Optimum Graph-Cuts for Pruning Binary Partition Trees of Polarimetric SAR images

Philippe Salembier, *Fellow IEEE,* and Samuel Foucher, *Member IEEE*

Abstract—This paper investigates several optimum graph-cuts techniques for pruning Binary Partition Trees (BPTs) and their usefulness for low-level processing of Polarimetric SAR (PolSAR) images. BPTs group pixels to form homogeneous regions, which are hierarchically structured by inclusion in a binary tree. They provide multiple resolutions of description and easy access to subsets of regions. Once constructed, BPTs can be used for a large number of applications. Many of these applications consist in populating the tree with a specific feature and in applying a graph-cut called *pruning* to extract a partition of the space. In this paper, different pruning examples involving the optimization of a global criterion are discussed and analyzed in the context of PolSAR images for segmentation. Through objective evaluation of the resulting partitions by means of Precision and Recall for boundaries curves, the best pruning technique is identified and the influence of the tree construction on the performances is assessed.

Index Terms—Binary Partition Tree, PolSAR image processing, graph-cut, tree pruning, speckle noise, segmentation, Precision and Recall for boundaries.

I. INTRODUCTION

The application of Binary Partition Trees (BPTs) [1] for remote sensing applications such as Polarimetric SAR (PolSAR) [2] and hyperspectral images [3], [4] is currently gaining interest. BPTs are hierarchical region-based representations in which pixels are grouped by similarity. Their construction is often based on an iterative region-merging algorithm: starting from an initial partition, the pair of most similar neighboring regions is iteratively merged until one region representing the entire image support is obtained. The BPT essentially stores the complete merging sequence in a binary tree structure. Once constructed, BPTs can be used for a large number of tasks including image filtering with connected operators, segmentation, object detection or classification [1], [5]. Many of these tasks involve the extraction of a partition from the BPT through a graph cut.

In this paper, we focus on low level PolSAR image processing tasks. We study in particular the interest of a specific graph cut called *pruning* in this context. We discuss and evaluate various pruning techniques formulated as the search in the BPT of a partition optimizing a certain criterion. The criteria we analyze take into account the specific nature of PolSAR data and the presence of speckle noise resulting from the coherent integration of the electromagnetic waves. The

Philippe Salembier is with the Signal Theory and Communications department of the Technical University of Catalonia, BarcelonaTech.

Samuel Foucher is with the Vision and Imaging Team of the Computer Research Institute of Montreal.

main contributions of this paper compared to [2], [5] are the proposal of new pruning strategy relying on optimization of a global criterion for PolSAR images as well as the objective evaluation of the resulting partitions thanks to a set of realistic simulated PolSAR images where the underlying ground-truth is available [6]. A subset of results reported here were previously published in [7], [8]. However, in this paper, we extend the set of optimum pruning criteria and make a detailed and objective analysis of the influence of the BPT construction on the results.

The paper is organized as follows: Section II gives a short introduction on PolSAR data. Section III discusses the BPT creation and its processing with graph cut. Several classical pruning strategies are presented in Section IV and more advanced approaches relying on the optimization of a global criterion are discussed in Section V. These techniques are evaluated objectively as well as subjectively in Section VI. Finally, conclusions are reported in Section VII.

II. POLSAR DATA

Synthetic Aperture Radars (SAR) are active microwave imaging systems. They are becoming increasingly popular for Earth observation because they work independently of the day and night cycle and of weather conditions. A SAR system essentially transmits an electromagnetic wave and records its echo to localize targets. In order to achieve a high spatial resolution, narrow beamwidth or equivalently large antennas are necessary. SAR systems deal with this issue by making use of the relative motion between the sensor and the target. As the radar moves, it repeatedly illuminates the target with electromagnetic pulses. The echoes are coherently recorded and combined in a post-processing that synthesizes a very large array and creates a high resolution image. The speckle noise results from the coherent addition of the scattered electromagnetic waves and is considered as one of the main problems for the exploitation of SAR data.

In the early 90's, multidimensional systems were developed. They provide complex SAR images $[S_1, S_2, \ldots, S_m]$ by introducing some sort of diversity. An important example is Polarimetric SAR (PolSAR) [9], [10] where the diversity is based on considering different polarization states for the transmitted and received electromagnetic waves. This makes SAR data sensitive to the target geometry, including vegetation, and to the dielectric properties of the target. For every resolution cell, a PolSAR system measures the scattering matrix:

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

where h and v represent the horizontal and vertical polarization states and S_{pq} for $p, q \in \{h, v\}$ denotes the complex SAR image where the reception (transmission) polarization states is p (q).

Since the dimensions of the resolution cell are normally larger that the wavelength of the electromagnetic wave, the scattered wave results from the coherent combination of many waves. This coherent addition process creates the speckle. Although the speckle represents a true electromagnetic measurement, its complexity is such that it is considered as a random process. The $\mathbf{S}$ scattering matrix is usually vectorized and written as a vector $\mathbf{k}$ [11]:

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

For fully developed speckle in a homogeneous (constant radar cross section) area, $\mathbf{k}$ is characterized by a three dimensional zero-mean complex Gaussian pdf:

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

Therefore, the distribution of $\mathbf{k}$ is completely described by the Hermitian positive definite covariance matrix:

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

where $E\{x\}$ is the statistical expectation of x. The Maximum Likelihood Estimation (MLE) of $\mathbf{C}$, i.e., the multilook, under the assumption of statistical ergodicity and homogeneity, is obtained by substituting the statistical expectation by an averaging:

$$\mathbf{Z} = \langle \mathbf{k}\mathbf{k}^H \rangle_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{k}_i \mathbf{k}_i^H \tag{5}$$

where n indicates the number of independent looks or samples employed to estimate $\mathbf{C}$ and $\mathbf{k}_i$ is the ith sample vector. The estimated covariance matrix $\mathbf{Z}$ is statistically characterized by a Wishart distribution [12].

III. BPT CREATION AND PROCESSING THROUGH GRAPH CUT

The BPT creation starts by the definition of an initial partition which can be composed of individual pixels as in [2], [5]. While this strategy guarantees a high precision as starting point of the merging process, it also implies high computational and memory costs as many regions have to be handled. As an alternative, the initial partition may correspond to an over-segmentation as a super-pixel partition. This initial partition issue was studied in [7] where several alternative strategies were evaluated. The main conclusion of this study is that the use of super-pixel partition as initial partition of the merging process can indeed drastically reduce the computational load of the BPT creation without any significant impact on the quality of the regions and partitions represented by the tree. One of the key point however is to use a denoising filter

adapted to PolSAR images such as [13], [14] before computing the super-pixel partitions. The best combination found in [7] involves the use of the σ-Lee denoising filter [13] followed by the SLIC algorithm [15] to compute the super-pixel. Only the three diagonal elements of the covariance matrices $\mathbf{Z}$ (after denoising) have been used to generate the super-pixels. For the time-being, we will assume that the initial partition is generated with this strategy but we will revisit some of these choices in the experimental section of this paper. The regions of this initial partition correspond to the leaves of the BPT.

