A COMPARISON STUDY BETWEEN WINDOWING AND BINARY PARTITION TREES FOR HYPERSPECTRAL IMAGE INFORMATION MINING

M.A. Veganzones[1], G. Tochon[1], M. Dalla Mura[1], A.J. Plaza[2], J. Chanussot[1,3]

[1]GIPSA-lab, Grenoble-INP, Saint Martin d'Hères, France
[2]Hyperspectral Computing Laboratory, University of Extremadura (UEX), Cáceres, Spain
[3]Faculty of Electrical and Computer Engineering, University of Iceland

ABSTRACT

Remote sensors capture large scenes that are conventionally split in smaller patches before being stored and analyzed. Traditionally, this has been done by dividing the scene in rectangular windows. Such windowing methodology could provoke the separation of spectrally homogeneous areas or objects of interest into two or more patches. This is due to the presence of objects of interest in correspondence to windows' borders, or because the fixed size of the windows does not adapt well to the scale of the objects. To alleviate this issue, the windows can be arranged in an overlapping way, incurring in some data redundancy storage. Recently, tree representations have been used as an alternative to windowing in order to structure and store large amounts of remote sensing data. In this work we explore the benefits of using Binary Partition Trees (BPT) instead of windowing to store hyperspectral large scenes. We are particularly interested in storing the information resulting of local spectral unmixing processes running over a large real hyperspectral scene. We show that under similar conditions BPT allows a better storage of the unmixing information in terms of reconstruction error.

***Index Terms*—** Hyperspectral images, BPT representation, windowing, spectral unmixing, image information mining

1. INTRODUCTION

Image Information Mining (IIM) implies the collection, storage and exploitation of large image databases. IIM has been a major issue in remote sensing community since the data collected for Earth Observation and Astronomical Observation have increased rapidly in the past decades. Large amounts of data provided by sensors such as Landsat, TerraSAR-X, SPOT and others have been collected during years. Despite some remarkable efforts, IIM for remote sensing data exploitation has not provided a satisfactory response and is still an active open field [1].

This work was supported by the Delegation Generale de l'Armement (DGA) under contract PDE2012600079.

A common procedure in image analysis is to divide large captured scenes in a regular grid before their storage and indexing. This division is done by regularly cutting the scene in rectangular patches of fixed size (windows). In order to alleviate problems derived by the possibility of breaking objects of interest into different windows, overlapping windows can be considered. However, this strategy incurs in data redundancy. Furthermore, it does not cope with the scale of the image regions, and the fixed size of the windows can yield to objects of interest being underrepresented by a patch if the window size is smaller than a given object, or by contrary the patch can contain a myriad of different objects with scale lower than the window size.

Recently, tree representations have been proposed as an alternative to windowing to structure and store image data [2]. Trees allow a hierarchical region-based representation of images which involves a number of regions that is much lower than the number of original pixels and it can be considered as a first level of abstraction with regard to the raw pixel-wise information [3]. Trees atomize the image into homogeneous components and embed multiple nested segmentations. From a tree structure it is easier to retrieve segments of interest and perform classification, which allows for interactive image content exploration [2].

In Binary Partition Trees (BPT) the leaf nodes of the tree represent the pixels in the original image, whereas the remaining nodes represent regions that are obtained by the merging of the two neighboring regions represented by two child nodes; the root node corresponds to the entire image [4]. In [5], the authors introduce the use of BPT for hyperspectral image representation and processing. We have recently proposed in [6] the use of BPT to locally unmix large hyperspectral scenes. The spectral unmixing of hyperspectral images is a process in which the original image is decomposed in the spectral signatures of the materials that constitute the image and their respective fractional abundances (their spatial distribution). This is usually a global process while in [6] we look for an optimal join of local spectral unmixing processes. We can think of windowing as an analogous process where a local unmixing is done over each resulting patch.

Here we propose the use of BPT to store and index hyperspectral data and compare it to traditional windowing for patches segmentation in terms of spectral unmixing quality. We show that under similar conditions, given by the number and size of the patches/regions, BPT outperforms windowing and so, better represents hyperspectral unmixing information.

The remainder of the paper is organized as follows. Section 2 gives a brief overview of spectral unmixing. Section 3 introduces the methodology to use BPT for hyperspectral unmixing. In section 4 we describe the experimental methodology and results, and finally we give some conclusions in section 5.

2. SPECTRAL UNMIXING

In the linear mixing model [7] a hyperspectral image can be seen as the result of the linear combination of the pure spectral signatures of spectrally pure material, named endmembers, with a fractional abundance matrix.

