Remote Sensing Image Processing with Graph Cut of Binary Partition Trees

Philippe Salembier[1] and Samuel Foucher[2]

[1] Technical University of Catalonia, Barcelona, Spain
[2] Computer Research Institute of Montreal, Vision Team, Montreal, Canada
Invited Paper

Abstract. This paper discusses the interest of hierarchical region-based representations of images such as Binary Partition Trees (BPTs) and the usefulness of graph cut to process them. BPTs can be considered as an initial abstraction from the signal in which raw pixels are grouped by similarity to form regions, which are hierarchically structured by inclusion in a tree. They provide multiple resolutions of description and easy access to subsets of regions. Their construction is often based on an iterative region-merging algorithm. Once constructed, BPTs can be used for many applications including filtering, segmentation, classification and object detection. Many processing strategies consist in populating the tree with features of interest for the application and in applying a specific graph cut called pruning. Different graph cut approaches are discussed and analyzed in the context of Polarimetric Synthetic Aperture Radar (PolSAR) images.

Keywords: Binary Partition Tree, graph cut, pruning, PolSAR images, super-pixel partition.

1 Introduction

Remote sensing technologies are currently undergoing an important evolution in terms of the quality and the quantity of information that is acquired. Sensors are able to capture data at an increasing resolution both in terms of spatial and spectral resolutions. This wealth of information generates real challenges with respect to signal processing tasks. One of the major issues is concerned with the handling of the signal correlation.

The traditional pixel-based image representation is not the most appropriate one to deal with the huge amount of information produced by high resolution remote sensing sensors while being able to deal with the signal correlation. A more appropriate representation should somehow group pixels of similar properties into elementary entities that should be easily handled and accessed. Moreover, the representation should be useful for different applications and therefore describe the data at multiple resolutions.

Recently, the interest of Binary Partition Trees (BPTs) [9] has been investigated for remote sensing including SAR [2] and hyperspectral images [10].

Mireya García-Vázquez, Grigori Sidorov, Sunil Kumar (Eds.).
Advances in computing science, control and communications.
Research in Computing Science 69, 2014, pp. 185- 196

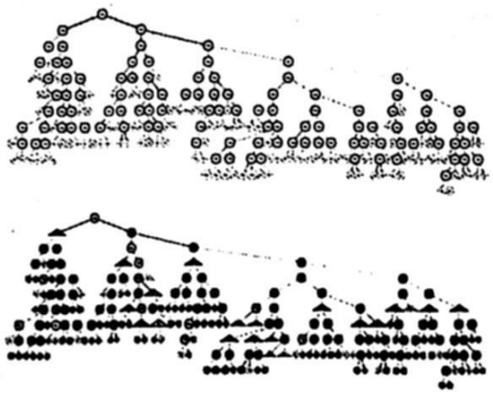

Fig. 1. Example of BPT (Top) and of pruning (Bottom).

BPTs are region-based representations in which pixels are grouped by similarity. They provide multiple resolutions of description and easy access to subsets of regions. 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 tree structure. Once constructed, BPTs can be used for a large number of tasks including image filtering, object detection or classification [3].

The processing of BPTs often relies on a specific type of graph cut called *pruning*. In this paper, we will formally define this notion and show how it can be used to extract partitions from the tree. Then, several examples of useful pruning are presented and analyzed in the context of PolSAR images.

The paper is organized as follows: Sec. 2 discusses the principles of BPT creation and their processing by means of graph cut. A possible way to evaluate the quality of a BPT is presented in Sec. 3 and used in Sec. 4 to study the influence of the initial partition on the BPT construction. Sec. 5 presents and analyzes four pruning techniques for low-level processing of PolSAR images. Finally, conclusions are reported in Sec. 6.

2 Binary partition Tree creation and processing through graph cut

The BPT creation starts by the definition of an initial partition. The initial partition can be composed of individual pixels as in [2, 3]. While this strategy guarantees a high precision as starting point of the merging process, it also implies high computational and memory costs as a huge number of regions have to be handled. As an alternative, the initial partition may correspond to an oversegmentation as a superpixel partition [1]. Once the initial partition is defined, the BPT construction is done by iteratively merging the pair of most similar neighboring regions.

