ADAPTIVE REGION-BASED COMPRESSION OF MULTISPECTRAL IMAGES

M.Cagnazzo, R.Gaetano, S.Parrilli, L.Verdoliva

Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni
Università Federico II di Napoli, via Claudio, 21 – 80125 Napoli, Italy
{cagnazzo,raffaele.gaetano,sara.parrilli,verdoliv}@unina.it

ABSTRACT

The region-based description of multispectral images enables important high-level tasks such as data mining and retrieval, and region-of-interest selection. In order to obtain an efficient representation of such images we resort to adaptive transform coding techniques. Such techniques, however, require a considerable information overhead, which must be carefully managed to obtain a satisfactory rate-distortion performance. In this work we develop several region-based coding schemes and compare them with conventional (non-adaptive) and class-based schemes, so as to single out the rate-distortion gains/losses of this approach.

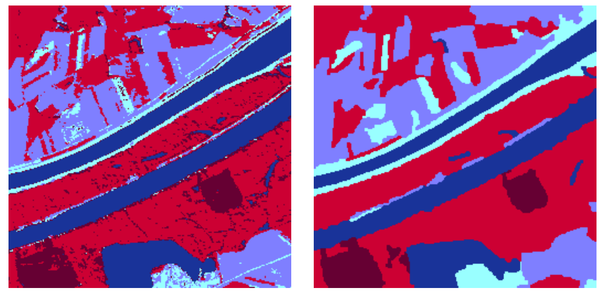
***Index Terms*—** Region-based, object-based, wavelet, multispectral, image compression

1. INTRODUCTION

Earth observing satellites produce a huge amount of images daily, to the point that the major problem for the final user is not so much the existence of images of interest, but the ability to locate them effectively.

This is a major reason towards the use of a region-based, rather than pixel-based, description of remote sensing images, so that the user can retrieve the image of interest based on the shape or the synthetic statistics of some of its regions. In other scenarios, the user might even be interested in downloading only the image segmentation map, or just a few regions of interest, with huge bandwidth saving in both cases. Given the typical size of remote-sensing images, it is important to use some form of compression, which means, in this case, some shape-adaptive coding techniques [1]. This research group has been working for some years on shape-adaptive transform coding [2, 3] focusing attention, however, on class-adaptive coding, because of its potential in terms of rate-distortion performance. In fact, when the image is segmented in a few classes with homogeneous statistics, the transforms are able to compact the signal energy in a small number of coefficients leading to a very good performance.

Although class-based and region-based coding technique share the same general approach, they differ under some important respects, that is, the number of objects (few classes vs. many regions), their statistical homogeneity and their spatial compactness (sparse classes vs compact regions).



(a) (b)

Fig. 1. Two example segmentation maps

The number of objects is important if one is interested in adaptive transform coding, because some parameters must be transmitted for each object, and therefore, there is a rapid growth of the side information in the region-based case, which could easily become prohibitive. The statistical homogeneity and the spatial compactness of objects are also important: while classes are singled out mostly on the basis of spectral homogeneity criteria, regions are required to satisfy some additional spatial criteria. For example, an isolated point can hardly be regarded as a region, and must be absorbed in some larger regions with different statistics. Therefore, regions tend be less "pure" than classes, and a spectral transform is less effective on them. On the other hand, a spatial transform applied to a compact region, rather than to a sparse class, might work better.

It is clear that the success of a region-based adaptive coding technique depends in large part on the segmentation. As an example, consider the two segmentation maps shown in Fig.1. The first one is obtained through a spectral clustering, with classes that are statistically homogeneous but very sparse and, therefore, is not suitable for region-based coding. The second one, instead, is obtained through a more sophisticated technique [4] with provides a much smaller number of connected regions, which are less homogeneous but can be certainly encoded individually.

