Remote Sensing of Environment 113 (2009) 1276–1292

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification

Fabio Pacifici [a,*], Marco Chini [b], William J. Emery [c]

[a] Earth Observation Laboratory, Tor Vergata University, Rome, Italy
[b] Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy
[c] Colorado University, Boulder, Colorado, USA

ARTICLE INFO

Article history:
Received 31 July 2008
Received in revised form 13 February 2009
Accepted 14 February 2009

Keywords:
Extended pruning
Multi-scale textural analysis
Neural networks
QuickBird
Urban land-use
Very high-resolution panchromatic data
WorldView-1

ABSTRACT

The successful launch of panchromatic WorldView-1 and the planned launch of WorldView-2 will make a major contribution towards the advancement of the commercial remote sensing industry by providing improved capabilities, more frequent revisits and greater imaging flexibility with respect to the precursor QuickBird satellite. Remote sensing data from panchromatic systems have a potential for more detailed and accurate mapping of the urban environment with details of sub-meter ground resolution, but at the same time, they present additional complexities for information mining.

In this study, very high-resolution panchromatic images from QuickBird and WorldView-1 have been used to accurately classify the land-use of four different urban environments: Las Vegas (U.S.A.), Rome (Italy), Washington D.C. (U.S.A.) and San Francisco (U.S.A.). The proposed method is based on the analysis of first- and second-order multi-scale textural features extracted from panchromatic data. For this purpose, textural parameters have been systematically investigated by computing the features over five different window sizes, three different directions and two different cell shifts for a total of 191 input features. Neural Network Pruning and saliency measurements made it possible to determine the most important textural features for sub-metric spatial resolution imagery of urban scenes.

The results show that with a multi-scale approach it is possible to discriminate different asphalt surfaces, such as roads, highways and parking lots due to the different textural information content. This approach also makes it possible to differentiate building architectures, sizes and heights, such as residential houses, apartment blocks and towers with classification accuracies above 0.90 in terms of Kappa coefficient computed over more than a million independent validation pixels.

1. Introduction

Human activity truly dominates the Earth's ecosystems with structural modifications. Rapid population growth over recent decades and the concentration of this population in and around urban areas have significantly impacted the environment (Pacifici et al., 2007). Although urban areas represent a small fraction of the land surface, they affect large areas due to the magnitude of the associated energy, food, water and raw material demands. Urban areas are undergoing dynamic changes and, as a consequence, are facing new spatial and organizational challenges as they seek to manage local urban development within a global community. Sub-urbanization refers to a highly dynamic process where rural areas, both close to, but also distant from, city centers become enveloped by, or transformed into, extended metropolitan regions. These areas are a key interface between urban and rural areas due to the provision of essential services in both (GLP, 2005).

During the last two decades, significant progress has been made in developing and launching satellites with instruments, in both the optical/IR and microwave regions of the spectra, well suited for Earth observation with an increasingly finer spatial and spectral resolutions (Bamler and Eineder, 2008; ESA, 2006; DigitalGlobe, 2008a; DigitalGlobe, 2008b). The successful launch of WorldView-1 in September 2007 and the planned launch of WorldView-2 have made and will make a major contribution towards the advancement of the commercial remote sensing industry by providing greater capability, more frequent revisits and greater imaging flexibility with respect to the precursor QuickBird satellite. WorldView-1 is the world's highest spatial resolution commercial imaging satellite with 50 cm resolution at nadir (DigitalGlobe, 2008a). Acquisitions are panchromatic in the spectral range of 400–900 nm. Operating at an altitude of 496 km, WorldView-1 has an average revisit time of 1.7 days (at 1 m Ground Sample Distance—GSD) and 5.9 days at 20° off-nadir (0.51 m GSD). QuickBird collects both multi-spectral and panchromatic imagery concurrently (DigitalGlobe, 2008b). The QuickBird panchromatic resolution ranges from 0.61 m at nadir to 0.72 m at 25° off-nadir with a bandwidth which spans from 450 to 900 nm.

* Corresponding author.
E-mail address: f.pacifici@disp.uniroma2.it (F. Pacifici).

0034-4257/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2009.02.014

Remote sensing data from these panchromatic systems has a potential for more detailed and accurate mapping of the urban environment with details of sub-meter ground resolution, but at the same time, they present additional problems for information mining. In fact, even though they have higher spatial resolutions, they are rarely exploited for urban classification due their complexity and the missing multi-spectral information. Taking into account this latter aspect, surfaces such as water and asphalt exhibit very similar values in panchromatic data, resulting in misclassification errors, as reported later in this paper.

Urban areas are composed of various materials (such as concrete, asphalt, metal, plastic, glass, shingles, water, grass, shrubs, trees and soil) arranged by humans in complex ways to build housing, transportation systems, utilities, commercial buildings and recreational areas (Carleer & Wolff, 2006). A simple building appears as a complex structure with many architectural details surrounded by gardens, trees, other buildings, roads, social and technical infrastructure and many temporary objects, such as cars, buses or daily markets. Therefore, the single panchromatic image is generally not thought suitable for land-use mapping of such a complex scenes. For these reasons it is necessary to extract additional information from panchromatic images in order to recognize objects within the scenes, such as texture or objects' shape.

The objective of this work is to systematically analyze the textural characteristics of very high-resolution panchromatic imagery to classify the land-use of different urban environments, helping end-users to choose more effectively the input parameters for successive classifications. Possible practical applications may be found in urban planning (Shackelford & Davis, 2003; Benediktsson et al., 2003), but also for crisis management or as support for civil protection activities (Pagot & Martino Pesaresi, 2008; Chini et al., 2009; Chini et al., 2008a). In fact, in destructive disasters, such as an earthquake, a prompt and accurate overview of the damage of the human settlements is very important to be able to manage the rescue efforts and, subsequently, to organize restoration activities (Voigt et al., 2007; Chini et al., 2008b).

The proposed method is based on the analysis of first- and second-order multi-scale textural features extracted from panchromatic data. To account for the spatial setting of cities, textural parameters have been computed over five different window sizes, three different directions and two different cell shifts for a total of 191 input features. Note, in this context we used the term feature to indicate any input, such as textural measures or the panchromatic band used to feed the network. Neural network pruning and saliency measurements made it possible to determine the most important textural features for sub-metric spatial resolution imagery of urban scenes.

The paper is organized as follows: related work on texture analysis and feature selection is outlined in Section 2, while the datasets are described in Section 3. Section 4 deals with methodology, introducing the multi-scale textural analysis and the concepts of neural pruning and extended pruning. Experimental results of the classification exercises and the analysis of the textural feature contributions are discussed in Section 5. Final conclusions follow in Section 6.

2. Related work

2.1. Texture analysis

In the past years, the value of adding textural features to the classification of satellite images has been clearly demonstrated as a method to overcome the lack of spectral resolution (Carleer & Wolff, 2006). 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. Tuceryan and Jain (1993) identify four major categories of texture feature analysis methods: i) statistical (such as those based on the computation of the gray-level co-occurrence matrix—GLCM), ii) geometrical (including structural), iii) model-based, such as Markov random fields (MRF), and iv) signal processing (such as Gabor filters). It was pointed out by Shanmugan et al. (1981) that textural features derived from GLCM are the most useful for analyzing the content of a variety of remote sensing imagery while, according to Treitz et al. (2000), statistical texture measures are more appropriate than structural for traditional land-cover classification. Recently, Clausi and Yue (2004) 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 in Haralick et al. (1973) and Haralick (1979), some of which are strongly correlated with each other (Cossu, 1988). In their investigation of the textural characteristics associated with gray-level co-occurrence matrix statistical parameters, Baraldi and Parmiggiani (1995) concluded that two parameters, energy and contrast, are the most significant in discriminating between different textural patterns.

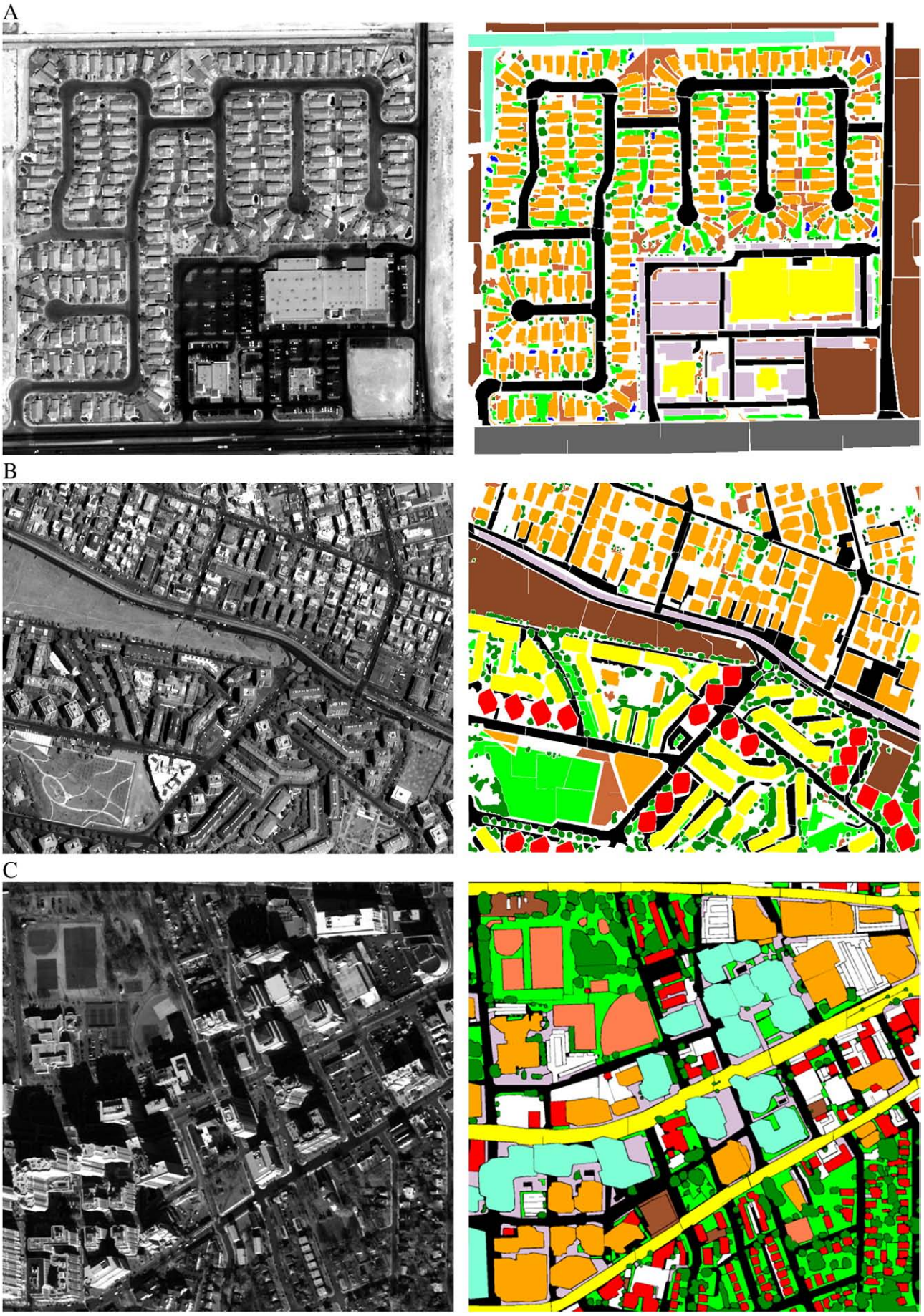
In the remote sensing literature, many examples of the use of textural parameters have been proposed for the extraction of quantitative information of building density (Karathanassi et al., 2000) or for the recognition of different urban patterns (Zhang et al., 2003). In most cases, texture increased the per-pixel classification accuracy, especially in urban areas where the images are more heterogeneous (Puissant et al., 2005). Chen et al. (2004) 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. The window size used for the texture computation plays an important role in texture extraction: larger window sizes are more appropriate for finer resolution than coarse resolution (Shaban & Dikshit, 2001). The inclusion of a single texture feature in the original input space appeared to improve the overall classification accuracy significantly (Shaban & Dikshit, 2002). Considerations about the relevance of textural parameters have been proposed by Dekker (2003) who found that variance, weighted-rank fill ratio, and semi-variogram performed better than other textural features. Furthermore, the latter were chosen as an alternative for the gray-level co-occurrence matrix because it gave better results with SAR data. At this point, it is important to point out that the majority of the studies that appeared in literature in the past years have dealt with decametric spatial resolution imagery. Consequently, the resulting classification maps are generally representative of different terrain patterns and not of single objects within the image, such as a building or a swimming pool (Pesaresi, 2000).