Let $\mathbf{E} = [\mathbf{e}_1, \ldots, \mathbf{e}_p]$ be the pure endmember signatures (normally corresponding to macroscopic objects in scene, such as water, soil, vegetation,...) where each $\mathbf{e}_i \in \mathbb{R}^q$ is a q-dimensional vector. Then, the hyperspectral signature $\mathbf{r}$ at each pixel in the image is defined by the expression:

$$\mathbf{r} = \mathbf{s} + \mathbf{n} = \sum_{i=1}^{p} \mathbf{e}_i \phi_i + \mathbf{n}, \tag{1}$$

where $\mathbf{r}$ is a q-dimensional signature given by the sum of the pixel's signal $\mathbf{s}$ and an independent additive noise component $\mathbf{n}$; and, ϕ is the p-dimensional vector of fractional abundances at the given pixel subject to constraints: $\phi_i \geq 0$, $\forall i = 1, \ldots, p$, and $\sum_{i=1}^{p} \phi_i = 1$. This equation can be extended to the full image as $\mathbf{H} = \mathbf{E}\mathbf{\Phi} + \boldsymbol{\eta}$, where $\mathbf{H}$ is the hyperspectral image, $\mathbf{\Phi}$ is a matrix of fractional abundances and $\boldsymbol{\eta}$ is independent additive noise.

The set of endmembers can be defined on the basis of a priori knowledge about the imaged scene. A library of known pure ground signatures or laboratory samples could be used. However, when large amounts of data must be processed some automatic procedure should be used instead. In such cases, the set of endmembers must be induced from the hyperspectral image data, for example, by means of Endmember Induction Algorithms (EIA). Once the set of endmembers, $\mathbf{E}$, has been induced from an hyperspectral image, their corresponding abundances can be estimated by Full-Constrained Least Squares Unmixing (FCLSU). The quality of the unmixing, the estimated $\hat{\mathbf{E}}$ and $\hat{\mathbf{\Phi}}$, at a given pixel $\mathbf{r}$ can be measured by the Root Mean Squared Error (RMSE), $\epsilon(\mathbf{r}, \hat{\mathbf{r}})$, of the original hyperspectral signature $\mathbf{r}$ respect to the reconstructed one, $\hat{\mathbf{r}} = \sum_{i=1}^{p} \hat{\mathbf{e}}_\mathbf{i} \hat{\phi}_i$:

$$\epsilon(\mathbf{r}, \hat{\mathbf{r}}) = \sqrt{\frac{1}{q} \sum_{j=1}^{q} (r_j - \hat{r}_j)^2}, \tag{2}$$

3. SCENE UNMIXING STORAGE BY BPT

Following the work in [6] we build a BPT of the scene and then, we prune it to achieve an optimal partition in terms of unmixing reconstruction error (2). First, a watershed oversegmentation of the scene is done using a morphological Beucher gradient of the scene. The regions on the watershed segmentation map identify with the leaf nodes of the tree. Then, the tree is built by iteratively merging those adjacent nodes that minimizes some pairwise similarity. In this case we compute the angular distance over the means of the regions associated to each pair of adjacent nodes:

$$s\left(\mathcal{R}_\alpha, \mathcal{R}_\beta\right) = d_{\text{SAM}}\left(\bar{\mathbf{r}}_\alpha, \bar{\mathbf{r}}_\beta\right) = \cos^{-1}\left(\frac{\sum_{j=1}^{q}\left(\bar{r}_\alpha^{(j)} \bar{r}_\beta^{(j)}\right)}{\sqrt{\sum_{j=1}^{q}\left(\bar{r}_\alpha^{(j)}\right)^2}\sqrt{\sum_{j=1}^{q}\left(\bar{r}_\beta^{(j)}\right)^2}}\right) \tag{3}$$

where $\bar{\mathbf{r}}_\alpha$ and $\bar{\mathbf{r}}_\beta$ denote the sample mean of the pixels in regions $\mathcal{R}_\alpha$ and $\mathcal{R}_\beta$ respectively.

The resulting tree, with the complete scene in the root node, is then pruned to find the optimal partition minimizing the maximum RMSE of the complete scene. In order to do that, the spectral unmixing process is run separately for each node. The result is that each node R_α contains a set of induced endmembers, E_α, and their corresponding fractional abundances. Then, the pruning is formulated as:

$$\mathcal{P}^\star = \arg\min_{\mathcal{P} \in \Omega} \max_{\mathbf{r}} \epsilon_{\mathcal{R}}(\mathbf{r}, \hat{\mathbf{r}}), \forall \mathcal{R} \in \mathcal{P}. \tag{4}$$

where P^* is the optimal partition among the set of all possible BPT partitions Ω. It is possible to constraint Ω to those partitions containing regions above a minimum spatial size. In order to do that, an additional term can be included in (4) modelling the size constraint:

$$\Omega = \{\mathcal{P}\}, \text{ s.t. } \forall \mathcal{P} \in \Omega,\ \forall \mathcal{R} \in \mathcal{P},\ \ |\mathcal{R}| \geq c, \tag{5}$$

where $|\mathcal{R}|$ denotes the cardinality (number of pixels) of region $\mathcal{R}$ and $c \geq 0$ is a threshold on the region size. If $c = 0$, the term (5) has no effect and the pruning criterion is considered to be unconstrained.