In the PolSAR case, the information carried by pixels of an image I corresponds to the covariance matrix $\mathbf{Z}_{ij}^{I}$ of the scattering vector: $\mathbf{k} = [S_{hh}, \sqrt{2}S_{hv}, S_{vv}]^T$

measured on the resolution cell at location (i, j). The subindices h and v indicate the horizontal and vertical polarization states and $S_{pq\in\{h,v\}}$ represents the complex SAR data where the polarization states employed in reception and transmission are given by p and q respectively. To construct the BPT, regions R can be modeled as in [2] 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. The distance between neighboring regions defining the merging order can be measured as in [3] by the geodesic similarity adapted to the cone of positive definite Hermitian matrices [4]:

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

where $log(.)$ is the matrix logarithm and $ln(.)$ the natural logarithm. Using this similarity measure, regions are iteratively merged until a unique region representing the entire image support is obtained. After each merging, a new region is created, its mean covariance matrix is computed and its similarity with its neighbors is updated.

An example of BPT can be seen on the top part of Fig. 1. The regions belonging to the initial partition form the leaves of the BPT (shown with green circles in Fig. 1). In this example, only a small image portion has been used. The initial partition involves 108 regions and the BPT is therefore composed of 215 nodes. During the merging process, the BPT is constructed by creating a parent node for each pair of merged regions. In Fig. 1, the edge color represents the similarity value between the two merged regions: blue (red) indicates very similar (dissimilar) regions.

Once the BPT has been constructed, it can be used for a wide range of applications including filtering, segmentation or classification. A large number of applications relies on the extraction of a partition from the BPT. This process can be seen as a particular graph cut called *pruning*: Assume the tree root is connected to a *source* node and that all the tree leaves are connected to a *sink* node. 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 component that includes the root node is itself a BPT and its leaves define a partition of the space. In the sequel, we discuss several examples of pruning in the context of PolSAR images. The first one will allow us to evaluate the quality of a BPT itself.

3 Graph cut to evaluate of the quality of a BPT

One of the major issues in PolSAR image is the speckle noise that results from the coherent integration of the scattered electromagnetic waves. Speckle filtering aims at reducing noise within homogeneous extended targets while preserving meaningful spatial details [5]. Insufficient noise reduction leads to important bias on derived polarimetric parameters therefore degrading the performance of

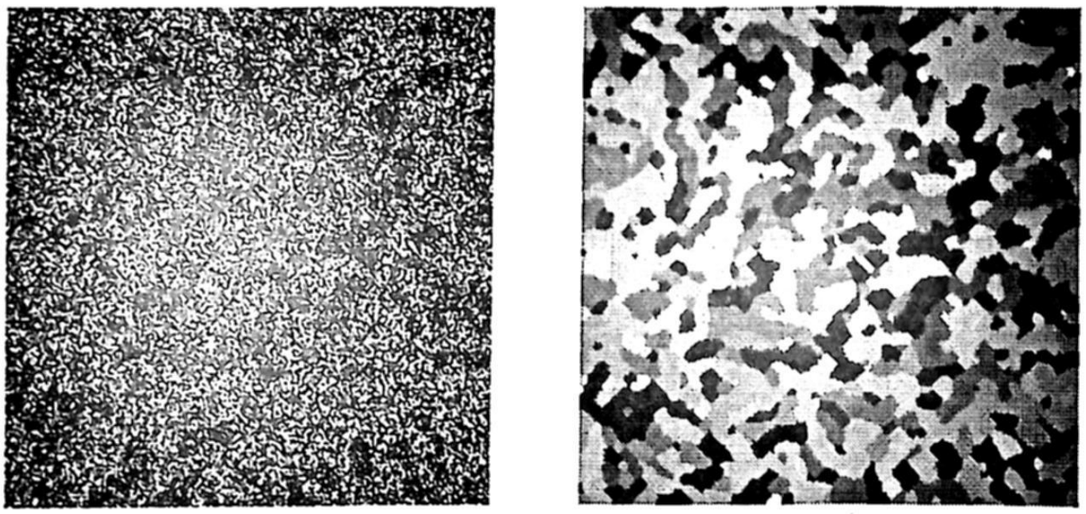

Fig. 2. Example of original PolSAR image (Left) and its corresponding ground-truth (Right). RGB-pauli color coding.

classification tasks or biophysical parameter retrieval. Most filters are based on adaptive strategies using a sliding square window with a fixed size [8, 7].

