OBJECT RECOGNITION IN URBAN HYPERSPECTRAL IMAGES USING BINARY PARTITION TREE REPRESENTATION

Silvia Valero[1,2,3], *Philippe Salembier*[2] *and Jocelyn Chanussot* [3,4]

[1]CESBIO - CNES, CNRS (UMR 5126), IRD, Universite de Toulouse, France
[2]Technical University of Catalonia (UPC), Barcelona, Catalonia, Spain
[3]GIPSA-lab, Signal & Image Dept., Grenoble Institute of Technology, Grenoble, France
[4]Faculty of Electrical and Computer Engineering, University of Iceland

ABSTRACT

In this work, an image representation based on Binary Partition Tree is proposed for object detection in hyperspectral images. The BPT representation defines a search space for constructing a robust object identification scheme. Spatial and spectral information are integrated in order to analyze hyperspectral images with a region-based perspective. Experimental results demonstrate the good performances of this BPT-based approach.

***Index Terms*—** Object detection, Region-based image analysis, BPT, hyperspectral

1. INTRODUCTION

Automatic object recognition to map urban areas is gaining increasing interest. In the hyperspectral literature, object detection techniques have been mainly developed in the context of pixel-wise spectral classification. The drawbacks of pixel-wise analysis is well-known in remote sensing [5]. Because of the pixel-based model limitations, research on region-based object detection algorithms has recently received much attention. In this context, the ECognition software [7] was developed. It relies on hierarchical segmentation and produces an image partition on which various region descriptors can be computed. These descriptors are then used as region features for the recognition of objets in the image. One of the main limitations of this strategy is that it assumes that the best partition corresponds to one level of the previously computed hierarchical segmentation. Unfortunately, this assumption is rarely true and, very often, coherent objects can be found at different levels of the hierarchy [8][2].

Ideally, a robust strategy should study the features in the complete hierarchy to detect the best regions representing the object. Instead of using a classical hierarchical segmentation approach which produces a single partition, a solution to address the need of multiscale analysis relies on image representations based on regions trees. These representations are useful because beside allowing the study of internal region properties (color, texture, shape, etc.), they also permit the study of external relations such as adjacency, inclusion, similarity of properties, etc. Furthermore, a tree is essentially a hierarchical structure and therefore supports multiscale analysis of regions. The multiscale nature of trees provides flexibility to situations where a given image has to be studied at different scales depending on the processing purpose.

The work presented here proposes to initially generate a hierarchical region-based representation of the image and, then, to use this representation as search space for the object detection (therefore avoiding the creation of a partition on which objects are searched as in [1, 7]). A Binary Partition Tree [2] (BPT) is used as hyperspectral image representation.

For object detection, the use of BPT has been introduced in [10] where a simple top-down analysis of the tree branches was done. During this analysis, the objects were detected by selecting the largest nodes having the appropriate features. However, the best region representing the object is not always the largest one with the appropriate features. Here, we present a more robust strategy that studies all BPT nodes to detect the best ones representing the sought object. The paper organization is as follows: Section 2 introduces the BPT and its construction. The BPT analysis for object detection is discussed in Section 3. Experimental results are reported in Section 4. Finally, conclusions are drawn in Section 5.

2. BPT CONSTRUCTION

This BPT is a structured representation of a set of hierarchical partitions which is usually obtained through a bottom-up region merging algorithm. Starting from individual pixels or any other initial partition, the tree is constructed by an iterative process in which regions are iteratively merged. Each iteration requires three different tasks: 1) the pair of most similar neighboring regions is merged, 2) a new region containing the union of the merged regions is formed, 3) the algorithm updates the distance between the newly created region with its neighboring regions. Fig.1 shows an example of BPT construction created from an initial partition.

Region merging algorithms are specified by: 1) a merging cri-

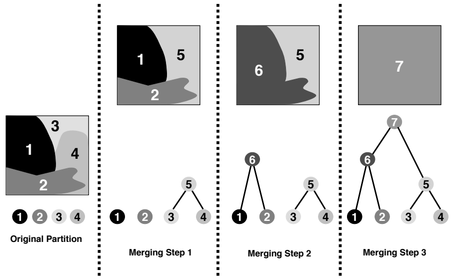

Fig. 1: Example of BPT image representation

terion, defining the similarity between pair of neighboring regions; and 2) a region model that determines how to represent a region and the union of two regions. Working with hyperspectral data, the definition of a region model and a merging criterion has been previously studied in [9]. Following this work, the BPT construction has been constructed here following the strategy presented in [10].

