

Urban monitoring using multi-temporal SAR and multi-spectral data

Luis Gomez-Chova^{a,*}, Diego Fernández-Prieto^b, Javier Calpe^a,
Emilio Soria^a, Joan Vila^a, Gustavo Camps-Valls^a

^a *Escuela Técnica Superior de Ingeniería, Dept. Ingeniería Electrónica, Universidad de Valencia,
C/Dr. Moliner, 50. 46100 Burjassot, Valencia, Spain*

^b *EO Science and Applications Department, European Space Agency, ESA-ESRIN, Via Galileo Galilei, 00044 Frascati, Rome, Italy*

Available online 29 September 2005

Abstract

In some key operational domains, the joint use of synthetic aperture radar (SAR) and multi-spectral sensors has shown to be a powerful tool for Earth observation. In this paper, we analyze the potentialities of combining interferometric SAR and multi-spectral data for urban area characterization and monitoring. This study is carried out following a standard multi-source processing chain. First, a pre-processing stage is performed taking into account the underlying physics, geometry, and statistical models for the data from each sensor. Second, two different methodologies, one for supervised and another for unsupervised approaches, are followed to obtain features that optimize the urban related information. Finally, classification of ‘Urban/Non-Urban’ areas is performed using standard algorithms. Multi-temporal data acquisition was carried out in the areas of Rome and Naples (Italy) in 1995 and 1999. The data set includes images from Landsat TM and 35-day interferometric pairs of ERS2 SAR images. We analyze the dependence of the classification accuracy on the selected input features. The good results obtained using selected features improve the overall classification accuracy, thus confirming the validity of our proposal.

© 2005 Elsevier B.V. All rights reserved.

Keywords: Remote sensing; Urban monitoring; Multi-spectral; SAR; Multi-source; Feature selection

1. Introduction

Monitoring urban areas at a regional scale, and even at a global scale, has become an increasingly important topic in the last decades in order to keep track of the loss of natural areas due to urban development. Earth observation (EO) using remotely sensed imagery is a relatively new tool to monitor urban growth. The combined use of optical and synthetic aperture radar (SAR) data, interferometric SAR processing, and development of new pattern recognition techniques, are expected to enable an accurate detection of the urban tissue. Moreover, the multi-temporal data acquisition allows the monitoring of urban expansion (Cas-tracane et al., 2003).

This paper analyzes the joint use of SAR and multi-spectral images to characterize and monitor urban areas. The final objective of the study is to improve the ‘Urban/Non-Urban’ classification results by optimizing data separability in the input space. In order to attain this objective, the next two issues are considered: (1) to propose a well suited multi-source approach to the optical and SAR data characteristics over urban areas; and (2) to consider different methodologies depending on the availability of a labelled training set.

On the one hand, the information retrieved by SAR and optical sensors differs to a great extent one from another. A multi-spectral image allows the reconstruction of the energy radiated by the Earth’s surface throughout the visible and infrared ranges of the electromagnetic spectrum (Clark, 1999). On the other hand, SAR complex images provide a measurement of the changes that microwaves

* Corresponding author. Tel.: +34 96 3900134; fax: +34 96 3160466.
E-mail address: luis.gomez-chova@uv.es (L. Gomez-Chova).

suffer in amplitude and phase when they interact with the Earth's surface (Oliver and Quegan, 1998). Therefore, SAR provides information that may not be obtained from optical sensors: SAR images are related to the dielectric properties of the target returning, are sensitive to micro-textures, and microwave radiation penetrates the soil and canopies depending on frequency and polarization. In repeat-pass SAR interferometry, the interferometric phase measures distance to the targets, and the correlation between the complex SAR image pair, known as *coherence*, measures the temporal stability. In addition, SAR images overcome two of the main drawbacks of optical sensors: they are illumination-independent since they come from active sensors, and weather-independent since the used wavelength is almost unaffected by clouds, fog, rain, etc. However, the coherent nature of the microwave signal gives rise to a phenomenon called *speckle*, which can be modelled as multiplicative random noise (Ulaby et al., 1986), and tends to mask the back-scattering characteristics of the observed objects (*granular noise*). As a consequence, the coherent image statistics of SAR data will be dominated by the presence of speckle. In this context, images of different nature, and hence with different properties, should be processed all together properly, something that cannot be readily done by means of standard processing methods. For these reasons, it is important to develop a methodology properly designed for the optimal exploitation of SAR and multi-spectral images.

On the other hand, this work is aimed at improving classification results using all the information and knowledge about the problem. However, in many cases, especially in continental or regional applications, a labelled training set is difficult to be collected on a regular basis. Thus, we

consider the peculiarities of supervised and unsupervised methods. Classical and advanced techniques have proven to be well suited in this kind of classification tasks but some problems are encountered: in supervised methods, the learning process heavily depends on the quality of the training data set (Fukunaga and Hayes, 1989); and in unsupervised methods, the relationship between clusters and classes of interest is not always ensured. In order to overcome these drawbacks, we propose to obtain potentially discriminative sets of features of 'Urban/Non-Urban' areas by using a different methodology for unsupervised and supervised classifiers, as follows:

- In unsupervised approaches, clusters are usually built attending to data similarity criteria, which could be not related to the underlying discriminative information among 'Urban/Non-Urban' areas. Hence, the classification process requires a previous extraction of the potentially discriminative features to improve the clustering process. The representative features for our problem are generated attending to SAR and optical data information content over urban areas (Strozzi et al., 2000; Gouinaud et al., 1996; Carrilero et al., 2001).
- When using supervised methods, a subset of features using feature selection algorithms to improve the classification results can be extracted (Gomez-Chova et al., 2003a,b).