In this context of PolSAR images and speckle noise, we would like to be able to evaluate the quality of a BPT and see the influence on the constructed BPT of specific choices concerning the region model, the similarity function, the initial partition, etc. This is not a trivial task since a huge number of partitions can be extracted from a given BPT. However, here, we are concerned by low-level processing and by removing the speckle noise as much as possible to allow a precise estimation of the polarimetric parameters. As a result, we are going to 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 [5] where the underlying ground-truth, i.e. the class regions, is modeled by Markov Random Fields. A set of typical polarimetric responses have been extracted from an AIRSAR image (L-band) so that they represent the 8 classes found in the H/α plane and randomly assigned to each class. Finally, single look complex images have been generated from the polarimetric responses using a Cholesky decomposition [6]. An example of image and its corresponding ground-truth is presented in Fig. 2. Thanks to this dataset involving ground-truth information, we can measure the quality of a BPT.

Let us define *the quality of a BPT as the quality of the best image, according to a given error measure E, that can be extracted from it.* Extracting an image from the BPT consists in selecting a set of nodes forming a partition of the image and in assigning the mean covariance matrix of the region to its pixels. So the question is to identify the *ideal* partition that can be extracted from the tree.

The error measure between an image $I(i,j)$ and the ground-truth image $I_{GT}(i,j)$ we use is defined by [2]:

$$E(I, I_{GT}) = \frac{1}{N} \sum_{i,j} \frac{\|\mathbf{Z}_{ij}^{I} - \mathbf{Z}_{ij}^{I_{GT}}\|_F}{\|\mathbf{Z}_{ij}^{I_{GT}}\|_F} \tag{2}$$

where N is the image number of pixels, $\mathbf{Z}_{ij}^{I}$ ($\mathbf{Z}_{ij}^{I_{GT}}$) is the pixel value of image I (I_{GT}) at location (i,j) and $\|.\|_F$ represents the Frobenius matrix norm. This measure is based on the average inverse signal to noise ratio.

As previously mentioned, the extraction of a partition from the BPT is defined by a pruning. To define the *ideal* pruning, let us use the following ideal criterion derived from Eq. 2 by noting that all pixels belonging to the same region R have the same covariance matrix Z_R:

$$C_{\text{ideal}} = \sum_R \phi_R \text{ with } \phi_R = \sum_{i,j \in R} \frac{\|\mathbf{Z}_R - \mathbf{Z}_{ij}^{I_{GT}}\|_F}{\|\mathbf{Z}_{ij}^{I_{GT}}\|_F}, \text{ s.t. } \{R\} \text{ is a partition} \quad (3)$$

This criterion is ideal because it uses the ground-truth image $Z^{I_{GT}}$ which will be unknown in practice. However, it is very useful to quantify the quality of a BPT and to define an upperbound on the performances of all possible pruning strategies.

This criterion can be efficiently minimized using an dynamic programing algorithm originally proposed in [9] for global optimization. The solution consists in propagating local decisions in a bottom-up fashion. To initialize the process, the leaves of the BPT are 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 costs ϕ_{R_1} and ϕ_{R_2}:

$$\text{If } \phi_R \leq \phi_{R_1} + \phi_{R_2} \begin{cases} \text{then } R \text{ belongs to the optimum partition} \\ \text{else } R_1 \text{ and } R_2 \text{ belong to the optimum partition} \end{cases} \quad (4)$$

The best choice (select either "R" or "R_1 plus R_2") is stored in the BPT node representing R together with the corresponding cost value (either ϕ_R or $\phi_{R_1} + \phi_{R_2}$). The procedure is iterated up to the root node and defines the best partition.

This algorithm guarantees to find the global optimum of the criterion on the tree because the criterion is additive with respect to regions. The bottom part of Fig. 1 shows the BPT after the nodes have been populated with the value $\phi_R/|R|$ (as for edges, low (high) values are represented in blue (red)). The triangle-shaped nodes show where the pruning has been done. They form the leaves of the pruned BPT and equivalently the regions of the extracted partition. Then, each region is represented by its mean covariance matrix to create the filtered image. Finally, this image is used to compute the BPT quality with Eq. 2. Based on this strategy to evaluate the BPT quality, we can now investigate the impact of specific choices related to the BPT construction. As an example, we analyze the influence of the initial partitions in the following section.

