

Processing Multidimensional SAR and Hyperspectral Images With Binary Partition Tree

A new hierarchical framework for the analysis of remote sensing data is discussed in this paper. A generic hierarchical strategy is adapted to handle VHR SAR data or hyperspectral data for various applications.

By ALBERTO ALONSO-GONZÁLEZ, *Student Member IEEE*, SILVIA VALERO, *Student Member IEEE*, JOCELYN CHANUSSOT, *Fellow IEEE*, CARLOS LÓPEZ-MARTÍNEZ, *Senior Member IEEE*, AND PHILIPPE SALEMBIER, *Fellow IEEE*

ABSTRACT | The current increase of spatial as well as spectral resolutions of modern remote sensing sensors represents a real opportunity for many practical applications but also generates important challenges in terms of image processing. In particular, the spatial correlation between pixels and/or the spectral correlation between spectral bands of a given pixel cannot be ignored. The traditional pixel-based representation of images does not facilitate the handling of these correlations. In this paper, we discuss the interest of a particular hierarchical region-based representation of images based on binary partition tree (BPT). This representation approach is very flexible as it can be applied to any type of image. Here both optical and radar images will be discussed. Moreover, once the image representation is computed, it can be used for many different applications. Filtering, segmentation, and classification will be detailed in this paper. In all cases, the interest of the BPT representation over the classical pixel-based representation will be highlighted.

KEYWORDS | Binary partition tree (BPT); classification; filtering; hyperspectral images; segmentation; synthetic aperture radar (SAR) images

I. INTRODUCTION

Optical as well as radar remote sensing technologies are currently undergoing a very significant evolution in terms of the quality and the quantity of information that is produced and made available for applications. In particular, sensors are becoming more precise and can capture data at a very high resolution. This can be observed both in terms of spatial and spectral resolutions. This wealth of information is, of course, very interesting from the application viewpoint but it also generates real challenges with respect to the signal processing tasks.

One of the major challenges can be formulated as the proper handling of the signal correlation. As the number of spatial samples for the same observed area or the number of samples in the spectral domain increases, neighboring samples cannot be considered as being independent. This is in stark contrast with traditional processing approaches that used to consider that neighboring pixels could be processed independently, for example, for classification or for detection purposes. Furthermore, most signal processing approaches for multispectral images did not try to assess and exploit the correlation between spectral bands of the same pixel. As the number of spectral bands dramatically increases with hyperspectral sensors, the spectral correlation cannot be ignored anymore.

Several strategies are currently being investigated to tackle this issue of increased resolution and the subsequent

Manuscript received February 7, 2012; revised April 30, 2012; accepted June 3, 2012. Date of publication August 13, 2012; date of current version February 14, 2013. This work was supported in part by Tecnologías Electrónicas y de Comunicaciones (TEC) Project MUSEO (TEC2011-28201-C02-01) and the Comissionat d'Universitats i Recerca (CUR) of the Departament d'Innovació, Universitats i Empresa (DIUE) of the Autonomous Government of Catalonia and the European Social Fund.

A. Alonso-González, C. López-Martínez and **P. Salembier** are with the Signal Theory and Communications Department, Technical University of Catalonia, 08034 Barcelona, Spain (e-mail: carlos.lopez@tsc.upc.edu).

S. Valero is with the Signal Theory and Communications Department, Technical University of Catalonia, 08034 Barcelona, Spain and also with the Grenoble Images Speech Signals and Automatics Laboratory, Grenoble Institute of Technology (INPG), 38016 Grenoble, France.

J. Chanussot is with the Grenoble Images Speech Signals and Automatics Laboratory, Grenoble Institute of Technology (INPG), 38016 Grenoble, France.

Digital Object Identifier: 10.1109/JPROC.2012.2205209

difficulties related to spatial and spectral correlation [59], [60]. Typical examples involve Markov models or multi-resolution representations such as wavelets. Here we follow a different approach that consists in redefining the signal representation. Indeed, the traditional pixel-based representation of images is not the most appropriate one to deal with the huge amount of information produced by high-resolution remote sensing sensors when seeking to capture and use the signal correlation. A more appropriate representation should somehow group pixels with high correlation or similar properties into elementary entities that could be precisely evaluated and on which one could easily act. Moreover, in the context of a given application, the signal representation should facilitate the access to the relevant subset of elementary entities which should ideally be of low cardinality. Finally, the signal representation should be generic and support a wide range of applications. As a result, it cannot target the description of the signal at a given resolution but should support multiple resolutions.

In this paper, we discuss the interest of one such representation known as the binary partition tree (BPT). It is a region-based representation in which pixels are grouped by similarity. It is a hierarchical representation providing multiple resolutions of description. Finally, it is also a structured representation supporting easy access to subsets of regions. Recently, it has been studied in the context of remote sensing data such as multidimensional synthetic aperture radar (SAR) and hyperspectral images. In both cases, the processing strategy is the same but the specific tools used to build the image representation depend on the signal nature. Once the representation has been computed, it can be used for many applications such as filtering, segmentation, classification, and object detection. The goal of this paper is to discuss approaches for creating the BPT for multidimensional SAR and hyperspectral images and to illustrate its general applicability over a wide range of problems.

The paper is organized as follows. Section II defines the BPT notion and its corresponding processing strategy. Section III focuses on SAR images. After a brief review of the state of the art in this area, the region model and the tree construction are detailed. Then, the representation is used for speckle noise filtering and segmentation. Hyperspectral images are discussed in Section IV. This section has the same structure as the previous one: state of the art, region model definition for hyperspectral data, BPT construction and applications which cover classification and segmentation. Finally, conclusions are drawn in Section V.

II. IMAGE REPRESENTATION AND PROCESSING WITH BPT

As discussed in the introduction, our goal is to discuss a hierarchical region-based representation of images based

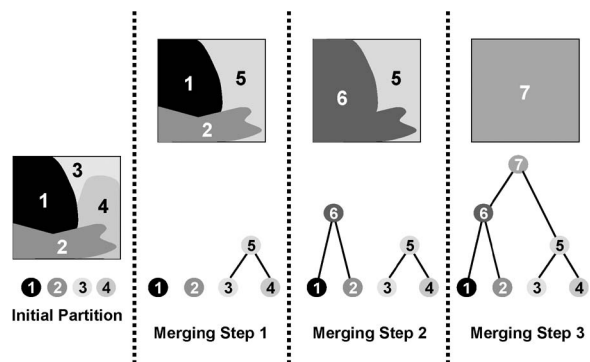


Fig. 1. Example of BPT construction.

on BPTs [45] and its application to two major image processing areas in remote sensing: multidimensional SAR and hyperspectral images. The BPT is a hierarchical region-based representation that has a rather generic construction (to a large extent, application independent). A BPT can be seen as a set of hierarchical regions stored in a tree structure. Fig. 1 is an illustration of a BPT as well as of its construction. The tree nodes represent image regions and the tree branches describe the inclusion relationship among the regions.

Three types of nodes can be distinguished: first, leaf nodes representing the regions of an initial partition; second, the root node representing the entire image support; and finally, the remaining tree nodes representing regions formed by the merging of their two child nodes corresponding to two adjacent regions. The BPT construction is often based on an iterative region-merging algorithm (although other strategies such as recursive splitting or combination of split and merge can be used). Starting from an initial partition, the region-merging algorithm is an iterative process in which the pair of most similar neighboring regions is merged. Note that if the preservation of the spatial resolution is an important issue, the initial partition may consist of regions involving only individual pixels. Each iteration implies three different steps: 1) find and merge the pair of most similar neighboring regions; 2) form a new node containing the union of the merged regions; and 3) update the similarity values between the newly created region and its neighboring regions.

In order to completely specify the merging algorithm, two important choices have to be made. The first one is the *region model*. On the one hand, the model should be precise enough in order to capture the important region features independently of their scale (regions close to the leaf nodes may represent a reduced set of pixels but regions close to the root describe very large portions of the image). But on the other hand, the model should be simple enough to allow efficient processing. In this paper, we will discuss two examples: the first one, used for SAR images,

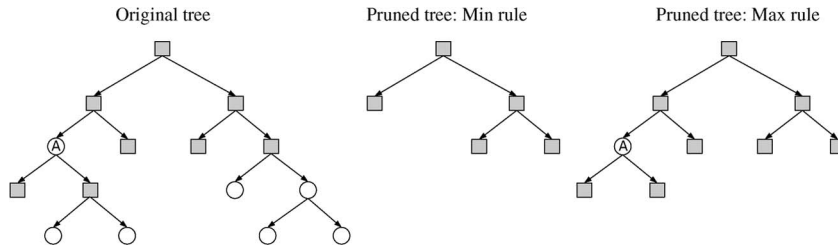


Fig. 2. Nonincreasing criterion and pruning strategies. Gray squares: nodes fulfilling the criterion; white circles: nodes not fulfilling the criterion.

will rely on a first-order statistical model, whereas the second one, applied for hyperspectral images, will be based on a nonparametric statistical model of the image frequency bands. As can be seen, one of the interesting features of the BPT approach is its flexibility in the sense that it is not limited to any particular model. Once the region model has been defined, the second important choice concerns the *region similarity*. It is used to define the order in which neighboring regions will merge and has a strong influence on the tree structure and the usefulness of the information represented by the tree nodes. This similarity measure has to take into account the nature of the data (for example, in multidimensional SAR data, pixel values are matrices where diagonal and nondiagonal elements are of different nature) and their possible correlation (for example, in hyperspectral data, information between adjacent spectral bands is often highly correlated).

Once the BPT representation has been computed, the tree is a generic and scalable image representation. It enables a very large range of applications. Most of the application-dependent processing techniques can be formulated as a tree pruning. The goal of the pruning is to remove subtrees composed of nodes which are considered to be irrelevant or homogeneous with respect to the application. The pruning is defined through a criterion ϕ_R which is evaluated on each node of the tree (which represents a region R). As an example, assume that ideally nodes (not) fulfilling the criterion have to be preserved (removed). Unfortunately, the individual evaluation of a criterion on the tree nodes does not generally define a pruning strategy. A typical example is illustrated in Fig. 2. The criterion is said to be *nonincreasing*.¹ This can be observed at node “A”: following the criterion evaluation, this node has to be removed but its two children have to be preserved. Therefore, the initial decisions do not directly define a pruning. Several strategies can be used to transform the individual criterion decisions into a pruning [46]. In this paper, two of them will be used.

- The *min rule* follows a top-down approach. Starting from the root node and going to the leaves, all tree branches are explored and the

algorithm searches for the first node such that all its ancestors have to be preserved and at least one of its child nodes has to be removed. Then, the two subtrees rooted at that node are pruned. As can be seen in Fig. 2, this strategy guarantees that all the leaves of the pruned tree correspond to regions which had to be preserved following the criterion.

- The *max rule* follows a bottom-up approach. It starts from the leaves and goes toward the root node and prunes a node if all its descendants have also to be pruned as well as its sibling (so that the pruning generates a new BPT). This rule is less strict than the *min rule* and preserves all details in the image for which the initial decision was to be preserved (see Fig. 2).

In order to show the usefulness and suitability of the BPT for a wide range of applications, we will discuss in this paper filtering, segmentation, classification, and object detection tasks.

The image processing framework based on BPT relies on two steps illustrated in Fig. 3. The first one corresponds to the construction of a generic BPT and the second one is the application-dependent pruning strategy. In the following section, we start by discussing the application of this framework to SAR images.

III. SYNTHETIC APERTURE RADAR

SARs that correspond to active microwave imaging systems are important instruments for the observation of the Earth

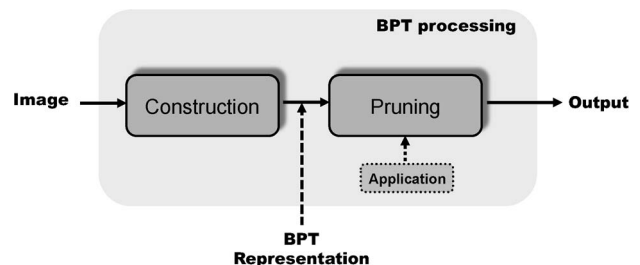


Fig. 3. Image processing framework with BPT.

¹Mathematically, a criterion ϕ_R is said to be increasing if the following property holds: $\forall R_1 \subseteq R_2 \Rightarrow \phi_{R_1} \leq \phi_{R_2}$.

surface, as its observation and analysis may be performed independently from the day and night cycle and the weather conditions, together with high spatial resolution [25], [26]. In addition, when boarded into space platforms, this type of systems allows global coverage.

In essence, a SAR system consists of a side-looking radar boarded into a moving platform to induce a relative motion between the radar and the target being observed by the system, producing finally 2-D complex images of the area of interest. In the dimension perpendicular to the system motion, referred to as *range*, a SAR system basically operates as a radar by transmitting an electromagnetic wave and recording its echo. Measuring the delay between them, a target may be located in this dimension. As in the case of a radar, the range resolution, that is, the capability to detect two close targets as being separated, depends on the wave bandwidth, in such a way that the larger this bandwidth, the higher the range resolution. On the contrary, the spatial resolution in the direction of the platform motion, known as *azimuth* dimension, depends on the antenna beamwidth in this dimension. In order to achieve a comparable space resolution to the one in range, narrow beamwidth antennas are necessary, but in the case of airborne or spaceborne systems, this requirement leads to very large and unpractical antennas. The way SAR systems circumvent this important limitation is by making use of the relative motion between the platform and the target. As the system moves, a target is repeatedly illuminated by the radar with electromagnetic pulses. The echoes are received and recorded coherently by the radar at different spatial positions. A postprocessing of the data allows to coherently combine these echoes to resolve targets by synthesizing a long array. From their conception in the 1950s to the early 1990s, SAR systems were basically 1-D systems. Despite the capability of this technology to perform Earth remote sensing which was largely demonstrated during this period, the speckle phenomenon [32], [33], [38], due to the coherent addition of the scattered electromagnetic waves, was also identified as one of the main problems for the exploitation of SAR data. The true potential of the SAR technology was realized in the early 1990s with the arrival of the multidimensional or multi-channel SAR systems, based on acquiring several complex SAR images under some sort of diversity. Among them, it is important to emphasize SAR interferometry (InSAR) [10], [35] and differential SAR interferometry (DInSAR) [4], SAR polarimetry (PolSAR) [20], [40], polarimetric differential SAR Interferometry [5], polarimetric SAR interferometry (PolInSAR) [18], or SAR tomography [43]. The importance of these techniques lies in the fact that multidimensional observations allow the quantitative estimation of a larger number of biophysical and geophysical parameters of the Earth surface. For instance, InSAR is sensitive to the terrain topography, differential techniques are able to retrieve terrain subsidence phenomena, PolSAR allows the study of the scatterer geometry and its dielectric

properties, whereas PolInSAR or SAR tomography allows the study of volumetric scatterers.

A multidimensional SAR system acquires m complex SAR images under some type of diversity $[S_1, S_2, \dots, S_m]$, where S_k for $k = 1, \dots, m$ indicates the individual complex SAR images. One of the most relevant types of diversity is based on considering different polarization states for the transmitted and received electromagnetic waves, making SAR data to be sensitive to the target geometry, including vegetation, and to the dielectric properties of the target. Hence, PolSAR data have been employed in a wide range of different applications: terrain classification and target identification, quantitative humidity, terrain roughness retrieval, etc. In addition, in conjunction with interferometric sensitivity, PolSAR data have been demonstrated to be useful for the observation and characterization of volumetric targets such as vegetation [9], forest [19], or even urban areas. In general, a PolSAR system measures the scattering matrix

$$\mathbf{S} = \begin{bmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{bmatrix} \quad (1)$$

for every resolution cell. The subindices h and v represent the horizontal and vertical orthogonal linear polarization states and S_{pq} for $p, q = h, v$ represents a complex SAR image where the polarization states employed in the reception and transmission processes are indicated by the first and second subindexes, respectively. The most prominent characteristic of PolSAR data is the capability to synthesize \mathbf{S} to any pair of orthogonal polarization states, giving its measurement in a single pair, which allows a detailed characterization of the target under observation [37], [58].