Once the initial partition is defined, the BPT construction is done by iteratively merging the pair of most similar neighboring regions. When two neighboring regions are merged, the resulting union is represented as a new node in the BPT and the inclusion relationship between the initial regions and their union is represented by two edges connecting the parent node to the two child nodes. The merging process is iterated until a unique region representing the entire image support is obtained. This region corresponds to the root node of the BPT.

In the PolSAR case of interest here, the information carried by pixels (i, j) of an image I corresponds to the covariance matrix $\mathbf{Z}_{ij}^I$ after denoising. To construct the BPT, we need to define a region model and a similarity measure between neighboring regions.

We use the strategy defined in [2], [7] where regions R are modeled by their mean covariance matrix $\mathbf{Z}_R = \frac{1}{|R|}\sum_{i,j \in R} \mathbf{Z}_{ij}^I$, where $|R|$ is the region number of pixels. If the region is homogeneous, this estimation corresponds to the MLE defined by Eq. 5. The distance between neighboring regions defining the merging order can be measured by the geodesic similarity adapted to the cone of positive definite Hermitian matrices [16]:

$$d_{Geod}(R_1, R_2) = \|log\left(\mathbf{Z}_{R_1}^{-1/2}\mathbf{Z}_{R_2}\mathbf{Z}_{R_1}^{-1/2}\right)\|_F$$
$$. \; ln\left(\frac{2|R_1||R_2|}{|R_1| + |R_2|}\right) \tag{6}$$

where $log(.)$ is the matrix logarithm, $ln(.)$ the natural logarithm and $\|.\|_F$ represents the Frobenius norm[1]. This similarity measure exploits the geometry of the space defined by the covariance matrices and also involves a multiplicative term, $ln\left(2|R_1||R_2|/|R_1| + |R_2|\right)$, encouraging the merging of small regions. In the experimental section, we will compare this solution to alternative similarity functions.

Once the BPT has been constructed, it can be used for a wide range of applications including filtering, segmentation or classification. In many cases, the application relies on the extraction of a partition from the BPT. This process can be seen as a particular graph cut called *pruning* that can be formally defined as follows: Assume the tree root is connected to a *source* node and that all the tree leaves are connected to a *sink* node (see Fig. 1). A *pruning* is a graph cut that separates the tree into two connected components, one connected to the source and the other to the sink, in such a way that any pair of siblings falls in the same connected component. The connected

---

[1]The Frobenius norm of matrix A with elements $[a(k,l)]$ is: $\|A\|_F = \sqrt{\sum_{k,l} a(k,l)^2}$.
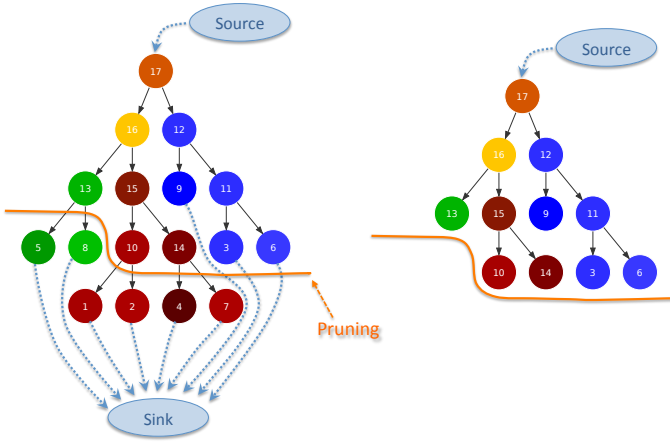
Fig. 1. Illustration of the pruning on a BPT. Left: The root of the BPT is connected to a *source* node and the leaves are connected to a *sink* node. The pruning creates two connected components, one connected to the *source* and one connected to the *sink*, such that all siblings fall into the same connected component. Right: the pruned BPT is the connected component that includes the *source*. Its leaves define the extracted partition.

component that includes the root node is itself a BPT and its leaves define a partition of the space which represents here our final segmentation result. This process is illustrated in Fig. 1. In the sequel, we discuss several examples of pruning for PolSAR image.

IV. SIMPLE STRATEGIES FOR BPT PRUNING

One of the simplest pruning approach consists in following the merging sequence and in stopping the iterative merging process when a predefined number N_R of regions is obtained [2]. Note that this can be viewed as a pruning of the BPT, but actually, there is no need to fully construct the BPT to compute the resulting partition. We will use this strategy as reference to highlight the interest of alternative approaches that actually rely on the complete BPT construction. In any case, this pruning strategy is quite difficult to use in practice as the appropriate number of regions to be extracted for a given image is generally unknown.

The second pruning [2] consists in populating the tree nodes with a feature measuring the region homogeneity through the norm of the difference between the pixel values and the region mean. In order to take into account the multiplicative nature of the speckle noise [11], this difference is normalized by the norm of the region mean. The homogeneity criterion is given by:

$$\phi_R^{Homog.} = \frac{1}{|R|} \sum_{i,j \in R} \|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F / \|\mathbf{Z}_R\|_F \qquad (7)$$

Once the tree has been populated, the feature value of each node is compared to a predefined threshold. Note that the feature value is expected to be rather high for large regions and low for small regions. Indeed, in the extreme case of single pixel regions, $\mathbf{Z}_R$ coincides with $\mathbf{Z}_{ij}^I$ and therefore $\phi_R^{Homog.} = 0$. However, the feature of a parent node is not always larger or equal to the features of its siblings. Theoretically, it is said that this criterion is not increasing.

Several strategies have been proposed in the literature to deal with these cases (see for example [17] and the references herein). To define the pruning here, we have used the so-called *Max* rule [5] which consists in selecting on each branch the closest node to the root for which the homogeneity criterion is below the threshold.

V. BPT PRUNING OPTIMIZING A GLOBAL CRITERION

The previous section has mentioned two simple pruning techniques that do not rely on any optimization strategy. In this section, we are interested in a pruning strategy that extracts partitions optimizing a certain criterion. More precisely, we restrict ourselves to additive criteria, that are criteria defined as:

$$C = \sum_R \phi_R \qquad (8)$$

where R is a set of regions, described as nodes in the BPT, that forms a partition and ϕ_R a measure depending on R.

This type of criterion can be efficiently minimized using a dynamic programming algorithm originally proposed in [1] for optimization of a global criterion on BPT. The solution consists in propagating local decisions in a bottom-up fashion. The BPT leaves are initially assumed to belong to the optimum partition. Then, one checks if it is better to represent the area covered by two sibling nodes as two independent regions $\{R_1, R_2\}$ or as a single region R (the common parent node of R_1 and R_2). The selection of the best choice is done by comparing the criterion ϕ_R evaluated on R with the sum of the criterion values ϕ_{R_1} and ϕ_{R_2}:

$$\text{If } \phi_R \leq \phi_{R_1} + \phi_{R_2} \begin{cases} \text{then} & \text{select } R \\ \text{else} & \text{select } R_1 \text{ and } R_2 \end{cases} \qquad (9)$$

The best choice (either "R" or "R_1 plus R_2") is stored in the node representing R with the corresponding criterion value (ϕ_R or $\phi_{R_1} + \phi_{R_2}$). The procedure is iterated up to the root and defines the best partition. All the criteria we discuss in this section are *h-increasing* [18] and this guarantees that the algorithm finds the global optimum of the criterion on the tree and the selected regions form a partition of the image.

