and TES method can be observed.

VI. CONCLUSION

In this letter, an image processing methodology stemming from the fusion of different and last generation remote sensing techniques, such as the automatic image classification with NNs and LST retrievals from hyperspectral data, has been presented. The final result is an original scheme for the production of building temperature maps over urban areas. On its turn, the image classification is achieved through an innovative mixed approach where the buildings are extracted automatically by combining Kohonen's SOM and an MLP network. In both topologies, either the spectral or textural features of the AHS images have been considered. With the thermal information acquired from the same airborne scanner, the building surface temperature map at high spatial resolution has been yielded in the optic to reduce, as much as possible, the use of ancillary information (i.e., radiosoundings and atmospheric correction) usually necessary for more standard techniques. The method has been applied to an image taken over the city of Madrid, and the obtained results show an rmse of 1.59 K if a comparison with the TES method is performed. Considering that an increased use of hyperspectral imagery for UHI management can be foreseen for the next years, the developed technique may contribute to improve the characterization of the building heat transfer processes in the more general context of energy balance models [21].

REFERENCES

[1] J. R. Jensen and D. C. Cowen, "Remote sensing of urban/suburban infrastructure and socio-economic attributes," *Photogramm. Eng. Remote Sens.*, vol. 65, no. 5, pp. 611–622, May 1999.
[2] S. Kato and Y. Yamaguchi, "Estimation of storage heat flux in an urban area using ASTER data," *Remote Sens. Environ.*, vol. 110, no. 1, pp. 1–17, 2007.
[3] J. A. Voogt and T. R. Oke, "Thermal remote sensing of urban climates," *Remote Sens. Environ.*, vol. 86, no. 3, pp. 370–384, Aug. 2003.
[4] H. Tran, D. Uchihama, S. Ochi, and Y. Yasuoka, "Assessment with satellite data of the urban heat island effects in Asian mega cities," *Int. J. Appl. Earth Observation Geoinformation*, vol. 8, no. 1, pp. 34–48, Jan. 2006.
[5] H. Bishof, W. Schneider, and A. J. Pinz, "Multispectral classification of Landsat-images using neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 30, no. 3, pp. 482–490, May 1992.
[6] C. Bishop, *Neural Networks for Pattern Recognition*. New York: Oxford Univ. Press, 1995.
[7] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, "Use of neural networks for automatic classification from high-resolution imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 800–809, Apr. 2007.
[8] F. Pacifici, M. Chini, and W. J. Emery, "A neural network approach using multi-scale textural metrics from very-high resolution panchromatic imagery for urban land-use classification," *Remote Sens. Environ.*, vol. 113, no. 6, pp. 1276–1292, Jun. 2009.
[9] T. Kohonen, *Self-Organizing Maps*, 3rd ed. Berlin, Germany : Springer-Verlag, 2001.
[10] D. Chanchal and B. Chakraborty, "Acoustic characterization of seafloor sediment employing a hybrid method of neural network architecture and fuzzy algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 743–747, Oct. 2009.
[11] M. Datcu, H. Daschiel, A. Pelizzari, M. Quartulli, A. Galoppo, A. Colapicchioni, M. Pastori, K. Seidel, P. G. Marchetti, and S. D'Elia, "Information mining in remote sensing image archives: System concepts," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 12, pp. 2923–2936, Dec. 2003.
[12] M. Lazzarini and F. Del Frate, "Features extraction from hyperspectral images: An approach based on spectral, textural and spatial information applied to urban environments," in *Proc. Hyperspectral Workshop, ESA-ESRIN*, Frascati, Italy, Mar. 17–19, 2010.
[13] M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst, Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.
[14] L. M. McMillin, "Estimation of sea surface temperatures from two infrared window measurements with different absorption," *J. Geophys. Res.*, vol. 80, no. C36, pp. 5113–5117, 1975.
[15] J. A. Sobrino, C. Jiménez-Muñoz, P. J. Zarco-Tejada, G. Sepulcre-Cantó, and E. De Miguel, "Land surface temperature derived from airborne hyperspectral scanner thermal infrared data," *Remote Sens. Environ.*, vol. 102, no. 1/2, pp. 99–115, 2006.
[16] J. A. Sobrino *et al.*, "The Dual-use European Security IR Experiment (Desirex 2008) a Study of the Urban Heat Island in Madrid," presented at the *ESA Living Planet Symp.*, Bergen, Norway, Jun. 28–Jul. 2, 2010.
[17] J. A. Sobrino, J. C. Jiménez-Muñoz, G. Sòria, M. Romaguera, L. Guanter, J. Moreno, A. Plaza, and P. Martiacute;nez, "Land surface emissivity retrieval from different VNIR and TIR sensors," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 2, pp. 316–327, Feb. 2008.
[18] A. Berk, G. P. Anderson, P. K. Acharya, M. L. Hoke, J. H. Chetwynd, L. S. Bernstein, E. P. Shettle, M. W. Matthew, and S. M. Adler-Golden, *MODTRAN4 Version 3 Revision 1 Users's Manual*. Bedford, MA: Air Force Res. Lab., Hanscom AFB, 2003.
[19] K. P. Czajkowski, S. N. Goward, D. Shirey, and A. Walz, "Thermal remote sensing of near-surface water vapor," *Remote Sens. Environ.*, vol. 79, no. 2/3, pp. 253–265, 2002.
[20] J. C. Jiménez-Muñoz and A. Sobrino, "Atmospheric water vapour content retrieval from visible and thermal data in the framework of the DAISEX campaigns," *Int. J. Remote Sens.*, vol. 26, no. 15, pp. 3163–3180, Aug. 2005.
[21] V. Masson, "A physically-based scheme for the urban energy balance in atmospheric models," *Bound.-Layer Meteorol.*, vol. 94, no. 3, pp. 357–397, 2000.