Since the dimensions of the resolution cell are normally larger than the wavelength of the electromagnetic wave, the scattered wave results from the coherent combination of many waves. When the scattered wave is due to a single target or a dominant target, \mathbf{S} may be employed to completely characterize the target under observation. Nevertheless, in the case of distributed targets, the scattered wave is due to the coherent addition of waves originated by many elementary targets within the resolution cell. This coherent addition process is referred to as speckle. Due to the complexity of the scattering process, the value of \mathbf{S} cannot be predicted in advance, as it depends on the type and internal arrangement of the individual targets in the resolution cell. Therefore, even though speckle represents a true electromagnetic measurement, its complexity makes it necessary to consider it as a random process. Considering \mathbf{k} as a vectorization of the \mathbf{S} matrix [17]

$$\mathbf{k} = [S_{hh}, \sqrt{2}S_{hv}, S_{vv}]^T \quad (2)$$

and on the basis of the central limit theorem (CLT), \mathbf{k} is characterized by 3-D, zero-mean, complex Gaussian probability density function (pdf)

$$p_{\mathbf{k}}(\mathbf{k}) = \frac{1}{\pi^3 |\mathbf{C}|} \exp(-\mathbf{k}^H \mathbf{C}^{-1} \mathbf{k}). \quad (3)$$

Equivalently to (2), other vectorizations of \mathbf{S} are possible [17]. As may be seen from (3), the distribution of \mathbf{k} is completely described by the Hermitian positive-definite covariance matrix

$$\begin{aligned} \mathbf{C} &= E\{\mathbf{k}\mathbf{k}^H\} \\ &= \begin{bmatrix} E\{S_{hh}S_{hh}^H\} & \sqrt{2}E\{S_{hh}S_{hv}^H\} & E\{S_{hh}S_{vv}^H\} \\ \sqrt{2}E\{S_{hv}S_{hh}^H\} & 2E\{S_{hv}S_{hv}^H\} & \sqrt{2}E\{S_{hv}S_{vv}^H\} \\ E\{S_{vv}S_{hh}^H\} & \sqrt{2}E\{S_{vv}S_{hv}^H\} & E\{S_{vv}S_{vv}^H\} \end{bmatrix} \end{aligned} \quad (4)$$

where $E\{x\}$ indicates the statistical expectation of the stochastic process x . The previous stochastic model is valid for describing stationary multidimensional SAR data, and indeed it represents a general purpose model. Nevertheless, (3) is not able to account for data texture or for situations where the CLT does not apply, for instance, high-frequency data [55]. From a general point of view, the presence of data texture may be modeled as a multiplicative factor [57]

$$\mathbf{k} = \sqrt{\tau} \mathbf{z} \quad (5)$$

where \mathbf{z} represents a Gaussian vector characterized by (3) and τ is a random variable that accounts for data texture. One of the most extended stochastic models to characterize \mathbf{k} is to assume τ to be described by a Gamma pdf, leading to the \mathcal{K} -distribution [57]. This model was extended in [30] to the family of distributions \mathcal{G}^0 with the objective to model highly heterogeneous data. In [55], under the hypothesis of the previous multiplicative model, a generalization of the \mathcal{K} -distribution based on the spherically invariant random vectors (SIRVs) was presented. Another generalization was presented by [28], and referred to as scale mixture of Gaussians (SMoG). As may be seen, all these generalizations assume the same texture in all the channels of the SAR data, therefore they are unable to model those situations in which the channels may present a different texture. Additionally, all these models describe the pdf of a sample. Hence, they are unable to characterize data texture considering spatial correlation. In general, all these generalized data models require additional moments, beyond the covariance matrix, to fully characterize the data distribution.

When considering the Gaussian model defined in (3), the covariance matrix represents the most important radar observable. Therefore, this matrix needs to be estimated from the data, or equivalently, the speckle noise must be filtered to have access to \mathbf{C} . The maximum-likelihood (ML) estimator of \mathbf{C} , also referred to as multilook filter, under the assumption of statistical ergodicity and homogeneity, is obtained by substituting the statistical expectation by a spatial averaging

$$\mathbf{Z} = \langle \mathbf{k}\mathbf{k}^H \rangle_n = \frac{1}{n} \sum_{i=1}^n \mathbf{k}_i \mathbf{k}_i^H \quad (6)$$

where n indicates the number of looks or samples employed to estimate \mathbf{C} and \mathbf{k}_i corresponds to the target vector of the i th sample. The estimated covariance matrix \mathbf{Z} is known as the sample covariance matrix, which is statistically determined by the complex Wishart pdf [31], [39], [50]

$$p_{\mathbf{Z}}(\mathbf{Z}) = \frac{n^{3n} |\mathbf{Z}|^{n-3}}{|\mathbf{C}|^n \tilde{\Gamma}_3(n)} \text{etr}(-n\mathbf{C}^{-1}\mathbf{Z}) \quad (7)$$

where $|\mathbf{X}|$ denotes the determinant of the matrix \mathbf{X} , $\text{etr}(\mathbf{X})$ is the exponential of the matrix trace, and

$$\tilde{\Gamma}_3(n) = \pi^3 \prod_{i=1}^3 \Gamma(n-i+1). \quad (8)$$

Finally, it is important to indicate that the pdf given by (7) is only valid for $n \geq 3$, that is, if the matrix \mathbf{Z} is not singular or $|\mathbf{Z}| \neq 0$.

A. Model Selection

As indicated in Section II, the use of a BPT structure requires the selection of a *region model* able to describe, in a precise way, the information of the image region represented by a particular node of the tree. In this paper, we will consider that the BPT represents stationary regions of the data according to (3). Consequently, the pixels of a particular region of the SAR image are fully characterized by the covariance matrix \mathbf{C} . Therefore, for every tree node, the sample covariance matrix \mathbf{Z} (6) will be estimated from the pixels of the region represented by the tree node [6], [7]. It is important to indicate that this matrix contains two types of information. On the one hand, there are the real diagonal elements of the matrix that represent the back-scattered power for a given combination of transmitted and received polarization states. On the other hand, there

are the complex off-diagonal elements that contain the correlation structure of the complex SAR images in \mathbf{k} .

When considering \mathbf{Z} as the region model, its application in the case of the leaf nodes, i.e., single pixels, must be carefully addressed. For an individual pixel, the covariance matrix is defined as the Hermitian product $\mathbf{k}\mathbf{k}^H$, resulting in a singular covariance matrix. Consequently, prior to the BPT construction, it is necessary to consider a regularization process to assure that for individual pixels $|\mathbf{Z}| \neq 0$. The way to regularize the data is to apply an initial filtering of a minimum of three independent pixels in order to assure the validity of (7). The most simple filtering that may be considered is a multilook filter [see (6)], within a 3×3 pixels window. In case of distributed targets, this filtering represents the ML estimator of the signal and no spatial resolution issues appear. However, in the case of point targets, this initial filtering leads to a reduction of the spatial resolution. The way to minimize this resolution loss is to apply a more elaborated initial filtering capable to adapt to the data morphology. The multilook filter exploits only the spatial locality of the data, assuming that nearby pixels will have a similar distribution. Nevertheless, the polarimetric locality may also be taken into account. In this sense, pixels with the same statistical distribution will be close in both the spatial and polarimetric domains. Bilateral filtering [49] applies this concept to compute a weighted average from the pixel vicinity having weights increasing with closeness in both domains. The filtered covariance matrix \mathbf{Z}^{ij} at position (i, j) can be expressed as

$$\mathbf{Z}^{ij} = k^{-1}(i, j) \sum_m \sum_n \tilde{\mathbf{Z}}^{mn} w_s(i, j, m, n) w_p(\tilde{\mathbf{Z}}^{mn}, \tilde{\mathbf{Z}}^{ij}) \quad (9)$$

where $m, n \in V(i, j)$ and $V(i, j)$ represents the local window around the pixel located at (i, j) , $\tilde{\mathbf{Z}}^{ij}$ represents the input image covariance matrix of that pixel, w_s is a weighting function depending on the spatial domain, w_p is a weighting function depending on the polarimetric domain, and k is the normalization factor to preserve the image power information

$$k(i, j) = \sum_m \sum_n w_s(i, j, m, n) w_p(\tilde{\mathbf{Z}}^{mn}, \tilde{\mathbf{Z}}^{ij}) \quad (10)$$

$$w_s(i, j, m, n) = \frac{1}{1 + \frac{d_s(m - i, n - j)}{\sigma_s^2}} \quad (11)$$

$$w_p(\tilde{\mathbf{Z}}^{mn}, \tilde{\mathbf{Z}}^{ij}) = \frac{1}{1 + \frac{d_p(\tilde{\mathbf{Z}}^{mn}, \tilde{\mathbf{Z}}^{ij})}{\sigma_p^2}} \quad (12)$$

where σ_s and σ_p represent the weights sensitivity in the spatial and polarimetric domains, respectively, and d_s is

the squared Euclidean distance in the spatial domain ($d_s(a, b) = a^2 + b^2$). In order to be coherent with the BPT construction which also relies on the estimation of covariance matrix similarity, d_p can be any of the similarity measures that will be defined in Section III-B.

In this paper, we consider PolSAR data as an example of multidimensional SAR data, whose characterization is performed via the covariance matrix. This data model may be extended to any other type of multidimensional SAR data, as for instance InSAR or PolInSAR. The main difference between the different multidimensional configurations lies in the physical interpretation of the information contained in the covariance matrix. Finally, as indicated, this model may present some limitations in order to represent the texture or high-frequency data. In these cases, the BPT nodes may be populated with the different parameters necessary to completely describe the data, showing the versatility of the BPT to account for any model able to describe the multidimensional SAR data.

B. BPT Construction for SAR Data Processing

Once the region model has been defined, it is necessary to establish a merging criterion that specifies the merging order of neighboring regions. This criterion must indicate which is the best merging among all the possible merging of neighboring regions that can be defined at every step of the BPT construction process. This criterion is referred to as *region similarity*.

In the case of PolSAR data, the similarity criterion must be defined in the space of the Hermitian positive-definite matrices, considering the information of the diagonal as well as off-diagonal elements. As mentioned before, the similarity criterion measures the similarity between two covariance matrices, where each one represents a particular BPT node.

In this paper, three different similarity measures are proposed and analyzed for the BPT construction process. These measures highlight the fact that different properties of the underlying region model may be exploited. Additionally, note that the similarity measures have to take into account the fact that regions may have arbitrary sizes.

The similarity measures are defined between two adjacent regions R_i and R_j , having estimated covariance matrices \mathbf{Z}_i and \mathbf{Z}_j and sizes of n_i and n_j pixels, respectively.

- Revised Wishart similarity (RWS): The revised Wishart similarity measure was defined in [36], and it is based on a statistical test assuming that the two regions under consideration follow a Wishart pdf and that one pdf is known. Thus, it is not symmetric as it depends on which region pdf is assumed to be known. In order to generate a similarity measure, a modified symmetric version is

proposed using $S_s(R_i, R_j) = S(R_i, R_j) + S(R_j, R_i)$ and multiplying by the region size term [6], [7]

$$S_{RWS}(R_i, R_j) = \left(\text{tr}(\mathbf{Z}_i^{-1} \mathbf{Z}_j) + \text{tr}(\mathbf{Z}_j^{-1} \mathbf{Z}_i) \right) (n_i + n_j) \quad (13)$$

where $\text{tr}(\mathbf{X})$ denotes the trace of the \mathbf{X} matrix and \mathbf{X}^{-1} its inverse. As one may observe, this similarity measure considers the stochastic nature of the PolSAR data.

- *Diagonal revised Wishart similarity (DWS)*: This similarity measure is based on the S_{RWS} previously defined in (13), but only considering the diagonal elements and setting all off-diagonal elements to 0. This simplifies the matrix inversion as being the inverse of the diagonal elements. After some mathematical manipulation, it may be expressed as [6], [7]

$$S_{DWS}(R_i, R_j) = \left(\sum_{k=1}^3 \left(\frac{Z_{i_{kk}}^2 + Z_{j_{kk}}^2}{Z_{i_{kk}} Z_{j_{kk}}} \right) \right) (n_i + n_j). \quad (14)$$

This measure is useful to determine the benefit of employing all the elements of \mathbf{Z} , in front of employing only the diagonal ones. Note that some of the PolSAR speckle filters defined in the literature only consider the diagonal information of \mathbf{Z} . Additionally, since this measure does not require a full-rank matrix, it can be directly applied over the original data. Consequently, this measure will be applied also to perform the initial bilateral filtering into (12).

- *Geodesic similarity (GS)*: The geodesic similarity is based on a distance adapted to the cone of positive-definite Hermitian matrices [11], that is, this measure exploits the geometry of the subset of the vector space defined by the covariance matrices that is closed under multiplication by positive scalars [8]

$$S_{GS}(R_i, R_j) = \left\| \log \left(\mathbf{Z}_i^{-1/2} \mathbf{Z}_j \mathbf{Z}_i^{-1/2} \right) \right\|_F + \ln \left(\frac{2n_i n_j}{n_i + n_j} \right) \quad (15)$$

where $\|\cdot\|_F$ represents the Frobenius matrix norm; $\log(\cdot)$ represents the matrix logarithm; and $\ln(\cdot)$ represents the natural logarithm [3].

C. BPT Pruning for SAR Data Applications

The BPT is a multiscale data representation in a set of hierarchical regions. Once the BPT has been constructed, one application may be interested in a particular set of

regions of the BPT. In this context, the BPT *pruning* process can be considered as the selection of this particular set of regions. This process is usually defined by the final application. In what follows, two applications will be considered: PolSAR speckle noise filtering and PolSAR data segmentation.

1) *Speckle Noise Filtering*: Speckle represents one of the main issues for the correct interpretation of SAR data. Hence, speckle noise filtering is crucial for the correct analysis of multidimensional SAR data. In the following, a BPT pruning strategy for the multidimensional speckle noise filtering application will be defined and evaluated.

The pruning criterion considered consists of selecting from the BPT a set of regions corresponding to the largest homogeneous regions. The region homogeneity is measured by ϕ_R , which may be interpreted as the average relative error when representing each region R_i by its model \mathbf{Z}_i

$$\begin{aligned} \phi_R(R_i) &= \frac{1}{n_i} \sum_{k=1}^{n_i} \frac{\|\mathbf{Z}_i^k - \mathbf{Z}_i\|_F^2}{\|\mathbf{Z}_i\|_F^2} \\ &= \frac{1}{n_i \|\mathbf{Z}_i\|_F^2} \sum_{k=1}^{n_i} \|\mathbf{Z}_i^k - \mathbf{Z}_i\|_F^2 \end{aligned} \quad (16)$$

where \mathbf{Z}_i^k represents the covariance matrix of the k th pixel of the region R_i .

The pruning process can be implemented by the *min rule*, which is a top-down approach that prunes first nodes R_i that fulfill the homogeneity criterion $\phi_R(R_i) < \delta$. Starting from the root node, the homogeneity of every node is tested according to (16). If the node is not homogeneous, having $\phi_R(R_i) \geq \delta$, the corresponding child nodes will be studied in order to detect the first homogeneous nodes. Finally, the leaves of the pruned tree are homogeneous regions according to ϕ_R that form a partition of the image. Note that this pruning criterion is independent of the region size. Hence, the resulting filtered image may be composed of large regions corresponding to distributed targets with an accurate estimation of the polarimetric information, but also very small regions due to point targets where the spatial resolution is preserved.