We discuss in this section five pruning criteria for low-level processing and grouping of PolSAR data. The main goal of theses pruning criteria is to segment the images so that a precise estimation of the region contours as well as of the polarimetric parameters can be done.

The first and most obvious pruning criteria relies on the adaptation of Square Error (SE) to the matrix case, here the covariance matrices. It simply consists in computing the matrix norm of the difference between the covariance matrices $\mathbf{Z}_{ij}^I$ of the pixels belonging to a given region R and the covariance matrix presenting the region model $\mathbf{Z_R}$:

$$\phi_R = \sum_{i,j \in R} \|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F \qquad (10)$$

This criterion essentially enforces the homogeneity of regions. However, on its own, it is useless because a partition made of the initial leaves of the BPT will be optimum as this is where the deviation of the individual pixels with respect to the region

mean will be minimized. Following classical approaches in functional optimization, ϕ_R can be interpreted as a data fidelity term and combined with a data regularization term which encourages the optimization to find partitions with a reduced number of regions. As simple data regularization, we use a constant value λ that penalizes the region presence. Therefore, the final homogeneity-based criterion to be minimized is given by $C = \sum_R \phi_R$ with ϕ_R defined as follows:

$$\phi_R^{SE} = \sum_{i,j \in R} \|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F + \lambda \tag{11}$$

This first pruning criterion may be interesting to extract homogeneous regions in terms of the data covariance matrix but it does not take into account the presence of the speckle noise. As in Section IV Eq. 7, the presence of the speckle noise can be taken into account by normalizing the homogeneity measure by the average norm of the region model. This reasoning leads to the second pruning criterion where ϕ_R can be written as:

$$\phi_R^{SAR\_SE} = \sum_{i,j \in R} \|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F / \|\mathbf{Z}_R\|_F + \lambda \tag{12}$$

The third pruning criterion relies on a Wishart similarity measure [19]. It is based on a statistical test that assumes that two matrices, Z_1 and Z_2, follow a Wishart pdf and that one pdf is known. It is not symmetric as it depends on which matrix pdf is assumed to be known. A symmetric version is given by: $tr(\mathbf{Z}_1^{-1}\mathbf{Z}_2) + tr(\mathbf{Z}_2^{-1}\mathbf{Z}_1)$ where $tr(.)$ denotes the trace of a matrix. This expression can be adapted to measure the similarity between pixels and the region model. It would lead to an expression such as: $tr((\mathbf{Z}_{i,j}^I)^{-1}\mathbf{Z}_R) + tr((\mathbf{Z}_R)^{-1}\mathbf{Z}_{i,j})$. However, the matrix inversion at the pixel level is computationally demanding and the matrix may even be singular. Therefore, we use a simplified formulation of this measure by taking into account only the diagonal elements of the matrices.

$$\phi_R^{Wishart} = \sum_{i,j \in R} \sqrt{\sum_{k=1,2,3} \left(\frac{\mathbf{Z}_{ij}^I(k,k)^2 + \mathbf{Z}_R(k,k)^2}{\mathbf{Z}_{ij}^I(k,k)\mathbf{Z}_R(k,k)} \right)} + \lambda \tag{13}$$

where $\mathbf{Z}_{ij}^I(k,k)$ and $\mathbf{Z}_R(k,k)$ respectively represent the diagonal elements of covariance matrices $\mathbf{Z}_{ij}^I$ and $\mathbf{Z}_R$.

The fourth pruning criterion relies on the geodesic distance already used for the BPT construction (Eq. 6). It is given by: $\|log\left(\mathbf{Z}_R^{-1/2}\mathbf{Z}_{i,j}^I\mathbf{Z}_R^{-1/2}\right)\|_F$. As in the Wishart case, since this measure is quite complex to compute, we use a simplified version taking into account only the diagonal elements of the matrices. The fourth pruning criterion is then given by:

$$\phi_R^{Geodesic} = \sum_{i,j \in R} \sqrt{\sum_{k=1,2,3} ln^2\left(\frac{\mathbf{Z}_{ij}^I(k,k)}{\mathbf{Z}_R(k,k)}\right)} + \lambda \tag{14}$$

where ln represents the natural logarithm.

Finally, the last pruning is also based on a pruning minimizing a global criterion but here the idea relies on ratio filters: if the ideal image structure is known (here represented by Z_R), then the ratio of the matrices diagonal values $\mathbf{Z}_{ij}^I(k,k)/\mathbf{Z}_R(k,k)$ should only contain noise of variance 1
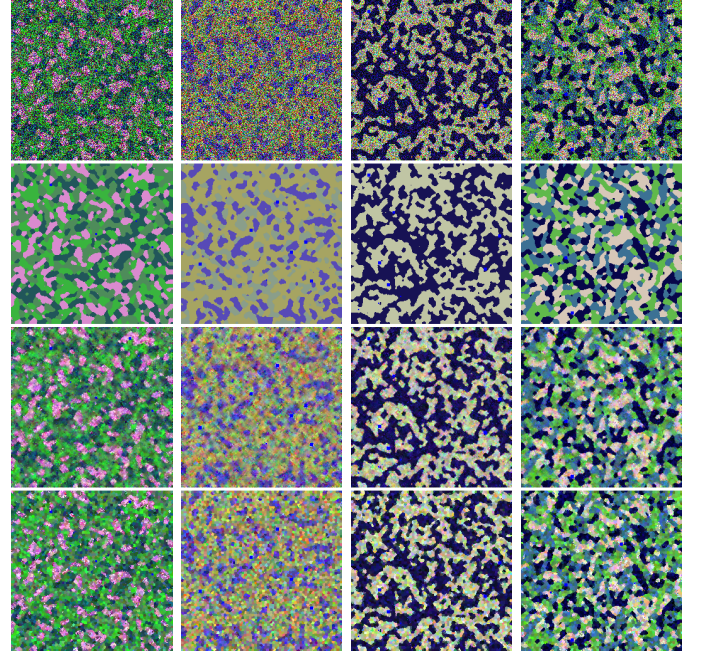

Fig. 2. Examples of original PolSAR images of the dataset (first row), their corresponding ground-truth (second row), denoised images with the σ-Lee filter (third row) and super-pixel partitions (fourth row). RGB Pauli color coding: the polarimetric channels $|S_{hh} - S_{vv}|^2$, $\sqrt{2}|S_{hv}|^2$ and $|S_{hh} + S_{vv}|^2$ are assigned to the RGB channels respectively.

and no structure information. If the structure information is absent, the energy of the ratio should be minimum. This reasoning leads to the following minimization criterion involving as before a data fidelity term and a data regularization term:

$$\phi_R^{Ratio} = \sum_{i,j \in R} \sum_{k=1,2,3} \|\frac{\mathbf{Z}_{ij}^I(k,k)}{\mathbf{Z}_R(k,k)}\| + \lambda \tag{15}$$