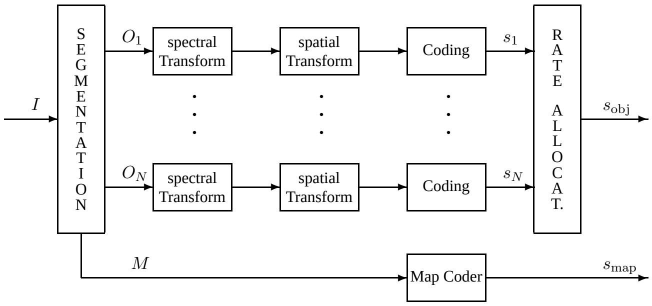


Fig. 2. Generic segmentation-based coding scheme

The goal of this work is to assess the potential of region-based transform coding. To this end we implement a number of region-based techniques and test them on two multispectral images. Rate-distortion performances are then compared with those of several reference techniques, both "flat" (without any image segmentation) and class-based. In the next Section, we describe all the coding schemes considered, while Section 3 is devoted to the experimental analysis. Section 4 will draw conclusions.

2. CODING SCHEMES

All coders considered in this work can be regarded as special cases of the generic segmentation-based scheme shown in Fig.2. There is first a segmentation of the image which singles out N objects (be they regions or classes), then each object O_i is subject to a spectral transform, a spatial transform, and finally quantization and coding, to produce an output stream s_i. A post-compression rate allocation block is used to select the optimal coding rate for each object, given the total resources available. Of course, the segmentation map must be itself coded and sent as a side information, together with the object coding parameters needed to synchronize the decoder at the receiver end.

To define a specific coding scheme we need to define precisely all six blocks of Fig.2 (replicas excluded), therefore we now review briefly these blocks and describe the alternative solutions we can consider.

We will carry out **segmentation** by means of the MRF-based algorithm proposed in [4] followed by a few morphological regularization steps in order to obtain a map with relatively few compact regions like that of Fig.2(b). The choice of a suitable segmentation map is central for the success of the overall scheme but is not the object of this paper and will not be discussed further, more details about this topic can be found in [5].

Similarly, we will not devote much attention to **map coding**, which is relatively straightforward and will be performed by the RAPP algorithm proposed in [6].

On the contrary, one of the major focus of this work is on the **spectral transform**, for which we will consider several alternatives: wavelet, KLT (Karhunen-Loeve transform), class-adaptive KLT, and region-adaptive KLT. The first two choices require no comment, but for the fact that KLT needs the transmission of the covariance matrix of the data (or the transform matrix itself) as a side information. In class-adaptive KLT, a different transform matrix is computed (and hence transmitted) for each class; this increases the cost of side information, but not very much, since only a few classes are typically present. In region-based KLT, instead, a different matrix is transmitted for each connected region; this allows for a better compaction of signal energy [5], but the cost of side information can become prohibitive. An intermediate solution, which saves some side information, is to send the mean vector for each region, but then use a single KLT matrix for all regions of the same class.

As **spatial transform** of the regions/classes we will consider always the shape-adaptive wavelet transform (SAWT) proposed in [7], which proved to be very efficient in a wide range of situations, works both on connected and sparse support, and falls back to an ordinary WT when the image is not segmented.

The **encoding** engine will be SPIHT [8] which is easily adapted to work on regions of arbitrary shape [2] and is also very efficient [9]. Independent of how the spectral transform was carried out, we can encode the image as a whole, as a set of classes or a set of individual regions, although some combinations make definitely more sense than others.

Finally, the **rate allocation** block takes the rate-distortion (RD) curves of all encoded objects as input and performs a lagrangian optimization to select the most suitable working points. We will distinguish between techniques that allocate bits to the whole object or individually to each band of the object. This distinction belongs more properly to the coding phase, but we consider it here because performance variations are due essentially to bit allocation.

3. EXPERIMENTAL RESULTS

Experiments have been carried out on two quite different test images, a 6-band, 512×512 pixel section of a Landsat TM image of a region near Lisbon, and a 32-band, 512×512 pixel section of an AVIRIS image of Moffett-field. We used a map with 3 classes and 59 regions for the TM image and 4 classes and 28 regions for the AVIRIS image.