In order to analyze the performance in terms of speckle filtering, the previously defined pruning strategy is analyzed in the following. One important aspect about speckle is that SAR data cannot be obtained without this component. Therefore, the evaluation of the performance of any type of application, based on real SAR data, relies on external information or ground truth. In many situations, the objective is the quantitative estimation of a certain biophysical or geophysical parameter that may be obtained through alternative techniques. Hence, a measure of error can be easily defined. On the contrary, when the objective is to obtain the response of the scene without noise, since

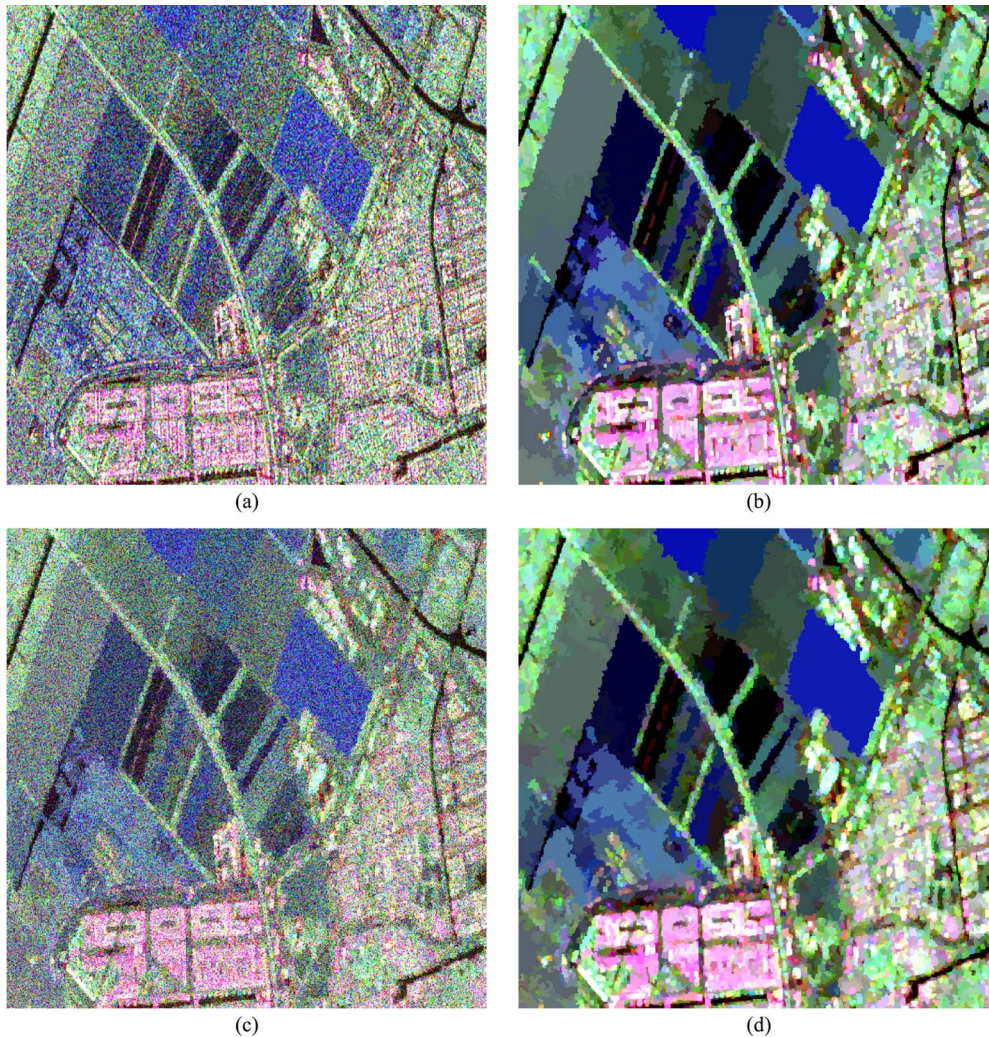


Fig. 4. Simulated data generation from a real image segmentation. (a) Original image. (b) Ground truth for simulation. (c) One generated realization. (d) Filtered realization.

this information is almost impossible to obtain externally, a measure of error is complex to implement. Therefore, a quantitative evaluation of a PolSAR speckle filter must be performed in terms of simulated PolSAR data.

To address this evaluation, simulated PolSAR data are generated from a filtered real PolSAR image. This reference image could be obtained from real data by manually segmenting it into homogeneous regions and estimating the ground truth employing the pixels within these regions. Nevertheless, the limitations of this approach, as for instance a limited number of regions or straight divisions among regions, would limit the generality of the evaluation. In this study, a 512×512 pixel portion of the RADARSAT-2 data set will be employed, shown in Fig. 4(a). The simulated ground truth, shown in Fig. 4(b), has been generated using a BPT pruning with the S_{GS} similarity measure, 3×3 multilook initial filtering, and $\delta = -3$ dB as a pruning factor. This ground truth represents 5851 different regions. It has been used to gener-

ate simulated PolSAR images, with realistic polarimetric and spatial information. Several different realizations of the speckle noise process are generated, for instance Fig. 4(c), according to the synthetic ground truth, and processed with the proposed BPT-based filtering, as shown in Fig. 4(d).

The filtering results are compared with the ground truth to assess numerically the goodness of the filtering technique. As an error measure, $E_R(X, Y)$ is defined between two images X and Y

$$E_R(X, Y) = \frac{1}{n_h n_w} \sum_{i=1}^{n_h} \sum_{j=1}^{n_w} \frac{\|Z_X^{ij} - Z_Y^{ij}\|_F}{\|Z_Y^{ij}\|_F} \quad (17)$$

where n_h and n_w are the image height and width in pixels, respectively, and Z_X^{ij} represents the (i, j) th pixel value of

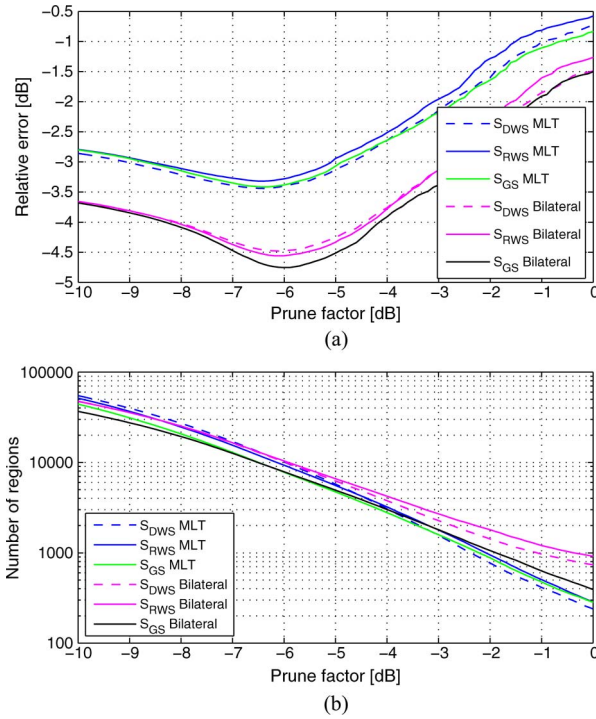


Fig. 5. (a) Relative error and (b) number of regions as a function of the pruning threshold for simulated PolSAR data.

image X . Note that the relative error measure defined in (17) is based on the inverse signal-to-noise ratio (SNR^{-1}) averaged for all the pixels in the image.

Fig. 5(a) details the evolution of $E_R(X, Y)$ with respect to the pruning factor δ , for the different similarity criteria employed in the BPT construction process and the two approaches considered for the initial regularization process: a 3×3 multilook filtering and three-iterations 3×3 bilateral filtering with $\sigma_s = 1.5$, $\sigma_p = 1$. With respect to the similarity measure employed for BPT construction, the behavior of the error as a function of δ is very similar for all of them. In particular, there is always a minimum in terms of $E_R(X, Y)$ located almost at the same value of δ . This closeness indicates that there is an optimum threshold almost independent from the similarity measure. In addition, Fig. 5(b) shows the number of leaves of the pruned BPT as a function of δ . In general, it may be observed that the use of the bilateral filtering approach leads to the lowest error for a similar number of regions of the filtered data. Even for very low values of δ , when the filtering process is mainly selecting nodes close to the BPT leaves, the use of this bilateral approach gives a lower error than the multilook approach, indicating a better performance for the preservation of spatial resolution and details. Looking at the results obtained by different similarity measures, some differences can be observed depending on the initial filtering step. Whereas in the case of the multilook, the lowest error is achieved by S_{DWS} , in the case of the

bilateral approach, the minimum is achieved by S_{GS} . Nevertheless, according to Fig. 5(b), a smaller number of regions for a fixed value of pruning factor are obtained by S_{GS} . That is, for a given level of region homogeneity, the geodesic similarity results in the largest homogeneous regions. For instance, when considering the multilook and S_{GS} , the minimum appears at $\delta = -6.4$ dB, leading to 9498 regions, whereas in the case of the bilateral approach, the minimum is at $\delta = -6$ dB, leading to 7848 regions.

In order to assess the proposed BPT-based speckle filtering technique with real data, it has been employed to process a RADARSAT-2 Fine Quad-Pol 5300×3100 pixel image corresponding to a test site in Flevoland, The Netherlands. The data set was acquired during the ESA AgriSAR 2009 campaign, devoted to analyze the agricultural fields temporal evolution with PolSAR data. The scene, covering an area of 25×25 km² with a spatial resolution of 5.2 m in range and 7.6 m in azimuth, is mainly composed of agricultural fields, sea, and urban areas. The PolSAR image, presented in Fig. 6(a), corresponds to an acquisition with a mean incidence angle of 33° , in ascending pass, acquired in April 4, 2009.

Fig. 6(b) presents the filtered image with the proposed method, employing the S_{GS} similarity measure for BPT construction and a pruning factor $\delta = -3$ dB. After a visual evaluation, the same contours can be seen over the image, while a large amount of speckle noise has been filtered out. The filtering process results in a contrast enhancement, but it must be highlighted that the colors are not altered indicating that polarimetric information is preserved. In order to observe the results in detail, the 512×512 pixels portion seen in Fig. 7(a) has been selected from the whole image. For comparison purposes, data have been filtered out with a 7×7 multilook [see Fig. 7(b)], the refined Lee filter [Fig. 7(c)] [1], and the IDAN filter [Fig. 7(d)] [2]. It is worth indicating that these filtering approaches are based on a pixel-based processing, whereas the BPT approach considers a region-based processing. The data set presented in Fig. 7(a) has been processed with the BPT approach considering different alternatives for the initial filtering, the similarity measure, and the homogeneity threshold, as shown in Fig. 8. As may be seen, the BPT approach is able to preserve the spatial resolution and spatial details associated to point targets, present in urban areas for instance, while eliminating the speckle noise over large agricultural fields. This is possible because the presented pruning strategy selects, for the former case, those nodes close to the leaves ones, whereas in the latter case, intermediate nodes according to the region homogeneity level selected by δ . A closer analysis of the images shows that if the multilook filter is used to regularize the covariance matrix, the results present a small resolution loss over small details compared with those obtained with the bilateral approach. In addition, a better preservation of the contours is shown in the second

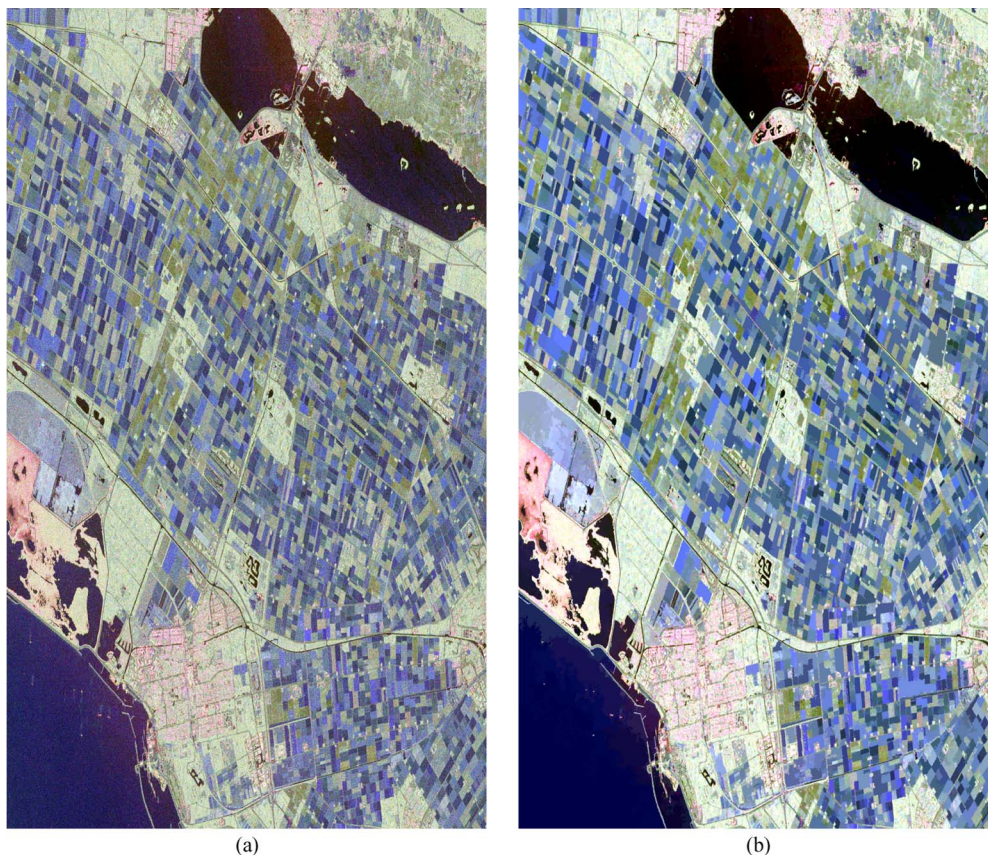


Fig. 6. (a) Original PolSAR image and (b) filtered image applying the BPT-based filtering employing the S_{GS} similarity measure for BPT construction and a pruning factor $\delta = -3$ dB. A 3×3 multilook has been employed as initial filtering.

case. These observations corroborate the results presented in Fig. 5(a) and (b), where it is shown that results employing a bilateral approach as initial filtering obtain lower error values. The closeness of the different plots in the previous figures also appears in Fig. 8 as all the images are very similar. Even in the case of selecting a similarity measure that considers only the diagonal elements of the covariance matrix S_{DWS} the filtering can be considered as correct. In our opinion, the performances of the BPT approach for speckle filtering must not be attributed exclusively to the use of the similarity concept, but also to its systematic use by the BPT, leading to a hierarchical representation of the information contained in the data. This hierarchical representation may be observed in Fig. 9, where filtering results for different values of δ are presented. As may be observed, as the value of this threshold increases, larger regions appear as the homogeneity criterion is relaxed. Concerning Fig. 9, it is important to highlight that despite increasing the amount of filtering, no new contours or artifacts appear in the filtered data. In particular, increasing δ reduces the number of contours by removing some of them while the others are preserved. In fact, the speckle filter presented here employing the BPT structure belongs to a larger class of filters having this

contour preservation property. They are called *connected operators* [44], [46].

Comparing the results obtained with simulated and real data, an important point arises. Similar results are obtained by means of the BPT-based filtering in Figs. 8 and 4 in real and simulated images. However, the pruning threshold in both cases differs significantly: -4 and -6 dB, respectively. This difference may be related to additional region features in real data not considered by the Gaussian model, which can be considered as the region texture. The homogeneity threshold has to be increased to absorb these modeling errors with real data. For the simulated data, since this texture is not reproduced, it is not necessary to increase δ and the same results are obtained with lower homogeneity thresholds.