VI. EVALUATION

A. Dataset

To objectively measure the performances of the pruning discussed in the previous section, we rely on a dataset of PolSAR images on which the ground-truth polarimetric information is available. More precisely, we use the set of simulated PolSAR images [6] where the underlying ground-truth, i.e. the class regions, is modeled by Markov Random Fields. A set of typical polarimetric responses has been extracted from an AIRSAR image (L-band) so that they represent the 8 classes found in the $H/\overline{\alpha}$ plane and randomly assigned to each class. Point scatterers of sizes between 2x2 and 5x5 pixels are also added at random spatial positions. Then, single look complex images have been generated from the polarimetric responses using a Cholesky decomposition [20] ensuring that the resulting data is Wishart distributed (Eq. 3). Examples of images and their corresponding ground-truth, denoised images (with the σ-Lee filter, window size = 7x7 and $\sigma = 9$) and super-pixel partitions are presented in Fig. 2. The dataset used for evaluation here is composed of ten images, five images with 256×256 pixels and five with 128×128 pixels.
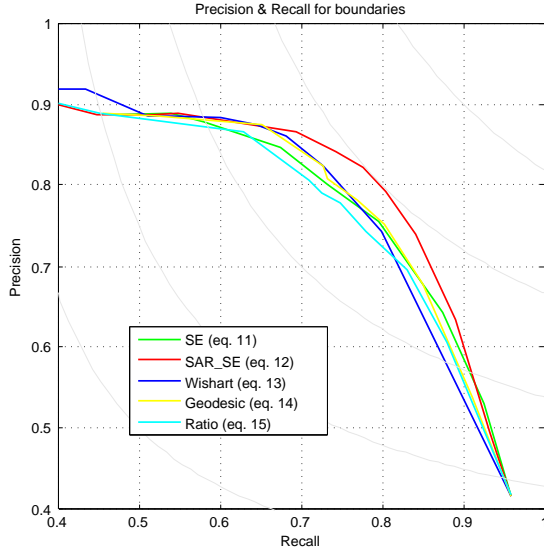
Fig. 3. Precision and Recall performances of the pruning techniques optimizing a global criterion: ϕ_R^{SE}, $\phi_R^{SAR\_SE}$, $\phi_R^{Wishart}$, $\phi_R^{Geodesic}$ and ϕ_R^{Ratio} (average over the entire dataset).
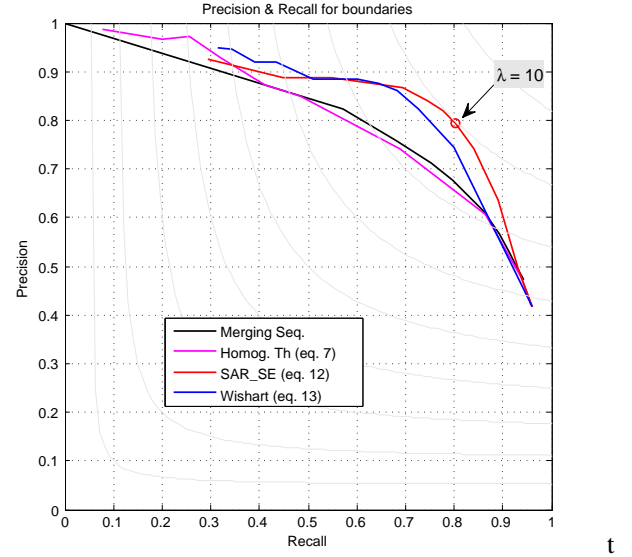

Fig. 4. Comparison of the Precision and Recall performances of pruning techniques involving the optimization of a global criterion (Sec.V): $\phi_R^{SAR\_SE}$ and $\phi_R^{Wishart}$, with pruning techniques not relying on optimization (Sec.IV): merging sequence and thresholding of $\phi_R^{Homog.}$ (average over the entire dataset).

Thanks to this simulated dataset with ground-truth, we can precisely measure the quality of the pruning techniques in the context of segmentation because we know the ideal partition.

B. Evaluation of pruning approaches

Fig. 3 and 4 show the evaluation of the segmentation results as classically done in the supervised case through *Precision and Recall for boundaries* curves [21]. The idea is to compare the boundary pixels of the partitions extracted from the BPT with that of the ground-truth partitions. To this end, an optimal bipartite graph matching is computed on the two sets of boundary pixels. All correspondences where localizations of matched boundary pixels are different by more than a low threshold are declared as unmatched leading to false positives or false negatives. More details on this procedure can be found in [21].

Each partition extracted from the BPT leads to a pair of Precision and Recall values. The Precision and Recall curves are formed by modifying the λ value to get coarser or finer partitions. When λ is small, the pruning remains close to the BPT leaves and the resulting partition is similar to the initial partition. This corresponds to high recall and low precision points in the Precision/Recall plane. When λ is large, the pruning is close to the BPT root and the resulting partition involves a few regions. This corresponds to low recall and high precision points. The ideal system has Precision and Recall values equal to one. Fig. 3 and 4 present the curves averaged over the 10 images of the datasets.

In Fig. 3, all techniques defined through the optimization of a global criterion (Section V) are compared. In order to clearly see the results, the dynamic range of Precision and Recall values is limited to 0.4 and 1. As can be seen, the best results are obtained with $\phi_R^{SAR\_SE}$ (Eq. 12) which allows to get at the same time Precision and Recall values equal to 0.8. Note that this is the only criterion that explicitly deals
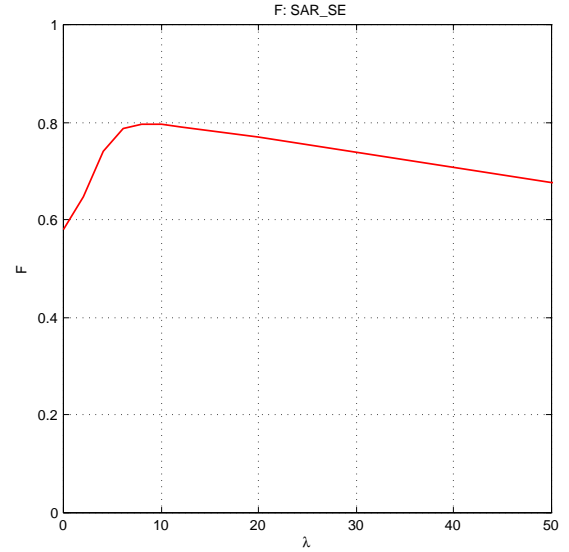
Fig. 5. F value as a function of the pruning parameter λ for the $\phi_R^{SAR\_SE}$ criterion.

with the multiplicative nature of the speckle noise at the level of individual pixel contribution. Dealing explicitly with the multiplicative nature of the speckle noise is particularly relevant in the definition of the pruning criteria as they all involve measuring in some sense the difference between the covariance matrix of the individual pixels Z_{ij}^I with that of the region Z_R. Even if the original image has been pre-filtered, the noise power corrupting the individual pixels is still related to the average norm of the covariance matrix.