With the increase of the spatial resolution of satellites such as QuickBird and Ikonos, texture features turn out to be valuable for the

Table 1
Characteristics of the images used.

Site information		Image information				
Location	Dimension (pixels)	Satellite	Date	Spatial res. (m)	View angle (°)	Sun elev. (°)
Las Vegas (U.S.A.)	755 × 722	QuickBird	May 10, 2002	0.6	12.8	65.9
Rome (Italy)	1188 × 973	QuickBird	July 19, 2004	0.6	23.0	63.8
Washington D.C. (U.S.A.)	1463 × 1395	WorldView-1	Dec. 18, 2007	0.5	27.8	24.9
San Francisco (U.S.A)	917 × 889	WorldView-1	Nov. 26, 2007	0.5	19.6	29.6

The dataset takes into account four different architectures from cities with diverse urban structures.


A
B
C

extraction of trees within a scene (Ouma et al., 2006) and for the retrieval of the leaf area index (Colomboa et al., 2003). In this latter case, textural information is related to the spatial canopy architecture. Multi-spectral Ikonos images have been studied for road identification with an approach that integrates an unsupervised classifier, fuzzy logic and the angular texture signature (Zhang & Couloigner, 2006), where a number of shape descriptors are proposed for the angular texture signature that are successfully used to separate roads from parking lots/buildings.

The GLCM method is widely accepted for classifying texture and several studies have used it for land-cover classification with decametric Synthetic Aperture Radar (SAR) data, such as Kurosu et al. (1999), Kurvonen and Hallikainen (1999) and Arzandeh and Wang (2002). Dellepiane et al. (1991) studied the combined use of backscattering intensity and its textural information, while Dell'Acqua and Gamba (2006) proposed a multi-scale textural approach for urban characterization. The coarse spatial resolution of the European Space Agency's European Remote Sensing satellites, ERS-1 and ERS-2, allowed the recognition of dense, residential, and sub-urban areas with sufficient precision and stability in the classification maps (Dell'Acqua & Gamba, 2003). Finally, the combined use of medium optical resolution imagery, SAR (and its texture features) has been exploited for mapping urbanization by Tatem et al. (2005).

2.2. Feature selection

If on one hand more information may be helpful for the classification process, on the other hand, the increasing number of input features may introduce additional complexity related to the increase of computational time and the "curse of dimensionality" (Bishop, 1995), overwhelming the expected increase in class separability associated with the inclusion of additional features. As a result, it is necessary to use a robust method to estimate parameters (Chini et al., 2008c) or to reduce the input dataset, as for example, using principal component analysis to diminish the number of inputs (Landgrebe, 2003). A large number of input features can be used in satellite image classification. Generally, not all of these input features are equally informative. Some of them may be redundant, noisy, meaningless, correlated or irrelevant for the specific task. Intuitively, any classification approach should include only those features that make significant contributions, which should result in better performance (Verikas & Bacauskiene, 2002). Under this scenario, Feature Selection (FS) aims at selecting a subset of the features, which is relevant for a given problem. The importance of FS is the potential for speeding up the processes of both learning and classifying since the amount of data or processes is reduced (Leray & Gallinari, 1998).

A wide number of approaches for FS has been proposed in the literature (Batitti, 1994; Cibas et al., 1996; Foroutan & Sklansky, 1987; Holz & Loew, 1994; Karthaus et al., 1995; Kittler, 1980; Kudo & Shimbo, 1993; Stahlberger & Riedmiller, 1997). We cannot be exhaustive in our review since the literature in this field is very rich and it is not the topic of this paper, but the major ideas are described here. Mainly, FS methods can be divided into two classes: those based on statistical information about features and those based on classification approaches. The former determines an optimal subset of features independently from the classification process, based on statistical criteria. The latter removes useless features or selects the most relevant of them based on specific criteria (Onnia et al., 2001).

When the only source of available information is provided by training data, the feature selection task can be performed using non-parametric approaches such as support vector machines (SVMs) or neural network (NNs). Several specific methods have been proposed for SVMs. A well-known embedded backward selection algorithm is the Recursive Feature Elimination (RFE) that uses the changes in the decision function of the SVM as criterion for the selection (Guyon et al., 2002; Weston et al., 2003). A genetic algorithm was proposed by Bazi and Melgani (2006) to select an optimal subset of features for a successive SVM classification. An example of feature selection with Multi-Layer Perceptron neural networks can be found in Del Frate et al. (2005). One potential advantage of neural networks for feature selection is that this approach can simultaneously "optimize" both the input feature set and the classifier itself, while other methods select the "best" subset of features with respect to a fixed classifier (Mao et al., 1994). More details on the use of Multi-Layer Perceptron neural networks for feature selection are provided in the next section.

2.2.1. Feature selection with Multi-Layer Perceptron neural network

The Multi-Layer Perceptron (MLP) neural network architecture can be conceptually divided into three logically separable parts: the input layer, the hidden layers and the output layer. The first and the last are generally application dependent. For example, in satellite image classification, the input dimension is the number of sensor bands, and the output dimension is the number of predefined classes (Zeng & Yeung, 2006), while the choice of the hidden layers (and numbers of neurons in each hidden layer) are generally based on some heuristic concept or user experience. Generally, a neural network with a small architecture may not be able to capture the underlying data structure. In fact, if the network has an insufficient number of free parameters (weights), it underfits the data, leading to excessive biases in the outputs (Chandrasekaran et al., 2000). When the number of neurons in the hidden layers increases, the network can learn more complex data patterns by locating more complex decision boundaries in feature space. However, a neural network with an architecture too large may fit the noise in the training data, leading to good performance on the training set but rather poor accuracy relative to the validation set, resulting in large variance in the predicted outputs and poor accuracy (Del Frate et al., 2007). Generally speaking, the larger the network, the lower its generalization capabilities, and the greater the time required for training networks (Zeng & Yeung, 2006).

Unfortunately, the optimal architecture is not known in advance for most real-world problems. Consequently, there are two common ways to find the desired network size: growing and pruning approaches. The first consists in starting with a small network and adding neurons until the optimization criteria are reached (Hirose et al., 1991). However, this approach may be sensitive to initial conditions and become trapped in local minima. The second approach consists of beginning with a large network and then removing connections or units that are identified as less relevant. This approach has the advantage that the network is less sensitive to initial conditions. Moreover by reducing network size, it improves its generalization capabilities when applied to new data (Kavzoglu & Mather, 1999).

Feature selection with neural networks can be seen as a special case of architecture pruning, where input units are pruned rather than hidden units or single weights. Many extended pruning procedures for removing input features have been proposed in the literature (Belue & Bauer, 1995; Cibas et al., 1996; Del Frate et al., 2005). Although many different pruning methods have been proposed in the literature, the main ideas underlying most of them are similar. They all establish a reasonable relevance measure, namely saliency, for the specific element (unit or weight) so that the pruning process has the least effect on the performance of the network. Pruning methods can be divided into three wide groups in terms of decision criteria for the removing of weights or nodes: i) sensitivity-based, ii) penalty-term approaches, and iii) others which may include interactive pruning, bottleneck

Fig. 1. Panchromatic images (left) and ground reference (right) of (A) Las Vegas, (B) Rome, and (C) Washington D.C. Color codes are in Tables 2–4, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method, or pruning by genetic algorithms. In the first method, the sensitivity is estimated by the error function after the removal of a weight or unit. Then, the less significant element (weight or unit) is removed. In penalty-term methods, one adds terms to the objective function that rewards the network for choosing efficient solutions. There is some overlap in these two groups since the objective function could include sensitivity terms (Chandrasekaran et al., 2000). The pruning problem has been also formulated in terms of solving a linear equation system (Castellano et al., 1997). Sietsma and Dow (1998) developed a two-stage procedure, which removes the less relevant hidden neurons and adjusts the remaining weights to maintain the behavior of the original network. Suzuki et al. (2001) proposed a pruning technique on the basis of the influence of removing units in order to determine the structures of both the input and the hidden layers. Reed (1993) and Engelbrecht and Cloete (1996) have given detailed surveys of pruning methods. More details on the pruning approach used in this paper are reported later in the paper.

3. Datasets

The datasets used take into account four different cities with diverse architectural urban structures: Las Vegas, Rome, Washington D.C. and San Francisco. The first two scenes were acquired by QuickBird in 2002 and 2004, respectively, the others by WorldView-1 in 2007. Details of sites and images are reported in Table 1.

3.1. Description of the scenes

The Las Vegas scene, shown in Fig. 1A, contains regular criss-crossed roads and examples of buildings with similar heights (about one or two floors) but different dimensions, from small residential houses to large commercial structures. This first scene was chosen for two reasons. First, its simplicity and regularity allowed an easier analysis and interpretation of the textural features. Second, it represents well a common American sub-urban landscape, including small houses and large roads, which is different from the European style of old cities built with more complex structures. To take into account this last situation, a second test area, shown in Fig. 1B, was used including a sub-urban scene of Rome composed of a more elaborate urban lattice with buildings showing a variety in heights (from four floors to twelve), dimensions and shapes including apartment blocks and towers. In particular, the Rome area has two completely different urban architectures separated by a railway. The area located in the upper right of the scene was built during the 60s: buildings are very close to each other and have a maximum of five floors, while roads are narrow and usually show traffic jams due to the presence of cars and busses. The other side of the railway was developed during the 80s and 90s: buildings have a variety of architectures, from apartment blocks (eight floors) to towers (twelve floors), while roads are wider than those on the other side of the railroad tracks. The Washington D.C. scene, shown in Fig. 1C, contains elements that characterize the other two, but imaged with a higher spatial resolution (0.5 m). Buildings have different heights, dimensions and shapes, varying from small residential houses to huge structures with multiple floors (more than twenty), while asphalt surfaces include roads with different widths (e.g. residential and highways) and parking lots. The image of San Francisco presents regular structures, such as a highway, residential roads, two different type of buildings, commercial/industrial and residential, some sparse trees and vegetated areas. This scene has been used for validation purposes only and it will be better described later in this paper.

3.2. Classes, training and validation set definition

Several different surfaces of interest have been identified many of which are particular to the specific scene. For the Las Vegas case, one goal was to distinguish the different uses of the asphalt surfaces, which included *Roads* (i.e. roads that link different residential houses), *Highways* (i.e. roads with more than two lanes) and *Parking Lots*. An unusual structure within the scene was a *Drainage Channel* located in the upper part of the image. This construction showed a shape similar to roads, but with higher brightness since it was built with concrete. A further discrimination was made between *Residential Houses* and *Commercial Buildings* due to the different size, and between *Bare Soil* (terrain with no use) and *Soil* (generally, backyards with no vegetation cover). Finally, more traditional classes, such as *Trees*, *Short Vegetation* and *Water* were added for a total of eleven classes of land-use. The areas of shadow were very limited in the scene since the modest heights of buildings and relative sun elevation.

Due to the dual nature of the architecture of the Rome test case and the high off-nadir angle, of about 23°, the selection of the classes was made to investigate the potential of discriminating between structures with different heights, including *Buildings* (structures with a maximum of 5 floors), *Apartment Blocks* (rectangular structures with a maximum of 8 floors) and *Towers* (more than 8 floors). As for the previous case, other surfaces of interest were recognized, including *Roads*, *Trees*, *Short Vegetation*, *Soil* and the peculiar *Railway* for a total of nine classes. Differently from the previous case, in this scene shadow occupies a larger portion of the image.