Initial Partition	Ideal pruning C_{ideal} (dB)	Number of regions of the initial partition	Acceleration factor w.r. to the pixel partition
Pixel	-16,12	65.536	1
Watershed	-13,53	7.720	46
SLIC (size=2)	-15,94	15.946	15
SLIC (size=3)	-15,38	7.257	56
SLIC (size=4)	-14,51	4.143	121

Table 1. Influence of the initial partition in terms of BPT quality and computational load. Results have been averaged over the entire dataset.

4 Superpixel initial partition and its influence on the BPT quality

As previously mentioned, the initial partition used in [2, 3] was composed of individual pixels. The main drawback of this choice is the high number of initial regions and the corresponding cost in memory usage and computational load. To see whether the number of initial regions can be drastically reduced without loosing too much in terms of quality, experiments with superpixel partitions have been conducted. Concretely, initial partitions have been generated either with a watershed applied on the vectorial image gradient and with the SLIC algorithm [1]. In both cases, only the diagonal elements of the covariance matrices have been used to generate the superpixel.

Thanks to the strategy presented in Sec. 3, we can compare the influence of various initial partitions on the BPT construction by extracting the ideal partition and measuring $E(I, I_{GT})$. The results are given in Table. 1. As can be seen, the best BPT is obtained with the pixel partition. However, the use of the SLIC superpixels almost preserves the BPT quality but can drastically reduce the complexity. It is therefore a very good alternative and, in the sequel, we use the SLIC (size 2) superpixels as initial partition.

5 Low level processing through BPT pruning

This section discusses pruning techniques for low-level processing and grouping of PolSAR data. The main goal is to allow for a precise estimation of the polarimetric parameters. In the previous section, we have used an ideal pruning technique to assess the quality of the BPT. It defines the upperbound on the quality of the partitions that can be extracted from the BPT but it cannot be used in practice as it relies on the ground-truth data.

Here we study the interest of four pruning techniques useful in practical situations. The two first ones were proposed in [2]. The first one simply consists in following the merging sequence and in stoping the iterative merging process when a predefined number N_R of regions is obtained. 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.

The second pruning [2] consists in populating the tree nodes with a feature measuring the region homogeneity (difference between the pixel values and the region mean) given by:

$$\phi_R = \frac{1}{|R|} \sum_{i,j \in R} \frac{\|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F}{\|\mathbf{Z}_R\|_F} \tag{5}$$

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. In the extreme case of regions made of a single pixel, $\mathbf{Z}^R$ coincides with $\mathbf{Z}_{ij}^I$ and therefore $\phi_R = 0$. However, the feature of a parent node is not always larger or equal to the features of its siblings. To define the pruning, we have used the so-called *Max* rule [3] which consists in selecting on each branch the closest node to the root for which the homogeneity criterion is below the threshold.

We introduce now two new pruning strategies based on the minimization of a global criterion as in Sec. 3. The initial idea is to use $C = \sum_R \phi_R$ with ϕ_R being the homogeneity criterion $\phi_R = \sum_{i,j \in R} \|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F / \|\mathbf{Z}_R\|_F$. Note that this criterion is the same as the one defined by Eq. 5 without the averaging parameter $|R|$. However, on its own, it is not useful because a partition made of individual pixels sets the criterion to 0. 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 term, we use here a constant value λ that penalizes the region presence. Therefore the final homogeneity-based criterion to be minimized is:

$$C_{\text{Homog.}} = \sum_R \phi_R \text{ with } \phi_R = \sum_{i,j \in R} \frac{\|\mathbf{Z}_{ij}^I - \mathbf{Z}_R\|_F}{\|\mathbf{Z}_R\|_F} + \lambda, \text{ s.t. } \{R\} \text{ is a partition} \tag{6}$$

Finally, the last pruning is also based on a graph cut 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 $\mathbf{Z}_{ij}^I / \mathbf{Z}_R$ should only contain noise of variance 1 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:

$$C_{\text{Ratio}} = \sum_R \phi_R \text{ with } \phi_R = \sum_{i,j \in R} \|\frac{\mathbf{Z}_{ij}^I}{\mathbf{Z}_R}\|_F + \lambda, \text{ s.t. } \{R\} \text{ is a partition} \tag{7}$$