Fig. 4 compares two pruning criteria involving the optimization of a global criterion ($\phi_R^{SAR\_SE}$ and $\phi_R^{Wishart}$) with the two simple techniques (not relying on optimization) presented in sec. IV: merging sequence and thresholding of $\phi_R^{Homog.}$
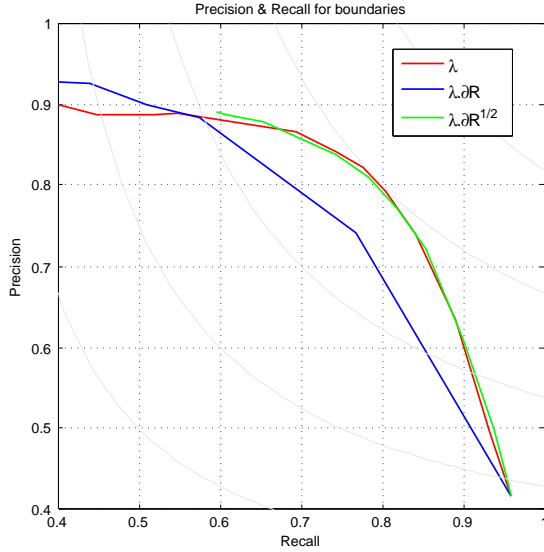
Fig. 6. Influence of the regularization term on the Precision and Recall performances for the $\phi_R^{SAR\_SE}$ pruning technique.

Fig. 7. Examples of partition extracted with the optimization criterion $\phi_R^{SAR\_SE}$ and $\lambda = 10$ for the dataset images of Fig. 2.

(Eq. 7). As can be seen, the pruning techniques based on optimization significantly outperform the remaining ones. The figure also highlights the specific point in the Precision and Recall curves where the best trade-off is obtained with the $\phi_R^{SAR\_SE}$ pruning. We will illustrate partitions corresponding to real images in the sequel with this pruning technique and λ value.

Precision and Recall curves describe the performances for the complete range of pruning parameter (λ) values. However, they do not efficiently describe the system sensitivity to the parameter value. To this end, Fig. 5 presents the F value as a function of the pruning parameter. The F value is classically used to summarize the Precision P and Recall R trade-off. It corresponds to the harmonic mean of P and R: $F = 2PR/(P + R)$. Fig. 5 shows that the pruning technique provides stable results for a rather wide range of λ values. For the best pruning approach $\phi_R^{SAR\_SE}$, this means that, in practice, λ values between 7 and 15 will extract similar partitions and there is no need to fine tune the parameter.

In Section V, the regularization term we have proposed is simply a constant value λ per region. In the context of image segmentation, many functionals inspired from the work of Mumford-Shah [22] use a regularization term that is proportional to the region perimeter ∂R. For the best pruning criterion, the one based on $\phi_R^{SAR\_SE}$, Fig. 6 compares three possible regularization terms: λ, $\lambda \partial R$ and $\lambda \sqrt{\partial R}$. As can be seen, λ and $\lambda \sqrt{\partial R}$ provides very similar results and $\lambda \partial R$ provides better results only for severely under-segmented partitions (low Recall values). As we are generally not interested by severely under-segmented partitions, we will use the simple λ regularization term in the sequel and avoid the need of computing the region perimeter.

Finally, the partitions extracted with $\phi_R^{SAR\_SE}$ and $\lambda = 10$ corresponding to the database image of Fig. 2 are shown in Fig. 7. Each region has been filled by its mean covariance matrix. These results have to be compared with the second
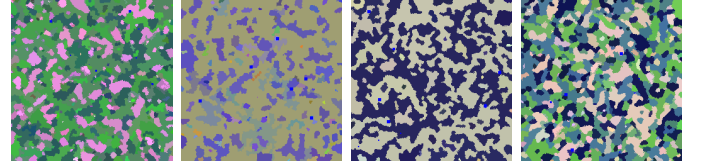
row of Fig. 2 and highlight the ability of the approach to generate useful partition preserving the various region contours and providing a good estimation of the real covariance matrix. Note that point scatterers appear in blue in these images and that they have been preserved and segmented as individual regions by the processing without any special treatment. In fact, these regions appear in the tree very close to the root node (as they are very different from their surrounding) and are preserved by the pruning.

C. Influence of the BPT construction on the pruning performances

1) Region similarity: The last set of experimental results we report here assesses the influence of the BPT construction on the pruning results. The first issue we analyze deals with the similarity criterion used to construct the BPT. In Section III, we mentioned that a geodesic distance was used and weighed by a multiplicative term corresponding to the logarithm of the harmonic mean of the region sizes:

$$d_{Geod}(R_1, R_2) = \|log\left(\mathbf{Z}_{R_1}^{-1/2}\mathbf{Z}_{R_2}\mathbf{Z}_{R_1}^{-1/2}\right)\|_F$$
$$\cdot \, ln\left(\frac{2|R_1||R_2|}{|R_1| + |R_2|}\right) \quad (16)$$

In [5], two other similarity expressions were also proposed. The first one relied on the geodesic distance but the size influence was handled by an additive term.

$$d_{Geod}^{Add}(R_1, R_2) = \|log\left(\mathbf{Z}_{R_1}^{-1/2}\mathbf{Z}_{R_2}\mathbf{Z}_{R_1}^{-1/2}\right)\|_F$$
$$+ \, ln\left(\frac{2|R_1||R_2|}{|R_1| + |R_2|}\right) \quad (17)$$

The second similarity measure was derived from the Wishart distance and the size influence was expressed as a multiplicative term corresponding to the mean of the region sizes:

$$d_{Wish}(R_1, R_2) = \left(tr(\mathbf{Z}_{R_1}^{-1}\mathbf{Z}_{R_2}) + tr(\mathbf{Z}_{R_2}^{-1}\mathbf{Z}_{R_1})\right)$$
$$\cdot \, (|R_1| + |R_2|) \quad (18)$$