The Washington D.C. scene has features in common with the previous two. Particularly, it is possible to distinguish different uses of asphalt surfaces, such a *Roads*, *Highway* and *Parking Lots*, while buildings show a variety of dimensions and heights, from small *Residential Houses* in the bottom-right of the scene, to *Tall Buildings* with multiple floors in the center of the area. An interesting feature is the role of the class *Trees*. The image was acquired in December and most of the plants were without leaves. Therefore, we did not associate these objects with the class *Trees*, but to a wider class *Vegetation* (including short vegetation and trees without leaves), and only trees with leaves were recognized belonging to *Trees*. Finally, the classes *Sport Facilities* and *Side Walks* were added for a total of 11 classes. It is important to highlight that similar to the Rome case, the image was acquired with a high off-nadir angle of about 28°, and with a sun elevation of about 25°, which caused large shadows.

The ground references for each scene, reported in Fig. 1, have been obtained by careful visual inspection of separate data sources, including aerial imagery, cadastral maps and *in situ* inspections (for the Rome scene only). An additional consideration regards objects within shadows that reflect little radiance because the incident illumination is occluded. Textural features can potentially be used to characterize these areas as if they were not within shadow. Therefore,

Table 2
Las Vegas case: classes, training and validation samples, and color legend.

Las Vegas classes	TR	VA
Bare soil	4255	44,675
Commercial buildings	1822	19,126
Drainage channel	1143	12,001
Highways	2836	29,774
Parking lots	2257	23,695
Residential houses	7007	73,563
Roads	6098	64,023
Short vegetation	1793	18,823
Soil	1472	15,437
Trees	1043	10,945
Water	118	1236
Total ROIs	29,844	313,298

Table 3
Rome case: classes, training and validation samples, and color legend.

Rome classes	TR	VA
Bare soil	4127	38,572
Apartment blocks	20,472	44,672
Buildings	27,188	77,034
Railway	2606	6727
Roads	35,531	69,002
Soil	3506	5776
Tower	9187	19,365
Trees	13,632	38,624
Short vegetation	10,443	29,587
Total ROIs	126,692	329,359

these surfaces were assigned to one of the corresponding classes of interest described above.

When classifying imagery at sub-meter spatial resolution, many of the errors may occur in the boundaries between objects. On the other hand, it is also true that the nature of the objects is fuzzy and often it is not possible to correctly identify an edge. To investigate this effect, we defined the first two ground references (Las Vegas and Rome) not including boundary areas, and the other (Washington D.C.) minimizing the extensions of these regions.

In order to select training and validation samples, having both statistical significance and avoiding the correlated neighboring pixels, we have adopted a stratified random sampling (SRS) method, ensuring that even small classes, such as water or trees, were adequately represented as reported by Chini et al. (2008c). In SRS, the population of N pixels is divided into k subpopulations of sampling units $N_1,N_2,\ldots,N_k$, which are termed "strata". Therefore, we have randomly sampled the pixels in each of those classes accordingly to their extension in area, based on the produced ground reference.

The number of pixels used for training may influence the final classification accuracy. To investigate this, we used about 5% and 10% of the total pixels for the Las Vegas and Rome scenes (same spatial resolution), respectively, and for comparison about 10% for the Washington case (higher spatial resolution). Details of the number of samples used as training (TR) and validation (VA) are reported in Tables 2–4 for the Las Vegas, Rome and Washington D.C. areas, respectively. Kappa coefficient and percentage overall error analysis (Foody, 2002) were used to evaluate the confusion matrices derived from the supervised classification.

4. Methodology

4.1. Multi-scale texture analysis

Spectral-based classification methods may fail with the increased geometrical resolution of the data available. As stated by Gong et al. (1992), improved spatial resolution data increases within-class variances, which results in high interclass spectral confusion. In many cases, several pixels are representative of objects, which are not part of land-use classes defined. Cars may be used as a representative example, as they do not belong to any land-use class, however cars may be present in related classes, such as roads and parking lots. Furthermore, cars may create textural patterns in imagery, for example in parking lots, and this effect may be measured and utilized for filtering them out. It is evident that this problem is intrinsically related to the sensor resolution and it cannot be solved by increasing the number of spectral channels. Therefore, a solution may be to include spatial information that can be used to characterize land-use classes by exploiting different concentrations and patterns of these non land-use pixels. We exploit these two last characteristics of land-use, using a multi-scale approach based on GLCM textural features.

In this work, six different second-order textural features derived from the GLCM and two first-order textural features have been considered. The formulation of these six second-order textural features is shown in the following:

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

$$\text{Contrast} = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} p(i,j)\cdot(i-j)^2 \tag{2}$$

$$\text{Dissimilarity} = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} p(i,j)\cdot|i-j| \tag{3}$$

$$\text{Entropy} = -\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} p(i,j)\cdot log(p(i,j)) \tag{4}$$

$$\text{Second moment} = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1} p(i-j)^2 \tag{5}$$

$$\text{Correlation} = \sum_{i=1}^{N-1}\sum_{j=1}^{N-1}\frac{(i\cdot j)\cdot p(i,j)-\mu_i\cdot\mu_j}{\sigma_i\cdot\sigma_j} \tag{6}$$

where σ and μ are the mean and standard deviation respectively, 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 assumes higher values for smaller digital number differences in pair elements. Therefore, this parameter is more sensitive to the presence of near diagonal elements in the GLCM. *Contrast* takes into account the spatial frequency, which is the difference in amplitude between the highest and the lowest values of a contiguous set of pixels. This implies that a low contrast image is not necessarily characterized by a low variance value, but the low contrast image corresponds to low spatial frequencies. Unlike *Contrast* where the weights increase exponentially as one moves away from the diagonal, for *Dissimilarity* the weights increase linearly. This parameter measures how different the elements of the co-occurrence

Table 4
Washington D.C. case: classes, training and validation samples, and color legend.

Washington D.C. classes	TR	VA
Buildings	24,178	76,159
Highways	17,985	56,653
Parking lots	17,019	53,611
Residential	14,195	44,714
Roads	20,618	64,946
Side walks	12,203	38,439
Soil	2553	8043
Sport facilities	8270	26,051
Tall buildings	21,047	66,297
Trees	18,535	58,386
Vegetation	23,403	73,720
Total ROIs	180,006	567,019

Table 5
Input space resulting from panchromatic band, first- and second- order textural features.

	Input features	Cell Size (pixel)	Step (pixel)	Direction (°)	# Inputs
	Panchromatic				1
First-order	Mean	3×3			10
	Variance	7×7			
		15×15			
		31×31			
		51×51			
Second-order	Homogeneity	3×3	15	0	180
	Contrast	7×7	30	45	
	Dissimilarity	15×15		90	
	Entropy	31×31			
	Second Moment	51×51			
	Correlation				
Total input features					191

matrix are from each other and it is high when the local region has a high contrast. *Entropy* measures the disorder in an image. When the image is not uniform, many GLCM elements have very small values, which implies that *Entropy* is very large. If we consider a window with completely random gray tones, the histogram for such a window is a constant function, i.e., all $p(i,j)$ are the same, and *Entropy* reaches its maximum. The *Second Moment* measures textural uniformity, i.e., pixel pairs repetitions. Indeed, when the image segment under consideration is homogeneous (only similar gray levels are present) or when it is texturally uniform (the shift vector always falls on the same (i,j) gray-level pair), a few elements of GLCM will be greater than 0 and close to 1, while many elements will be close to 0. *Correlation* is expressed by the correlation coefficient between two random variables i and j, and it is a measure of the linear-dependencies between values within the image. High *Correlation* values imply a linear relationship between the gray levels of pixel pairs. Thus, *Correlation* is uncorrelated with *Energy* and *Entropy*, i.e. to pixel pair repetitions (Baraldi & Parmiggiani, 1995).

First-order statistics can be computed from the histogram of pixel intensities in the image. These depend only on individual pixel values and not on the interaction or co-occurrence of neighbouring pixel values. The first-order parameters used in this paper are *Mean* and *Variance*. The *Mean* is the average gray-level in the local window and the *Variance* is the gray-level variance in the local window. The latter is high when there is a large gray-level standard deviation in the local region.

Textural information is valuable for the discrimination of different classes that have similar spectral responses. At the same time, it is also necessary to exploit a multi-scale approach to better deal with objects having different spatial coverage in an area. For this purpose, the eight features defined before have been computed with five different window sizes 3×3, 7×7, 15×15, 31×31, 51×51 pixels (about 1.5–1.8 m, 3.5–4.2 m, 7.5–9.0 m, 15.5–18.6 m and 25.5–30.6 m, with WorldView-1 and QuickBird, respectively), three different directions 0°, 45° and 90° and two different cell shift values of 15 and 30 pixels for a total of 191 textural features, as reported in Table 5. The dimensions of the windows and the values of the shift have been based on the analysis of a previous work of Small (2003) who estimated the characteristic length scale of 6357 sites in 14 urban areas around the World, showing that the majority of sites have characteristic length scales between about 8.0 and 24.0 m.

4.2. Magnitude-based pruning

As pointed out in the previous sections, enlarging the size of the input space increases resource requirements and causes the "curse of dimensionality". Therefore, it is necessary to eliminate redundant or useless features. This introduces the concept of saliency, or contribution, of a feature. Some inputs may be useful only in the presence of other features, while being useless on their own. Thus, the saliency computation generally involves all network parameters simultaneously.

In the following, we first describe the feature selection method used in this paper, known as Neural Network Pruning. Then, we illustrate the criteria adopted for computing the feature contribution of each input using a trained (and pruned) neural network.

As stated, among the Neural Network Pruning techniques, a sensitivity-based method proved to be the most popular. Magnitude-based (MB) pruning is the simplest weight-pruning algorithm which is based on removing links with the smallest magnitude value. Thus the saliency of a link is simply the absolute value of its weight. Tarr (1991) explained this concept considering that when a weight is updated, the learning algorithm moves the weight to a lower value based on the classification error. Thus, given that a particular feature is relevant to the problem solution, the weight would be moved in a constant direction until a solution with no error is reached. If the error

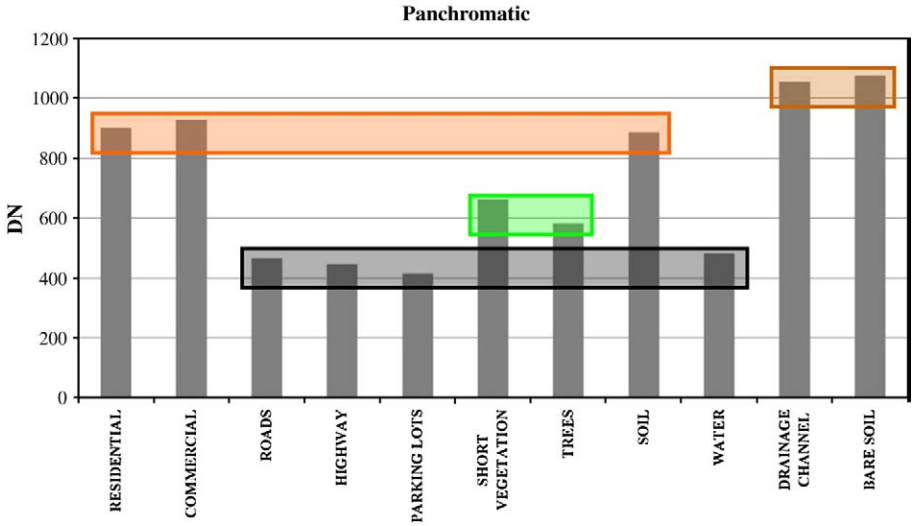


Fig. 2. Mean digital number of the eleven classes of Las Vegas. Only four groups were identified using the panchromatic band and represented by horizontal blocks (colors are consistent with Table 2). Particularly, water was grouped with roads, highways and parking lots, while soil appeared to be closer to residential and commercial buildings than to bare soil. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

term is consistent, the direction of the movement of the weight vector will also be consistent (a consistent error term is the result of all points in a local region of the decision space belonging to the same output class). If the error term is not consistent, which can be the case of a single feature of the input vector, the movement of the weight attached to that node will also be inconsistent. In a similar fashion, if the feature did not contribute to a solution, the weight updates would be random. In other words, useful features would cause the weights to grow, while weights attached to non-salient features simply fluctuate around zero. Consequently, the magnitude of the weight vector serves as a reasonable saliency metric. Although this method is very simple, it rarely yields worse results than the more sophisticated algorithms (Zell et al., 1995; Le Cun et al., 1990; Hassibi & Stork, 1993).

