HAL
open science

New hyperspectral data representation using binary partition tree

Silvia Valero, Philippe Salembier, Jocelyn Chanussot

To cite this version:

Silvia Valero, Philippe Salembier, Jocelyn Chanussot. New hyperspectral data representation using binary partition tree. IGARSS 2010 - IEEE International Geoscience and Remote Sensing Symposium, Jul 2010, Honolulu, Hawaii, United States. conference proceedings. hal-00578960

HAL Id: hal-00578960

https://hal.science/hal-00578960v1

Submitted on 27 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NEW HYPERSPECTRAL DATA REPRESENTATION USING BINARY PARTITION TREE

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

[1]Technical University of Catalonia (UPC), Barcelona, Catalonia, Spain
[2]GIPSA-lab, Signal & Image Dept., Grenoble Institute of Technology, Grenoble, France.

ABSTRACT

The optimal exploitation of the information provided by hyperspectral images requires the development of advanced image processing tools. This paper introduces a new hierarchical structure representation for such images using binary partition trees (BPT). Based on region merging techniques using statistical measures, this region-based representation reduces the number of elementary primitives and allows a more robust filtering, segmentation, classification or information retrieval. To demonstrate BPT capabilites, we first discuss the construction of BPT in the specific framework of hyperspectral data. We then propose a pruning strategy in order to perform a classification. Labelling each BPT node with SVM classifiers outputs, a pruning decision based on an impurity measure is addressed. Experimental results on two different hyperspectral data sets have demonstrated the good performances of a BPT-based representation

Index Terms— Hyperspectral imaging, Binary Partition Tree, segmentation, classification

1. INTRODUCTION

Recent advances in remote sensing and geographic information has led the way for the development of hyperspectral sensors which produce a data cube of hundreds of contiguous waveband images. Therefore, each pixel is represented by a spectrum related to the light absorbing and/or scattering properties of the spatial region that it represents.

Given the wide range of real-life applications, great deal of research is invested in the field of hyperspectral image segmentation. The segmentation of these images is a key step in their analysis. Unfortunately, hyperspectral image processing is still a difficult endeavor due to the huge amount of data involved. Consequently, most of the standard segmentation methods fail.

In the literature, different segmentation algorithms based on morphological profiles [1], endmember extraction [2], Markov random fields [3], Bayesian segmentation [4] and hierarchical segmentation [5] have been proposed. The goal of segmentation (in particular for all the algorithms mentioned before) is to compute a partition from a pixel-based representation of the image.

This approach has two drawbacks: 1) The segmentation cannot be generic and also reliable. In fact, it has to depend on the application. 2) The initial pixel-based representation is too low level which implies that the segmentation algorithm is quite complex or not very robust. To tackle these issues, we would like to define a new data representation which represents a first abstraction from the pixel-based representation and that is multiscale to be able to cover a wide range of applications.

Binary Partition Tree (BPT) is one example of such representations. Having a rather generic construction (more or less application independent), they can be interpreted as a set of hierarchical regions stored in a tree structure.

Note that from the tree representation, many partitions can be extracted for various applications. The processing of BPT will then involve an application dependant pruning strategy.

Hence, we propose BPT as a new region-based hierarchical representation [6] for hyperspectral images.

In the case of remote sensing hyperspectral data, different prunings can be suitable for filtering, classification and segmentation purposes. As a first instance, we present here a pruning strategy aiming at a classification of the image.

The organization of this paper is given as follows: Section 2 gives a brief introduction on BPT, explaining the details of its construction. The BPT pruning for classification is discussed in section 3. Experimental results are shown in section 4. Finally, conclusions are drawn in section 5.

2. CONSTRUCTION OF THE BPT

Binary Partition Tree (BPT) is a hierarchical representation of a set of regions obtained from an initial partition. The tree leaves correspond the regions of the initial partition and the remaining tree nodes represent regions formed by the merging of two children regions.The root node represents the entire image support.

The tree construction is performed by keeping track of merging steps of an iterative region merging algorithm (see Fig. 1). The creation of BPT implies two important notions. On one hand, the merging criterion $O(Ri, Rj)$ between two adjacent regions Ri and Rj, on the other hand, the region model M_{Ri}. The merging criterion defines the similarity of neighboring regions and hence determines the order in which regions are

going to be merged. The region model specifies how regions are represented and how to model the union of two regions. Nevertheless, the definition of $O(Ri, Rj)$ as a similarity measure between two hyperspectral regions nodes is not an easy issue.