Up to now, we have only considered an evaluation of the filtering results in morphological terms. One of the most important aspects when considering PolSAR speckle filtering is the correct estimation of the polarimetric information, i.e., the covariance matrix, without introducing any type of bias on the estimated data. In order to perform this analysis, the entropy H and the mean alpha angle $\bar{\alpha}$ are examined, obtained through the eigenanalysis of the estimated covariance matrix [17]. Entropy gives information

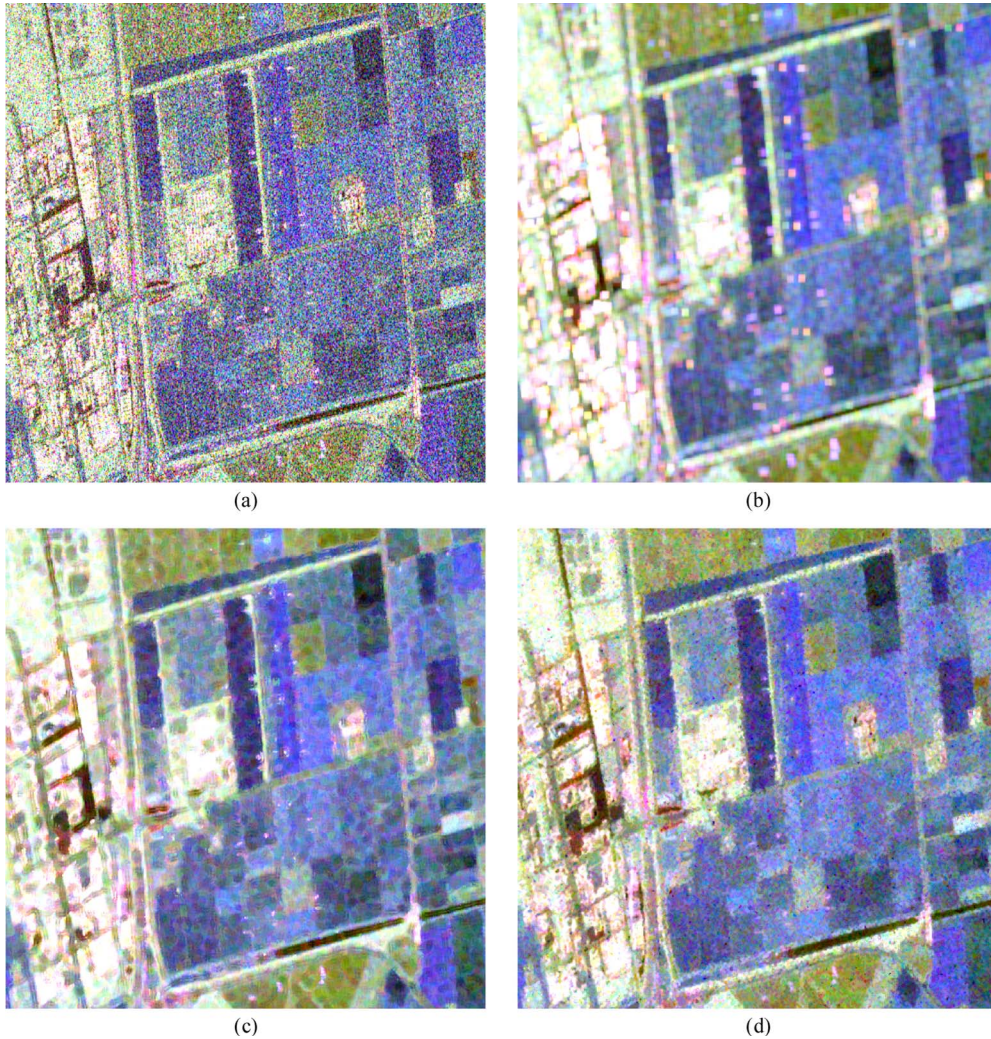


Fig. 7. (a) Original detail of the PolSAR image and the same detail after applying (b) the 7×7 multilook, (c) the refined Lee filter, and (d) the IDAN filter.

concerning the degree of randomness of the scattering process (the entropy $H = 0$ for deterministic and $H = 1$ for completely random scattering mechanisms), whereas $\bar{\alpha}$ contains information about the type of scattering: $\bar{\alpha} = 0$ for surface scattering, $\bar{\alpha} = \pi/4$ for volume scattering, and $\bar{\alpha} = \pi/2$ for dihedral scattering. Fig. 10(a) and (b) shows the estimated values of H and $\bar{\alpha}$, respectively, considering a 7×7 multilook filter. The rest of the plots in this figure show the estimated values for the BPT-based proposed filtering employing S_{GS} and considering a multilook approach and a bilateral approach as initial filters. In general, it may be observed that the estimated values are basically the same as in the case of the 7×7 multilook filter, which corresponds to the MLE of the data. Nevertheless, it may be observed that when considering the BPT approach these parameters are estimated with a higher spatial accuracy, which is specially noticeable when considering the bilateral approach as initial filtering.

2) *SAR Data Segmentation*: The speckle filtering application detailed in Section III-C1 was focused primarily on homogeneous regions. They usually correspond to BPT nodes relatively close to the leaves. Nevertheless, the hierarchical structure of the BPT may contain a lot more information about the structure and morphology of the image. In what follows, the information contained in the nodes close to the root node, which may contain information concerning large structures of the image, will be considered. In order to perform this selection, a second tree pruning based on selecting a fixed number of regions n_r corresponding to the most different regions is proposed. If the difference between regions is evaluated using the same similarity measure as for the construction of the BPT, the pruning process is equivalent to stop the tree construction process when n_r regions are obtained. Fig. 11 shows an example where this pruning process has been employed to generate a segmentation result having 20 regions. The S_{GS}

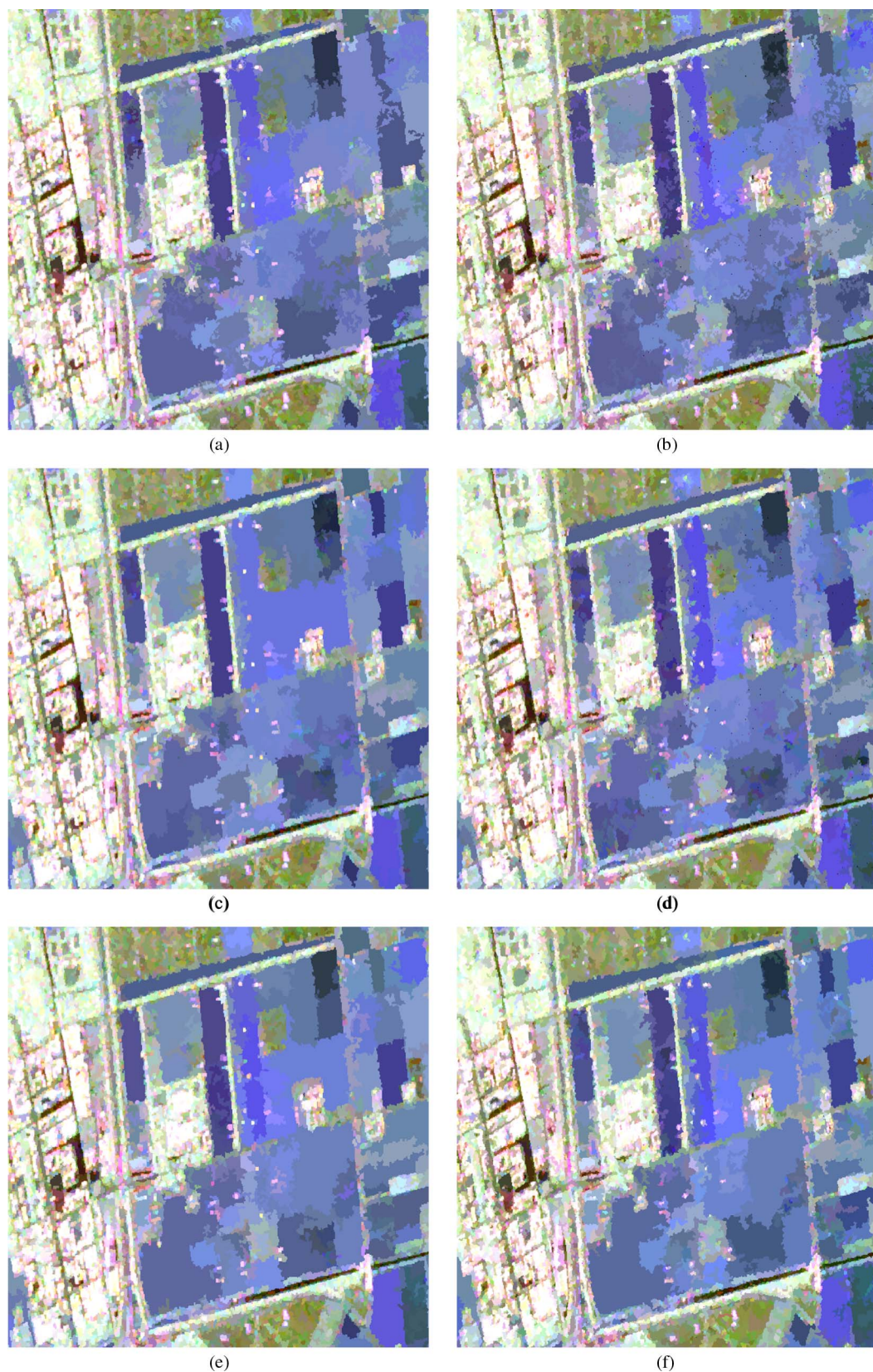


Fig. 8. Details of the PolSAR image processed with different similarity measures and initial filtering. (a) S_{DWS} , $\delta = -4$ dB, multilook; (b) S_{DWS} , $\delta = -4$ dB, bilateral; (c) S_{RWS} , $\delta = -4$ dB, multilook; (d) S_{RWS} , $\delta = -4$ dB, bilateral; (e) S_{GS} , $\delta = -4$ dB, multilook; (f) S_{GS} , $\delta = -4$ dB, bilateral.

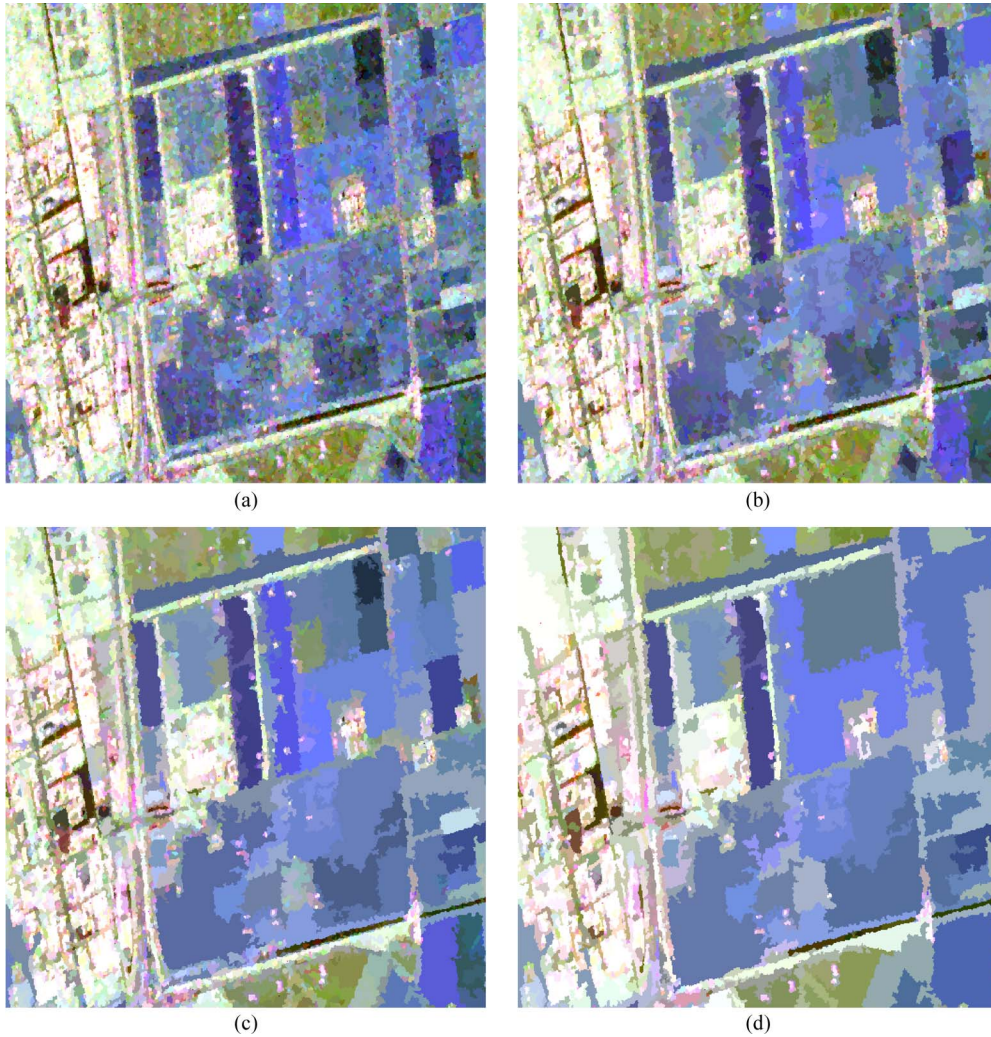


Fig. 9. Details of the PolSAR image processed with different pruning factors. (a) S_{GS} , $\delta = -7$ dB, bilateral; (b) S_{GS} , $\delta = -5$ dB, bilateral; (c) S_{GS} , $\delta = -3$ dB, bilateral; (d) S_{GS} , $\delta = -1$ dB, bilateral.

similarity measure has been employed for the BPT construction process and a 3×3 multilook has been used as initial filtering. In this case, it may be observed that the resulting regions contain information regarding the sea, the lagoons, several regions grouping the agricultural fields, and the urban areas. It is worth noting that the information of these regions may be employed to segment the different parts of the image with a very high spatial resolution. For instance, the detail image presented in Fig. 12 shows that the different regions could be used to detect the coastline, preserving the small details and the structures on it.

IV. HYPERSPECTRAL IMAGES

Hyperspectral passive remote sensing is based on the illumination of Earth objects by the sun. Solar photons travel through the atmosphere, reflect on the surface and reach the sensor after a second path through the atmos-

phere. The physical interactions that happen along this path modulate the signal sensed by passive scanners providing clues on the composition of materials at the surface and the atmospheric components. Several decades ago, technological improvements in the field of spectroscopy and traditional imaging gave rise to the birth of a passive technique named as hyperspectral imaging coupling spectroscopy and imaging to determine the spatial distribution of the materials and of their properties. Hyperspectral imaging expands traditional imagery to the range of the electromagnetic spectrum that is invisible to the human eye. In this way, hyperspectral sensors have the ability to discriminate very precisely between different materials that may look similar in the visible range.