Fig. 3 shows, for the image of Fig. 2, the evolution of the error measure $E(I_{Processed}, I_{GT})$ as a function of the pruning parameters: the number of regions for the first pruning, the threshold on the homogeneity value for the second pruning and the λ value for the remaining pruning. In terms of global error measure, we can see that the best pruning techniques are the two based on

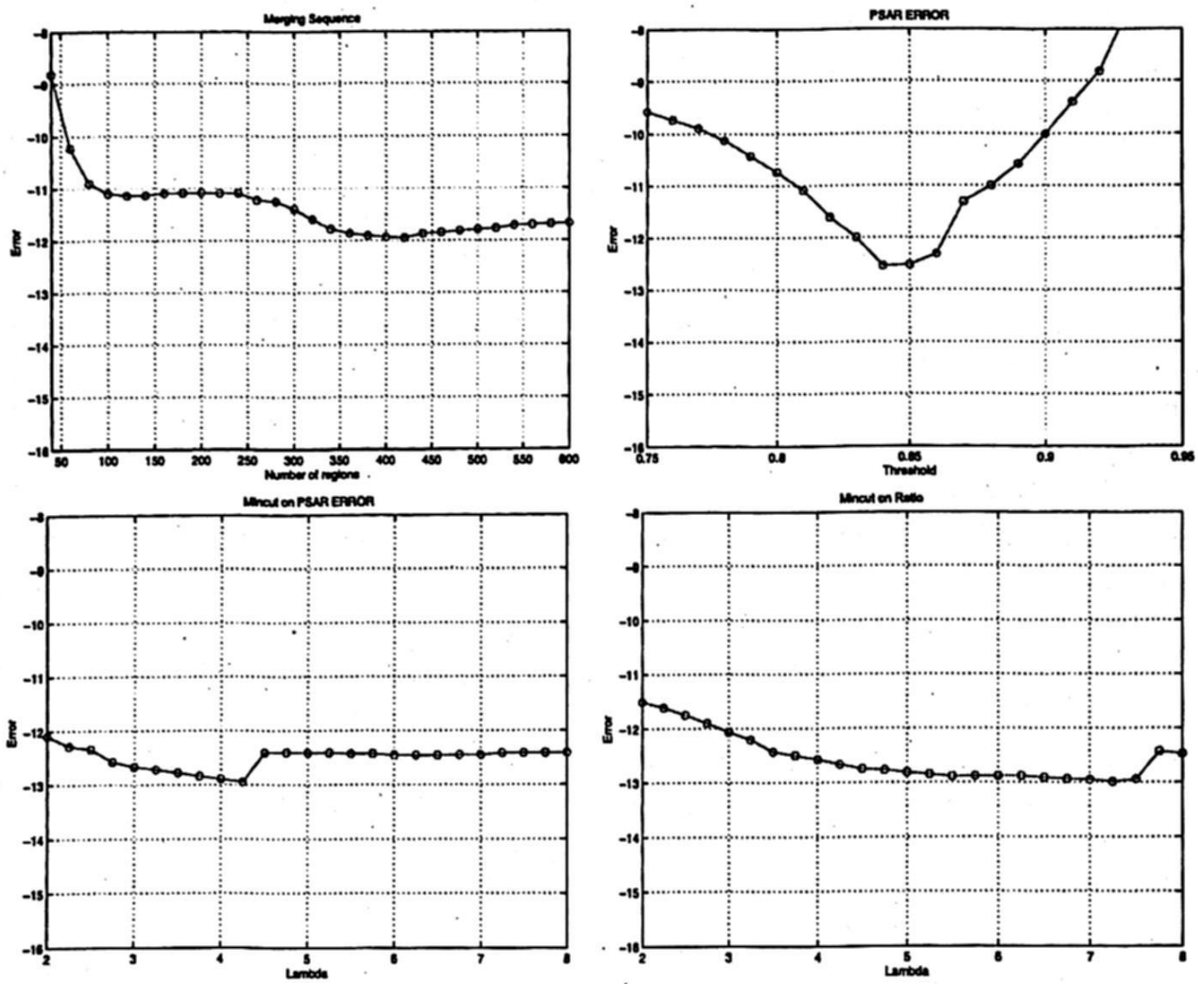

Fig. 3. Error measure $E(I_{Processed}, I_{GT})$ in dB as a function of the pruning parameters. Top row: Merging sequence and threshold on homogeneity (Eq. 5). Bottom row: mincut on homogeneity criterion (Eq. 6) and on ratio image (Eq. 7).