2. Material

Data used in this work were collected in the Urban Expansion Monitoring (UrbEx) ESA-ESRIN DUP project (Castracane et al., 2003). Results from UrbEx project were

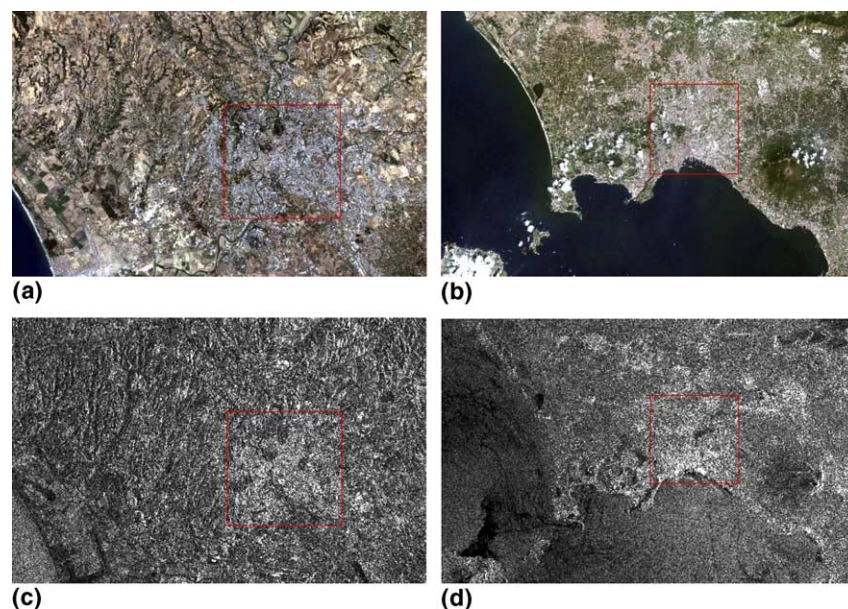


Fig. 1. Images of the test areas of Rome (a,c) and Naples (b,d) acquired at 1999: RGB composite from L3, L2, and L1 bands (a,b); and SAR *log*-intensities (c,d).

used to perform the analysis of the selected test sites and for validation purposes as well (for further details, visit <http://dup.esrin.esa.int/ionia/projects/summary30.asp>). The test sites were Rome and Naples (Italy), where images from ERS2 SAR and Landsat TM sensors were acquired in 1995 and 1999 (Fig. 1). For each one of the four *scenarios* the data set consisted of two multi-temporal 35-day ERS2 SAR images and one multi-spectral 7-band Landsat image. Moreover, an external Digital Elevation Model (DEM) and a reference land cover map provided by the Italian Institute of Statistics (ISTAT) were also available. The ERS2 SAR pairs were selected with perpendicular baselines between 20 and 150 m in order to obtain the interferometric coherence from each complex SAR image pair. The available images were labelled as: $L1$ – $L7$ for Landsat bands; $In1$ – $In2$ for the SAR backscattering intensities (0–35 days); and Co for coherence. Image acquisition from the same area in two dates under different conditions guarantees robustness and stability of the results. Therefore, for each one of the two sites and dates, the *ISTAT ground truth* classification was used for a detailed training and validation sets selection. The total number of considered labelled samples were 297,199 (approximately 25% from each *scenario*). Two thirds of these samples were randomly selected to build a training set, and the remaining was reserved for validation purposes. Assessment on balanced training-validation sets was performed in terms of mean and variance of the input features. This allows us to test the different capabilities of our data in a systematic way, and to select the best models. In addition, the *ISTAT ground truth* classification can be used as a reference to estimate the classification accuracy over full images (1900×1200 pixels).

3. Methodology

3.1. Pre-processing of optical and SAR data

Since images come from different sensors, the first step is to perform a specific processing and conditioning of optical and SAR data, and to co-register all the images. The seven bands of Landsat TM were co-registered with the ISTAT classification data, and resampled to 30×30 m with the Nearest-Neighbor algorithm. This processing was automatically performed taking into account the geo-location information available from Landsat data and using a cross-correlation technique over edge-detected images. The DEM had a pixel spacing of 20 m and was used to correct the topographic effects in SAR images over rugged terrains in two different ways: (i) subtracting the topographic phase of the interferometric phase in the coherence estimation; and (ii) normalizing the backscattering intensity for the true pixel size to reduce slope effects. All this SAR processing was performed over a five-azimuth-looks interferograms and intensity images, in which we increased the number of looks up to five by averaging five times in the azimuth direction in order to have approximately square pixels and to reduce speckle. As SAR images were regis-

tered over the geo-coded DEM, the optical and SAR data were automatically co-registered and resampled resulting in a pixel size of 30×30 m. The registration for the multi-source images was performed at sub-pixel level obtaining a root mean squared (RMS) error of about 15 m, which potentially enables good urban classification ability.

The second step consists of combining both sensor data in order to have a suitable set of input features that facilitate the work of the subsequent pattern recognition methods. The adopted approach in this work is to use consolidated, general-purpose techniques, which are valid for a restrictive but simple statistical data model, and adapt the input data to their assumptions through a pre-processing phase. In particular, if the collected data followed a Gaussian distribution, one could use a wide range of methods and techniques developed under this hypothesis. The drawback is that the Gaussian probability density function (pdf) is ill-suited for SAR data. A lot of research has been done to find a suitable distribution for SAR images (Frery et al., 1999). A modern unification of distributions for the backscatter (Frery et al., 1995) states that the Generalized Inverse Gaussian distribution is quite a good general model. However, this pdf model does not allow a straightforward embedding into advanced statistical methods, such as adaptive filters, principal component transformation, maximum likelihood classification, clustering algorithms, etc. In addition, the SAR intensity image is affected by speckle as a multiplicative noise, which disturbs an overall pdf model assumption. However, we can model SAR intensity images with the K -distribution and, when the number of looks is high enough, it approximates to a *log*-normal distribution (Oliver and Quegan, 1998). Thus, a Gaussian distribution can be obtained by *log*-transforming SAR intensities (Pellizzeri et al., 2003). Then, we can assume that, in homogeneous areas of the scene, all variables present a Gaussian distribution approximately. With this hypothesis in mind, we can combine the optical and SAR features using a stacked vector approach (Pellizzeri et al., 2003; Weydahl et al., 1995) and model the joint pdf with the well-known multi-variate normal distribution (Duda et al., 2000).