Kavzoglu and Mather (1999) investigated different pruning techniques using SAR and optical data for land-cover mapping. They found that the MB pruning technique generally yielded better results despite the simplicity of the algorithm. Moreover, their results show that pruning not only reduces the size of neural networks, but also increases overall network performance. MB pruning appeared to be robust enough and was chosen in our work to eliminate the weakest inputs.

The network pruning technique provides a reduced set of features and at the same time optimizes the network topology. However, the resultant input space may have more than a reasonable number of input features. A trade-off between classification accuracy and computational time should be determined. The so-called Extended Pruning technique, Del Frate et al. (2005), is the process of eliminating iteratively (by successive pruning phases) the least contributing inputs until the training error reaches a specified limit. This process identifies a sub-optimal feature set; sub-optimal from the classification accuracy point of view. In fact, this further input reduction results in a decrease in the classification accuracy. Thus, after the optimization (from the classification accuracy point of view) of the network topology by pruning, the Extended Pruning technique has been used in order to identify a reduced set of input features.

4.2.1. Computing feature saliency

Considering how weights in a neural network are updated, the weights can be used to calculate the saliency. Once the network has been pruned (simple or extended pruning), a general method for determining the relative significance of the remaining input features has been suggested by Tarr (1991). Input units whose weighted connections have a large absolute value are considered to be the most important. He proposes the following saliency metric to define the

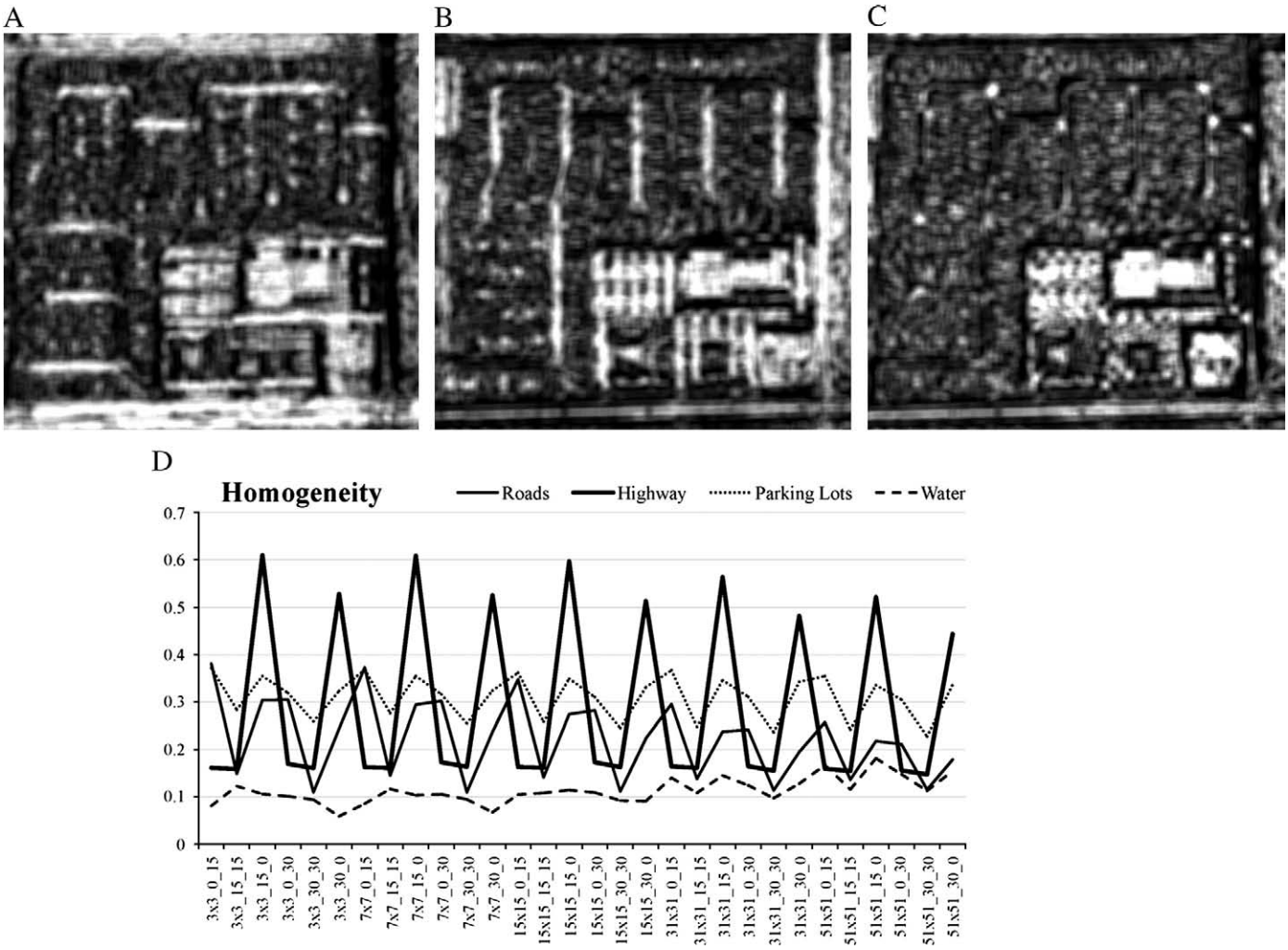


Fig. 3. In (A), (B) and (C) is shown the homogeneity textural parameter computed over Las Vegas using the same step and window size, but different directions (horizontal, vertical and diagonal, respectively). The directional information highlights different structural patterns within the area, such as vertical or horizontal roads. In (D) is shown the main homogeneity values for classes roads, highway, paring lots and water computed for all directions, window sizes and shifts. In particular, the highway (which is a horizontal structure) assumes the highest homogeneity values with respect to the other classes using horizontal texture parameters while parking lots show a distinguishing behavior (with respect to the other classes) in the diagonal directions since it is a wide structure. The notation "$A \times B\_x\_y$" has the following meaning: (A,B) are the dimensions of the window size, while (x,y) represent the Cartesian components of direction and shift. For example, the feature $3 \times 3\_15\_0$ is computed with a 3×3 window size, horizontal direction and 15 pixels of shift.

Table 6
Classification accuracies for the Las Vegas, Rome and Washington D.C. cases at the three classification stages.

	Las Vegas			Rome			Washington D.C.		
	Cl. err. (%)	Kappa coeff.	# Inputs	Cl. err. (%)	Kappa coeff.	# Inputs	Cl. err. (%)	Kappa coeff.	# Inputs
Panchromatic	50.2	0.378	1	66.0	0.184	1	68.6	0.187	1
Full NN	7.1	0.916	191	16.9	0.798	191	14.5	0.838	191
Pruned NN	6.8	0.920	169	5.0	0.941	140	8.6	0.904	152

relevance for every weight between the input i and hidden unit j of the network:

$$S_i = \sum_{j}^{N_h} w_{ij}^2 \tag{7}$$

which is simply the sum of the squared weights between the input layer and the first hidden layer. This formulation may not be completely representative when dealing with pruned networks, since several connections between the input and the first hidden layer may be missing. Yacoub and Bennani (1997) exploited both weight values and network structure of a Multi-Layer Perceptron network in the case of one hidden layer. They derived the following criterion:

$$S_i = \sum_{j\in H}\left(\frac{|w_{ji}|}{\sum_{i'\in I}|w_{ji'}|}\sum_{k\in 0}\frac{|w_{kj}|}{\sum_{j'\in H}|w_{kj'}|}\right) \tag{8}$$

where I, H, O denote the input, hidden and output layer, respectively. For the two hidden layers case, which is the network topology used in our work, we adopted a slight variation of Eq. (8). In particular, the importance of variable i for output j is the sum of the absolute values of the weight products over all paths from unit i to unit j, and it is given by:

$$S_{ij} = \sum_{k\in \mathrm{H1}}\left[\frac{|w_{ik}|}{\sum_{k'\in \mathrm{H1}}|w_{ik'}|}\cdot\sum_{x\in \mathrm{H2}}\left(\frac{|w_{kx}|\,|w_{xj}|}{\sum_{x'\in \mathrm{H2}}|w_{kx'}|\cdot\sum_{x'\in \mathrm{H2}}|w_{x'j}|}\right)\right] \tag{9}$$

where H1 and H2 denote the first and the second hidden layers, respectively. Then, the importance of variable i is defined as the sum of these values over all the outputs classes N_{cl}.

$$S_i = \sum_{j=1}^{N_{\mathrm{cl}}} S_{ij}. \tag{10}$$

5. Results

In this section we investigate the capability of panchromatic data to produce land-use maps of the scenes described previously. As stated, to overcome the spectral deficit of panchromatic imagery, it may be necessary to extract additional information to recognize objects within the scene. Neural Network Pruning has been adopted to optimize the input features space and network topology, as described in Section 5.1. Then, the analysis of most effective input features has been carried out by the Extended Pruning technique and described in Section 5.2. A detailed analysis of the best 10 features is illustrated in Section 5.3. In addition, an independent test site belonging to the city of San Francisco, has been used to show the value of the selected features for mapping an urban scenario not included in the features extraction phase. The analysis of the texture properties of shadowed areas follows Section 5.4.

Fig. 4. Classification maps of (A) Las Vegas, (B) Rome and (C) Washington D.C. obtained after the pruning phase. The maps discriminated different asphalt surfaces, such as roads, highways and parking lots due to the different textural information content. This approach also allowed to differentiate building architectures, such as residential houses, apartment blocks and towers. Shadowed areas did not influence any of the maps obtained.