3. OBJECT DETECTION STRATEGY

As instantiations of the object of interest, O, may have many different visual appearances, the detection relies on a set of features, Ω_F, characterizing O. Based on these features, the likelihood of each BPT node $P(O|R_i)$ to be an instantiation of O is assessed and assigned to the node. Once the BPT has been populated with these likelihood, a search is performed to detect the most probable instantiations of the object of interest.

3.1. Populating of BPT

For each node R_i, the likelihood $P(O|R_i)$ is computed by using a set of spectral and spatial features $\Omega_F = \{F_1, F_2, ..., F_K\}$. Based on these features, the likelihood of each node to be an instantiation of O can be estimated by the Bayes rule. The a priori probability of the object $P(O)$ is being equally probable to observed (uniformed prior) and the probability of the evidence $P(\Omega_F)$ can be viewed as a normalizing constant. Thus, considering independent the K local features computed at each R_i, the $P(O|R_i)$ can be defined by

$$P(O|R_i) \simeq P(\Omega_F|O) \simeq \prod_{n=1}^{K} P(F_n|O) \quad (1)$$

The specific choice of features depends on the reference object. Here, four features are used: the region class membership homogeneity, the spectral class probability distribution $\mathcal{P}_{R_i}$ [2] , the region area and the area of the smallest oriented bounding box containing the region. For the features related to class membership, we assume that a certain number of spectral classes c_s defining different types of materials have been defined and that a SVM classifier has been trained for these classes and used on the region mean spectrum. As a result, the class probability distribution $\{\mathcal{P}_{R_i}(c_s)\}_{1\leq s\leq N_c}$ is available for each node. The four feature probabilities are detailed in the following:

3.1.1. Class membership homogeneity

This feature evaluates the region homogeneity in terms of class membership. Note that if a region is an object, all its pixels ideally belong to the same class. This term is important in the BPT context, as nodes close to the root node represent regions combining many different classes. It is defined as:

$$P(F_1|R_i) = \sum_{s=1}^{N_c} \sqrt{\mathcal{P}_{R_i}^{R_r}(c_s)\mathcal{P}_{R_i}^{R_l}(c_s)} \quad (2)$$

where $\mathcal{P}_{R_i}^{R_l}$ and $\mathcal{P}_{R_i}^{R_r}$ are the class probability distributions of the left and the right child nodes of R_i. N_c represents the number of different classes c_s used to train the SVM classifier. Note that if two sibling nodes have similar class probability distributions, their union will also have a similar distribution, i.e. the object is in the process of being formed.

3.1.2. Region area

This feature corresponds to the number of pixels forming the region contained in each BPT node. The goal of this feature is to prevent the detection of small or large meaningless regions. It is done by assuming that the area interval $[\mathcal{A}_{min}, \mathcal{A}_{max}]$ of the object of interest is known. $P(F_2|O)$ is then defined as a uniform distribution between $[\mathcal{A}_{min}, \mathcal{A}_{max}]$. The definition of $\mathcal{A}_{max}$ is important to detect individual objects as the union of two identical objects can result into a similar object of larger size.

3.1.3. Spectral class probability

This term $P(F_3|O)$ corresponds to the probability $\mathcal{P}_{R_i}(C_s)$ that the region R_i has to belong to the material class C_s of the object of interest. For instance, for the road detection application, this probability is the likelihood that the region belongs to the asphalt class. This probability is directly extracted from the class probability distribution $\mathcal{P}_{R_i}$ estimated by the SVM.

3.1.4. Area of the smallest oriented bounding box

This last feature is used to compute a probability related to the region shape. In this work, two different $P(F_4|O)$ have been used to deal with two different object detection applications. Both are based on the same assumption : the use of a measure normalized between $[0, 1]$ as a shape probability distribution. In the case of building detection, $P(F_4|O)$ measures the region compactness and is the ratio between the area of the region and the area of the smallest oriented bounding box including the region.