These alternatives are difficult to compare through subjective evaluation of a set of resulting partitions. However, thanks to the use of a dataset with ground-truth and the objective assessment methodology involving Precision and Recall curves, a precise comparison can be made. The results are shown on Fig. 8 for the best pruning technique ($\phi_R^{SAR\_SE}$). As can be seen, the geodesic distance combined with a multiplicative size term provides the best results and significantly outperforms the other alternatives. In this figure, we also show the results obtained with a simplification of
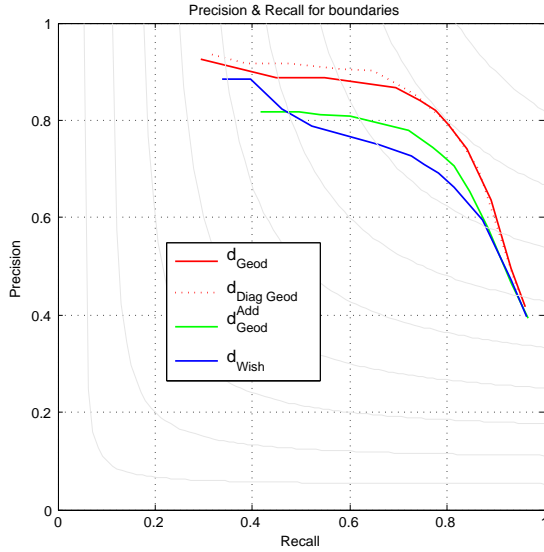
Fig. 8. Influence of the region similarity in the BPT creation on the Precision and Recall performances for the $\phi_R^{SAR\_SE}$ pruning technique.
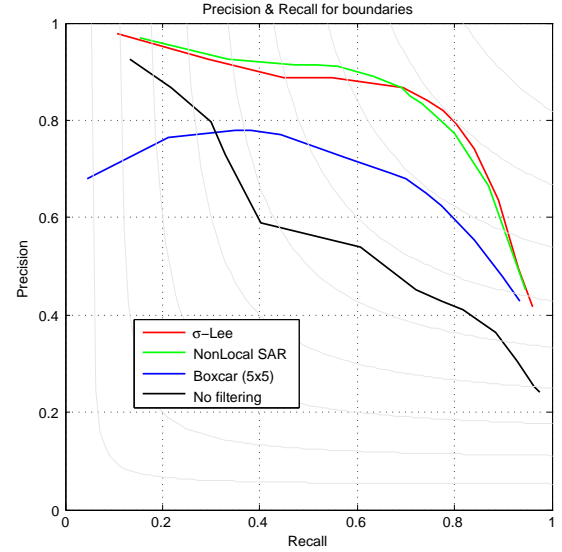

Fig. 9. Influence of the denoising on the Precision and Recall performances for the $\phi_R^{SAR\_SE}$ pruning technique.

the geodesic distance considering only the diagonal terms of the matrices (as done in Section V to simplify the pruning criterion). The resulting criterion is defined by:

$$d_{DiagGeod}(R_1, R_2) = \sqrt{\sum_{k=1,2,3} ln^2\left(\frac{\mathbf{Z}_{R_1}(k,k)}{\mathbf{Z}_{R_2}(k,k)}\right)}$$
$$\cdot\, ln\left(\frac{2|R_1||R_2|}{|R_1| + |R_2|}\right)$$

As can be seen, this criterion is not only simpler to compute, it also gives very similar results compared to the unsimplified geodesic distance. It seems even more robust for the extraction of coarse partitions (obtained for high λ values).

Two main conclusions can be drawn from this section: The first one is the superiority of the Geodesic distance compared to the Wishart distance for the construction of the BPT and the second one concerns the way the region size is used to weight the distance. The superiority of the geodesic distance highlights the importance of considering the covariance matrices in the space of Hermitian definite positive matrices which is not a Euclidean space but a Riemannian manifold. Note that this issue has already been observed and discussed in several references such as [16], [23], [24].

Concerning the way to deal with the region size, it has to be noted that a good merging process has to encourage the merging of small regions so that regions that are close to the root provide a broad view on the data (except if these regions are extremely different from their surrounding such as point scatterers as shown in Fig. 7). The distance defined in Eq. 16 results from the multiplication between two distances, one based on radiometric information and another one dealing with the region sizes. The size distance relies on the harmonic mean of the individual sizes and it is therefore dominated by the minimum of its arguments, which favors the merging of small regions. Furthermore, a possible way to interpret Eq. 16 is to see it as an aggregation in Fuzzy logic [25]. In this context,

the aggregation relies on a conjunctive operator. Consequently, the distance will always favor the merging of small regions. If the aggregation relies on a disjunction as in Eq. 17, this behavior is less clear as the resulting expression computes a compromise between the radiometric and size distances. One additional issue with this disjunction is that the two terms may have different ranges. When combined, one term may dominate the other.

2) Denoising: In Section III, we also stressed the importance of denoising before the computation of the initial partition and the construction of the BPT. Fig. 9 shows results corresponding to four different options: no filtering, 5x5 Boxcar, Non-Local SAR [14] and σ-Lee [13] (window size = 7x7, $\sigma = 9$). As can be observed, the Non-Local SAR and σ-Lee filters are very good options. The σ-Lee filter allows to reach the best compromise between Precision and Recall (0.8 on both values).

Once the initial partition has been computed, one may wonder whether is still important to keep working with the filtered image or if the original image can be used. Indeed, the computation of the region model for the initial partition averages several pixels and can be considered as a denoising step. To answer this question, three different settings are compared in the following experiment:

1) Both the BPT construction and pruning are performed on the filtered image. This is the setting used up to now in this paper.
2) Both the BPT construction and pruning are performed on the original image.
3) The BPT construction is done on the filtered image but the pruning is performed on the original image.

The results are shown in Fig. 10 for the $\phi_R^{SAR\_SE}$ criterion (although not shown here, the same behaviour has been observed for the other criteria: ϕ_R^{SE}, $\phi_R^{Wishart}$ and $\phi_R^{Geodesic}$). As can be seen, processing the unfiltered image leads to worst results. This shows that, although the initial partition
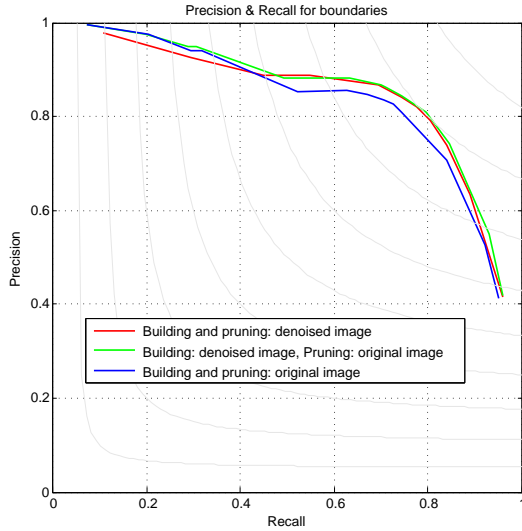
Fig. 10. Influence of the denoising on the Precision and Recall performances for the $\phi_R^{SAR\_SE}$ pruning technique.

is composed of regions involving several pixels, the number of pixels is too small to efficiently deal with the noise involved in the original image. However, once the BPT is constructed, the optimal pruning extracts in average much larger regions and, in this case, averaging the original or the filtered image covariance matrices to define the region model involved in the pruning criterion leads to very similar results. In fact, the option using the original image gives marginally better results.

D. Processing of real images

Finally, the pruning with $\phi_R^{SAR\_SE}$ (Eq. 12) has been used on two real images. The first one is a L-band fully polarimetric data set acquired in 2003 by the Deutsches Zentrum für Luftund Raumfahrt (DLR) ESAR system over the area of the Oberpfaffenhofen airport near Munich, Germany. The images are Single Look Complex with a pixel size of 1.5x1.5m. The second dataset is a C-band fully-polarimetric SAR image collected by the RadarSat-2 system at the fine quad mode over the area of Flevoland, Netherland.