In addition, we can take advantage of a priori knowledge about the optical and SAR data characteristics. Before the classification, we can use the multi-spectral information: first, by masking covers without any relation with our study (e.g. water bodies); and second, by marking the areas where optical information is not reliable and only SAR data will be used (e.g. cloud covers).

3.2. Feature extraction in urban monitoring

When a labelled training set is not available, one is forced to use unsupervised methods or methods based on previous knowledge about the problem. Usually, unsupervised classification methods build clusters attending to data similarity criteria that might not be related to the underlying and potential discriminative information between

‘Urban/Non-Urban’ areas. If we include features that are not related to this information, classification results could be wrong or misleading. Therefore, analysis of the optical and SAR data characteristics over urban areas, and definition and extraction of discriminant features are mandatory when using unsupervised or knowledge-based methods.

3.2.1. Synthetic aperture radar data

The primary factors influencing an object’s return signal intensity are their geometric and electrical characteristics. Due to the reflectivity and angular structure of buildings, bridges, and other human-made objects, these targets tend to behave as corner reflectors and show up as bright spots in a SAR image. Therefore, in the SAR intensity images, the urban areas are characterized by very bright reflections (given by the presence of planes at 90° angles) which are maintained in time and under varying angles (Strozzi et al., 2000). In repeat-pass interferometry, when the satellite is on an exact repeat orbit—baseline zero—and the phase contribution due to topography (DEM) is also removed, the absolute value of coherence provides information on measurement stability over time. Therefore, since urban areas do not have strong and fast changes, it should lead to a high coherence (Fanelli et al., 2000).

The main problem is that images acquired by SAR are accompanied with speckle, which disturbs the image interpretation since individual pixels provide very inaccurate measurements. In order to reduce speckle in SAR images, the simplest approach is to apply multi-look processing during the image formation (average over several resolution cells). Numerous filtering techniques, based on digital image processing techniques or statistical models, have been proposed to reduce and smooth the speckle noise, e.g. the Lee (Lee, 1981), Frost (Frost et al., 1982), Kuan (Kuan et al., 1985), and Gamma MAP (Lopes et al., 1993) filters. In all these speckle noise reduction techniques, the statistical distribution of SAR data plays an important role. For multi-look images, speckle noise obeys a Gamma distribution and can be modelled as a multiplicative noise (Ulaby et al., 1986): $\text{In}(x, y) = S(x, y) \cdot \mathcal{N}(x, y)$, where $\text{In}(x, y)$ is the intensity of an observed image pixel, $S(x, y)$ is the noise-free image pixel and $\mathcal{N}(x, y)$ is the noise, characterized by a distribution with unit mean. However, by applying the logarithmic transformation to SAR intensities, the multiplicative nature of noise is changed to additive one. Taking logarithms to both sides of the equation, we obtain: $\log \text{In}(x, y) = \log S(x, y) + \log \mathcal{N}(x, y)$, where $\log \mathcal{N}(x, y)$ is the additive noise with zero mean distribution. A direct consequence of this approach is that now the filters to reduce image noise are not limited to the multiplicative noise model.

In consequence, the original images cannot be analyzed at a pixel level (as occurs in optical imagery), and thus, each pixel has to be analyzed within its neighborhood. In this work, we extract two features from the 35-day interferometric pairs of Single Complex Look SAR images: a spatially filtered coherence map and a spatially filtered image

obtained from the two multi-temporal intensity images. The objective of this processing is not limited to reduce image noise in order to obtain the true backscattering (the mentioned standard filters could do this task). We propose an application-oriented spatial filtering based on previous knowledge and optimized to differentiate urban areas. The resulting features must improve ‘Urban/Non-Urban’ classification results even though we cannot discriminate other kind of classes or phenomena.

Since high intensity is an indicator of urban areas, taking the minimum between the intensity values of two multi-temporal passes emphasizes them with respect to other areas, which could show high intensities by chance. Therefore, before filtering, we combine the two intensity images of the interferometric pair. The next step consists of a non-linear spatial filtering. This processing aims to convert a given image I containing intensity patterns of urban areas into a transformed image I' which can be directly interpreted as a *probability of urbanization*. That is: (i) the resulting image will present high homogeneous values in urban areas and low values in all the other cover types; (ii) ‘Urban/Non-Urban’ classes must present high separability; and (iii) it must be resampled without information loss. The later issue is important since we are interested in detecting all urban agglomerations rather than its internal structure. For this purpose, three type of filters are used:

- A maximum filtering of window size $N \times M$ applied to an image $I(x, y)$ yields an output image $\mathcal{I}(x, y) = I_{N \times M}^{\max}(x, y)$, which is determined by the maximum of the neighborhood pixels $N \times M$. We use this filter to maximize the already high values over urban areas, i.e. local window with at least one high intensity pixel (e.g. corner reflector), avoiding low intensity pixels inside urban areas.
- Similarly, a median filtering yields an output image $\mathcal{I}(x, y) = I_{N \times M}^{\text{median}}(x, y)$, where the already high values over urban areas are maximized. The median filter is much less sensitive than the mean to extreme values. We use this filter to keep the value that represents the majority number of pixels in a local window.
- A pixel-wise adaptive Wiener filter (Lim, 1990) performs a smoothing of an image $I(x, y)$ by estimating the local mean $\mu_{N \times M}$ and variance $\sigma_{N \times M}^2$ within an $N \times M$ local neighborhood of each pixel (x, y) , and gives the filtered image as $\mathcal{I}(x, y) = I_{N \times M}^{\text{wiener}}(x, y) = \mu_{N \times M} + \frac{\sigma_{N \times M}^2 - v^2}{\sigma_{N \times M}^2} \times [I(x, y) - \mu_{N \times M}]$, where v^2 is the noise variance (average of all the local estimated variances). The maximum and median filter are used with small windows to enhance the urban areas. We use the Wiener filter in a second filtering stage with a wider window to reduce additive noise (e.g. speckle of a *log*-transformed SAR intensity) while preserving sharpness and detail.