3.2. Processing populated BPT

At this stage, the BPT processing consists in detecting the nodes which are the most likely to be the sought objects. This strategy assumes that the objects of interest appear as individual nodes. The goal is to use the $P(O|R)$ values to discard nodes that significantly differ from the object of interest and to detect the best object representations. At this point, it should be remembered that the BPT structure represents inclusion relationships between regions. As a result, it is likely that nodes belonging to the same tree branch have similar $P(O|R)$ values than their parent or child nodes. As our goal is to detect non overlapping regions representing instantiations of the object of interest, only the best node R^* on the branch should be detected.

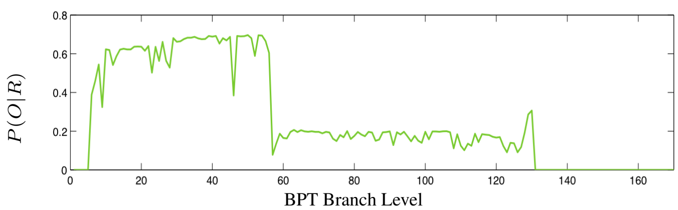

Fig. 2: Example of $P(O|R)$ evolution along a BPT branch

Taking into account these considerations, the approach used here is based on the analysis of the $P(O|R)$ evolution during the object formation along the branch. If we draw the $P(O|R)$ values along a BPT branch containing an object of interest starting from the leaf node, the first interesting point of the curve arrives when the smaller regions start having a high $P(O|R)$ value. After this, a stable range of values where no important change concerning $P(O|R)$ is generally observed. Finally, the last important step occurs when $P(O|R)$ suffers an important decrease after a specific merging step. At this point, the resulting region usually corresponds to a non-meaningful object of the image. In these situations, the best object representation R^* is found just before the important decrease. An example of this typical evolution can be observed in Fig. 2 where the curve of $P(O|R)$ values from a leaf to the root is represented. The horizontal axis indicates the level on the BPT branch the left side corresponds to the leaf and the right side to the root node) whereas the vertical axis indicates to probability values. We have observed that this behavior is really typical of branches containing the object of interest. Accordingly, the detection of R^* in a BPT branch is given by

$$R^* = \min_{R} \; P(O|R^+) - P(O|R) \tag{3}$$

$$with \; P(O|R) > \delta_T \tag{4}$$

where R^+ is the parent node of R and δ_T is the threshold used to decide if a region may be considered as a candidate of the sought object. Because of the inclusion relationship described by the BPT, the detection process described above may result in several detections of R^* along unique BPT branch. Hence, a decision should be taken in order to avoid overlapping regions in the final result.
Here, it has been considered that the region analysis is more reliable for large regions. Accordingly, in case of overlap, the R^* corresponding to the closest region to the root is kept. Following this pruning strategy, the selection of the R^* corresponding to the sought objects is done in a top-down fashion: the BPT is analyzed from the root to the leaves by selecting the first nodes found as R^*.

4. EXPERIMENTAL RESULTS

This section addresses the evaluation of the object detection strategy proposed in Section 3. The goal of the experiments is to compare the results of the proposed strategy with a classical pixel-wise method such as SVM classification. The evaluation is performed on two different portions shown in Fig. 3. Both hyperspectral images were acquired over Pavia (Italy) by the ROSIS sensor having a 1.3m spatial resolution. It corresponds to a urban area and the hyperspectral data involves 102 spectral bands. The experiment targets the detection of buildings. On these hyperspectral images, the BPTs are computed with the procedure described in Section 2.

Fig. 3: False color composition of two portions of the Pavia urban hyperspectral data used for building detection

Once the BPT has been computed, the four features presented in Section 3 are computed. The class probability distribution $\{\mathcal{P}_{R_i}(c_s)\}_{1 \le s \le N_c}$ is estimated with a SVM Gaussian kernel function constructed through a training step. This step follows the classical cross-validation strategy. The SVM training step is done by selecting randomly 20% of samples for each class from the available reference data. Once the kernel function is constructed, it is used to assign to each BPT node their class probability distribution $\{\mathcal{P}_{R_i}(c_s)\}$. In order to classify nodes corresponding to regions with several pixels, the region mean spectrum is used as input to the SVM classifier.
The constructed SVM kernel function is also used to perform a pixel-wise classification. The corresponding classification results obtained for the class of buildings are shown in Fig. 4. As can be seen, these classification results are rather noisy.