A hyperspectral image $I_\lambda(p)$ is typically a compilation of a few hundred spectral images, each one expressing the sensed radiance coming from a given scene at a given wavelength. The stacking of spectral images produces

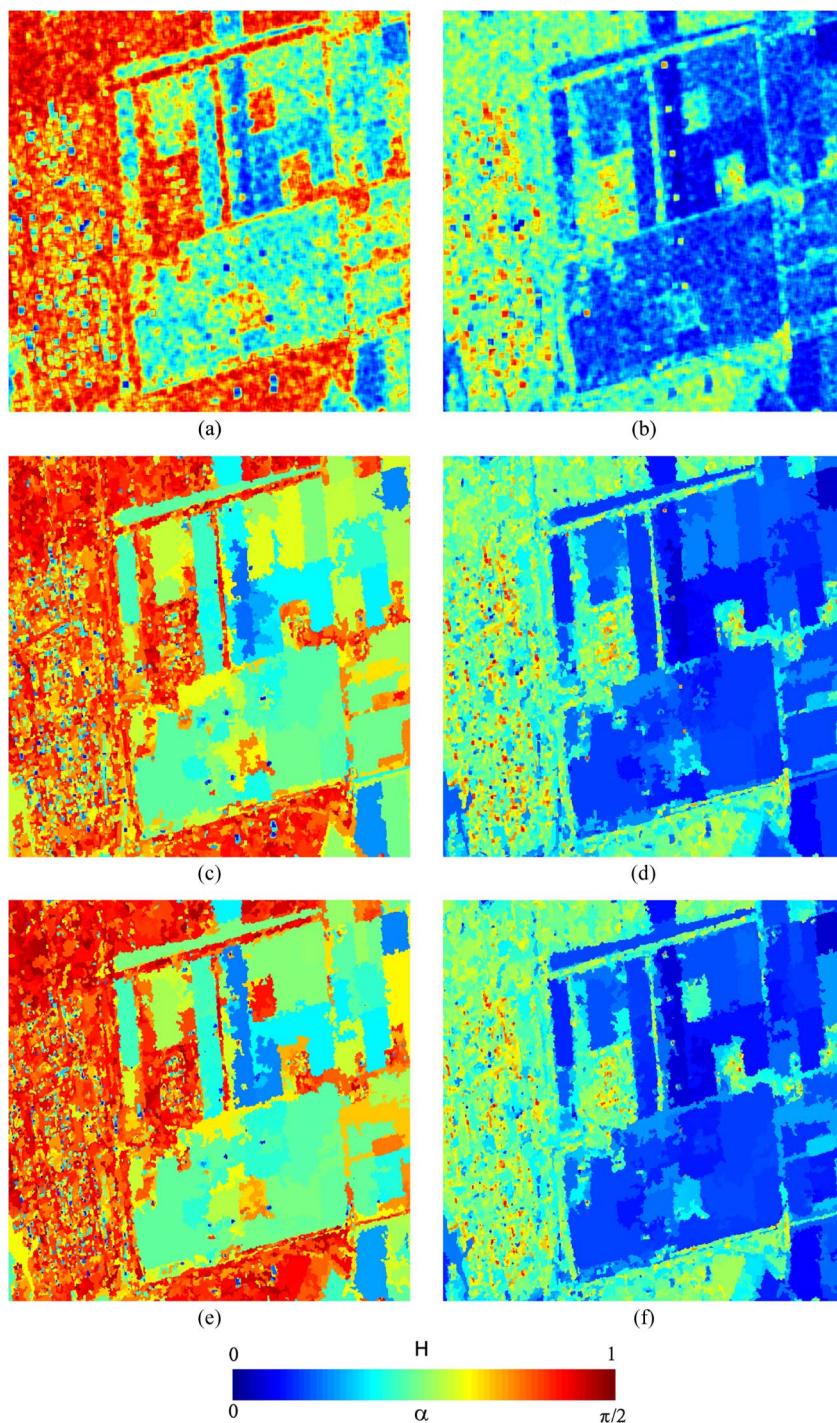


Fig. 10. H/α decomposition for different filtering strategies. (a) H , 7×7 multilook; (b) α , 7×7 multilook; (c) H , S_{GS} , $\delta = -3$ dB, multilook; (d) α , S_{GS} , $\delta = -3$ dB, multilook; (e) H , S_{GS} , $\delta = -3$ dB, bilateral; (f) α , S_{GS} , $\delta = -3$ dB, bilateral.

hyperspectral images, generally seen as a cube, in which two axes correspond to the two spatial dimensions and the third axis to the spectral dimension according to the wavelength. Generally, hyperspectral images cover the visible ($400 < \lambda < 700$ nm), the near-infrared range ($0.7 < \lambda < 3.5$ μm), and the shorter wavelengths of the thermal

infrared ($3.5 < \lambda < 5$ μm). Each image composed of a set of pixels corresponding to a given wavelength is called a spectral band. The term spectrum is used for the spectral data contained in a given spatial position. A spectrum expresses the variation according to the wavelength of the measured radiance.

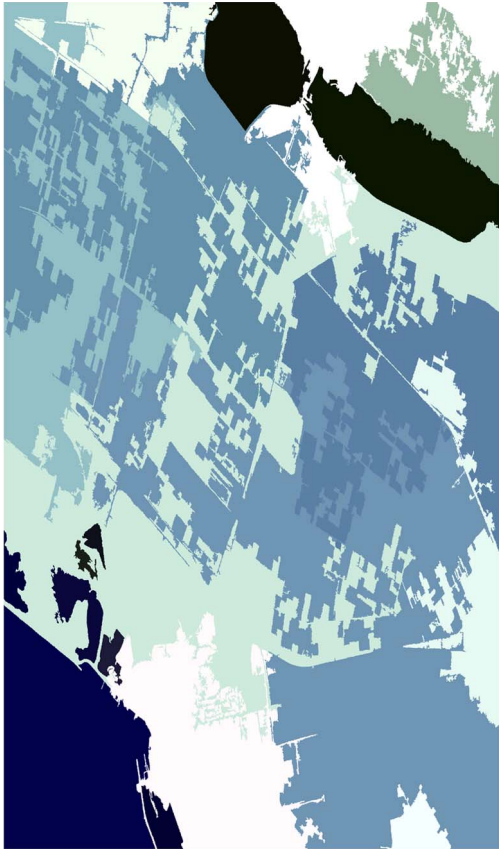


Fig. 11. Full PolSAR image segmentation results in 20 regions.

The price of the wealth of information provided by hyperspectral images is a huge amount of data that cannot be fully exploited using traditional image analysis tools. On the one hand, the spectral space is important because it

contains much more information about the surface of target objects than what can be perceived by human vision. One of the main difficulties to exploit this space is the important redundancy or equivalently the high correlation between consecutive values of the spectra explained by the small difference between two consecutive wavelengths. Another significant drawback is the large spectral variability introduced by several factors such as the noise resulting from atmospheric conditions, the sensor influence, indirect reflection, or the illumination effects. On the other hand, the spatial space is also important because it describes the spatial variations and correlation in the image and this information is essential to interpret objects in natural scenes. Accordingly, hyperspectral analysis tools should take into account both the spatial and spectral spaces in order to be robust and efficient. However, the number of wavelengths per pixel and the number of pixels per image, as well as the complexity of jointly handling spatial and spectral correlation, explain why this approach is still a largely open research issue for effective and efficient hyperspectral data processing.

A. Model Selection

In the case of hyperspectral images, a spectrum $I_\lambda(p) = \{I_{\lambda_1}(p), \dots, I_{\lambda_{N_\lambda}}(p)\}$ of hundreds of bands is associated to each pixel p . As the sampling of the frequency axis is rather dense, most spectral bands are highly correlated. Moreover, in many situations, the relevant information is the shape of the spectrum more than its absolute value. This phenomenon is often due to illumination variations that reduce the material reflectance but preserve the shape of the spectral response. Furthermore, one of the difficulties in hyperspectral imaging is that often the difference between specific materials is only noticeable in a reduced number of bands. Finally, areas of homogeneous material

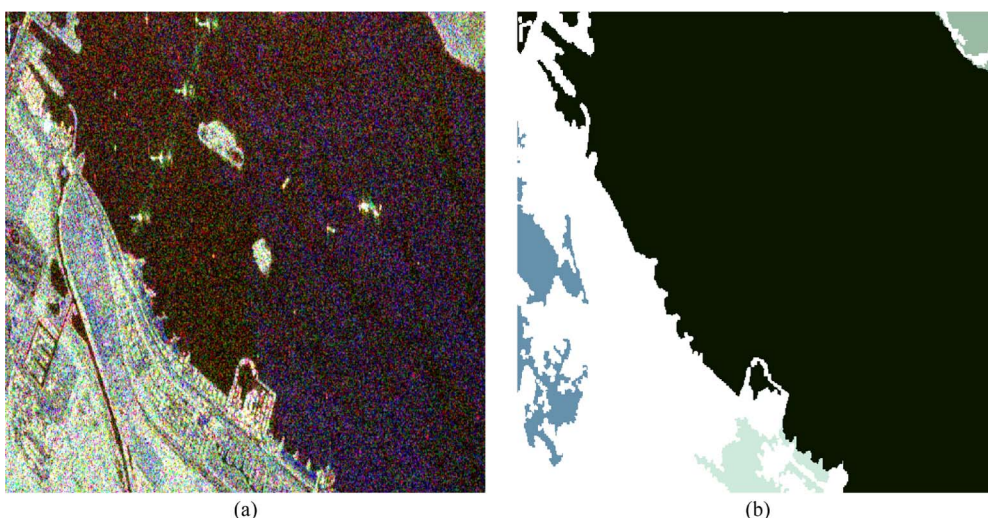


Fig. 12. Detail of the PolSAR image segmentation result where the coastline contours can be clearly identified. (a) Original. (b) Segmentation.

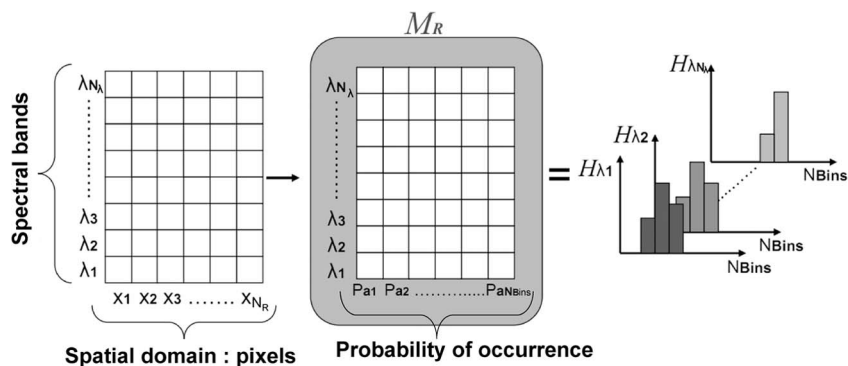


Fig. 13. Hyperspectral region model.

cannot be assumed to be homogenous due to texture and intraclass variability. As a result, a first-order model, which would correspond to the average spectrum, is not accurate enough to precisely describe hyperspectral regions and to discriminate between different materials [51], [52]. In order to get a more precise region description, nonparametric models of the pixel pdf such as histograms must be used.

The first alternative would be to use the full multidimensional histogram. Although, it would be able to capture the signal correlation, it is very difficult to estimate because of its sparseness and to handle because of the very high number of dimensions. As a compromise following the contribution of [13] for color images, a hyperspectral region model made of the set of spatial histograms for each spectral band was proposed in [51]

$$\mathbf{M}_R = \{H_R^{\lambda_1}, \dots, H_R^{\lambda_{N_\lambda}}\} \quad (18)$$

where each $H_R^{\lambda_i}$ represents the histogram of the reflectance value at wavelength λ_i of pixels belonging to region R . As histograms are used, no assumption is made about the nature of the pdfs. Fig. 13 illustrates this model. In Fig. 13, N_R and N_{Bins} , respectively, represent the number of pixels in the region R and the number of bins of each histogram.

At the beginning of the merging process, the pdf of the regions of the initial partition should be estimated and assigned to the tree leaves. As in Section III-A, the initial partition is composed of single pixel regions. One simple solution is to define the pdf of individual pixels as being equal to zero except for the observed value for each spectral band. This solution gives already very good results, however, the pdf estimation can be improved using self-similarity ideas [12]. More precisely, assume that a patch is a local neighborhood around the pixel whose pdf has to be estimated. If similar patches are found in the image, it can be assumed that the central pixel of these patches tells

something about other potential realizations of the pixel and these realization can be used to improve the pixel pdf estimation [27], [53]. Finally, when two regions are merged, a new region is created and its corresponding model is simply computed by adding the histograms of the two regions.

B. BPT Construction for Hyperspectral Image Processing

Once the region model has been selected, a region similarity measure should be defined. This issue has been studied in [52] and [53], where various region models and similarity measures for hyperspectral images are compared and analyzed. The main conclusions are that similarity measures relying on the first-order model (average spectrum) do not provide accurate region contours and precise BPT structures. Concerning the similarity using the region model defined by (18), many approaches measuring the difference between two distributions or histograms have been proposed in the literature. They can be classified in two different categories: *bin-by-bin* and *cross-bin* similarity measures. Bin-by-bin similarity measures only compare contents of corresponding histogram bins, that is, they compare the probabilities of occurrence of the same value in both histograms but do not study the probabilities of occurrence of different values. By contrast, the cross-bin measures contain terms that compare noncorresponding bins. One of the important conclusions of [52] and [53] is that cross-bin similarity measures provide better BPT structure and region contours. Here, two such similarity measures are presented: the diffusion distance S_{Diff} and the association measure via multidimensional scaling S_{MDS} .

1) *Similarity Measure Based on Diffusion Distance*: The main idea behind the diffusion distance [41] is to study the relationship between histograms with the help of a diffusion process. The diffusion process allows the evaluation of the similarity across various scales and is robust to histogram quantization and deformation. The diffusion

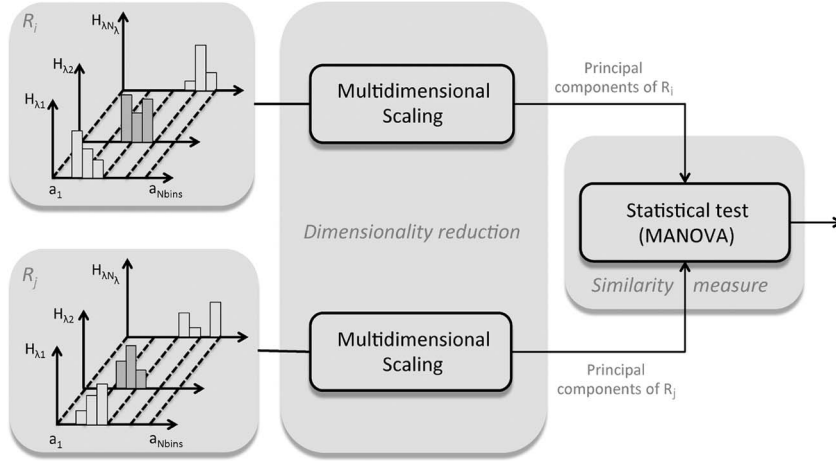


Fig. 14. Similarity measure based on multidimensional scaling.

distance is computed as the sum of dissimilarities over scales.

Let us denote by $H_{R_i}^{\lambda_k}(a_n)$ and $H_{R_j}^{\lambda_k}(a_n)$ the two histograms of regions R_i and R_j in the spectral band λ_k that have to be compared. The histogram bins correspond to the values a_n for $1 \leq n \leq N_{\text{Bins}}$. The diffusion process is implemented through the iterative convolution with a Gaussian filter $h_{\text{Gaus}}(a_n)$ and downsampling applied on the histogram difference $d_l(a_n)$

$$\begin{aligned} d_0(a_n) &= H_{R_i}^{\lambda_k}(a_n) - H_{R_j}^{\lambda_k}(a_n) \\ d_l(a_n) &= [d_{l-1}(a_n) * h_{\text{Gaus}}(a_n)] \downarrow_2, \quad l \in [1, \dots, L]. \end{aligned} \quad (19)$$

The notation \downarrow_2 denotes downsampling by a factor of two. l is the number of scales which corresponds to the first level where $\sum_n |d_l(a_n)|$ is lower than 0.01. The standard deviation for the Gaussian filter has been set equal to 0.5 according to [41]. From the diffusion process, the distance D_{Diff} between the histograms is defined as follows:

$$D_{\text{Diff}}(H_{R_i}^{\lambda_k}, H_{R_j}^{\lambda_k}) = \sum_{l=0}^L \sum_n |d_l(a_n)|. \quad (20)$$

Finally, the overall region similarity is derived as [52]

$$S_{\text{Diff}}(R_i, R_j) = \sum_{k=1}^{N_\lambda} D_{\text{Diff}}(H_{R_i}^{\lambda_k}, H_{R_j}^{\lambda_k}). \quad (21)$$

Note that the hyperspectral bands are processed separately. As a result, the correlation between bands is not taken into account by this similarity measure. In order to

solve this problem, a new similarity measure that exploits the distances between wavebands to remove redundant information contained in each region model was proposed in [53]. This new similarity measure is based on distances between observations and canonical correlations [23].