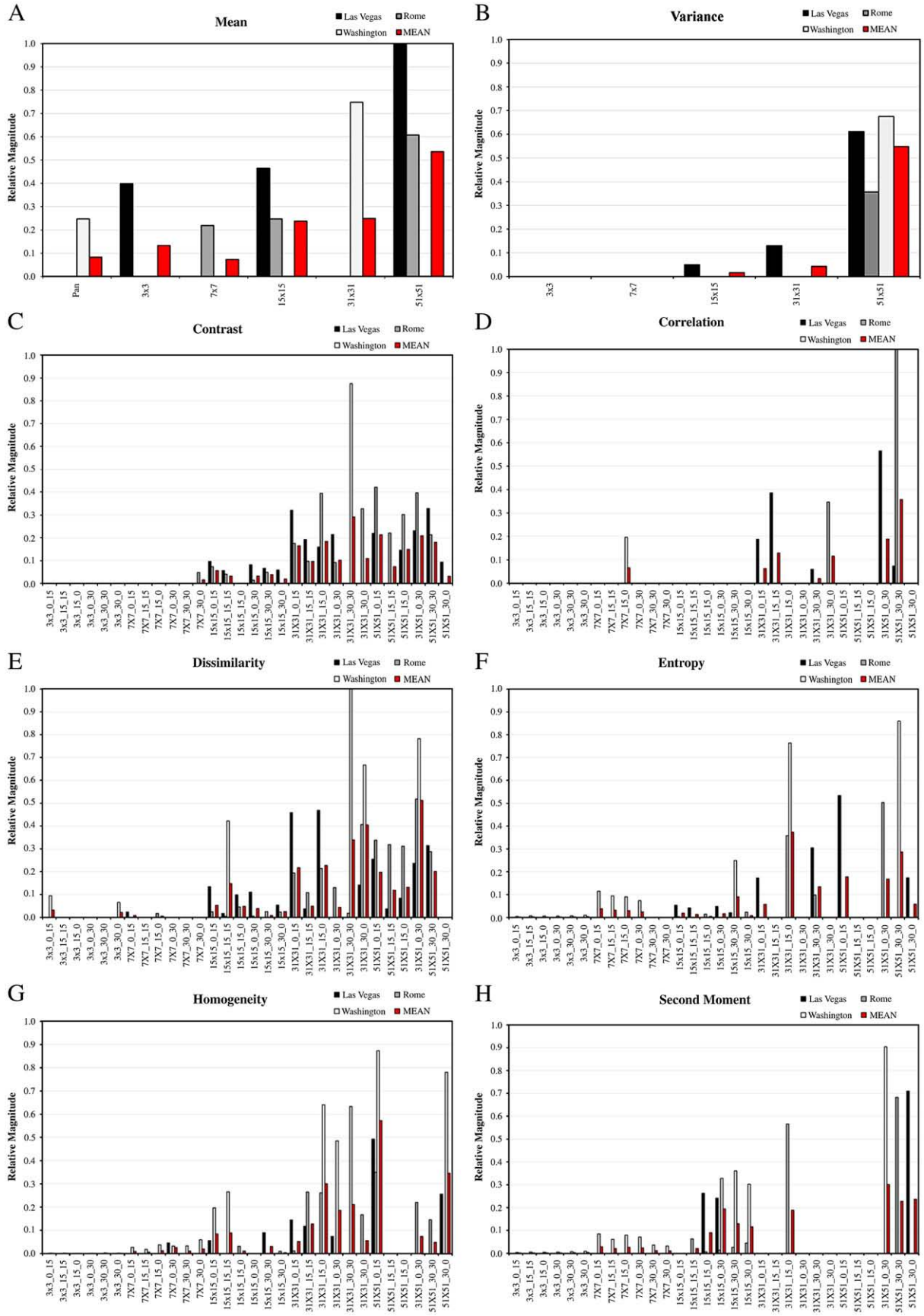


Fig. 5. Relative feature contribution of the input features not eliminated by the extended pruning of (A) Panchromatic and Mean, (B) Variance, (C) Contrast, (D) Correlation, (E) Dissimilarity, (F) Entropy, (G) Homogeneity and (H) Second Moment computed over all window size, directions and shifts. The mean values of these contributions, shown in red, highlights that many of the remaining inputs have a smaller influence on the classification process compared to other features. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.1. Optimization of the feature space and network topology

The first effort has been to produce land-use maps using only panchromatic information with neural network. As expected, the results obtained appeared to be really poor for the three test cases in terms of classification accuracy. The obtained Kappa coefficient values are 0.378 for Las Vegas, 0.184 for Rome and 0.187 for Washington D.C. As expected, several of the defined classes were not recognized. For example, for the Las Vegas scene, only Bare Soil, Residential Houses, Roads and Short Vegetation have been identified, which are representative of the materials present in the area. This is shown in Fig. 2 where it is possible to group the digital number values into four sets. Even though water and asphalt are different, it is well known from the literature that they may assume similar values in panchromatic data. This has led to some confusion where water is grouped with roads, highway and parking lots.

As already explained in Section 3.1, a multi-scale textural input space has been extracted from panchromatic images (see Table 5), including first- and second-order textural parameters computed with 5 window sizes, 3 directions and 2 step sizes. Note that directions and steps are related only to second-order statistics. Fig. 3 shows an example of the homogeneity parameter computed over Las Vegas with three different directions (same step and window size). As expected, the direction calculation highlights different structural patterns within the area, such as vertical or horizontal roads, and parking lots. In fact, the highway, which is a horizontal structure, has its highest value in the 15_0 and 30_0 directions, while parking lots show a distinct behavior (with respect to the other classes) in the diagonal direction since they are wide structures. Note, the notation "$A\times B\_x\_y$" has the following meaning: (A,B) are the dimensions of the window size, while (x,y) represent the Cartesian components of direction and shift. For example, the feature $3\times3\_15\_0$ is computed with a 3×3 window size, horizontal direction and 15 pixels of shift.

A considerable increase in classification accuracy has been obtained using the entire input dataset of 191 textural features. More precisely, Kappa coefficient values of 0.916 for Las Vegas, 0.798 for Rome and 0.838 for Washington D.C. have been obtained. With respect to the previous implementation, all classes were identified in the classification maps. On the other hand, as mentioned, a large input space rarely yields high classification accuracies due to information redundancy. This results in a necessity to estimate the contribution of each parameter in order to reduce and optimize the input space. To this end, we adopted Neural Network Pruning to eliminate the weakest connections, optimizing at the same time the network topology. Generally, this process increases the classification accuracy by eliminating features that do not contribute to the classification process, but instead only introduce redundancy. After the pruning phase, the remaining inputs are 169 for Las Vegas, 140 for Rome and 152 for Washington D.C., respectively. This relatively small feature elimination resulted in a further increase of classification accuracy. In particular, the Kappa coefficient values increased to 0.920 for Las Vegas, 0.941 for Rome and 0.904 for Washington D.C., whose classification maps are illustrated in Fig. 4 and the accuracies are summarized in Table 6 for the reader's convenience. Taking into account the extension of boundary areas between objects, we noticed a slight decrease of the classification accuracy for the Washington D. C. case, whose ground reference included boundary areas.

The obtained classification maps discriminated different asphalt surfaces, such as roads, highways and parking lots due to the different textural information content. This approach also made it possible to differentiate building architectures, sizes and heights, such as residential houses, apartment blocks and towers. Here, it is important to note that shadowed areas did not influence any of the maps obtained. The reason for this will be described later in the paper. The accuracies obtained with the optimization of the network topology are considered here as an upper bound of the classification accuracies derived from the multi-scale approach.

5.2. Features selection by extended pruning technique

In the previous section, the network pruning technique provided a reduced set of textural features and an optimized network topology in order to obtain the most accurate classification map. However, the input space is not close to a reduced number of features relative to the

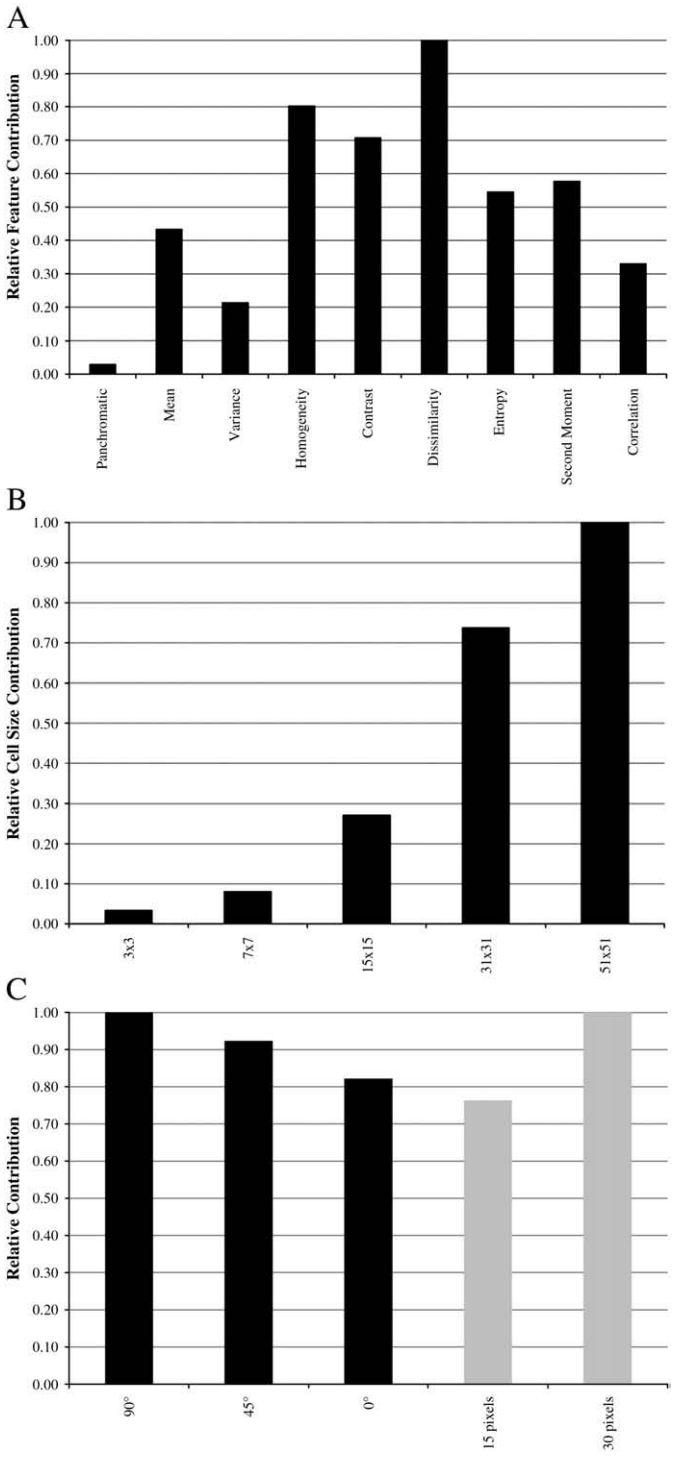


Fig. 6. Feature contributions with respect to textural parameters, window sizes and directions. In (A) is shown the contributions of textural parameters regardless of the choice of cell sizes and directions. Dissimilarity appears to be the most informative texture parameter. In (B) is shown the importance of the cell sizes, regardless of the choice of textural parameters and directions. In (C), the three directions (in black) and the two-step size (in gray) show high and relatively similar contributions, meaning that both directions and step sizes are relevant for the classification phase.

computational time, and a trade-off between classification accuracy and computational time should be found. The so-called Extended Pruning technique is the process of eliminating the least contributing inputs, even from this optimal classification, in order to identify a minimal sub-optimal textural feature set. The resulting classification is sub-optimal from the classification accuracy point of view, since this further input reduction results in a decrease in the classification accuracy. Particularly, the criterion chosen to stop the extended pruning phase was to reach a classification accuracy of about 0.800 in terms of Kappa coefficient.

After the extended pruning phase, the remaining inputs are 59 for Las Vegas, 61 for Rome and 59 for Washington D.C., with accuracies (Kappa coeff.) decreased to 0.859, 0.820 and 0.796, respectively. In Fig. 5 we present the relative contributions of the input features (including the panchromatic image itself), which are not eliminated by the extended pruning. The saliency metric used to compute the feature contributions has been illustrated in Sec. 4.1, Eqs. (9) and (10). Moreover, we normalized all the contribution values between 0 and 1, where 0 means no contribution to the classification phase. As shown, the contribution of each input varies from city to city, due to the architectural peculiarities (and diversity) of them. The analysis of the mean values of these contributions, shown in red, clearly indicates that many of the remaining inputs have a smaller influence on the classification process compared to other features. This means that using certain textural features (including different cell sizes and directions) may have more significance than others.

To analyze the importance of textural parameters regardless of the choice of cell sizes and directions, we computed, for each of them, the feature contributions as sums over the different cell sizes and directions. As shown in Fig. 6A, the panchromatic band, which does not contain any information on cell sizes and directions, has the smallest contribution. Fist-order textural features, which do not contain any information on directions, have smaller contributions than second-order features. Particularly, Dissimilarity appears to be the most informative texture parameter, even if it is similar to Contrast. This may be related to the linear weighting of the gray-tone levels of the scene (to be compared to the exponentially weight of Contrast). In the same way, we analyzed the importance of the cell sizes, regardless of the choice of textural parameters and directions. This is illustrated in Fig. 6B, where larger cell sizes of 31×31 and 51×51 show higher contributions. This may be related to the very high-resolution data used in this paper. Since QuickBird and WorldView-1 sensors have resolutions close to half a meter, it is reasonable that the textural information is contained in the spatial range of 15.0–25.0 m, a result which is consistent with Small (2003).