The combination of filters applied to the SAR images is as follows:

$$I'(x, y) = \frac{1}{2} \left\{ I(x, y)_{5 \times 5}^{\max} \right\}_{9 \times 9}^{\text{wiener}} + \frac{1}{2} \left\{ I(x, y)_{7 \times 7}^{\text{median}} \right\}_{9 \times 9}^{\text{wiener}}$$

where I represents either the minimum intensity image, $I = \min(\log I_{n1}, \log I_{n2})$, or the coherence map, $I = \text{Co}$. These constitute the two main features extracted from SAR image processing.

3.2.2. Multi-spectral optical data

The information contained in multi-spectral images allows the reconstruction of the energy radiated by the Earth's surface throughout the visible (VIS) and infrared (IR) ranges of the electromagnetic spectrum. The properties of reflection, absorption, and emission of the materials in certain wavelengths, due to their composition and molecular structure, make possible the characterization and identification of the observed materials from their spectral signature. In particular, urban areas present no chlorophyll absorption, like soil and dry organic matter, with under-developed red edge, and absorption features of iron bearing minerals (Carrilero et al., 2001). Seven spectral bands that Landsat TM provides in each multi-spectral image (three VIS, one Near IR, two Short-Wave IR, and one Thermal IR) constitute the initial optical features.

3.3. Feature selection methods

When dealing with supervised methods, a high number of input features related to the number and quality of samples can induce the model to overfit the data or to the well-known problem of the *curse of dimensionality* (Hughes, 1968). We alleviate this problem by applying a feature selection stage in order to reduce the input space dimension. Two methods are explored: sequential floating forward selection (SFFS), and inspection of main and surrogate splits in a classification and regression tree (CART).

- The SFFS algorithm identifies the variables that better discriminate among classes in a two-stage feature selection process. Firstly, a search strategy for feature group selection is carried out, and secondly, the objective function that evaluates the different subgroups is calculated. The SFFS algorithm produces a feature selection giving the best subset of features for each dimension. In (Gomez-Chova et al., 2003a), we proposed a methodology that uses an objective function that maximizes the mean probabilistic distance between classes. We use the Bhattacharyya metric as the objective function to be maximized, since it has better properties than Euclidean metric.
- CART is a binary decision tree algorithm (Breiman et al., 1984) which has two branches in each internal node. It is based on the classification success and not on a similarity criterion, it is non-parametric, and does not impose a specific functional form. In a previous work (Gomez-Chova et al., 2003b), we proposed to

select those features that better discriminate among classes by analyzing nodes and main splits of the optimal classification tree. The SFFS method can use the accuracy of a classifier as an objective function too, but CART allows knowledge discovery and full interpretability by analyzing the tree structure.

4. Results

4.1. Knowledge-based feature extraction

In Section 3.2, some features extracted from SAR and multi-spectral data were introduced to improve the results of an unsupervised classifier by optimizing the differences between 'Urban/Non-Urban' areas. A first approach to test the goodness of these features for our test sites is the calculation of the histograms of 'Urban/Non-Urban' classes over the whole image. We generate the histograms for each class in order to automatically find the thresholds between classes and their corresponding classification error (i.e. class overlapping). The histograms are represented with the *optimum resolution cell* (number of bins) in order to compensate the effects of a finite number of samples. This procedure consists on iteratively reducing the number of bins, and checking stability of the number and position of thresholds. This yields a smoother histogram with small error. In addition, the effect of a priori probabilities is compensated by normalizing the histograms by the number of samples of each class.

In the histograms plotted in Fig. 2, one can see the performance of the *log*-transformation of the SAR intensities. In Fig. 2a, the *log*-normal distribution is well suited to the data histograms, and in Fig. 2b the histograms of the *log*-arithmetic of the intensities are modelled with the corresponding Gaussian distribution. When the proposed spatial filtering is applied, both cases experiment an improvement in the separability of the two classes. However, the resulting class distribution persists Gaussian when filtering the *log*-transformed intensities (Fig. 2d) and is distorted when the original ones are filtered (Fig. 2c). This confirms the change from multiplicative to additive noise model when applying logarithms since the Wiener adaptive spatial filtering is appropriate for additive noise but not for speckle in the original image (in fact the Frost filter is a Wiener filter adapted to multiplicative noise (Frost et al., 1982)).

Following the procedure explained in Section 3.2.1, to calculate the minimum of both multi-temporal intensities for each pixel, and then applying the spatial filtering, drastically increases accuracy from 54% to 75%. However, the spatial resolution after filtering is lower. In addition, this filter is useful not only to combine intensities but to filter the coherence map. In some coherence maps with poor discrimination between classes (Fig. 3), classification accuracy increases from 51% to 89% when applying the spatial filter. Since coherence of urban areas is high, the proposed filter maximizes the value of local areas with high coherence. In