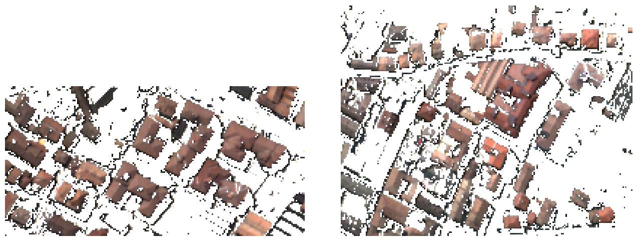

Fig. 4: Pixel-wise SVM classification on Pavia urban data

The results obtained by the proposed BPT-based strategy are shown in Fig. 5. In this case, the δ_T parameter is set to 0.65 and the range $[\mathcal{A}_{min}, \mathcal{A}_{max}]$ is set to $[30, 1000]$. As can be seen, most of the rectangular buildings have been precisely detected. These results corroborate the advantage of using the BPT representation. The use of spectral as well as spatial descriptors of BPT nodes clearly outperforms the classical pixel-wise detection using only spectral information. On the other hand, it should be also remarked that the results shown in Fig. 5 are also comparable with the results obtained by [1], where a building detection map is also presented by the using the same hyperspectral Pavia image.

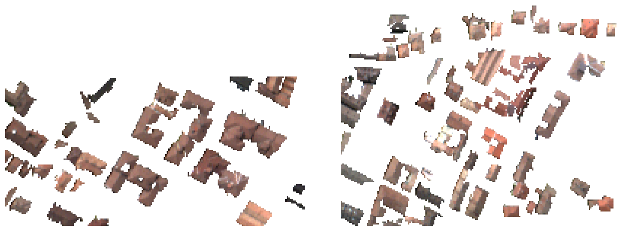

Fig. 5: BPT-based detection of building on Pavia urban data

5. CONCLUSIONS

An automatic hyperspectral object detection methodology using a BPT image representation has been detailed in this work. It has been illustrated how BPT can be a powerful image representation which provides a hierarchically structured search space for object recognition applications where the spectral and the spatial information can be incorporated in the search of a reference object. The obtained results show the interest of studying the objects of the scene with a region-based perspective. Future works will be conducted on the detection of other urban structures using the presented methodology.

6. REFERENCES

[1] H.G. Akay and S. Aksoy. Automatic detection of geospatial objects using multiple hierarchical segmentations. *In IEEE Trans. on Geoscience and Remote Sensing*, vol. 46(7), pp. 2097-2111, July 2008.

[2] A. Alonso-Gonzlez, S. Valero, J. Chanussot, C. Lopez-Martinez and P. Salembier. Processing Multidimensional SAR and Hyperspectral Images With Binary Partition Tree. *In Proceedings of the IEEE*, vol. 101(3), pp. 723-747, 2013.

[3] F. Calderero and F. Marqués. Region-merging Techniques Using Information Theory Statistical Measures. *In IEEE Trans. on Image Processing*, vol.19(6), pp. 1567-1586, 2010.

[4] T.F. Cox and M.A. Cox. Multidimensional Scaling. *In K. Fernandez and A. Morineau (Ed.),Chapman & Hal*, London,1994.

[5] T. Blaschke and J. Strobl. What's wrong with pixels? some recent developments interfacing remote sensing and gis. *Proceedings of GIS, Zeitschrift fur Geoinformationsystemte*, pp. 12-17, 2001.

[6] C.M. Cuadras, S. Valero, D. Cuadras, P. Salembier and J. Chanussot. Distance-based measures of association with applications in relating hyperspectral images. *Communications in Statistics - Theory and Method*, vol.41, pp. 2342-2355, 2012.

[7] A. Darwish, K. Leukert, and W. Reinhardt. Image segmentation for the purpose of object-based classification. *In IEEE Proc. of IGARSS*, pp. 2039-2041, 2003.

[8] A. Plaza and J. Tilton. Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images. *In IEEE Proc. of IGARSS*, pp. 4946-4949, 2005.

[9] S. Valero, P. Salembier and J. Chanussot. Hyperspectral image representation and processing with Binary Partition Trees. *In the IEEE Trans. on Image Processing*, vol. 22(4), pp. 1430-1443, 2013.

[10] S. Valero, P. Salembier, J. Chanussot and C.M. Cuadras. Improved Binary Partition Tree Construction for hyperspectral images: Application to object detection. *In IEEE Proc. of IGARSS*, pp. 2515-2518, 2011.

[11] 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 JSTARS, vol. 5, pp: 262 - 272 , 2012.