4. EXPERIMENTS AND RESULTS

4.1. Pavia University scene

The Pavia University hyperspectral image was collected by the ROSIS-03 sensor over the facilities of the University of Pavia in Italy. After discarding pixels with no information and noisy spectral bands, the image has a spatial size of 610×340 pixels with a spatial resolution of 1.3 m per pixel, and 103 spectral bands comprised in the range of 430-860 nm. The scene shows an urban area comprised of different buildings, parking lots, roads and other typical human-made constructions, together with trees, green areas and bare soil.

4.2. Methodology

On one hand, we cut the Pavia University scene in patches using a non-overlapping windowing of increasing sizes: 16×16, 32×32, 64×64, 128×128 and 192×192. We also cut the scene using an overlapping windowing with increasing overlapping rate: 3, 5, 10, 20 and 30 for the same respective window sizes than in the non-overlapping windowing. On the other hand, we represented the scene by a BPT (3), and we found optimal partitions by pruning it according to the pruning criterion (4) and the maximum region size constraint (5), with threshold values c in the range $[0, 5000]$. The unmixing results have been obtained by the Vertex Component Analysis (VCA) [8] endmember induction algorithm and the FCLSU unmixing algorithm. As the VCA is an stochastic algorithm, we run it 20 times for each patch/node, keeping the results of the one achieving the minimum RMSE error (2).

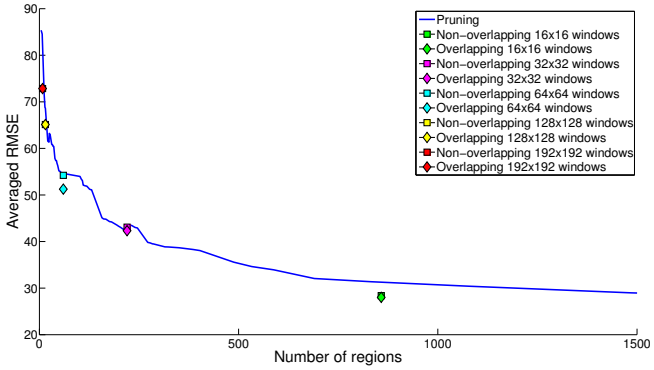
4.3. Results

Fig. 1 shows the averaged and maximum RMSE reconstruction errors obtained using the non-overlapping and overlapping windowing compared to the BPT pruning approach. The averaged RMSE error is similar for both, windowing and pruning approaches, although a slightly improve can be found for the overlapping windowing with sizes 16×16 and 64×64. However, the pruning approach clearly outperforms the windowing approaches in maximum RMSE error.

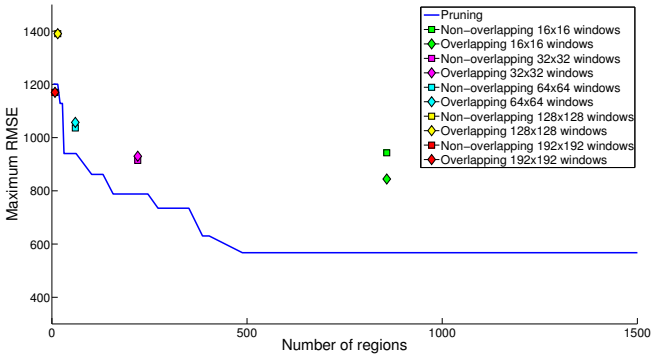
Fig. 2 depicts the RMSE maps obtained for the windowing and pruning approaches. In the case of the pruning approaches, we have shown the results of the optimal partitions with the closer number of regions respect to the corresponding windowing approach. The windowing effect can be clearly observed in Fig. 2(a-b). The RMSE maps obtained by the pruning approach are smoother and more visually appealing. Also, it can be appreciated that the number of pixels with high RMSE errors (in red) is higher for the windowing approaches than for the pruning.

5. CONCLUSIONS

We propose the use of BPT representations to store hyperspectral data in large databases instead of the traditional windowing patching, either with or without overlapping. Specifically, we want to store the results of an unmixing process. We compare both approaches, windowing and pruning, in terms of the reconstruction error obtained in each case, using the real Pavia University hyperspectral scene. We have shown that the pruning approach outperforms the windowing in terms of maximum RMSE, and is similar in terms of averaged RMSE. Also, the RMSE maps are more visually appealing in the pruning approach. Further work will focus on the role of the spectral information, the endmembers obtained by the unmixing process, in the BPT representation.