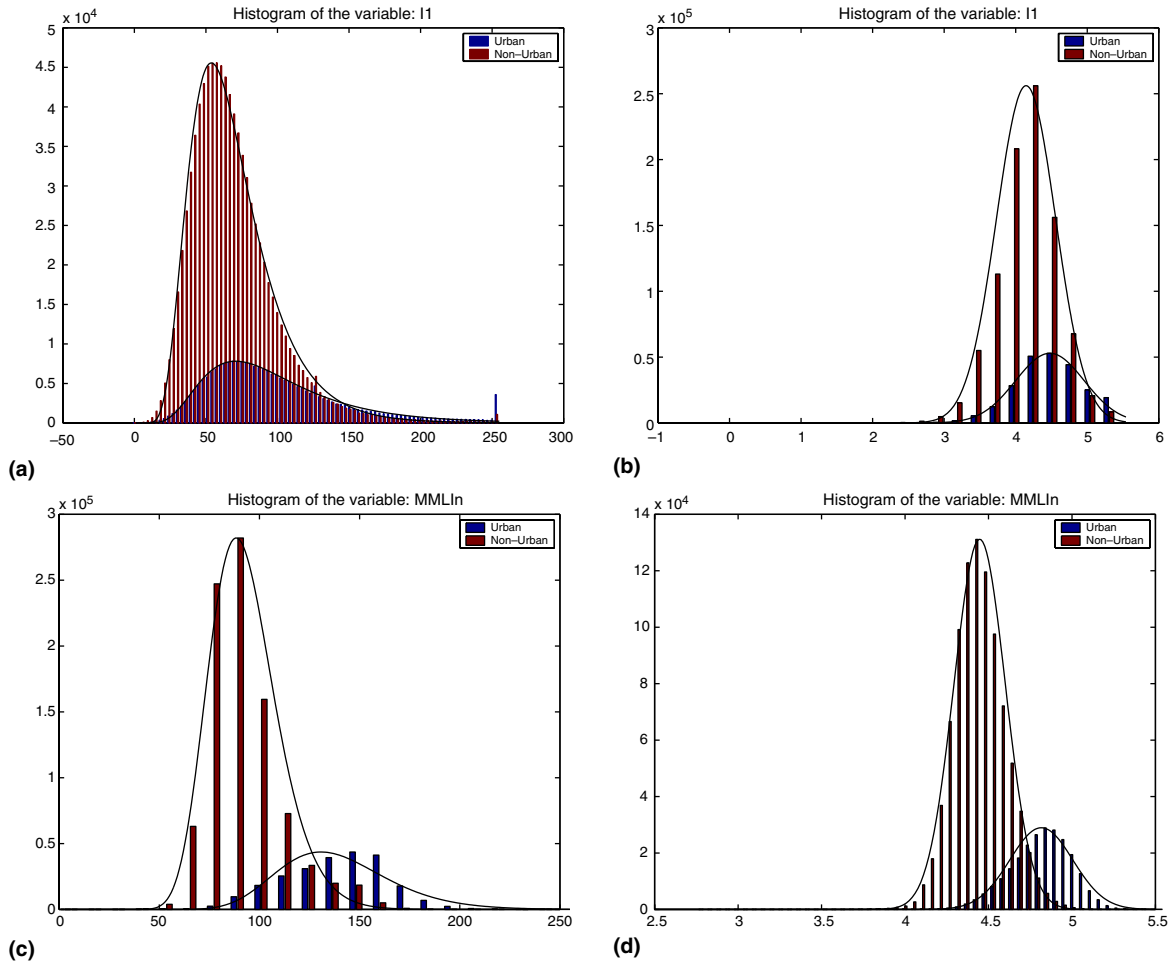


Fig. 2. Histograms of the original (a) and \log -transformed (b) intensities, and the respective outputs of the adaptive spatial filter (c,d).

the histograms, one can appreciate that the coherence distribution differs from a Gaussian (Fig. 3b), but after filtering (Fig. 3d) this assumption is acceptable (because of the high equivalent number of looks achieved (Strozzi et al., 2000)).

The combination of optical and SAR features through a stacked vector approach can improve results, which can be further enhanced by taking advantage of a priori knowledge about optical and SAR data characteristics. For instance, the spectral information can be useful to simplify the analysis by masking covers without any relation to our study (e.g. water bodies difficult to see in SAR data), and to identify cloud covers by labelling the areas where the optical information is not reliable and only SAR data can be used.

When using multi-spectral data, difference measures are preferred (*continuum* removal, band ratios, absolute indexes) since they are robust to variations in illumination and atmospheric conditions. In our case, we computed the normalized vegetation index (NDVI), band ratios between Landsat bands 1 (*VIS*) and 4 (*Near IR*), the *cloud mask* generated using a global reflectance threshold and a temperature threshold level in the thermal band 6 (cold clouds and warmer ground surface), and *water mask* gener-

ated using Landsat bands 1, 4, and 5, due to the high water absorption over 800 nm wavelengths.

4.2. Clustering heterogeneous classes

So far, we have made an effort to ensure that our data follow approximately a Gaussian distribution. In this case, we can model our data with their second order statistics (mean and covariance matrix) and use a wide range of well-established statistical methods (in our case the probabilistic Bhattacharyya distance for normal distributions and the Gaussian Maximum Likelihood classifier). However, to obtain good results when employing these methods, we must ensure the Gaussian distribution of classes (and not only of features). This assumption could be false when dealing with heterogeneous classes as 'Urban/Non-Urban', but in any case the distribution of each class can be modelled as a mixture of Gaussian components (Duda et al., 2000). For this reason, we first perform an unsupervised clustering for each training data class to detect the data subclasses that should be better handled separately. Since the number of clusters in each class is unknown, we repeat a k -means procedure for a set of different number of clusters, k . The Davies–Bouldin index (Davies and