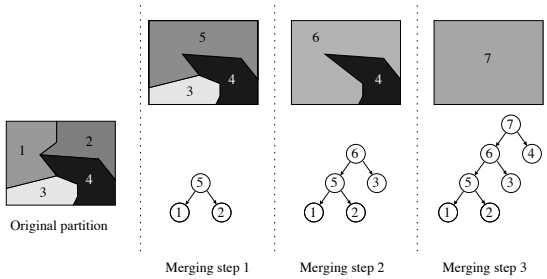

Fig. 1. Example of BPT construction

In the literature, some distances such as Spectral Angle Mapper or Spectral Information Divergence have been proposed to measure spectral similarity. However, their use as $O(Ri, Rj)$ is not straightforward as each region is made of several pixels and therefore several spectra. To overcome this problem, past approaches [5] have assumed that M_{Ri} is a constant, representing the regions by their mean spectrum. With this approach, the interclass spectral variability induced by natural variations, noise and mixed pixels is overlooked. In order to take into account this spectral variability within regions, we propose to model each band of the region spectrum by its probability density function [7].

2.1. Region Model

Working with N bands, the region model consists of N histograms representing for each band the empirical distribution of the pixels belonging to the region.
Consequently, the region model M_{Ri} is given by

$$M_{Ri} = \{P^1_{Ri}, P^2_{Ri}, ..., P^N_{Ri}\} \quad (1)$$

where P^k_{Ri} is the empirical distribution of the region Ri in the band k which is formed by

$$P^k_{Ri} = \{P^k_{Ri}(a_1)P^k_{Ri}(a_2), ...P^k_{Ri}(a_{|\chi|})\} \quad (2)$$

being a_i the possible values of the pixels in each band k.
We must remark that this region model can also be defined when tree leaves are single pixels by exploiting the image self-similarity. Indeed, the probability density function for individual pixels can be estimated and the precise modeling of the pixels pdf is important in order to get very precise region contours [8].

2.2. Merging criterion: Bhattacharyya coefficient

For each band k of each region R, the model P^k_R is an empirical discrete probability distribution. Accordingly, the Bhattacharyya coefficient [7] can be used to measure the similarity between two adjacent regions R_i and R_j of a given band k. Theorically, this measure is defined by:

$$BC(P^k_{Ri}, P^k_{Rj}) = -\log(\sum_{j=1}^{|\chi|} P_{Ri(k)}(a_j)^{\frac{1}{2}} P_{Rj(k)}(a_j)^{\frac{1}{2}}) \quad (3)$$

Existing a perfect overlap between both probability distributions, the Bhattacharyya coefficient will be 0. Consequently, a merging criterion of a pair of adjacent regions can be defined as the *minimum* sum of the N dissimilarity measures obtained for the different bands.

$$O(R_i, R_j) = \underset{Ri,Rj}{\operatorname{argmin}} \sum_{k=0}^{N-1} BC(P^k_{Ri}, P^k_{Rj}) \quad (4)$$

Experimentally, we have observed that the criterion of Eq. 4 does not assure that the areas of the regions tend to increase as the number of regions into the partition decreases. Then, in order to avoid small and meaningless regions into the generated partitions, the merging of very small regions has to be favored. To this goal we introduce a regularization term based on the size of the regions.

$$O(R_i, R_j) = min(\sqrt{N_{R_i}}, \sqrt{N_{R_{ij}}})O(R_i, R_j) \quad (5)$$

Note that we propose to use the square root of the *minimum* area.
To conclude this section, we must let us mention that the merging criterion defined by Eq. 4 simply adds the contribution of the various bands without exploiting their mutual information. Future works will analyze how this mutual information between bands can be used in the merging criterion.

3. BPT PRUNING

In this section, we discuss an example of tree processing for a classification application. The processing can be seen as a tree pruning step the goal of which is to remove subtrees composed of nodes belonging to the same class. To perform this task, we analyze the tree starting from the leaves and moving along the branches to select the nodes of largest area that involve pixels belonging to a unique class.
As a first step, we measure a specific region descriptors for each node R_i along the tree structure. These values are used to compute an increasing cost C associated to each BPT node. The increasingness of C along the branches guarantees that removing nodes having a cost lower than given threshold leads to a pruning.