2) *Similarity Measure Based on Multidimensional Scaling:* The proposed similarity measure is based on distances between wavebands and canonical correlations [22]. Two different steps highlighted in Fig. 14 can be distinguished in the approach. The first step corresponds to a local dimensionality reduction. For each region, the frequency interband similarity relationship is analyzed in order to remove the redundant hyperspectral information. A multidimensional scaling technique is performed which summarizes the important region information in a reduced number of components called the *principal components*. The second step defines the similarity measure based on the principal components of the two regions R_i and R_j to be compared. This similarity measure is derived from a statistical test based on the multivariate analysis of variance (MANOVA). The idea is to test whether a dependence or correlation exists between the principal components of the regions. Let us describe more precisely each step [24], [53].

Step 1—Dimensionality Reduction: Multidimensional scaling (MDS) [21] represents a set of objects as a set of points in a map of chosen dimensionality, based on their interpoint distances. The goal is to maximize the agreement between the displayed interpoint distances and the given ones. Thus, MDS attempts to locate N_o objects as points in a Euclidean space E where the geometric differences between pairs of points in E agree, as closely as possible, with the true differences between the N_o objects. In our case, the N_o objects are the N_λ probability distributions of each region.

Hence, the probability distribution similarities of a region R can be represented by an $N_\lambda \times N_\lambda$ distance matrix $\Delta_R = (\delta_{kl})$, where $\delta_{kl} = \delta_{lk} \geq 0$ represents the difference between histograms of wavelength λ_k and λ_l . In our case, the diffusion distance defined by (20) to measure this difference is used, that is, $\delta_{kl} = D_{\text{Diff}}(H_R^{\lambda_k}, H_R^{\lambda_l})$.

Then, let \mathbf{A} denote the matrix with entries $\mathbf{A} = -(1/2)\delta_{kl}^2$ and \mathbf{H} the centering matrix $\mathbf{H} = \mathbf{I}_N - (1/N)\mathbf{1}\mathbf{1}^t$, where $\mathbf{1}$ denotes a multidimensional vector with unitary components. The so-called inner product matrix \mathbf{B}_R associated to Δ_R can be computed by $\mathbf{B}_R = \mathbf{H}\mathbf{A}\mathbf{H}$ for each region R [21]. The inner product matrix \mathbf{B}_R is an $N_\lambda \times N_\lambda$ symmetric matrix which can be spectrally decomposed as $\mathbf{B}_R = \mathbf{U}_R \Lambda_R^2 \mathbf{U}_R^t$. Assuming that the eigenvalues $e_{R,t}$ in Λ_R^2 are arranged in descending order, the matrices $\mathbf{U}_R \Lambda_R$ and \mathbf{U}_R contain, respectively, the principal and the standard coordinates of region R . As mentioned above, the MDS aims at representing the region with a reduced number of dimensions. This is achieved by taking the first D_s most representative principal or standard coordinates.

Now, given two regions R_i and R_j , the goal of the first step of Fig. 14 is to compute the first D_s standard coordinates of R_i and R_j . Therefore, two distance matrices Δ_{R_i} and Δ_{R_j} are computed to find $\mathbf{B}_{R_i} = \mathbf{U}_{R_i} \Lambda_{R_i}^2 \mathbf{U}_{R_i}^t$ and $\mathbf{B}_{R_j} = \mathbf{U}_{R_j} \Lambda_{R_j}^2 \mathbf{U}_{R_j}^t$. Finally, two sets of principal components are obtained.

The number D_s of dimensions is an important aspect in most multivariate analysis methods. In MDS, the number of dimensions is based on the percentage of variability accounted for by the first dimensions. Here, a criterion which extends a sequence c defined and studied in [23] is used to set the value of D_s . First, a maximum dimension number N_s suggested by the data is defined. It is the dimension for which $\sum_{t=1}^{N_s} e_{R,t}^2 / \sum_{t=1}^{N_\lambda} e_{R,t}^2 \approx 1$ if the eigenvalues are denoted by $e_{R,t}$.

Then, if $u_{R_i,t}$ for $t = 1, \dots, N_s$ and $u_{R_j,p}$ for $p = 1, \dots, N_s$ are, respectively, the first N_s columns of \mathbf{U}_{R_i} and \mathbf{U}_{R_j} , the sequence C_k is defined by

$$C_k = \frac{\sum_{t=1}^k \sum_{p=1}^k e_{R_i,t}^2 \left(u'_{R_i,t} u_{R_j,p} \right)^2 e_{R_j,t}^2}{\sum_{t=1}^{N_s} \sum_{p=1}^{N_s} e_{R_i,t}^2 \left(u'_{R_i,t} u_{R_j,p} \right)^2 e_{R_j,t}^2}, \quad k \in [1, \dots, N_s] \quad (22)$$

where $e_{R_i,t}^2$, $e_{R_j,t}^2$ are the eigenvalues of \mathbf{B}_{R_i} and \mathbf{B}_{R_j} , which are proportional to the variances of the corresponding principal axes. Note that $(u'_{R_i,t} u_{R_j,p})^2$ is the correlation coefficient between the t th and p th coordinates. Thus, the numerator of c_k is a weighted average of the relationships between principal axes. Clearly, $0 \leq C_1 \leq \dots \leq C_{D_s} \leq \dots \leq C_{N_s} = 1$. The dimension D_s is then chosen such that C_{D_s} is high, for instance, $C_{D_s} = 0.9$.

Step 2—Similarity Measure Based on the Principal Axis of R_i and R_j : Once the number of dimensions representing R_i and R_j has been reduced, our goal is to measure the correlation, also called dependency, between their principal coordinates. To this end, the D_s columns of $\mathbf{U}_{R_i} \Lambda_{R_i}$ and $\mathbf{U}_{R_j} \Lambda_{R_j}$ are, respectively, considered as a predictor \mathbf{X} and a response variable \mathbf{Y} of a multivariate linear regression model expressed as

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{e} \quad (23)$$

where β is the matrix of regression coefficients and \mathbf{e} is the matrix of residual errors. The β matrix reflects whether \mathbf{X} and \mathbf{Y} are correlated, in other words, if \mathbf{Y} is a linear combination of \mathbf{X} . If there is no relationship between \mathbf{X} and \mathbf{Y} , the β matrix is equal to 0. Our approach consists in testing the multivariate general linear hypothesis $\beta = 0$.

One of the simplest methods to estimate the matrix $\hat{\beta}$ is the least square method which leads to $\hat{\beta} = (\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{Y}$. In order to test whether the hypothesis $\beta = 0$ is true or false, a classical statistical test known as Wilks' lambda test W is used here. It can be written as

$$W(\mathbf{X}, \mathbf{Y}) = \frac{\det((\mathbf{Y} - \mathbf{X}\hat{\beta})^t (\mathbf{Y} - \mathbf{X}\hat{\beta}))}{\det(\mathbf{Y}^t \mathbf{Y})}. \quad (24)$$

The Wilks test computed by (24) gives a value between 0 and 1. The zero value indicates that the hypothesis $\beta = 0$ is false and that \mathbf{X} and \mathbf{Y} are highly correlated. In our case, $\mathbf{X} = \mathbf{U}_{R_i} \Lambda_{R_i}$ and $\mathbf{Y} = \mathbf{U}_{R_j} \Lambda_{R_j}$. After some manipulations, the test, used as a similarity measure between regions, can be expressed as

$$S_{\text{MDS}}(R_i, R_j) = W(\mathbf{U}_{R_i} \Lambda_{R_i}, \mathbf{U}_{R_j} \Lambda_{R_j}) = \det(\mathbf{I} - \mathbf{U}_{R_j}^t \mathbf{U}_{R_i} \mathbf{U}_{R_i}^t \mathbf{U}_{R_j}). \quad (25)$$

To conclude this section, it should be remarked that the area of the regions is not included in the merging order defined by (21) and (25). As a result, some small regions may only be merged at a very late stage of the construction. This is most of the time inappropriate as mainly large regions should appear on the upper part of the tree. To solve this issue, the strategy proposed in [13] is used. The idea consists in forcing, in-between the merging iterations, the merging of regions that have an area smaller than a given percentage (typically 15%) of the average size of the regions currently created by the merging process.

C. BPT Punning for Hyperspectral Image Applications

1) *Classification*: In this section, the first application example dealing with supervised classification is described. The strategy is formally similar to the one described in Section III-C in the sense that the first step consists in measuring a certain feature for each tree node and then in deciding where to prune the tree.

In the classical pixel-wise supervised classification [16], representative training samples are used to specify the spectral attributes of the different classes composing the scene. Then, a classification algorithm compares the image pixels with the class spectral attributes to decide on the most likely class membership. The drawback of this approach is its lack of robustness because the process is pixel-based and spatial correlation is not taken into account. For this kind of application, the potential advantage of BPT is that the spatial correlation of pixels has already been exploited during the construction. This results in a set of regions, hierarchically structured, which can be classified based on their mean spectrum.

As an example, the support vector machine (SVM) [14] classifier is used. It is trained with a set of pixels for which the ground truth is known as in the classical approach. The features on which the classification pruning is defined are based on the class probability distribution of each BPT node. That is, each node R_i of the BPT is represented by its mean spectrum and the SVM classifier is used to estimate the probabilities that this spectrum corresponds to all possible classes C_j . The resulting class distribution is denoted by $\mathcal{P}_{R_i}(C_j)$ and it is estimated by [56]. Once the class probability distribution has been computed for each node, the *misclassification rate* (\mathcal{MR}) is computed. For leaf nodes, it corresponds to the error committed when the most probable class is assigned to the node. It is defined by

$$\mathcal{MR}(R_i) = \left(1 - \max_j \{\mathcal{P}_{R_i}(C_j)\}\right). \quad (26)$$

For nonleaf nodes, this expression is not appropriate because of the potential region size differences. For example, if a region results from the merging of a large region with a small region, the small region has little influence in the mean spectrum even if its class is different from the class of the large region. In this case, the misclassification rate of a region can be estimated by measuring the difference between the two class distributions of its child regions. The classical Bhattacharyya coefficient can be used to measure the distribution difference

$$\mathcal{BC}(\mathcal{P}_{R_i}^L, \mathcal{P}_{R_i}^R) = \sum_j \sqrt{\mathcal{P}_{R_i}^L(C_j) \mathcal{P}_{R_i}^R(C_j)} \quad (27)$$

where $\mathcal{P}_{R_i}^L$ and $\mathcal{P}_{R_i}^R$, respectively, represent the left and right child regions of R_i . Based on this measure, the misclassification rate for nonleaf regions is defined by

$$\mathcal{MR}(R_i) = \left(1 - \mathcal{BC}(\mathcal{P}_{R_i}^L, \mathcal{P}_{R_i}^R)\right). \quad (28)$$

Once the misclassification rate has been evaluated for all nodes, a decision has to be taken on where to prune the tree. To this end, we study the evolution of the misclassification rate between node R_i and its leaves R_i^{leaves} (that is, the leaves that belong to the subtree rooted at R_i). More precisely, the pruning decision is defined by

$$\phi_R(R_i) = \mathcal{MR}(R_i) - \overline{\mathcal{MR}(R_i^{\text{leaves}})} \quad (29)$$

where $\overline{\mathcal{MR}(R_i^{\text{leaves}})}$ represents the average misclassification rates of the leaves of R_i . Note that this rule is not increasing. We use here the *max rule*, that is, a bottom-up strategy that consists in pruning all nodes for which $\mathcal{MR}(R_i)$ is below a given threshold α_C and such that this statement is also true for all their descendants. Experimental evaluations have shown that α_C values around 0.3 are appropriate. Finally, once the tree has been pruned, a classification map is constructed by assigning the most probable class (as defined by the SVM classifier) to the leaf regions of the pruned tree.

In order to illustrate the utility of the approach, the classical Indian Pines AVIRIS hyperspectral data are used. It involves 200 spectral bands. An RGB composition of this data set is shown in Fig. 15 as well as the ground truth composed of 16 different classes. Two BPTs have been created for this image, one with the diffusion distance S_{Diff} and one with the multidimensional scaling approach S_{MDS} . In both cases, 150 bins have been used to represent the region model histograms.

The two resulting classification maps are shown in Fig. 16 and compared to the classical pixel-wise classification (also performed with the same SVM classifier). As can be seen, the BPT approach gives much better results than the pixel-based approach. Regions of the classification map are less noisy and the quality of contours between regions is very good. Table 1 illustrates the class-specific as well as the global classification accuracies. The benefit of the BPT approach is quite clear. Moreover, it can be concluded that the multidimensional scaling similarity S_{MDS} leads to the best results. Those results have been confirmed on other data sets [54].

2) *Segmentation*: The goal of this section is to discuss an unsupervised pruning technique for segmentation. Section III-C2 has suggested a supervised segmentation approach in which the number of regions is *a priori* known

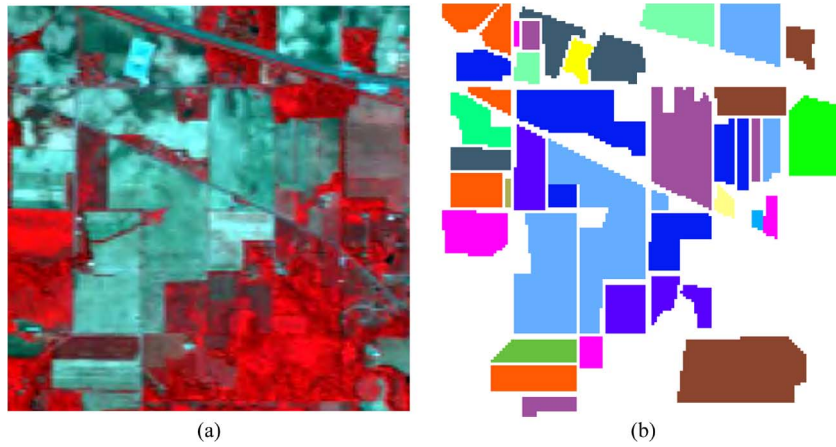


Fig. 15. Indian Pines AVIRIS data set. (a) RGB composition. (b) Ground truth.

or fixed by the user. Here, an approach based on graph partitioning that does not require any *a priori* information about the segmentation result is described. Normalized and min-cut algorithms are popular approaches to partition graph structures and they can potentially be used on a BPT as it is a particular graph. Here, we focus on the normalized cut algorithm [47]. Let us briefly recall this approach on a generic graph \mathcal{G} formed by $\mathcal{N}_{\mathcal{N}}$ nodes, whose edges between nodes i and j are weighted by values $w_{i,j}$. The optimum partitioning generates two subsets of graph nodes \mathcal{U} and \mathcal{V} , such that the following cost is minimized:

$$Ncut(\mathcal{U}, \mathcal{V}) = \frac{cut(\mathcal{U}, \mathcal{V})}{assoc(\mathcal{U}, \mathcal{N}_{\mathcal{N}})} + \frac{cut(\mathcal{V}, \mathcal{U})}{assoc(\mathcal{V}, \mathcal{N}_{\mathcal{N}})} \quad (30)$$

where $cut(\mathcal{U}, \mathcal{V}) = \sum_{i \in \mathcal{U}, j \in \mathcal{V}} w_{i,j}$ and $assoc(\mathcal{U}, \mathcal{N}_{\mathcal{N}}) = \sum_{i \in \mathcal{U}, j \in \mathcal{N}_{\mathcal{N}}} w_{i,j}$.