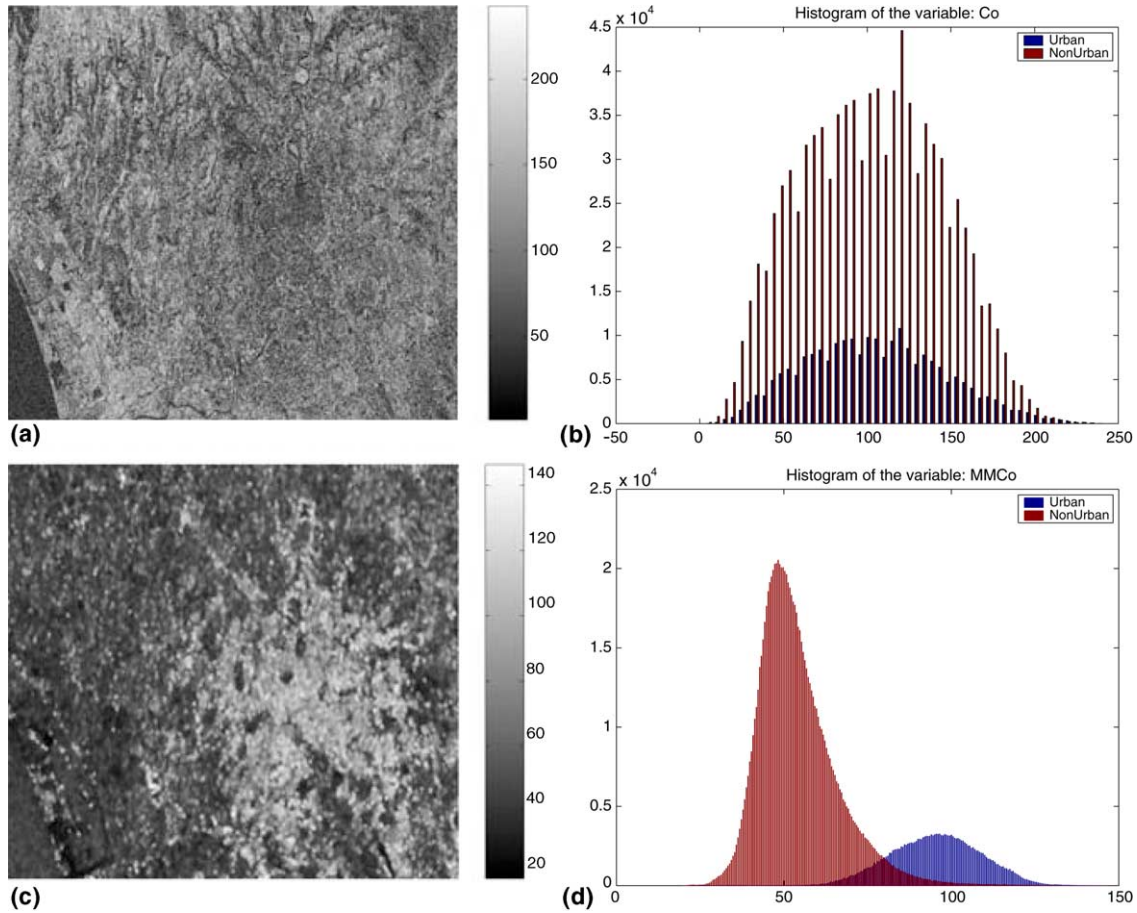


Fig. 3. Image and histograms of the coherence (up) and the spatially filtered coherence (down) which improves the classification accuracy from 51% to 89% (Rome'99).

Bouldin, 1979) is calculated for the identified clusters with each k , and the one with lowest index is selected. Due to the random initialization of the k -means, if we run the clustering algorithm only once, the results could be suboptimal and skewed. Therefore, for each specific number k of clusters, the k -means is run 10 times and the best clustering is

selected on the basis of the obtained sum of squared errors. With this procedure, we cluster all labelled samples (training and validation) of both 'Urban/Non-Urban' classes in nine and eight subclasses respectively. These number of clusters is defined by the minimum Davies–Bouldin index as is shown by the vertical bars in Fig. 4.

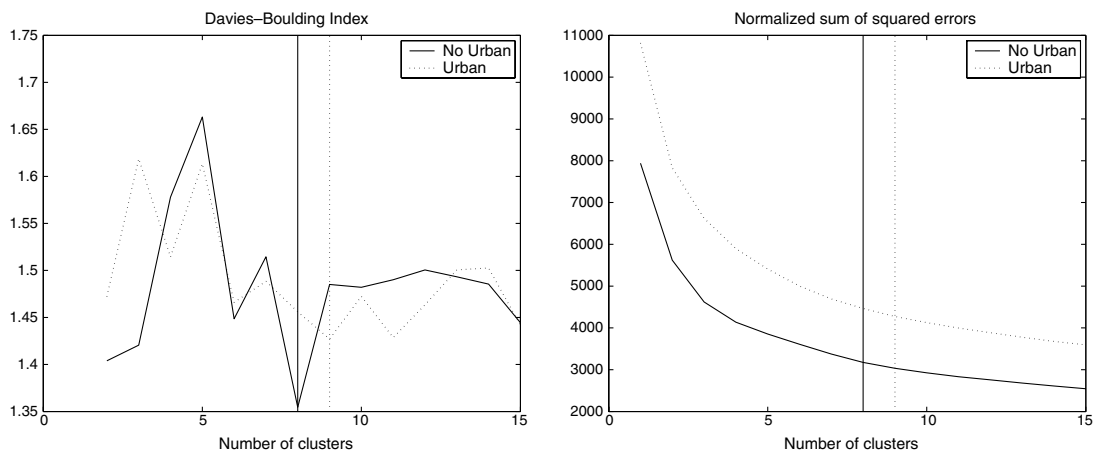


Fig. 4. Davies–Bouldin index (left) and sum of squared errors (right) as a function of the number of clusters of the Urban and Non-Urban classes. Vertical bars mark the best number of clusters selected by the minimum Davies–Bouldin index.

These subclasses do not have a priori meaning and we consider them only for subsequent feature selection and classification. In the SFFS algorithm, we treat them separately to compute the probabilistic Bhattacharyya distance among subclasses, but later we compute a weighted average distance between the urban and non-urban classes. Similarly, in the Gaussian maximum likelihood (GML) classifier, we perform the classification for all the subclasses, but at the end we obtain the overall classification accuracy (CA%) only for the two classes of interest, i.e. classification accuracy weighted by the number of samples in each class. We confirm the validity of this approach since the classification accuracy using all features increases from 75.93% (when considering only the two classes) to 87.49% (using subclasses).

4.3. Selected features by the SFFS algorithm

We apply the SFFS algorithm in order to find the set of features that maximizes the average probabilistic Bhattacharyya distance between subclasses of Urban and Non-Urban class. In order to test the sets of selected features, we compute the overall classification accuracy of the Gaussian maximum likelihood (GML) classifier. Results are computed using the clustering of the classes (9 and 8 subclasses), so avoiding the non-Gaussian heterogeneous classes.