In Fig. 6C, the three directions (in black) and the two-step sizes (in gray) have high and similar contributions, meaning that both directions and step sizes are relevant for the classification phase. This last result points out the necessity of having directional information to better capture differences in textural patterns.

5.3. Analysis of the best 10 textural features

We have shown that many of the remaining inputs, after the extending pruning phase, have a smaller influence on the classification process compared to other features. To further investigate individual feature contributions, the frequency of the input features with respect to the input feature contribution is shown in Fig. 7, which highlights that only very few inputs show a relative contribution greater than 0.30. The best ten features are reported in Table 7 with the corresponding values of contribution relative to the mean values over three cities and several different land-uses.

To understand the contribution of a single class to these ten features, we merged together the different land-use classes into five groups, which are common to the three scenes. The resultant common five classes are: Buildings, Roads, Soil, Trees and Vegetation. The contributions per single class of the ten best features with respect to these five classes are illustrated in Fig. 8. For example, the first-order Mean 51×51 seems to be appropriate for the discrimination of roads, while Dissimilarity 51×51_0_30 appears to be valuable for the detection of trees.

With this drastic reduction of input features (from about 60 to 10), we expected a further decrease of the classification accuracy with respect to the extended pruning results. On the other hand, this effect was compensated for by the reduction (from about 11 to 5) of the output classes. In particular, we re-classified the Washington D.C. scene using only the ten best textural inputs, obtaining an accuracy of

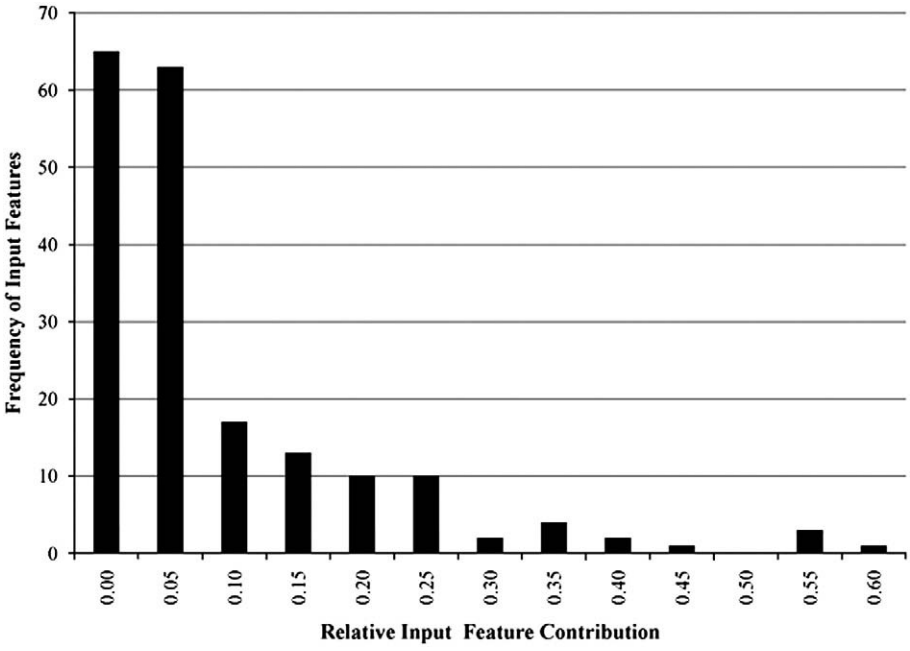


Fig. 7. Frequency of the input features with respect to the input feature contribution highlights that only very few inputs show a relative contribution greater than 0.30.

Table 7
Best ten features and corresponding feature contribution averaged over three cities and several different land-uses classes.

Best 10 features	Relative input feature contribution
Mean 51×51	0.535
Variance 51×51	0.547
Homogeneity 51×51_0_15	0.572
Homogeneity 51×51_30_0	0.345
Dissimilarity 31×31_30_30	0.339
Dissimilarity 31×31_30_0	0.404
Dissimilarity 51×51_0_30	0.512
Entropy 31×31_15_0	0.374
Second Moment 51×51_0_30	0.301
Correlation 51×51_30_30	0.357

0.861. This result is particularly relevant since it was obtained after a generalization (mean) of results obtained over all of the cities with different architectures.

Even though these considerations show similarities, especially in terms of computational time and generalization capability for different urban scenarios, it is necessary to emphasize the importance of exploiting the entire textural features dataset, including all different spatial scales and directions. Starting from the same textural features and pruning the neural networks made it possible to classify different urban scenarios with high accuracies, showing both the efficiency and the robustness of the multi-scale approach used here.

5.3.1. The San Francisco test case

In the previous section, we discussed the feature contributions as mean values over three different test sets corresponding to Las Vegas, Rome and Washington D.C. A reduced set of ten features has been identified as valuable for urban classification. Now the question is: how well do these ten features classify a new urban scene? To answer this question, we used these ten features to classify an independent data set of a portion of San Francisco. Note that different combinations of texture metrics might produce more accurate results for this particular test case. However, we only want to investigate the potentialities of the selected features (obtained as average of 3 totally different conditions) when applied to a new scenario. In this sense, these ten features are not an optimal combination, but simply a set of inputs that may potentially increase the classification accuracy when applied to very high-resolution imagery.

The scene, shown in Fig. 9A, has been imaged by WorldView-1 with an off-nadir angle of 19.6°. Further, long shadows are caused by a sun elevation of 29.6°. This neighborhood may be considered a representative architecture for many cities, making it suitable for validation purposes. The same five common classes defined previously have been used. Training and validation pixels have been selected and are summarized in Table 8, while the entire ground reference map is shown in Fig. 9B.

We first classified the panchromatic image alone, obtaining a Kappa coefficient of about 0.224. Then, using only the reduced set of ten textural features, we obtained a map, shown in Fig. 9C, whose accuracy was about 0.917.

5.4. Analysis of texture properties of shadowed areas

Shadow effects are often ignored when using decametric spatial resolution images, such as Landsat. In these cases, shadowed pixels may be located on an object's boundaries where there is a mixture of radiances caused by different surfaces. Vice versa, shadows have a huge impact on classification with metric or sub-metric spatial resolution images, such as QuickBird or WorldView-1. In urban areas, shadows are mainly caused by buildings, tress or bridges and may potentially provide additional geometric and semantic information on the shape and orientation of objects. On the other hand, shadowed surfaces require more consideration since they may cause a loss of information or a shape distortion of objects.

In the literature, shadow is generally dealt with as an additional class (Benediktsson et al., 2005; Bruzzone and Carlin, 2006). In this paper, we have defined the ground reference regardless of the presence of shadowed areas, leading to classification maps, which do not contain any pixels of shadow. To further investigate this, we extracted the shadowed pixels of the Rome scene and analyzed their textural values (normalized between 0 and 1) with respect to the non-shadowed pixels of the same class. As illustrated in Fig. 10, where continuous-lines represent pixels of shadow and dotted-lines the non-shadowed pixels of the corresponding class. Only panchromatic information appears to not be able to adequately separate the classes whose pixels are covered by shadow. In fact, they are mainly concentrated in the lower part of the scale values, acting more as a unique class. Mean values (and their standard deviation) are also reported in Table 9 for blocks, roads and vegetation. Even though these

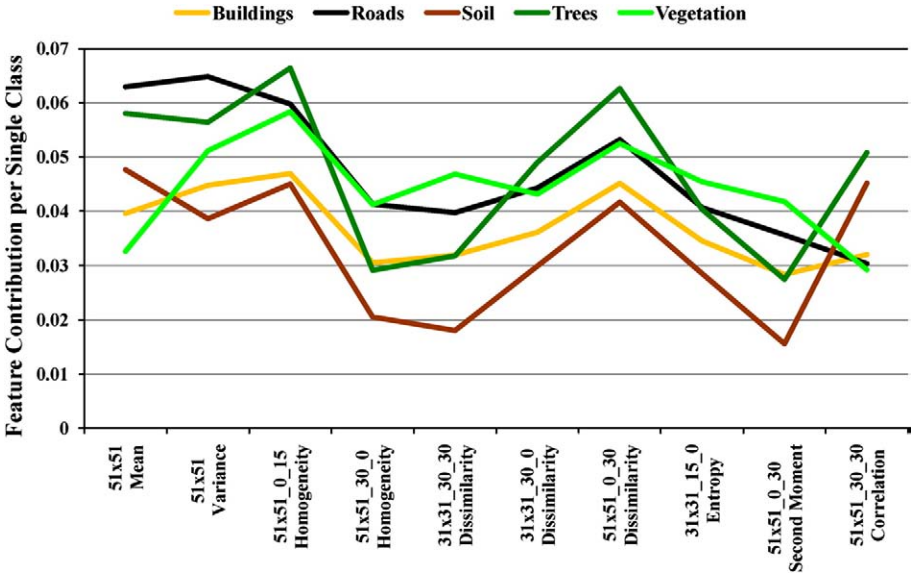


Fig. 8. Contributions per single class of the best ten features with respect to the five common classes defined. The first-order Mean and Variance 51×51 seem to be more appropriate for the discrimination of roads than Correlation or Second Moment, while Dissimilarity 51×51_0_30 appears to be valuable for the detection of trees.

Table 8
San Francisco case: classes, training and validation samples, and color legend.

Classes	TR	VS
Buildings	36,689	85,240
Roads	49,436	114,857
Soil	6524	15,158
Trees	7745	17,993
Vegetation	1465	3404
Total ROIs	101,859	236,652

surfaces are different, the (normalized) panchromatic values of shadowed areas are very similar: 0.08, 0.05 and 0.07, respectively. By contrast, shadowed areas have their own texture properties, allowing the discrimination between different shadowed classes.

To make Fig. 10 clearer, we only consider the class buildings (including both apartment blocks and towers), which is one of the most relevant for urban classification. Shadows of buildings are mainly due to two different effects: i) shadows of the building on itself and ii) shadows of other objects, such as other buildings or trees, on a building. As shown in Fig. 11, buildings show a sort of characteristic textural signature. For example, the variance of shadowed buildings is slightly higher than non-shadowed buildings. This may be interpreted by considering the smaller extension in the area of shadow with respect to buildings (for the scene considered) and the higher contrast in terms of gray-tone levels on the boundary between shadowed and non-shadowed pixels. Smaller objects with higher contrast lead to higher spatial variability, thus larger variance values. Similar considerations may be given to the homogeneity feature, which shows an inversion of the trend between shadowed and non-shadowed buildings due to the larger step size (30 instead of 15 pixels), but equal window size of 51×51 pixels.

6. Conclusion

In this work, we investigated the potential of very high-resolution panchromatic QuickBird and WorldView-1 imagery to classify the land-use of urban environments. Spectral-based classification methods may fail with the increased geometrical resolution of the data available. In fact, improved spatial resolution data increases within-class variances, which results in high interclass spectral confusion. In many cases, several pixels are representative of objects, which are not part of land-use classes defined. This problem is intrinsically related to the sensor resolution and it cannot be solved by increasing the number of spectral channels. To overcome the spectral information deficit of panchromatic imagery, it is necessary to extract additional information to recognize objects within the scene.

The multi-scale approach discussed in this paper exploits the contextual information of first- and second-order textural features to characterize the land-use. For this purpose, textural parameters have been systematically investigated computing the features over five different window sizes, three different directions and two different cell shifts for a total of 191 inputs. Neural Network Pruning and saliency measurements allowed the optimization of the network topology and give an indication of the most important textural features for sub-metric spatial resolution imagery and urban scenarios.