This minimization problem can be formulated as a spectral graph theory problem [47], which can be solved by finding the eigenvectors of $(\mathbf{D} - \mathbf{W})\mathbf{x} = \lambda\mathbf{D}\mathbf{x}$, where $\mathbf{W} = [w_{i,j}]$ is the affinity matrix and \mathbf{D} the diagonal matrix whose diagonal elements are $d_i = \sum_{j \in \mathcal{V}} w_{i,j}$. The sign of the components of the eigenvector with the second smallest eigenvalue creates a partition of the graph minimizing the normalized cut given by (30).

We use the normalized cut approach to partition *individually* each BPT branch. Here, a branch is a path that connects the BPT root node R_{Root} to one of its leaf node R_{Leaf} . As the branch represents a very simple graph with a linear structure and very few nonzero weights between nodes, as an alternative to the spectral graph theory solution, an exhaustive search of the minimum normalized cut can also be performed.

The part of the branch which has to be pruned should involve nodes that are similar to the leaf node. By contrast, the nodes that are going to be preserved should be quite different from the leaf node. If $P(R_i = R_{Leaf})$ denotes the

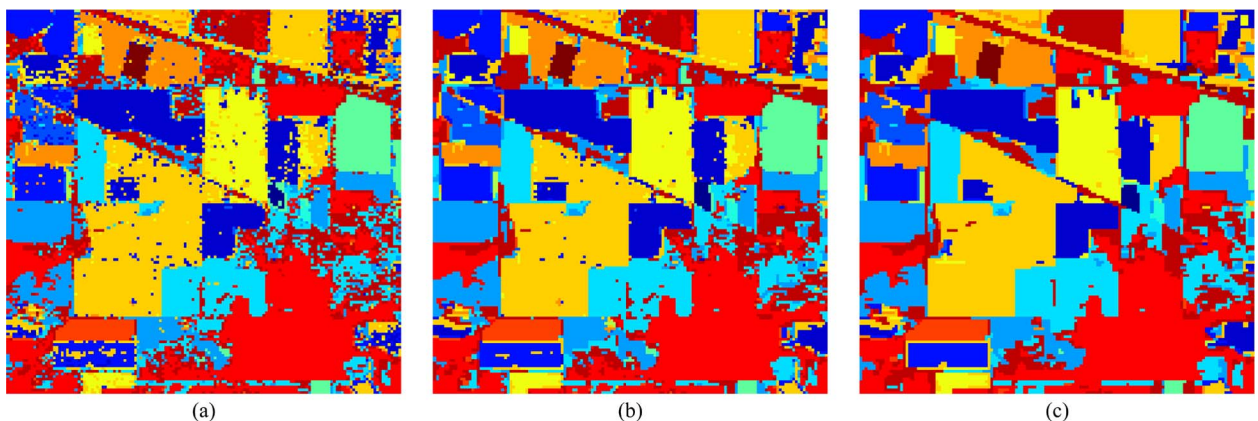


Fig. 16. Classification maps. (a) Pixel-wise classification. (b) BPT, S_{Diff} . (c) BPT, S_{MDS} .

Table 1 Class-Specific Accuracy

Class	Pixel-wise	BPT S_{DIF}	BPT S_{MDS}
1	75.61	82.93	80.49
2	83.46	92.19	92.75
3	84.35	96.81	96.01
4	76.14	92.05	91.48
5	94.37	95.17	95.71
6	97.15	98.40	97.50
7	92.31	88.89	88.89
8	98.09	99.18	99.73
9	90.00	85.71	100.00
10	83.06	89.94	88.57
11	91.52	96.06	99.73
12	86.55	91.76	94.36
13	96.22	98.11	98.76
14	95.57	97.89	97.94
15	67.72	80.35	97.89
16	91.67	93.06	97.22
Average	87.74	92.40	94.69

probability that R_i is similar to the leaf R_{Leaf} , the weights defined by

$$w_{ij} = e^{\sqrt{P(R_i=R_{Leaf})P(R_j=R_{Leaf})} + \sqrt{P(R_i \neq R_{Leaf})P(R_j \neq R_{Leaf})}} - 1 \quad (31)$$

should state that nodes on both sides of the cut are similar (because they are either very similar or very dissimilar to the leaf). Several solutions can be used to estimate $P(R_i = R_L)$. The MDS criterion $P(R_i = R_L) = S_{MDS}(R_i, R_L)$ defined for the BPT construction was used in the following. In order to build the affinity matrix, the weights w_{ij} are not computed between all nodes in the branch but only between a node and its two children, its sibling, and its direct ancestor. The weight is zero for its similarity with the remaining branch nodes.

In practice, the first cut may not be the best one. For instance, if the root node is very different from the rest of the tree, the first cut will separate, for all branches, the root from the rest of the branch. For this reason, instead of a single cut, a recursive spectral graph partitioning algorithm is performed as suggested in [47]. The idea is to iterate the partitioning of the branch on the subset of nodes that includes the leaf R_L until the N_{cut} value is smaller than a given threshold \mathcal{T}_{Ncut} (set to 0.3 for all experiments).

Once each branch has been investigated with the normalized cut algorithm, the pruning can be performed, but of course, the information coming from all the individual branch partitions should be merged. In fact, except leaves,

each BPT node belongs to several branches. In order to decide whether a node has to be pruned, a majority vote based on the results of all branches is performed: If a node belongs to M different branches and more than $M/2$ branches have stated that the node has to be pruned, it will be pruned. Finally, as in the case of the classification pruning, the pruning decision may not be increasing, so a node is pruned only if all its descendants have to also be pruned (*max rule*).

Fig. 17 illustrates the results of the segmentation approach on four portions of classical Hyperspectral data sets. Fig. 17(a) and (b) has been extracted from the Pavia University data set, which was acquired over the city of Pavia, Italy, by the Reflective Optics Systems Imaging Spectrometer (ROSIS-03) hyperspectral sensor. The spatial resolution is 1.3 m. The original data set is composed of 115 spectral bands, ranging from 0.43 to 0.86 μm with a bandwidth of 4 nm. However, typical noisy bands corresponding to the water absorption were previously discarded leading to 103 bands. Fig. 17(c) has also been acquired by the hyperspectral ROSIS-3 sensor and corresponds to the Pavia city center. It contains 102 spectral bands after removal of noisy bands. Finally, Fig. 17(d) was extracted from the HYDICE data set showing an urban area. Originally, it had 210 spectral bands from 0.4 to 2.5 μm . After removing water absorption and noisy bands, the data contain 167 spectral bands with a spatial resolution of approximately 3 m.

The second and third rows of Fig. 17 compare the partitions obtained with the BPT pruning approach with these obtained with a classical region merging algorithm for hyperspectral images: recursive hierarchical segmentation (RHSEG) [34]. Note that the RHSEG algorithm is supervised in the sense that it is an iterative region merging algorithm and the user has to decide when to stop the merging process. Even though the BPT approach is unsupervised, it obtains much better results. This statement can be corroborated both subjectively on the partitions of Fig. 17 as well as objectively. To this end, ground truth partitions have been manually created. They are presented in the fourth row of Fig. 17. In order to assess the quality of the obtained partitions with respect to the ground truth, we have used the asymmetric distance d_{asym} proposed in [15]. $d_{\text{asym}}(\mathcal{P}, \mathcal{P}_{GT})$ measures the minimum number of pixels whose labels should be changed so that a partition \mathcal{P} becomes *finer* than the ground truth partition \mathcal{P}_{GT} . A partition \mathcal{P} is said to be *finer* than a partition \mathcal{Q} , if each region of \mathcal{P} is included in a region of \mathcal{Q} . This distance is normalized by the total image pixel number minus one. As a result, d_{asym} is ranging between 0 and 1. As the distance is not symmetrical, the final distance is given by $(d_{\text{asym}}(\mathcal{P}, \mathcal{P}_{GT}) + d_{\text{asym}}(\mathcal{P}_{GT}, \mathcal{P}))/2$. Table 2 reports the distance values for the four images as well as the average value. Here also the benefits of using the BPT approach are quite clear.

Before concluding, in order to assess the complexity of the BPT construction, several tests have been performed.

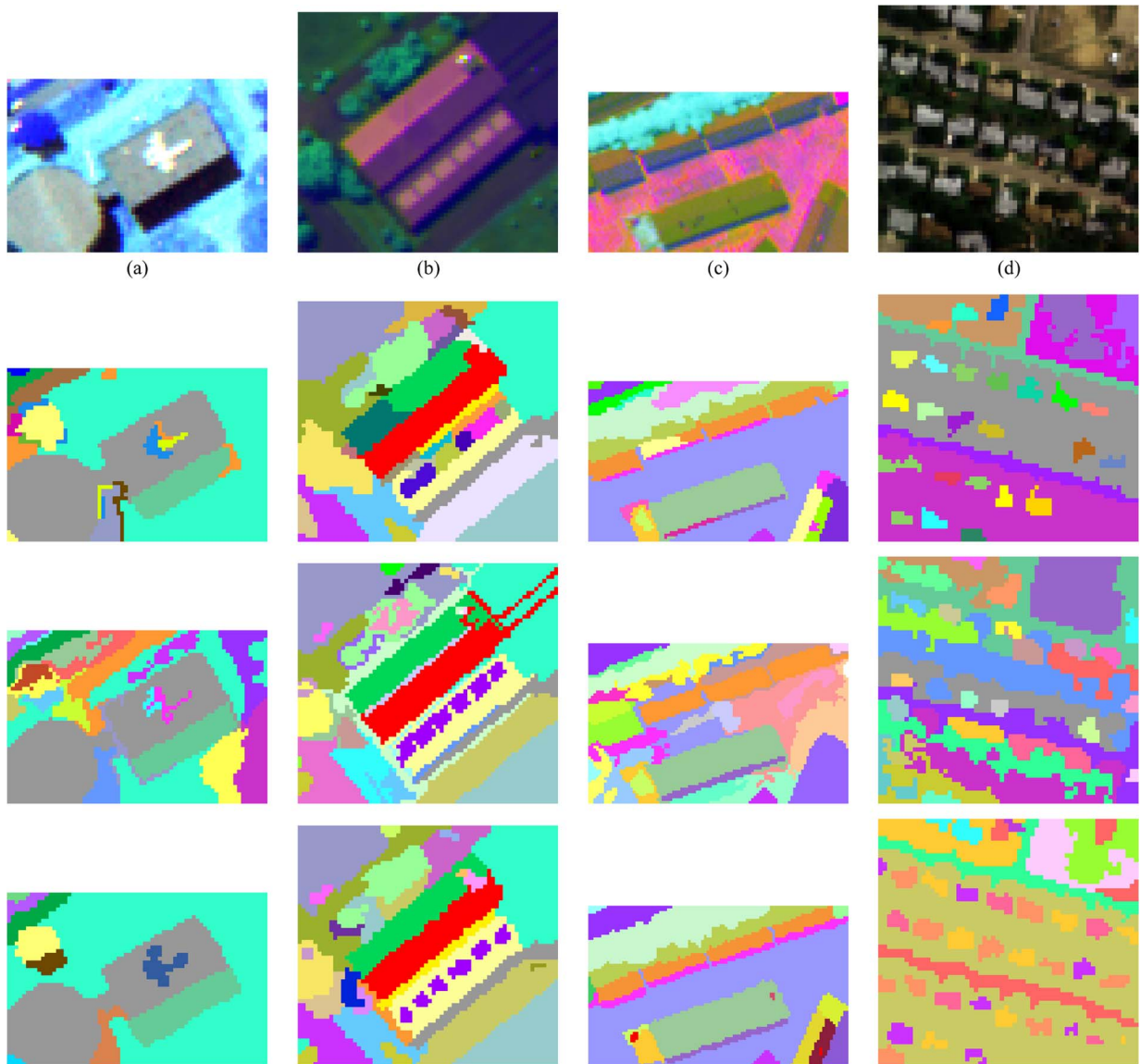


Fig. 17. Segmentation results. First row: Original RGB compositions. Second row: Segmentation results obtained with BPT. Third row: Segmentation results obtained with [34]. Fourth row: Ground Truth partitions. (a) Pavia Univ.1. (b) Pavia Univ.2. (c) Pavia center. (d) HYDICE.

The BPT construction algorithm has been tested using an Intel Xeon processor running under Linux with a clock speed of 2.40 GHz and 20 GB of random access memory (RAM). The central processing unit (CPU) time and the required memory for BPT construction for the images Pavia Univ.1 and Pavia Univ.2 are shown in Table 3. In this last table, the two different merging criteria S_{Diff} and S_{MDS} are compared. In this case, the number of bins for the nonparametrical statistical region model is set to 150 and the intrinsic dimension of the image is equal to $D_s = 3$.

Table 3 corroborates that the merging criterion O_{MDS} is clearly the most time consuming. This is completely explained by the computation of the matrix distance used to compute the principal coordinates. However, it must be

noticed that the BPT is viewed in this work as an image representation that has to be constructed only once and that opens the door to a wide range of applications. Therefore, the computational load of the BPT construction is not a serious drawback in terms of the applicability of the approach.

V. CONCLUSION

This paper has focused on the interest of BPTs as image representation for remote sensing. This representation possesses some interesting features to deal with high-resolution images. In particular, similar pixels are grouped into regions which form the entities on which the further processing is based. Moreover, the representation is

Table 2 Partition Distance

Data Set	BPT Segmentation	RHSEG [34]
Pavia Univ.1	0.091	0.249
Pavia Univ.2	0.191	0.193
Pavia Center	0.167	0.345
Hydice	0.211	0.374
Average	0.165	0.290

Table 3 Performances BPT Construction

Merging Criterion	Time (minutes)	Memory (Mbytes)
Pavia Univ.1 (3705 pixels)		
O_{DIF}	1.01	309.24
O_{MDS}	9.24	337.92
Pavia Univ.2 (3174 pixels)		
O_{DIF}	0.58	290.81
O_{MDS}	7.50	327.68

structured in a tree representing the inclusion relationship between regions. This structure leads to an easy access to the image content. Finally, the representation describes the content at various scales of resolution.

Two BPT constructions have been presented in this paper. They rely on iterative merging algorithms. The first one was described in the context of multidimensional SAR images. The region model was based on the estimated mean covariance matrix of the pixels, and several region

similarity measures, such as the Wishart and the geodesic similarities, have been defined. Then, the tree was used for speckle noise reduction. The complete noise filtering approach turned out to be a connected operator that has the ability to remove noise while preserving the contour information and spatial details. Its advantage over classical filtering techniques has been highlighted. We have also illustrated how the tree can be used for specific segmentation tasks leading to very precisely defined regions.

The second context in which the BPT has been studied deals with hyperspectral images. Here, the mean model may not be accurate enough and the effectiveness of non-parametric statistical model was discussed. A powerful solution based on multidimensional scaling to capture the correlation between spectral bands was proposed as the similarity measure. In terms of applications, the tree was used for supervised classification. The resulting hierarchical region-based classification has demonstrated its superiority compared to pixel-based classification. Unsupervised segmentation was also studied. The BPT was pruned with the help of a graph cut algorithm (normalized cut).

Although the use of BPT for remote sensing images is still in an early stage and many aspects can be improved, this kind of image representation already provides very interesting results compared to the current state of the art. ■

Acknowledgment

The authors would like to thank the European Space Agency for providing the AgrisAR2009 Campaign data.