Since the number of added or removed features is adapted during the process, the SFFS algorithm gives the best subset of features for each dimension. As shown in Table 1, the selection process is not a mere addition of the best next feature (see best subsets of three and four features). Using these selected features improves the classification accuracy over the validation set and yields valuable information on the relevance of input features.

4.4. Selected features by the CART algorithm

CART processing is carried out using a shareware implementation from Salford Systems (Steingberg and Colla, 1997), in which the complexity and accuracy of the final tree can be easily controlled. Analysis of surrogate and main splits in CART yields valuable information on

Table 1
Best subset of features selected by means of the SFFS algorithm for each number of variables (DIM) and overall classification accuracy (CA) over the validation set for the selected features

DIM	Best feature subset	CA (%)
1	L1	74.61
2	L1, L5	78.65
3	L1, L5, Co	83.21
4	L1, Co, L4, L6	86.91
5	L1, L5, Co, L4, L6	87.04
6	L1, L5, Co, L4, L6, L7	87.37
7	L1, L5, Co, L4, L6, L7, In2	87.55
8	L1, L5, Co, L4, L6, L7, In2, L2	87.57
9	L1, L5, Co, L4, L6, L7, In2, L2, L3	87.72
10	L1, L5, Co, L4, L6, L7, In2, L2, L3, In1	87.49

Table 2
Ranking of features obtained by means of the CART algorithm for each dimension (DIM), number of nodes of the CART which includes the new feature, and overall classification accuracy (CA%) over the validation set when adding the new feature

DIM	Feature	Nodes	CART-CA (%)	GML-CA (%)
1	L1	1	75.40	74.61
2	L6	2	78.64	77.54
3	Co	4	81.49	82.86
4	L4	8	84.63	86.91
5	L2	9	85.44	86.69
6	L5	10	85.77	87.26
7	L3	12	86.37	87.44
8	In2	18	87.00	87.56
9	In1	19	87.30	87.35
10	L7	89	88.80	87.49

the relevance of input features. Each variable in the tree has an importance score based on how often and with what significance it serves as primary or surrogate splitter throughout the tree. With CART is not necessary to use the subclasses since it is a non-parametric method and does not assume a Gaussian distribution. Table 2 shows a ranking of variables according to this measure. Confidence on this analysis can be ensured since classification rates of the best tree (all features and 157 nodes) achieves CA higher than 88% in both classes for the validation set, suggesting that the underlying differences between classes has been captured. In order to show the goodness of the obtained results, Table 2 summarizes the CA when different sets of features are used with CART and GML classifiers.

4.5. Dimensionality reduction

Finally, in order to compare and test the validity of the proposed supervised methodologies, we analyze the classification accuracy as a function of the input dimension. The

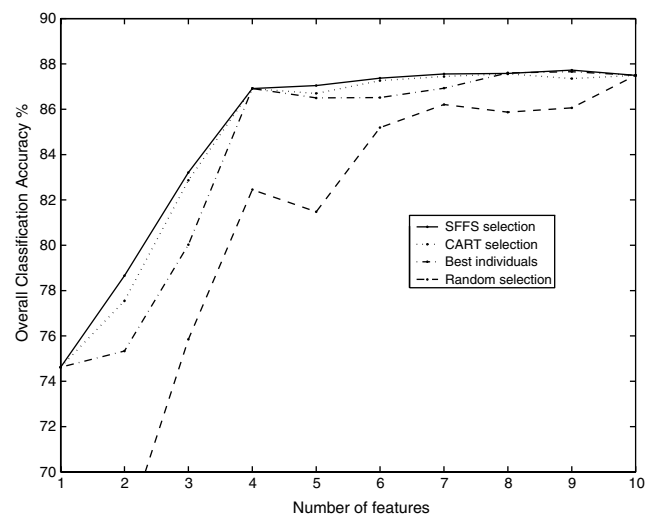


Fig. 5. Classification accuracy for different selection procedures as a function of the number of selected features: ‘SFFS’, ‘CART’, ‘best individuals’, ‘random selection’.

results obtained using the GML classifier with different feature selection methods show that the classification success rate does not improve beyond five features, or even falls when adding more features less related with information characteristic of urban areas. Fig. 5 shows the CA as a function of the number of features for different selection procedures: ‘SFFS’, selected features by the SFFS method in Table 1; ‘CART’, selected features by the CART method in Table 2; ‘best individuals’, addition of the best not-included individual feature attending to the Bhattacharyya distance (from most to less discriminative: L1, L4, Co, L6, L7, L2, L3, L5, In1, In2); ‘random selection’, random selected features (plotted case: L2, In2, L4, L6, In1, Co, L3, L7, L5, L1). It is noticeable that the SFFS represents an upper bound of the CA due to the searching of the best feature subset for each dimension.

5. Conclusions

In this communication, we have analyzed the potential of combining SAR and multi-spectral images to characterize and monitor urban areas. Multi-source strategy has been proposed taking into account the optical and SAR data characteristics over urban areas. A pre-processing of both sensor data has been performed to obtain a suitable set of input features which confirm the Gaussian assumption. We have proposed two procedures, depending on the availability of a labelled training set, to reduce dimensionality while preserving relevant information for urban classification. The first one is a knowledge-based approach that enhances the discriminative information between ‘Urban/Non-Urban’ areas. For supervised problems, we have proposed feature selection procedures based on SFFS and CART analysis algorithms. These feature selection procedures yield the following benefits: (i) increase the performance of the classifier by mitigating the Hughes phenomenon; (ii) reduce the amount of data allowing faster calculations, which in turn, enables training iterative methods, such as artificial neural networks; (iii) using the selected features in the classifier optimizes the class separation, allowing easy recognition of the natural structure in data; and (iv) the feature ranking provides a criterion to further reduce the input dimension. In our experiments, the classification results are quite satisfactory. The procedure is able to reduce the dimension up to five features without significantly affecting the classification accuracy. Finally, we can conclude that the extracted features are directly related to the urban cover properties.