(a)



(b)

Fig. 1. (a) Averaged and (b) maximum RMSE errors for the Pavia University scene, using the windowing and pruning approaches.

6. REFERENCES

[1] M. Datcu, "Image information mining for eo payload ground segments," in *Proceedings of the ESA-EUSC-JRC Eight Conference on Image Information Mining*, Prof. Mihai Datcu, Ed. 2012, JRC.

[2] L. Gueguen, G.K. Ouzounis, M. Pesaresi, and P. Soille, "Tree based representations for fast information mining from vhr images," in *Proceedings of the ESA-EUSC-JRC Eight Conference on Image Information Mining*, Prof. Mihai Datcu, Ed. 2012, JRC.

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

[4] 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*, vol. 50, no. 2, pp. 593 –605, feb. 2012.

[5] S. Valero, P. Salembier, and J. Chanussot, "Hyperspectral image representation and processing with binary partition

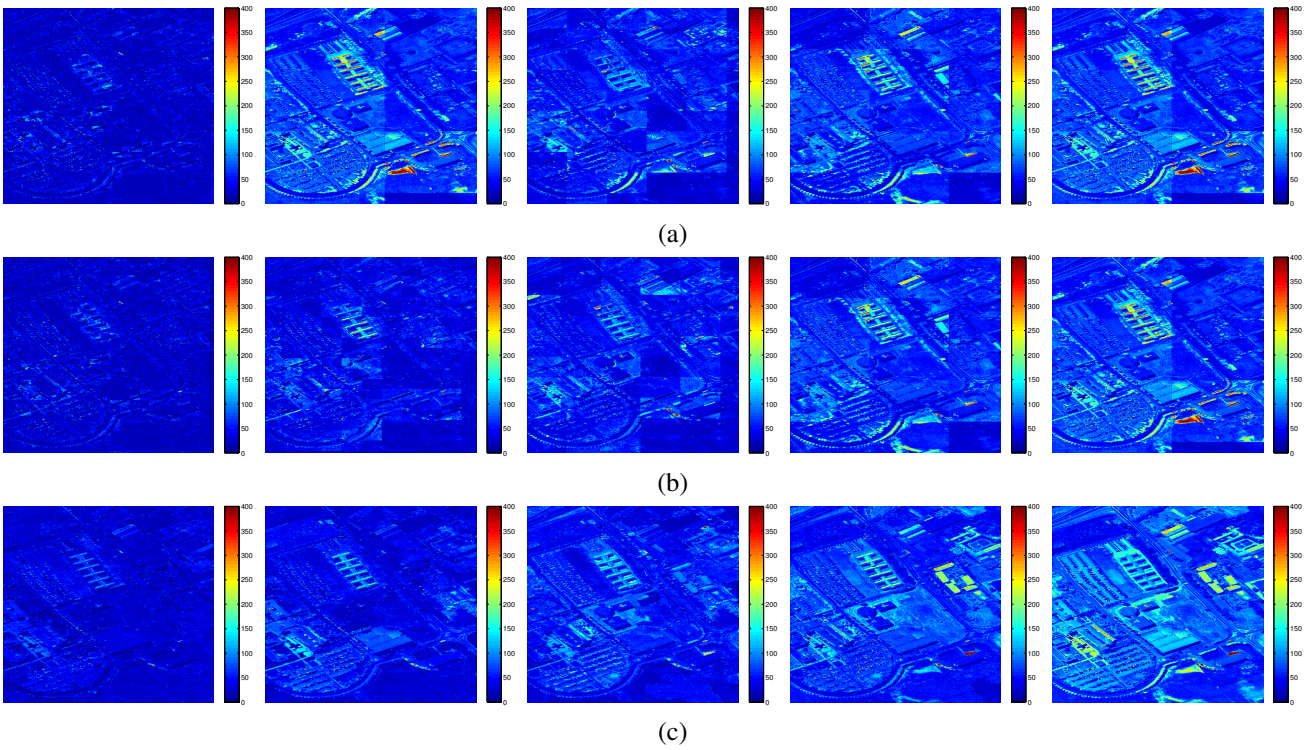


Fig. 2. RMSE maps of the Pavia University scene using (a) Non-overlapping windowing, (b) overlapping windowing and (c) pruning approach. Colums from left to right indicate an increasing window size.

trees," *IEEE Transactions on Image Processing*, vol. PP, no. 99, pp. 1, 2012.

[6] M.A. Veganzones, G. Tochon, M. Dalla Mura, A. Plaza, and J. Chanussot, "On the use of binary partition trees for hyperspectral unmixing," in *Proceedings of the IEEE 2013 International Conference on Image Processing (ICIP)*, In press.

[7] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Qian Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354 –379, april 2012.

[8] J.M.P. Nascimento and J.M. Bioucas Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.