The choice of region descriptors is determined by the application. In our case, the BPT pruning is focussed on the hyperspectral data classification. Hence, we propose a pruning strategy populating the nodes with the density probability function of belonging to each class.
Such a task can be achieved using a multi-class classifier output. Here, we use Support Vector Machine as a classifier which has demostrated its advantages in high dimensional data. We note that being supervised, SVM needs firstly to construct a model to be able to classify the data. Then, we start constructing the model by training the SVM classifier using some leaves nodes according to the available ground truth. After the model construction, modelling each R_i by its mean spectrum, all nodes are populated by their class probability estimation Cp_{R_i} and their predicted class $Class_{R_i}$. Using Cp_{R_i} values, we define an increasing iterative cost C along tree branches using a node impurity measure. The impurity of a node is interpreted by how mixed is the node, that is, the proportion of elements of different classes in the same region. To measure that, we propose a popular impurity function such as the entropy. Therefore, merging R_i at level l, the cost associated to R_i is computed using the following equation:

$$C(R_i) = C' - \sum_{t=0}^{N_c} Cp_{R_i}(t) \log(Cp_{R_i}(t)) \qquad (6)$$

where N_c is the number of classes and C' is the *maximum* cumulative cost until the $l-1$ branch level.
It should be noticed that measuring the sum of all the impurities, a maximum threshold λ should be set to determine the last pure node.
Thus, a node R_i is removed if $C(R_i) < \lambda$ and if all its ancestors also satisfy this condition. After tree pruning, we construct the classification map by selecting the lower nodes of the resulting pruned tree. Regions contained in these nodes are labelled by the $Class_{R_i}$ which has been assigned by the SVM classifier in the tree population.

4. EXPERIMENTAL RESULTS

4.1. Experiment with AVIRIS Indian Pines

In our first experiment, Indian Pines AVIRIS hyperspectral data containing 200 spectral bands having a spatial dimension of 145 X 145 pixels is used. The whole image is formed by 16 different classes having an available ground truth.
Before constructing BPT as detailed in Section 2, some parameters such as the number of bins N_{bins} used to represent P_R^k should be set. In our case, having different ranges of values in each chanel, we set N_{bins} as the *minimum* range difference found in the image (N_{bins}=46). Once the BPT has been created, we train the SVM classifier selecting randomly 30% of samples for each class from the reference data. After that, Cp_R and $Class_R$ values are assigned to each node to perform the pruning task.
In this pruning step, we should set λ in order to define the *maximum* impurity cost allowed along BPT branches. After some experimental tests, we set λ=20. Fig. 2 compares the obtained results using the BPT pruning against a classical SVM pixel classification. The same training samples are used for both classification methods.

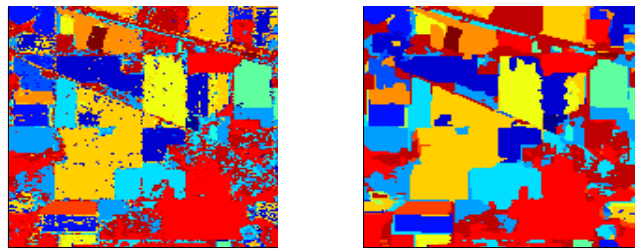

Fig. 2. Obtained classification map. Left: SVM result. Right: Pruning BPT result

Looking at BPT pruning results, we observe that the classification map is formed by quite homogeneous regions. In particular, the BPT nodes selection according to the proposed pruning criterion provide a less noisy classification. The obtained results also corroborate the BPT performances since extracted nodes reflect semantic real-world regions of the image. We should remark that Indian Pines has a high spectra variability due to its low spatial resolution.
Table I illustrates the class-specific and the global classification accuracies. Observing these results, we verify that the proposed BPT classification improves the classification accuracies for almost all the classes.