Results, with two λ values: 10 and 30, are shown in Fig. 11 and Fig. 12 together with the original image. They visually highlight the interest of the BPT to perform low-level processing of PolSAR images. The pruning with $\lambda = 30$ provides a coarser partition than the one with $\lambda = 10$. These pruning define many large homogeneous regions on which the estimation of the covariance matrix can be precisely done while preserving the spatial resolution of the content. Note in particular on Fig. 11 that the five corner reflectors at the center of the scene are well preserved for the different values of λ.

E. Local estimation of the covariance matrix

Beside the preservation of the spatial resolution which has already been discussed and assessed in previous sections, the optimum partition extracted from the BPT can be used as a basis for the local estimation of the covariance matrix. Note
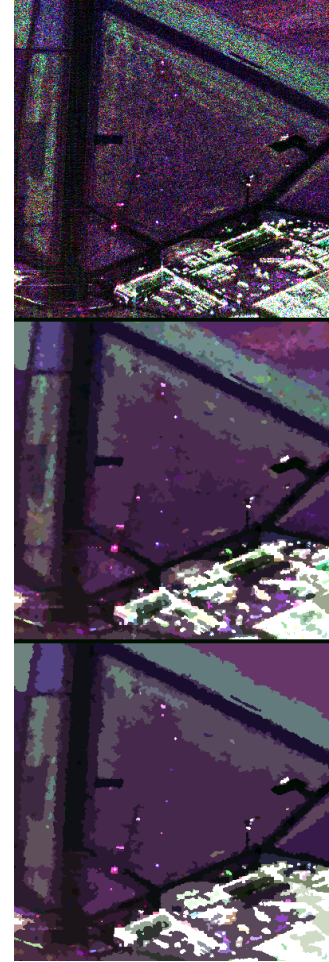

Fig. 11. Results on the ESAR image of Oberpfaffenhofen airport with the $\phi_R^{SAR\_SE}$. Top: Original images (RGB Pauli composition). Middle: Pruning with $\lambda = 10$ resulting in 858 regions. Bottom: Coarser partition obtained with $\lambda = 30$, 325 regions.

that the region model based on the average correlation matrix has been used up to now as a simple mean to represent regions for the computation of distances in the context of BPT creation and pruning. The primary goal of this model was not to create a precise estimation of the covariance matrix at the level of individual pixels. If we consider for example large regions which are extracted on real images (Fig. 11&12), they are indeed homogeneous in particular with respect to their surrounding regions. It is however unrealistic to consider that the covariance matrix is actually constant on such large areas.

Once a partition has been extracted from the BPT, many strategies can be used to estimate the covariance matrix at the level of individual pixels. One can obviously consider the covariance matrices of the pixels belonging to the region but other information related to the shape of the regions, their size, their contours or even information from regions belonging to the same branch in the BPT can be used. This topic deserves a research on its own and will be the basis of a future publication. However, we present here a very simple strategy using the best partition extracted from the BPT to locally estimate the covariance matrix and evaluate its performance with respect to state of the art PolSAR denoising filters. The
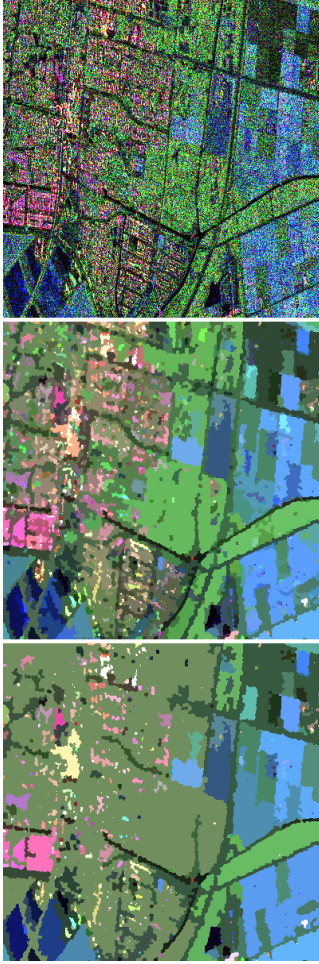
Fig. 12. Results on the RadarSat-2 image of Flevoland area with the $\phi_R^{SAR\_SE}$. Top: Original images (RGB Pauli composition). Middle: Pruning with $\lambda = 10$ resulting in 1372 regions. Bottom: Coarser partition obtained with $\lambda = 30$, 579 regions.
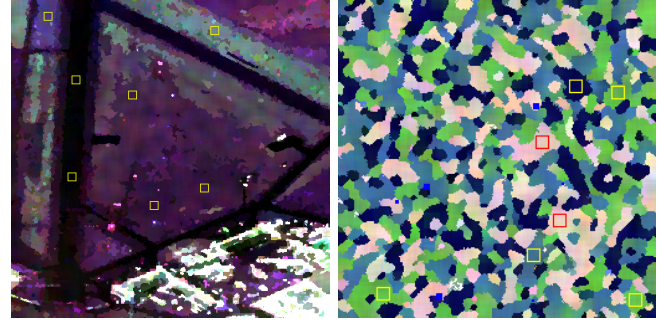

Fig. 13. Examples of local estimation of the covariance matrix on the ESAR image (Left) and on a simulated image (Right). Squares indicate homogeneous areas on which results reported in table I were evaluated.

the σ-Lee, the NonLocal SAR and the BPT approach for three values of the N parameter.

	ESAR image		Simulated image	
	Rel. Bias %	ENL	Rel. Bias %	ENL
σ-Lee [13]	4,46	13,9	5,79	12,3
NonLocal SAR [14]	5,72	65,8	3,89	56,8
BPT ($N = 9$)	4,17	67,5	3,92	114,0
BPT ($N = 11$)	4,31	95,9	3,77	166,9
BPT ($N = 13$)	4,37	127,1	3,73	229,0

TABLE I
ESTIMATION OF THE RELATIVE BIAS AND THE ENL ON SEVEN SQUARES
OF THE IMAGES OF FIG. 13

As previously mentioned, the BPT approach can be extended and probably improved as the approach is very straightforward and preliminary. However, these results already highlight the interest of the BPT approach for locally estimating the covariance matrix of individual pixels, as it provides lower bias for a higher number of ENL compared to state of the art filters.

estimation strategy simply consists of computing for each pixel the average of its neighboring pixels in the original image that actually belong to the same region. In our experiment, the neighborhood was defined as a square of NxN. When pixels are in the center of large regions, the average uses all the pixels included in the NxN squares. But as the pixel gets close to the region border or belongs to a small region, the average uses less pixels and do not take into account information coming from neighboring regions. Examples obtained with this strategy with $N = 13$ are show in Fig. 13.