global optimization (Eq. 6 and Eq. 7). They provide a lower value of the error and moreover their dependency on the λ value is smooth. The pruning following the merging sequence does not lead to the best estimation of the polarimetric parameters. This result highlights the interest of constructing the BPT to extract partitions that have not been observed during the merging process. Moreover, in practice, it is difficult to define a priori the appropriate number of regions. Finally, the pruning involving the thresholding on the homogeneity criterion provides intermediate results but we may also note that the value of the threshold has a very strong impact on the results.

Processing technique	Original image	Boxcar	Refined Lee	BPT: homog. mincut	BPT: ratio mincut
$E(I_{Processed}, I_{GT})$ in dB	-1,87	-9,11	-12,17	-14.57	-14,43

Table 2. Results of low-level processing (average over the entire dataset).

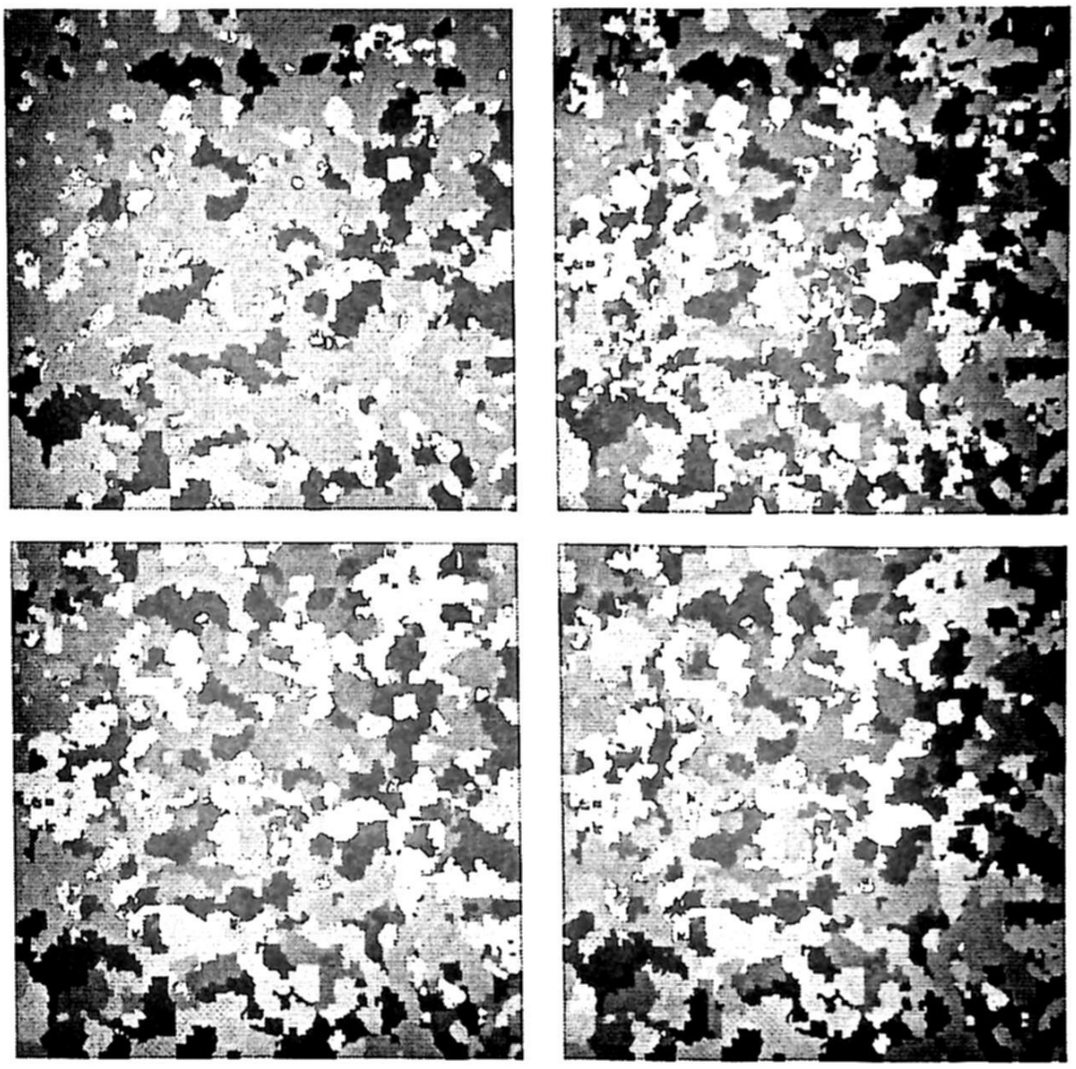

Fig. 4. Processed images with the optimum pruning parameters. Top row: Merging sequence and threshold on homogeneity (Eq. 5). Bottom row: mincut on homogeneity criterion (Eq. 6) and on ratio image (Eq. 7).

Fig. 4 shows the images resulting from the four pruning techniques using the optimum parameter in each case. The visual inspection of these images corroborates the analysis done on Fig. 3. The pruning following the merging sequence merges too many regions represented in dark and light green in the ground-truth image (Fig. 2). The image corresponding to the thresholding on the homogeneity criterion is still noisier that the one involving the global optimization.

The results obtained with best pruning techniques are given in Table 2 as well as results for classical filtering strategies. These results have been obtained by averaging $E(I_{Processed}, I_{GT})$ over the entire database of 10 images. The table shows the interest of the pruning techniques involving global optimization. It also allows us to quantify the gap between these pruning techniques reaching about -14.5dB and the ideal pruning corresponding to almost -16dB (see Table 1). It can be concluded that there is still some room from improvement of the pruning.