REFERENCES

- [1] J.-S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999.
- [2] G. Vasile, E. Trounev, J. S. Lee, and V. Buzuloiu, "Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 6, pp. 1609–1621, Jun. 2006.
- [3] G. H. Golub and C. F. van Van Loan, *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [4] O. Mora, J. J. Mallorquí, and A. Broquetas, "Linear and nonlinear terrain deformation maps from a reduced set of interferometric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2243–2253, Oct. 2003.
- [5] L. Pipia, X. Fàbregas, A. Aguasca, C. López-Martínez, S. Duque, J. J. Mallorquí, and J. Marturia, "Polarimetric differential SAR interferometry: First results with ground-based measurements," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 11, pp. 167–171, Jan. 2009.
- [6] A. Alonso-Gonzalez, C. López-Martínez, and P. Salembier, "Filtering and segmentation of polarimetric SAR images with binary partition trees," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 25–30, 2010, pp. 4043–4046.
- [7] A. Alonso-Gonzalez, C. López-Martínez, and P. Salembier, "Filtering and segmentation of polarimetric SAR data based on binary partition trees," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 593–605, Feb. 2012.
- [8] A. Alonso-Gonzalez, C. López-Martínez, and P. Salembier, "PolSAR speckle filtering and segmentation based on binary partition tree representation," in *Proc. 5th Int. Workshop Sci. Appl. SAR Polarimetry Polarimetric Interferometry*, Frascati, Italy, Jan. 24–28, 2011, (ESA SP-695, Mar. 2011).
- [9] J. D. Ballester Berman, "Retrieval of biophysical parameters of agricultural crops using polarimetric SAR interferometry," Ph.D. dissertation, Dept. Física, Ingeniería de Sistemas y Teoría, Universitat D'Alacant, San Vicente del Raspeig, Spain, 2007.
- [10] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problems*, vol. 14, pp. R1–R54, 1998.
- [11] F. Barbaresco, "Interactions between symmetric cone and information geometries: Bruhat-tits and siegel spaces models for high resolution autoregressive doppler imagery," in *Emerging Trends in Visual Computing*, vol. 5416, *Lecture Notes in Computer Science*, F. Nielsen, Ed. Berlin, Germany: Springer-Verlag, 2009, pp. 124–163.
- [12] A. Buades, B. Coll, and J. M. Morel, "A non local algorithm for image denoising," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, 2005, vol. 2, pp. 60–65.
- [13] F. Calderero and F. Marques, "Region-merging techniques using information theory statistical measures," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1567–1586, Jun. 2010.
- [14] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
- [15] J. S. Cardoso and L. Corte-Real, "Toward a generic evaluation of image segmentation," *IEEE Trans. Image Process.*, vol. 14, no. 11, pp. 1773–1782, Nov. 2005.
- [16] C. I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Dordrecht, The Netherlands: Kluwer, 2003.
- [17] S. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 498–518, Mar. 1996.
- [18] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 5, pp. 1551–1565, Sep. 1998.
- [19] S. Cloude and K. Papathanassiou, "Three-stage inversion process for polarimetric SAR interferometry," *Proc. Inst. Electr. Eng.—Radar Sonar Navig.*, vol. 150, no. 3, pp. 125–134, 2003.

- [20] S. Cloude, *Polarisation Applications in Remote Sensing*. Oxford, U.K.: Oxford Univ. Press, 2009.
- [21] T. F. Cox and M. A. Cox, *Multidimensional Scaling*. London, U.K.: Chapman & Hall, 1994.
- [22] E. Cramer and W. A. Nicewander, "Some symmetric, invariant measures of multivariate association," *Psychometrika*, vol. 44, pp. 43–54, 1979.
- [23] C. M. Cuadras, A. Arenas, and J. Fortiana, "Some computational aspects of a distance-based model for prediction," *Commun. Stat., Simul. Comput.*, vol. 25, pp. 593–609, 1996.
- [24] C. M. Cuadras, S. Valero, P. Salembier, and J. Chanussot, "Some measures of multivariate association relating two spectral data sets," presented at the 19th Int. Conf. Comput. Stat., Paris, France, Aug. 22–27, 2010.
- [25] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Boston, MA: Artech House, 2005.
- [26] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*. New York: Wiley, 1991.
- [27] M. Dimiccoli and P. Salembier, "Hierarchical region-based representation for segmentation and filtering with depth in single images," in *Proc. IEEE Int. Conf. Image Process.*, El Cairo, Egypt, pp. 3533–3536.
- [28] A. P. Doulgeris, S. N. Anfinson, and T. Eltoft, "Classification with a non-Gaussian model for PolSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 10, pp. 2999–3009, Oct. 2008.
- [29] M. Fauvel, J. Chanussot, and J. A. Benediktsson, "A spatial-spectral kernel based approach for the classification of remote sensing images," *Pattern Recognit.*, vol. 45, no. 1, pp. 281–392, 2011.
- [30] A. Frery, H.-J. Muller, C. Yanasse, and S. Sant'Anna, "A model for extremely heterogeneous clutter," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 3, pp. 648–659, May 1997.
- [31] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Stat.*, vol. 34, pp. 152–177, 1963.
- [32] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, no. 11, pp. 1145–1149, Nov. 1976.
- [33] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*. Greenwood Village, CO: Roberts & Company, 2007.
- [34] J. A. Gualtieri and J. C. Tilton, "Hierarchical segmentation of hyperspectral data," presented at the AVIRIS Earth Sci. Appl. Workshop, Pasadena, CA, Mar. 5–8, 2002.
- [35] R. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*, Norwell, MA: Kluwer, 2001.
- [36] P. Kersten, J.-S. Lee, and T. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 519–527, Mar. 2005.
- [37] A. B. Kostinski and W. M. Boerner, "On foundations of radar polarimetry," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 12, pp. 1395–1403, Dec. 1986.
- [38] J. S. Lee, "Speckle analysis and smoothing of synthetic aperture radar images," *Comput. Graph. Image Process.*, vol. 17, pp. 24–32, 1981.
- [39] J.-S. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1017–1028, Sep. 1994.
- [40] J. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL: CRC Press, 2009.
- [41] H. Ling and K. Okada, "Diffusion distance for histogram comparison," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, New York, Jun. 17–22, 2006, pp. 246–253.
- [42] J. C. Platt, "Probabilistic outputs for support vector machines and comparison to regularized likelihood methods," in *Advances in Large Margin Classifiers*, A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans, Eds. Cambridge, MA: MIT Press, 2000.
- [43] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline l-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
- [44] P. Salembier and J. Serra, "Flat zones filtering, connected operators and filters by reconstruction," *IEEE Trans. Image Process.*, vol. 3, no. 8, pp. 1153–1160, Aug. 1995.
- [45] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation and information retrieval," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 561–576, Apr. 2000.
- [46] P. Salembier and M. H. F. Wilkinson, "Connected operators," *IEEE Signal Process. Mag.*, vol. 26, no. 6, pp. 136–157, Nov. 2009.
- [47] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, Aug. 2000.
- [48] Y. Tarabalka, J. A. Benediktsson, and J. Chanussot, "Spectral-spatial classification of hyperspectral imagery based on partitioned clustering techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.
- [49] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Comput. Vis. Conf.*, 1998, pp. 839–846.
- [50] R. J. A. Tough, D. Blacknell, and S. Quegan, "A statistical description of polarimetric and interferometric synthetic aperture radar data," *Proc. R. Soc. Lond. A*, vol. 449, pp. 567–589, 1995.
- [51] S. Valero, P. Salembier, and J. Chanussot, "New hyperspectral data representation using binary partition tree," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Honolulu, HI, Aug. 2010, pp. 80–83.
- [52] S. Valero, P. Salembier, and J. Chanussot, "Comparison of merging orders and pruning strategies for binary partition tree in hyperspectral data," in *Proc. IEEE Int. Conf. Image Process.*, Hong Kong, China, Sep. 2010, pp. 2565–2568.
- [53] S. Valero, P. Salembier, J. Chanussot, and C. Cuadras, "Improved binary partition tree construction for hyperspectral images: Application to object detection," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Vancouver, BC, Canada, Jul. 2011, pp. 2515–2518.
- [54] S. Valero, "A new tool for hierarchical representation and analysis of hyperspectral images," Ph.D. dissertation, GIPSA-lab, Grenoble Images Parole Signal Automatique/Signal Theory and Communications Dept., Grenoble Univ. Tech./Univ. Catalonia, Barcelona, Spain, 2011.
- [55] G. Vasile, J.-P. Ovarlez, F. Pascal, and C. Tison, "Coherency matrix estimation of heterogeneous clutter in high-resolution polarimetric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 1809–1826, Apr. 2010.
- [56] T.-F. Wu, C.-J. Lin, and R. C. Weng, "Probability estimates for multi-class classification by pairwise coupling," *J. Mach. Learn. Res.*, vol. 5, pp. 975–1005, 2004.
- [57] S. Yueh, J. Kong, J. Jao, R. Shin, and L. Novak, "K-distribution and polarimetric terrain radar clutter," *J. Electromagn. Waves Appl.*, vol. 3, pp. 747–768, 1989.
- [58] J. J. V. Zyl, C. H. Papas, and C. Elachi, "On the optimum polarizations of incoherently reflected waves," *IEEE Trans. Antennas Propag.*, vol. AP-35, no. 7, pp. 818–825, Jul. 1987.
- [59] D. H. Hoekman, M. A. M. Vissers, and T. N. Tran, "Unsupervised full-polarimetric SAR data segmentation as a tool for classification of agricultural areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 2, pp. 402–411, 2011.
- [60] Y. Tarabalka, J. C. Tilton, J. A. Benediktsson, and J. Chanussot, "A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 1, pp. 262–272, 2012.

ABOUT THE AUTHORS

Alberto Alonso-González (Student Member, IEEE) was born in León, Spain, in 1984. He received the B.Sc. degree in computer science and the M.Sc. degree in telecommunication engineering from the Technical University of Catalonia (UPC), Barcelona, Spain, in 2007 and 2009, respectively, where he is currently working toward the Ph.D. degree in telecommunication engineering. His Ph.D. dissertation is focused on multidimensional synthetic aperture radar (SAR) data modeling and processing.



In 2009, he joined the Signal Theory and Communications Department (TSC), Technical University of Catalonia. His research interests include multidimensional synthetic aperture radar (SAR), SAR interferometry and polarimetry, digital signal and image processing, and data segmentation and simplification techniques.

Mr. Alonso-González received the First Place Student Paper Award at the 2012 European Conference on Synthetic Aperture Radar (EUSAR), Nuremberg, Germany.

Silvia Valero (Student Member, IEEE) received the M.S. degree in electrical engineering from the Technical University of Catalonia, Barcelona, Spain, in 2007 and the M.S. degree in computer science from the Grenoble Institute of Technology (Grenoble-INP), Grenoble, France, in 2008. She received the conjoint Ph.D. degree from the Grenoble-INP and the Technical University of Catalonia in 2011. Her Ph.D. work was devoted to developing advanced image processing techniques for hyperspectral images.



She is working in the construction and exploitation of hierarchical region-based representation for hyperspectral data. Her research interests are image processing, hyperspectral imaging, pattern recognition, information retrieval, and tree processing techniques.

Dr. Valero received the Symposium Prize Paper Award at the IEEE International Geoscience and Remote Sensing Symposium.

Jocelyn Chanussot (Fellow, IEEE) received the M.Sc. degree in electrical engineering from the Grenoble Institute of Technology (Grenoble INP), Grenoble, France, in 1995 and the Ph.D. degree from Savoie University, Annecy, France, in 1998.



In 1999, he was with the Geography Imagery Perception Laboratory for the Delegation Generale de l'Armement (DGA—French National Defense Department). Since 1999, he has been with Grenoble INP, where he was an Assistant Professor from 1999 to 2005, an Associate Professor from 2005 to 2007, and is currently a Professor of Signal and Image Processing. He is currently conducting his research at the Grenoble Images Speech Signals and Automatics Laboratory (GIPSA-Lab). His research interests include image analysis, multicomponent image processing, nonlinear filtering, and data fusion in remote sensing.

Dr. Chanussot is the founding President of the IEEE Geoscience and Remote Sensing French chapter (2007–2010), which received the 2010 IEEE Geoscience and Remote Sensing Society (GRSS) Chapter Excellence Award “for excellence as a Geoscience and Remote Sensing Society chapter demonstrated by exemplary activities during 2009.” He was the recipient of the 2006 NORSIG Best Student Paper Award, the 2011 IEEE GRSS Symposium Best Paper Award, and the 2012 IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING Prize Paper Award. He was a member of the IEEE GRSS AdCom (2009–2010), in charge of membership development. He was the General Chair of the first IEEE GRSS Workshop on Hyperspectral Image and Signal Processing, Evolution in Remote sensing (WHISPERS). He was the Chair (2009–2011) and the Cochair of the GRS Data Fusion Technical Committee (2005–2008). He was a member of the Machine Learning for Signal Processing Technical Committee of the IEEE Signal Processing Society (2006–2008) and the Program Chair of the IEEE International Workshop on Machine Learning for Signal Processing (2009). He was an Associate Editor for the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS (2005–2007) and for *Pattern Recognition* (2006–2008). Since 2007, he has been an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING. Since 2011, he has been the Editor-in-Chief of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

Carlos López-Martínez (Senior Member, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree in multidimensional synthetic aperture radar speckle noise modeling and filtering from the Technical University of Catalonia, Barcelona, Spain, in 1999 and 2003, respectively.



From October 2000 to March 2002, he was with the Frequency and Radar Systems Department, HR, German Aerospace Center, DLR, Oberpfaffenhofen, Germany. From June 2003 to December 2005, he was with the Image and Remote Sensing Group-SAPHIR Team, in the Institute of Electronics and Telecommunications of Rennes (I.E.T.R.—CNRS UMR 6164), Rennes, France. In January 2006, he joined the Technical University of Catalonia as a Ramón-y-Cajal Researcher, where he is currently an Associate Professor in the area of remote sensing and microwave technology. His research interests include synthetic aperture radar (SAR) and multidimensional SAR, radar polarimetry, physical parameter inversion, digital signal processing, estimation theory, and harmonic analysis.

Prof. López-Martínez is an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING and has served as Guest Editor of the *EURASIP Journal on Advances in Signal Processing*. He has organized many invited sessions in international conferences on radar and SAR polarimetry. He has presented advanced courses and seminars on radar polarimetry to a wide range of organizations and events. He received the Student Prize Paper Award at the 2002 European Conference on Synthetic Aperture Radar (EUSAR), Cologne, Germany.

Philippe Salembier (Fellow, IEEE) received a degree from the Ecole Polytechnique, Paris, France, in 1983 and a degree from the Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1985. He received the Ph.D. from the Swiss Federal Institute of Technology (EPFL), Zürich, Switzerland, in 1991.



He was a Postdoctoral Fellow at the Harvard Robotics Laboratory, Cambridge, MA, in 1991. From 1985 to 1989, he worked at Laboratoires d'Electronique Philips, Limeil-Brevannes, France, in the fields of digital communications and signal processing for HDTV. In 1989, he joined the Signal Processing Laboratory of the Swiss Federal Institute of Technology, Lausanne, Switzerland, to work on image processing. At the end of 1991, after a stay at the Harvard Robotics Laboratory, he joined the Technical University of Catalonia, Barcelona, Spain, where he is currently Professor lecturing on the area of digital signal and image processing. His current research interests include image and sequence coding, compression and indexing, segmentation, video sequence analysis, mathematical morphology, level sets, and nonlinear filtering. In terms of standardization activities, he has been particularly involved in the definition of the MPEG-7 standard (Multimedia Content Description Interface) as Chair of the Multimedia Description Scheme group between 1999 and 2001.

Prof. Salembier served as an Area Editor of the *Journal of Visual Communication and Image Representation*, a Co-Editor-in-Chief of *Signal Processing*, and an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE SIGNAL PROCESSING LETTERS. Currently, he is an Associate Editor of the *Journal on Image and Video Processing*, *Signal Processing: Image Communication*, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY. He was a member of the Image and Multidimensional Signal Processing Technical Committee of the IEEE Signal Processing Society between 2000 and 2006.