Table 1. Class Specifical Accuracy in percentage

Class	Simple SVM	Pruned BPT
1	86.11	94.44
2	88.39	93.41
3	83.45	89.03
4	77.56	80.77
5	95.18	92.77
6	97.39	98.39
7	83.33	88.89
8	97.85	99.08
9	64.29	100
10	84.98	88.70
11	91.19	96.72
12	92.93	93.66
13	100	98.59
14	96.99	99.42
15	67.72	98.43
16	95.30	100
Overall	87.67	94.52

4.2. Experiment with ROSIS-03 over the University of Pavia, Italy

In this second experiment, data from the ROSIS-03 optical sensor over the University of Pavia is presented. The image is formed by 103 denoised channels possesing 610 X 340 pixels. In this work, due to space limitations, only the top-down corner of this image is considered.
For this example, we should increase N_{bins} to 100 considering that this second data has a smaller spatial resolution (1.3 m per pixel). Although the merging criterion is not strongly dependent of the N_{bins}, it is better to take it into account.
Fig. shows the results obtained after applying BPT pruning.

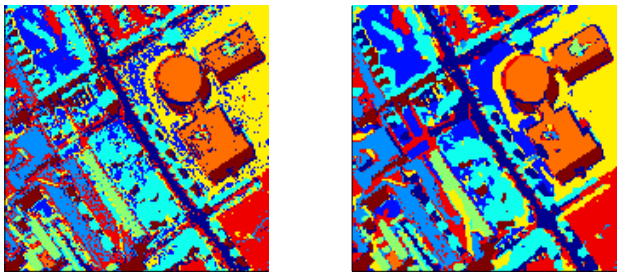

Fig. 3. Obtained classification map. Left: SVM result. Right: Pruning BPT result

It can be observed that, using the BPT, a better classification map is also obtained for this second data set . Despite of the improvement, some noise is still present in the results. This implies that our pruning criterion can be improved.
Regarding the global accuracy, the simple SVM classifier reaches 90.68 % whereas our proposed BPT pruning achieves 95.19%.
It should be observed that BPT pruning improves the classification accuracy preserving most of the edges and shapes.

5. CONCLUSIONS

In this work, Binary Partition Trees have been proposed as a new representation for hyperspectral images. Obtained through a recursive region merging algorithm, they can be interpreted as a new region-based and hierarchical representation of the hyperspectral data. The main advantatge of BPT is that it can be considered as a generic representation. Hence, it can be constructed once and used for many applications. Many tree processing techniques can be formulated as pruning strategies.
Concerning the BPT construction, a solution for the problem of the spectra variability for clustering hyperspectral data has been proposed using statistical region models. BPT enables the extraction of a hierarchically structured set of regions representing a semantic content of the image.
As a first example of BPT processing, we have proposed and illustrated a pruning strategy to classify the hyperspectral data. Experimental results have shown that the proposed method improves the classification accuracies of a classical SVM, providing classification maps with a reduced amount of noise.
Future work will be conducted for improving the merging criterion given that information between bands is not introduced in our similarity measure. Regarding the pruning strategy, new techniques are currently being studied to improve the accuracy and the robustness of the segmentation results.

6. REFERENCES

[1] M. Pesaresi, J.A. Benediktsson, and K. Arnason, "Classification and feature extraction for remote sensing", IEEE Trans. Geoscience and Remote Sensing, 2003, vol. 41, pp. 19401949.

[2] A. Plaza, J.A.Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, A. Gualtieri, J. Tilton and G. Trianni."Advanced Processing of Hyperspectral Images", Remote Sensing of Environment, DOI: 10.1016/j.rse.2007.07.028, 2009.

[3] J.L. Marroquin, E.A. Santana, S. Botello "Hidden Markov Measure Field Models for Image Segmentation", IEEE Trans. on Pattern Analysis and Machine Intelligence, 2003, vol.25(11), pp. 13801387.

[4] J. Li, J.M. Bioucas-Dias and A. Plaza, "Semi-supervised hyperspectral image segmentation", IEEE GRSS Workshop on Hyperspectral Image-WHISPERS'2009, Grenoble 2009.

[5] J. A. Gualtieri and J.C. Tilton," Hierarchical Segmentation of Hyperspectral Data", 2002 AVIRIS Earth Science and Applications Workshop Proceedings, 2002, pp 58.

[6] P. Salembier and L. Garrido, " Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval" IEEE Trans. Image Processing, 2000, vol. 9, no. 4, pp. 561576,

[7] F. Calderero and F. Marques," General region merging approaches based on information theory statistical measures", in Proc.of ICIP 2008, San Diego (CA), pp. 3016-3019

[8] M. Dimicolli and P. Salembier," Hierarchical region-based representation for segmentation and filtering with depth in single images", in ICIP 2009, November 2009, Cairo, EGYPT.