Fig. 9. (A) Panchromatic image and (B) ground reference of San Francisco. In (C) is shown the classification map obtained using the reduced set of ten textural features. Color codes are in Table 8. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

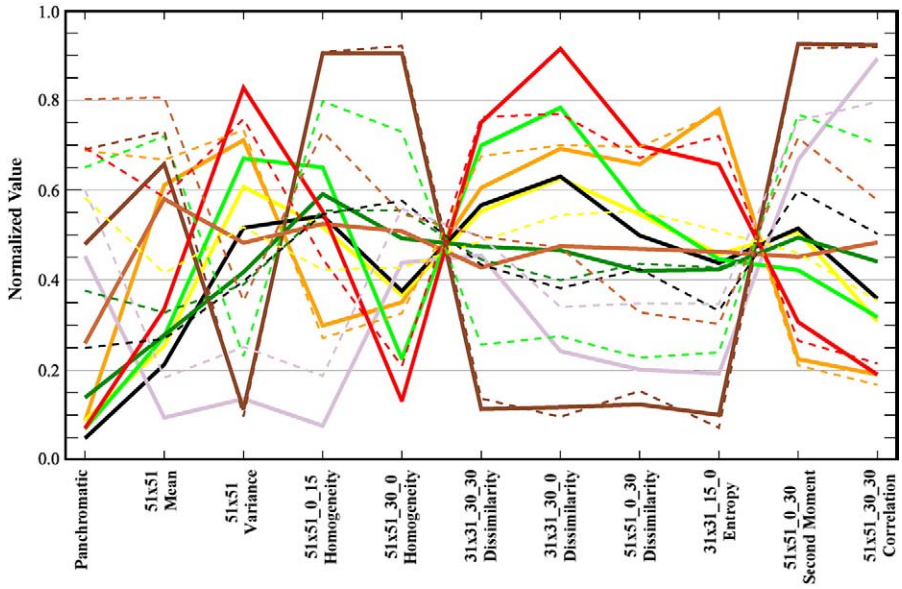


Fig. 10. Normalized textural values of the most contributing features and panchromatic band of the Rome case for shadowed (continuous-lines) and non-shadowed (dotted-lines) pixels. The only panchromatic information appears to not separate sufficiently the classes whose pixels are covered by shadow, since the result mainly concentrated in the lower part of the scale values. As opposed, shadowed areas show their own texture properties, allowing the discrimination between different shadowed classes. Color codes are in Table 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We believe that the network pruning should be mandatory in any neural-net based classification. As summarized in Table 6, using the full neural network, the classification accuracies are modest (~0.8 in terms of Kappa coeff.). After network pruning, the maps of the three dataset exploited for the texture analysis, i.e. Las Vegas, Rome and Washington D.C., show higher land-use classification accuracies, above 0.90 in terms of Kappa coefficient computed over more than a million independent validation pixels. The network pruning greatly improved the classification accuracies due to two main effects: (1) the network topology is optimized; (2) the smaller set of input features reduced the effect of the "curse of dimensionality".

The identification of the most effective textural features has been carried out using the Extended Pruning technique with saliency measurements. Fist-order textural features, which do not contain any information on directions, have smaller contributions than second-order features. Dissimilarity appears to be the dominant texture parameter. For the spatial resolution and test cases considered, bigger cell sizes, such as 31×31 and 51×51 pixels, show higher contributions than smaller cell sizes (regardless of the choice of textural parameters and directions), making it clear that there is a need to exploit the entire directional information. However, many of the remaining inputs, after the extending pruning phase, have a smaller influence on the classification process compared to other features. Specifically, very few inputs showed a relative contribution greater than 0.30, as reported in Table 7. As expected, the drastic reduction of input features (from about 60 to 10) decreased the classification accuracy of the Washington D.C. scene to 0.861 (Kappa coeff.) with respect to the extended pruning result. Nevertheless, the simple selection of these features resulted in remarkable results compared to those obtained using the only panchromatic information. Furthermore, the result obtained for the independent test case of San Francisco, i.e. not included in the multi-scale textural analysis carried out in the first part of the paper, indicates the potentialities of the reduced set of textural features for mapping a common urban scenario.

Since the textural analysis was carried out as an average of three different environments (and validated over more than a million independent samples), we believe that this approach can efficiently be extended to large areas, such a whole city. In this sense, the San Francisco scene may represent a non exhaustive, but significant, example.

To conclude, we notice that only panchromatic information appears to not be able to adequately separate classes whose pixels are covered by shadow. In our analysis, these pixels' values were

Table 9
Normalized textural values (and their standard deviation) of the panchromatic band and the most contributing features of the Rome case for shadowed (SH) and non-shadowed (NON-SH) pixels.

Panchromatic band + best 10 features	Blocks		Roads		Vegetation	
	SH	NON-SH	SH	NON-SH	SH	NON-SH
Panchromatic	0.08 (0.13)	0.58 (0.26)	0.05 (0.07)	0.25 (0.16)	0.07 (0.11)	0.65 (0.14)
Mean 51×51	0.25 (0.22)	0.41 (0.24)	0.21 (0.21)	0.27 (0.20)	0.28 (0.21)	0.72 (0.23)
Variance 51×51	0.61 (0.17)	0.52 (0.20)	0.52 (0.26)	0.40 (0.25)	0.67 (0.19)	0.23 (0.23)
Homogeneity $51\times51\_0\_15$	0.52 (0.25)	0.42 (0.26)	0.54 (0.27)	0.55 (0.25)	0.65 (0.21)	0.80 (0.19)
Homogeneity $51\times51\_30\_0$	0.37 (0.21)	0.43 (0.26)	0.38 (0.28)	0.58 (0.26)	0.22 (0.20)	0.73 (0.27)
Dissimilarity $31\times31\_30\_30$	0.55 (0.27)	0.49 (0.24)	0.57 (0.28)	0.43 (0.28)	0.70 (0.23)	0.26 (0.23)
Dissimilarity $31\times31\_30\_0$	0.63 (0.18)	0.54 (0.23)	0.63 (0.26)	0.38 (0.25)	0.78 (0.21)	0.28 (0.25)
Dissimilarity $51\times51\_0\_30$	0.55 (0.25)	0.55 (0.23)	0.50 (0.28)	0.42 (0.27)	0.56 (0.24)	0.23 (0.23)
Entropy $31\times31\_15\_0$	0.46 (0.19)	0.51 (0.19)	0.44 (0.23)	0.33 (0.20)	0.45 (0.23)	0.24 (0.21)
Second Moment $51\times51\_0\_30$	0.51 (0.20)	0.46 (0.20)	0.51 (0.24)	0.60 (0.24)	0.42 (0.21)	0.77 (0.24)
Correlation $51\times51\_30\_30$	0.31 (0.17)	0.35 (0.22)	0.36 (0.27)	0.50 (0.28)	0.32 (0.21)	0.70 (0.23)

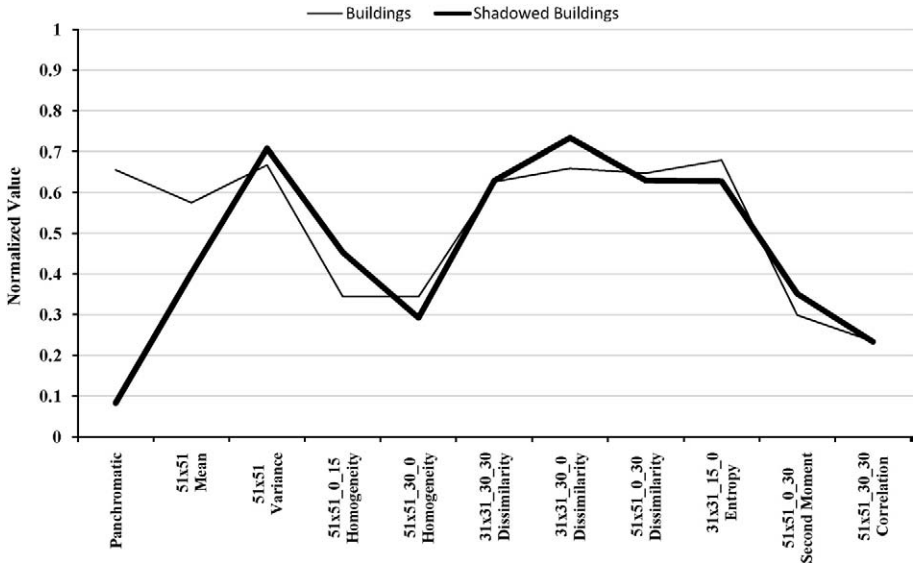


Fig. 11. Normalized textural values of the most contributing features and panchromatic band of the Rome case for shadowed and non-shadowed pixels of buildings (including also apartment blocks and towers). The variance of shadowed buildings is slightly higher than non-shadowed buildings, while homogeneity shows an inversion of the trend between shadowed and non-shadowed buildings using a larger step size.

grouped together, acting as a unique class. In contrast, multi-scale textural analysis proved that it is possible to distinguish different shadowed areas, since they have their own texture properties.

Future research will address a comparison between the results found in this paper and the use of isotropic textural features applied to very high-resolution optical imagery for urban land-use classification.

Acknowledgment

The authors would like to thank DigitalGlobe for providing the data used in this paper and also for partially funding Fabio Pacifici. William J. Emery was supported by NASA's Earth Science Division. Marco Chini has been partially funded by the Italian Space Agency through the SIGRIS project. The authors are grateful to Dan Baldwin for the helpful discussions.

References

Arzandeh, S., & Wang, J. (2002). Texture evaluation of RADARSAT imagery for wetland mapping. *Canadian Journal of Remote Sensing*, *28*, 653—666.

Bamler, R., & Eineder, M. (2008). The Pyramids of Gizeh seen by TerraSAR-X — A prime example for unexpected scattering mechanisms in SAR. *IEEE Geosciences and Remote Sensing Letters*, *5*, 468—470.

Baraldi, A., & Parmiggiani, F. (1995). An investigation of the textural characteristics associated with gray level co-occurrence matrix statistical parameters. *IEEE Transaction on Geosciences and Remote Sensing*, *33*, 293—304.

Batitti, R. (1994). Using mutual information for selecting features in supervised neural net learning. *IEEE Transactions on Neural Networks*, *5*, 537—550.

Bazi, A., & Melgani, F. (2006). Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *11*, 3374—3376.

Belue, L. M., & Bauer, K. W. (1995). Determining input features for multilayer perceptrons. *Neurocomputing*, *7*, 111—121.

Benediktsson, J. A., Pesaresi, M., & Arnason, K. (2003). Classification and feature extraction for remote sensing images from urban areas based on morphological transformations. *IEEE Transaction on Geosciences and Remote Sensing*, *9*, 1940—1949.

Benediktsson, J. A., Palmason, J. A., & Sveinsson, J. R. (2005). Classification of hyperspectral data from urban areas based on extended morphological profiles. *IEEE Transaction on Geosciences and Remote Sensing*, *43*, 480—491.

Bishop, C. (1995). *Neural networks for pattern recognition*. New York: Oxford Univ, Press.

Bruzzone, L., & Carlin, L. (2006). A Multilevel context-based system for classification of very high spatial resolution images. *IEEE Transaction on Geosciences and Remote Sensing*, *44*, 2587—2600.

Carleer, A. P., & Wolff, E. (2006). Urban land cover multi-level region-based classification of VHR data by selecting relevant features. *International Journal of Remote Sensing*, *27*, 1035—1051.

Castellano, G., Fanelli, A., & Pelillo, M. (1997). An iterative pruning algorithm for feedforward neural networks. *IEEE Transaction on Neural Networks*, *8*, 519—531.

Chandrasekaran, H., Chen, H. H., & Manry, M. T. (2000). Pruning of basis functions in nonlinear approximators. *Neurocomputing*, *34*, 29—53.

Chen, D., Stow, D. A., & Gong, P. (2004). Examining the effect of spatial resolution and texture window size on classification accuracy: An urban environment case. *International Journal of Remote Sensing*, *25*, 2177—2192.

Chini, M., Bignami, C., Brunori, C., Atzori, S., Tolomei, C., Trasatti, E., et al. (2008). The SIGRIS project: A remote sensing system for seismic risk management. *Proc. IEEE/IGARSS 2008*.