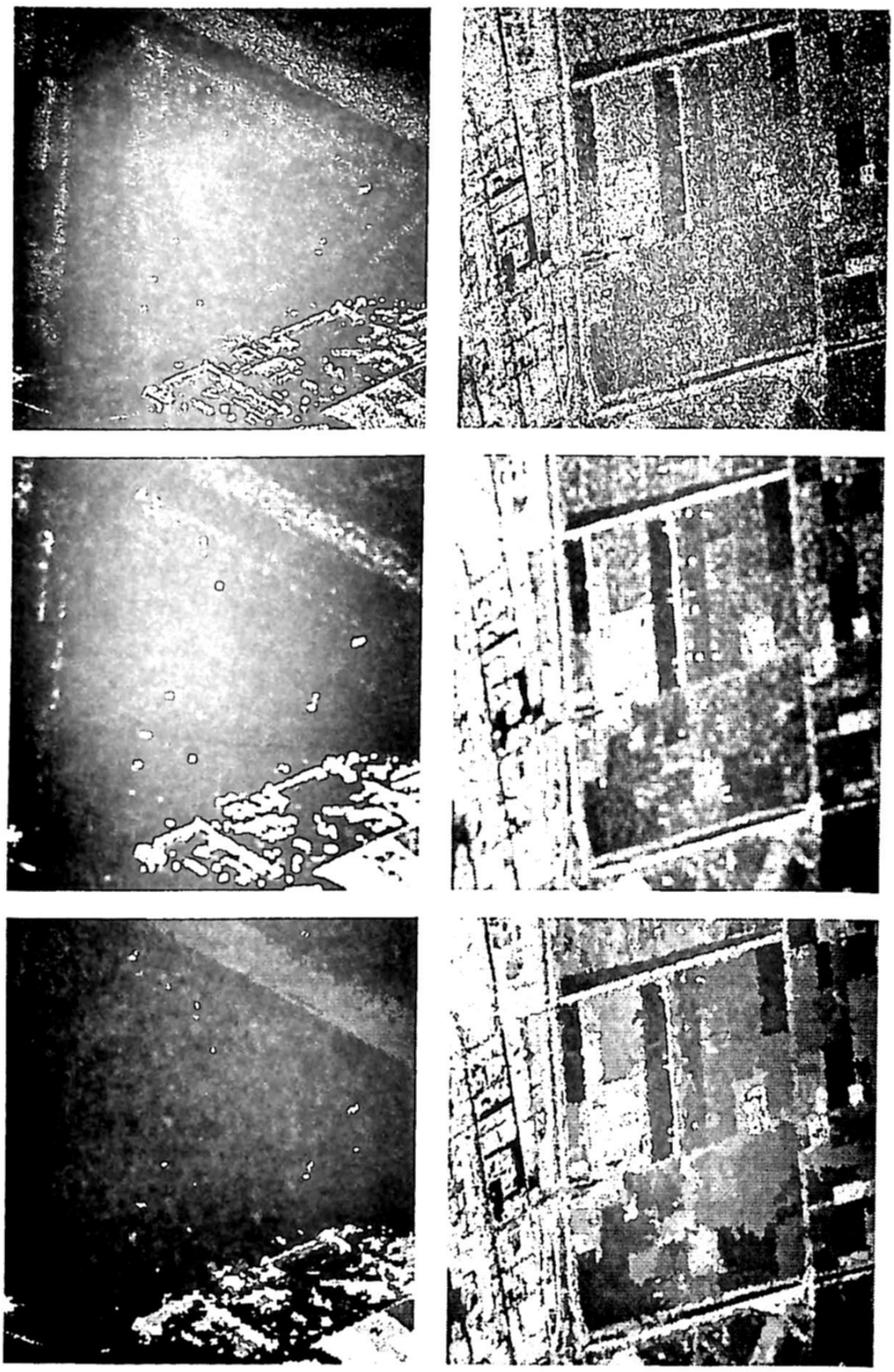

Fig. 5. Results on real images. Top row: Original images (RGB Pauli composition). Middle row: Multilook filtering, Bottom row: Mincut on homogeneity criterion (Eq. 6).

Finally, results on applying the pruning with global optimization of the homogeneity (Eq. 6) are shown in Fig. 5 together with the original image and the result of the classical multilook filter. These results visually highlight the interest

of the BPT to perform a low-level processing of PolSAR images while preserving the spatial resolution of the content.

6 Conclusions

This paper has discussed the interest of Binary Partition Trees (BPTs) for remote sensing applications such as PolSAR. These hierarchical region-based representations of images are useful for many tasks. Here, we have mainly focussed on low-level processing of PolSAR covariance matrices. The paper has highlighted the usefulness of a specific type of graph cut called pruning that extracts partitions from the BPT. Specific pruning techniques have been defined to evaluate the quality of BPT and to perform low-level grouping allowing a precise estimation of the polarimetric information to be done without losing in terms of spatial resolution. In this context, the pruning techniques resulting from the global optimization of a criterion minimizing the region homogeneity or the energy of the ratio image have proved to be very efficient and robust.

References

1. R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274 – 2282, May 2012.