Acknowledgments

This research has been supported by the *Ministerio de Educación y Ciencia* (ESP2004-06255-C05-02) and the *European Space Agency*. Results of this work will be applied in the ESA Category-1 project (CIP-ID2489) titled “Development of an Specialized Classification System for

Urban Monitoring at Regional Scale Based on ASAR and MERIS data”. All data were acquired in the Urban Expansion Monitoring (UrbEx) ESA project.

References

- Breiman, L., Friedman, J., Olshen, R., Stone, C., 1984. Classification and Regression Trees, third ed. Chapman & Hall, New York.
- Carrilero, A., Maitre, H., Roux, M., 2001. Material determination from reflectance properties in aerial urban images. In: 11th Int. Conf. on Image Analysis and Processing, vol. I, pp. 553–558.
- Castracane, P., Iavarone, F., Mica, S., Sottile, E., Vignola, C., Arino, O., et al., 2003. Monitoring urban sprawl and its trends with EO data. UrbEx, a prototype national service from a WWF-ESA joint effort. In: 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, pp. 245–248.
- Clark, R.N., 1999. Spectroscopy and principles of spectroscopy, manual of remote sensing. John Wiley and Sons, Inc., Available from: <<http://speclab.cr.usgs.gov>>.
- Davies, D., Bouldin, D., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Machine Intell., PAMI 1 (2), 224–227.
- Duda, R., Hart, P., Stork, D., 2000. Pattern Classification, second ed. John Wiley & Sons, NY.
- Fanelli, A., Santoro, M., Vitale, A., Murino, P., Askne, J., 2000. Understanding ERS coherence over urban areas. In: ESA-SP-461 (Ed.), ERS-Envisat Symposium (p. CDROM).
- Frery, A., Freitas, C., Sant’Anna, S., 1995. Alternative distributions for the multiplicative model in SAR images. In: IEEE Int. Geosci. and Remote Sensing Symp., IGARSS95, vol. I, pp. 169–171.
- Frery, A., Freitas, C., Sant’Anna, S., Rennó, C., 1999. Statistical properties of SAR data and their consequences. Seminars of the United Nations Programme on Space Applications, vol. 10, pp. 53–62.
- Frost, V., Stiles, J., Shanmugan, K., Holtzman, J., 1982. A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Trans. Pattern Anal. Machine Intell., PAMI 4 (2), 157–165.
- Fukunaga, K., Hayes, R.R., 1989. Effects of sample size in classifier design. IEEE Trans. Pattern Anal. Machine Intell. 11 (8), 873–885.
- Gomez-Chova, L., Calpe, J., Camps-Valls, G., Martín, J., Soria, E., Vila, J., et al., 2003. Feature selection of hyperspectral data through local correlation and SFFS for crop classification. In: IEEE Int. Geosci. and Remote Sensing Symposium, IGARSS03, vol. I, pp. 555–557.
- Gomez-Chova, L., Calpe, J., Soria, E., Camps-Valls, G., Martín, J., Moreno, J., 2003. CART-based feature selection of hyperspectral images for crop cover classification. In: IEEE International Conference on Image Processing, vol. III, pp. 589–592.
- Gouinaud, C., Tupin, F., Maitre, H., 1996. Potential and use of radar images for characterization and detection of urban areas. In: IEEE Int. Geosci. and Remote Sensing Symposium, IGARSS96, vol. I, pp. 474–476.
- Hughes, G.F., 1968. On the mean accuracy of statistical pattern recognizers. IEEE Trans. Inform. Theory 14 (1), 55–63.
- Kuan, D., Sawchuk, A., Strand, T., Chavel, P., 1985. Adaptive noise smoothing filter for images with signal-dependent noise. IEEE Trans. Pattern Anal. Machine Intell. 7 (2), 165–177.
- Lee, J., 1981. Speckle suppression and analysis for synthetic aperture radar images. Opt. Eng. 25 (5), 636–643.
- Lim, J., 1990. Two-dimensional Signal and Image Processing. Prentice Hall, Englewood Cliffs, NJ.
- Lopes, A., Nezry, E., Touzi, R., Laur, H., 1993. Structure detection and statistical adaptive speckle filtering in SAR images. Int. J. Remote Sensing 14, 1735–1758.
- Oliver, C., Quegan, S., 1998. Understanding SAR Images. Artech House, Boston, MA.
- Pellizzeri, T., Gamba, P., Lombardo, P., Dell’Acqua, F., 2003. Multitemporal/multiband SAR classification of urban areas using spatial

- analysis: Statistical versus neural kernel-based approach. *IEEE Trans. Geosci. Remote Sensing* 41 (10), 2338–2353.
- Pellizzeri, T., Lombardo, P., Gamba, P., Dell'Acqua, F., 2003. Multi-source urban classification: Joint processing of optical and SAR data for land cover mapping. In: *IEEE Int. Geosci. Remote Sensing Symp., IGARSS03*, vol. II, pp. 1044–1046.
- Steingberg, D., Colla, P., 1997. *CART. Classification and regression trees*, San Diego, CA. Available from: <<http://www.salford-systems.com>>, Salford Systems.
- Strozzi, T., Dammert, P., Wegmüller, U., Martinez, J., Askne, J., Beaudoin, A., et al., 2000. Landuse mapping with ERS SAR interferometry. *IEEE Trans. Geosci. Remote Sensing* 38 (2), 766–775.
- Ulaby, F., Moore, R., Fung, A., 1986. *Microwave Remote Sensing*. Artech House Inc, Boston, MA.
- Weydahl, D., Becquey, X., Tollefsen, T., 1995. Combining ERS-1 SAR with optical satellite data over urban areas. In: *IEEE Int. Geosci. and Remote Sensing Symp., IGARSS95*, vol. III, pp. 2161–2163.