Chini, M., Bignami, C., Emery, W. J., Pierdicca, N., & Stramondo, S. (2008). QuickBird panchromatic images for mapping damages at building scale caused by the 2003 Bam earthquake. *Proc. IEEE/IGARSS 2008*.

Chini, M., Pacifici, F., Emery, W. J., Pierdicca, N., & Del Frate, F. (2008). Comparing statistical and neural network methods applied to very high resolution satellite images showing changes in man-made structures at rocky flats. *IEEE Transactions on Geoscience And Remote Sensing*, *46*, 1812—1821.

Chini, M., Pierdicca, N., & Emery, W. J. (2009). Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake. *IEEE Transaction on Geosciences and Remote Sensing*, *47*(1), pp. 145–152 (January).

Cibas, T., Fogelman Soulié, F., Gallinari, P., & Raudys, S. (1996). Variable selection with neural networks. *Neurocomputing*, *12*, 223—248.

Clausi, D. A., & Yue, B. (2004). Comparing co-occurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery. *IEEE Transaction on Geoscience and Remote Sensing*, *42*, 215—228.

Colomboa, R., Bellingerib, D., Fasolinic, D., & Marino, C. M. (2003). Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sensing of Environment*, *86*, 120—131.

Cossu, R. (1988). Segmentation by means of textural analysis. *Pixel*, *1*, 21—24.

Dekker, R. J. (2003). Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands. *IEEE Transactions on Geoscience And Remote Sensing*, *41*, 1950—1958.

Del Frate, F., Iapaolo, M., Casadio, S., Godin-Beekmann, S., & Petitdidier, M. (2005). Neural networks for the dimensionality reduction of GOME measurement vector in the estimation of ozone profiles. *Journal of Quantitative Spectroscopy & Radiative Transfer*, *92*, 275—291.

Del Frate, F., Pacifici, F., Schiavon, G., & Solimini, C. (2007). Use of neural networks for automatic classification from high-resolution images. *IEEE Transactions on Geoscience And Remote Sensing*, *45*, 800—809.

Dell'Acqua, F., & Gamba, P. (2003). Texture-based characterization of urban environments on satellite SAR images. *IEEE Transactions on Geoscience And Remote Sensing*, *41*, 153—159.

Dell'Acqua, F., & Gamba, P. (2006). Discriminating urban environments using multiscale texture and multiple SAR images. *International Journal of Remote Sensing*, *27*, 3797—3812.

Dellepiane, S., Giusto, D. D., Serpico, S. B., & Vernazza, G. (1991). SAR image recognition by integration of intensity and textural information. *International Journal of Remote Sensing*, *12*, 1915—1932.

DigitalGlobe (2008). *WorldView-1 products quick reference guide.*: Longmont, CO [Online]: http://digitalglobe.com

DigitalGlobe (2008). *QuickBird imagery products: Product guide.*: Longmont, CO [Online]: http://digitalglobe.com

Engelbrecht, A. P., & Cloete, I. (1996). A sensitivity analysis algorithm for pruning feedforward neural networks. *Proceedings of IEEE International Conference on Neural Networks*, *2*, 1274—1277.

European Space Agency (2006). *Annual Report 2006* [Online]: http://www.esa.int/

Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*, 185−201.
Foroutan, I., & Sklansky, J. (1987). Feature selection for automatic classification of non-Gaussian data. *IEEE Transactions on Systems, Man, and Cybernetics*, *17*, 187−198.
GLP (2005). Science plan and implementation strategy. *IGBP Report No. 53/IHDP Report No. 19*. Stockholm: IGBP Secretariat.
Gong, P., Marceau, D. J., & Howarth, P. J. (1992). A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sensing of Environment*, *40*, 137−151.
Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*, 389−422.
Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceeding of the IEEE*, *67*, 786−804.
Haralick, R. M., Shanmuga, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics, SMC*, *3*, 610−621.
Hassibi, B., & Stork, D. G. (1993). Second-order derivatives for network pruning: Optimal brain surgeon. In S. J. Hanson, J. D. Cowan, & C. L. Giles (Eds.), *Advances in neural information processing systems 5* (pp. 164−171). San Mateo: Morgan Kaufmann.
Hirose, Y., Yamashita, K., & Hijiya, S. (1991). Back-propagation algorithm which varies the number of hidden units. *Neural Networks*, *4*, 61−66.
Holz, H. J., & Loew, M. H. (1994). Relative feature importance: A classifier independent approach to feature selection. In E. S. Gelsema & L. N. Kanal (Eds.), *Proceedings of pattern recognition in practice IV: Multiple paradigms, comparative studies and hybrid systems* (pp. 473–1823).
Karathanassi, V., Iossifidis, C., & Rokos, D. (2000). A texture-based classification method for classifying built areas according to their density. *International Journal of Remote Sensing*, *21*, 1807−1823.
Karthaus, V., Thygesen, H., Egmont-Petersen, M., Talmon, J., Brender, J., & McNair, P. (1995). User-requirements driven learning. *Computer Methods and Programs in Biomedicine*, *48*, 39−44.
Kavzoglu, T., & Mather, P. M. (1999). Pruning artificial neural networks: An example using land cover classification of multi-sensor images. *International Journal of Remote Sensing*, *20*, 2787−2803.
Kittler, J. (1980). Computational problems of feature selection pertaining to large data sets. In E. S. Gelsema & L. N. Kanal (Eds.), *Pattern recognition in practice* (pp. 405−414). : North Holland.
Kudo, M., & Shimbo, M. (1993). Feature selection based on the structural indices of categories. *Pattern Recognition*, *26*, 891−901.
Kurosu, T., Uratsuka, S., Maeno, H., & Kozu, T. (1999). Texture statistics for classification of land use with multitemporal JERS-1 SAR single-look imagery. *IEEE Transaction on Geoscience and Remote Sensing*, *37*, 227−235.
Kurvonen, L., & Hallikainen, M. (1999). Textural information of multitemporal ERS-1 and JERS-1 SAR images with application to land and forest type classification in boreal zone". *IEEE Transaction on Geoscience and Remote Sensing*, *37*, 680−689.
Landgrebe, D. A. (2003). *Signal theory methods in multispectral remote sensing*. Hoboken, New Jersey: John Wiley and Sons.
Le Cun, Y., Denker, J. S., & Solla, S. A. (1990). Optimal brain damage.. In D. S. Touretsky (Ed.), *Advances in neural information processing systems 2* (pp. 598−605). San Mateo: Morgan Kaufmann.
Leray, P., & Gallinari, P. (1998). Feature selection with neural networks. *Behaviormetrica Université Paris* (pp. 6).
Mao, J., Mohiuddin, K., & Jain, A. K. (1994). Parsimonious network design and feature selection through node pruning. *IEEE Proceedings of the 12th IAPR International. Conference on Pattern Recognition*, *Vol. 2*. (pp. 622−624).
Onnia, V., Tico, M., & Saarinen, J. (2001). Feature selection method using neural network. *IEEE Proceedings of the International Conference on Image Processing*, *Vol 1*. (pp. 513−516).
Ouma, Y. O., Ngigi, T. G., & Tateishi, R. (2006). On the optimization and selection of wavelet texture for feature extraction from high-resolution satellite imagery with application towards urban-tree delineation. *International Journal of Remote Sensing*, *27*, 73−104.
Pacifici, F., Del Frate, F., Solimini, C., & Emery, W. J. (2007). An innovative neural-net method to detect temporal changes in high-resolution optical satellite imagery. *IEEE Transaction on Geoscience and Remote Sensing*, *45*, 2940−2952.
Pagot, E., & Martino Pesaresi, M. (2008). 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*, *2*, 120−128.
Pesaresi, M. (2000). Texture analysis for urban pattern recognition using fine-resolution panchromatic satellite imagery. *Geographical & Environmental Modelling*, *4*, 43−63.
Puissant, A., Hirsch, J., & Weber, C. (2005). The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. *International Journal of Remote Sensing*, *26*, 733−745.
Reed, R. (1993). Pruning algorithms — A survey. *IEEE Transactions on Neural Networks*, *4*, 740−747.
Shaban, M. A., & Dikshit, O. (2001). Improvement of classification in urban areas by the use of textural features: The case study of Lucknow city, Uttar Pradesh. *International Journal of Remote Sensing*, *22*, 565−593.
Shaban, M. A., & Dikshit, O. (2002). Evaluation of the merging of SPOT multispectral and panchromatic data for classification of an urban environment. *International Journal of Remote Sensing*, *23*, 249−262.
Shackelford, A. K., & Davis, C. H. (2003). A Hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, *9*, 1920−1932.
Shanmugan, K. S., Narayanan, V., Frost, V. S., Stiles, J. A., & Holtzman, J. C. (1981). Textural features for Dadar image analysis. *IEEE Transactions on Geoscience and Remote Sensing*, *19*, 153−156.
Sietsma, J., & Dow, R. J. F. (1998). Neural net pruning: Why and how. *IEEE Proceedings of the International Conference on Neural Networks*, *Vol 1*. (pp. 325−333).
Small, C. (2003). High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment*, *88*, 170−186.
Stahlberger, A., & Riedmiller, M. (1997). Fast network pruning and feature extraction using the unit-OBS algorithm. In M. C. Mozer, M. I. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems* (pp. 655−660). MA: Morgan Kaufmann.
Suzuki, K., Horiba, I., & Sugie, N. (2001). A simple neural network pruning algorithm with application to filter synthesis. *Neural Processing Letters*, *13*, 43−53.
Tarr, G., (1991). Multi-layered feedforward neural networks for image segmentation, Ph. D. dissertation, School of Engineering, Air Force Institute of Technology. Wright-Patterson AFB OH.
Tatem, A. J. T., Noorb, A. M., & Hay, S. I. (2005). Assessing the accuracy of satellite derived global and national urban maps in Kenya. *Remote Sensing of Environment*, *96*, 87−97.
Treitz, P. M., Howarth, P. J., Filho, P. J., & Soulis, E. D. (2000). Agricultural crop classification using SAR tone and texture statistics. *Canadian Journal of Remote Sensing*, *26*, 18−29.
Tuceryan, M., & Jain, A. K. (1993). Texture analysis. In C. Chen, L. Pau, & P. Wang (Eds.), *Handbook of pattern recognition and computer vision* Singapore: World Scientific.
Verikas, A., & Bacauskiene, M. (2002). Feature selection with neural networks. *Pattern Recognition Letters*, *23*, 1323−1335.
Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., & Mehl, H. (2007). Satellite image analysis for disaster and crisis-management support. *IEEE Transactions on Geoscience and Remote Sensing*, *6*, 1520−1528.
Weston, J., Elisseeff, A., Schölkopf, B., & Tipping, M. (2003). Use of the zero-norm with linear models and kernel methods. *Journal of Machine Learning Research*, *3*, 1439−1461.
Yacoub, M., & Bennani, Y. (1997). HVS: A heuristic for variable selection in multilayer artificial neural network classifier. In C. Dagli, M. Akay, O. Ersoy, B. Fernandez, & A. Smith (Eds.), *Intelligent engineering systems through artificial neural networks* (pp. 527−532). Missouri, USA: St Louis.
Zell, A., et al. (1995). SNNS, Stuttgart Neural Network Simulator, user manual. *University of Stuttgart* Germany: Stuttgart www.ra.informatik.uni-tuebingen.de/SNNS
Zeng, X., & Yeung, D. S. (2006). Hidden neuron pruning of multilayer perceptrons using a quantified sensitivity measure. *Neurocomputing*, *69*, 825−837.
Zhang, Q., & Couloigner, I. (2006). Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery. *Pattern Recognition Letters*, *27*, 937−946.
Zhang, Q., Wang, J., Gong, P., & Shi, P. (2003). Study of urban spatial patterns from SPOT panchromatic imagery using textural analysis. *International Journal of Remote Sensing*, *24*, 4137−4160.