2. A. Alonso-Gonzalez, C. Lopez-Martinez, and P. Salembier. Filtering and segmentation of polarimetric SAR data based on binary partition trees. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):593–605, 2012.
3. 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 the IEEE*, 101(3):723–747, 2013.
4. 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*, volume 5416, pages 124–163. Lecture Notes in Computer Science, F. Nielsen, Ed., Springer Berlin / Heidelberg, 2009.
5. S. Foucher and C. López-Martínez. An evaluation of PolSAR speckle filters. In *IEEE International Geoscience and Remote Sensing Symposium, IGARSS*, 2009.
6. J.-L. Lee, T.L. Ainsworth, J.P. Kelly, and C. López-Martínez. Evaluation and bias removal of multilook effect on entropy/alpha/anisotropy in polarimetric SAR decomposition. *IEEE Trans. Geoscience and Remote Sensing*, 46(10):3039–3051, 2008.
7. J.S. Lee, M.R. Grunes, and G. De Grandi. Polarimetric SAR speckle filtering and its implication for classification. *IEEE Trans. Geoscience and Remote Sensing*, 37(5):2363–2373, 1999.
8. J.S. Lee, M.R. Grunes, and S.A. Mango. Speckle reduction in multipolarization and multifrequency SAR imagery. *IEEE Trans. Geoscience and Remote Sensing*, 29(4):535–544, 1991.

9. P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. on Image Processing*, 9(4):561 –576, apr 2000.
10. S. Valero, P. Salembier, and J. Chanussot. Hyperspectral image representation and processing with binary partition trees. *IEEE Transactions on Image Processing*, 22(4):1430 – 1443, 2013.