As classically done (see [6] and the references therein), the relative bias and the Equivalent Number of Look (ENL) are evaluated on homogeneous areas (here seven 11x11 rectangles have been used. They are identified by squares in the images of Fig. 13). The relative bias is defined by $|\hat{\mu} - \mu|/\mu$ where $\hat{\mu}$ is the estimated mean value in the square after the processing by the BPT or by state of the art denoising filters such as the σ-Lee [13] or the NonLocal SAR [14] and μ is the ideal average matrix. This ideal μ value is already known in the case of the simulated image and it is estimated as the mean of the 11x11 square in the original ESAR image. The ENL is defined by: μ^2/σ^2. Table I provides the results comparing

VII. CONCLUSIONS

This paper has discussed the interest of Binary Partition Trees (BPTs) for PolSAR images and highlighted the usefulness of a particular type of graph cut called pruning to extract partitions from the BPT. Several pruning techniques involving the optimization of a global criterion related the region homogeneity have been evaluated. This evaluation has been done objectively through Precision and Recall for boundaries curves on a dataset where the ground-truth is available. The best pruning strategy relies on a normalized version of squared error where the normalization takes into account the specific multiplicative nature of the speckle noise. The influence of decisions regarding the BPT construction, in particular the region similarity measure and the denoising filter has also been evaluated objectively. Finally, the best pruning strategy has been applied to real images showing the interest of the BPT optimal pruning approach to produce precise partitions for PolSAR data preserving the various region contours and allowing good estimation of the covariance matrix to be done.

ACKNOWLEDGMENT

REFERENCES

[1] P. Salembier and L. Garrido, "Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval," *IEEE Trans. on Image Processing*, vol. 9, no. 4, pp. 561–576, 2000.

[2] A. Alonso-Gonzalez, C. Lopez-Martinez, and P. Salembier, "Filtering and segmentation of polarimetric SAR data based on binary partition trees," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 50, no. 2, pp. 593–605, 2012.

[3] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition trees," *IEEE Trans. on Image Processing*, vol. 22, no. 4, pp. 1430 – 1443, 2013.

[4] M. Veganzones, G. Tochon, M. Dalla-Mura, A. Plaza, and J. Chanussot, "Hyperspectral image segmentation using a new spectral unmixing-based binary partition tree representation," *IEEE Trans. on Image Processing*, vol. 23, no. 8, pp. 3574 – 3589, 2014.

[5] A. Alonso-Gonzalez, S. Valero, J. Chanussot, C. Lopez-Martinez, and P. Salembier, "Processing multidimensional SAR and hyperspectral images with binary partition tree," *Proceedings of IEEE*, vol. 101, no. 3, pp. 723–747, 2013.

[6] S. Foucher and C. Lopez-Martinez, "Analysis, evaluation, and comparison of polarimetric SAR speckle filtering techniques," *IEEE Trans. on Image Processing*, vol. 23, no. 4, pp. 1751–1764, 2014.

[7] P. Salembier, S. Foucher, and C. Lopez-Martinez, "Low-level processing of PolSAR images with binary partition trees," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2014*, Quebec, Canada, July 2014.

[8] P. Salembier, "Study of binary partition tree pruning techniques for polarimetric SAR images," in *Proc. of the International Symposium on Mathematical Morphology, ISMM 2015,*, Reykjavik, Iceland, May 2015.

[9] S. Cloude, *Polarisation Applications in Remote Sensing*. Oxford Univ. Press, 2009.

[10] J.-S. Lee and E. Pottier, *Polarimetric Radar Imaging: From Basics to Applications*. Boca Raton, FL: CRC Press, 2009.

[11] S. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 34, no. 2, pp. 498–518, March 1996.

[12] J.-S. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 32, no. 5, pp. 1017–1028, Sept. 1994.

[13] J.-S. Lee, J. . Wen, T. Ainsworth, K. Chen, and A. Chen, "Improved sigma filter for speckle filtering of SAR imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 202 – 213, 2009.

[14] C. Deledalle, F. Tupin, and L. Denis, "Polarimetric SAR estimation based on non-local means," in *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2010*, 2010.

[15] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274 – 2282, 2012.

[16] 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*. LNCS, 2009, vol. 5416, pp. 124–163.

[17] P. Salembier and M. H. F. Wilkinson, "Connected operators: A review of region-based morphological image processing techniques," *IEEE Signal Processing Magazine*, vol. 6, no. 129, pp. 136–157, 2009.

[18] J. Serra, "Hierarchy and optima," in *Discrete Geometry for Computer Imagery*, ser. LNCS 6007, D.-R. et al., Ed. Springer, 2011, pp. 35–46.

[19] P. R. Kersten, J.-S. Lee, and T. L. Ainsworth, "Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 519–527, Mar. 2005.

[20] J.-S. Lee, T. Ainsworth, J. Kelly, and C. López-Martínez, "Evaluation and bias removal of multilook effect on entropy/alpha/anisotropy in polarimetric SAR decomposition," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 46, no. 10, pp. 3039–3051, 2008.

[21] D. Martin, C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.

[22] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989. [Online]. Available: http://dx.doi.org/10.1002/cpa.3160420503

[23] Y.-H. Wang and C.-Z. Han, "PolSAR image segmentation by mean shift clustering in the tensor space," *Acta Automatica Sinica*, vol. 36, no. 6, pp. 798–806, 2010.

[24] P. Formont, J. Ovarlez, and F. Pascal, "On the use of matrix information geometry for polarimetric SAR image classification," in *Matrix Information Geometry*, F. Nielsen and R. Bhatia, Eds. Springer, 2011.

[25] G. Beliakov and J. Warren, "Appropriate choice of aggregation operators in fuzzy decision support systems," *IEEE Transaction on Fuzzy Systems*, vol. 9, no. 6, pp. 773–784, Dec. 2001.

Philippe Salembier (M'96-SM'09-F'11) received an engineering degree from the Ecole Polytechnique, Paris, France, in 1983 and an engineering degree from the Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1985. 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 in Lausanne, Switzerland, to work on image processing and received a PhD in 1991. At the beginning of 1992, after a stay at the Harvard Robotics Laboratory as a Postdoctoral Fellow, 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 research interests include remote sensing image processing, image and video sequence processing, compression and indexing, mathematical morphology, level sets and nonlinear filtering, as well as signal processing tools for genomics. In terms of standardization activities, he has been 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.

Philippe Salembier is a Fellow of the IEEE. He served as a Associate Editor of various journals including the Journal of Visual Communication and Image Representation (Academic Press), Signal Processing (Elsevier), Signal Processing: Image Communication (Elsevier), the Eurasip Journal on Image and Video Processing, the IEEE Transactions on Image Processing, the IEEE Transactions on Circuits and Systems for Video Technology and the IEEE Signal Processing Letters. He was member of the Image and Multidimensional Signal Processing Technical Committee of the IEEE Signal Processing Society between 2000-2006, an AdCom officer of the European Association for Signal Processing (EURASIP) between 1994-1999.

Samuel Foucher received the degree in telecommunications engineering from the Télécom Bretagne (Brest, France) and the Ph.D. degree in radar imaging from Sherbrooke University (Quebec, Canada) and Université de Rennes I (Rennes, France) in 2001. Since 2002, he's a senior researcher and director of the vision and imaging team at the Computer Research Institute of Montreal (Montreal, QC, Canada). His research interests are in image processing, polarimetric SAR image processing, multi-resolution encoding techniques (wavelets), and content-based video technologies.