Even though some blocks of Fig.2 have been kept fixed in this work, a large number of alternative coding schemes remain possible, most of which have been actually implemented and studied in the experiments. Two results emerge very clearly from all the experiments: 1) WT is not competitive with any form of KLT as spectral transform and 2) an explicit band-by-band bit allocation is much better than the implicit allocation performed by SPIHT. Fig. 3 and Fig. 4 re-

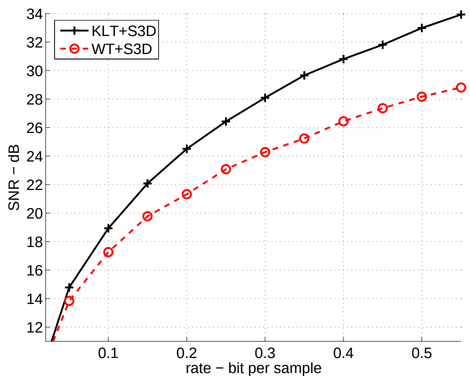


Fig. 3. Comparison of WT and KLT as spectral transform on the AVIRIS image. The spectral transform is followed by 2d WT and 3D SPIHT encoding.

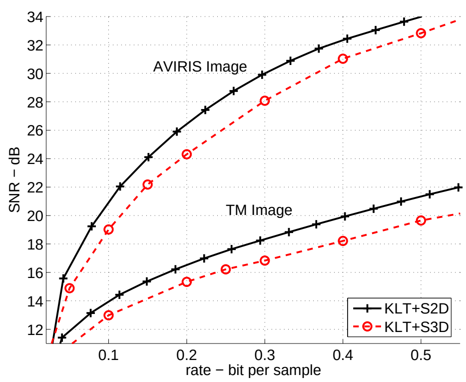


Fig. 4. Comparison of explicit (with 2D SPIHT) and implicit (3D SPIHT) rate allocation. The spectral transform is KLT and the spatial one is WT.

port just a few RD curves obtained on the test images which clearly substantiate our claims. Similar results have been obtained for other images. In the following, therefore, we will always consider KLT as spectral transform and use band-by-band allocation.

In order to study the potential of region-based approaches, we select three such schemes and three reference schemes, as summarized in Tab. 1. The first reference scheme is just a conventional "flat coder", with spectral KLT, spatial WT, and band-by-band SPIHT with explicit rate allocation. The following two schemes are class-based coders, which differ in the amount of adaptivity offered and side information required. The first one (#2) uses global KLT after removing the global mean vector; #3 uses a different KLT matrix for each class, after removing the class mean vectors. Coding scheme #6 is the only "pure" region-based coder, as it carries out all transform and coding steps on a per-region basis. The other region-based coders resort to some compromise to reduce side information: #4 performs a global KLT, while #5 performs a class-based KLT, removing means class-by-class. The last

#	No	Mean	KLT	Cod.	SI TM	SI AVI
1	Yes	G	G	G	≈ 0	0.003
2	Yes	G	G	C	0.012	0.003
3	Yes	C	C	C	0.012	0.006
4	Yes	G	G	R	0.012	0.003
5	Yes	C	C	R	0.012	0.006
6	Yes	R	R	R	0.026	0.031

Table 1. Reference and region-based coding schemes, and side information (bit/sample) for Landsat and AVIRIS images. G=global, C=class-based, R=region-based.

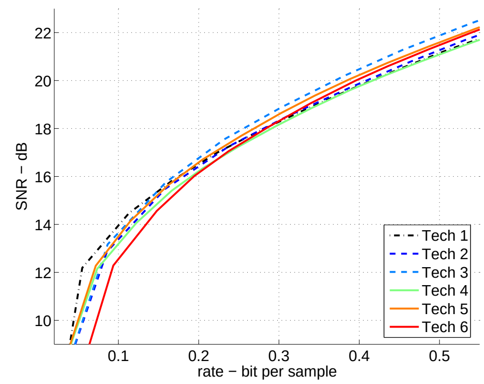


Fig. 5. Performance of test techniques on the TM image

two columns of the table show the corresponding side information for the Landsat TM and the AVIRIS image, expressed in bit per sample. The side information mainly accounts for the KLT matrices and the segmentation map information. For the 6-band TM images the map cost is prominent, as it is shared among a relatively small number of bands. On the contrary, for the 32-band AVIRIS image, the main contribution to the side information is due to the KLT matrices. Anyway it becomes significant only for the last technique, when we need a different KLT matrix for each region, but even in this case this cost is quite low – about 0.03 bit per sample.

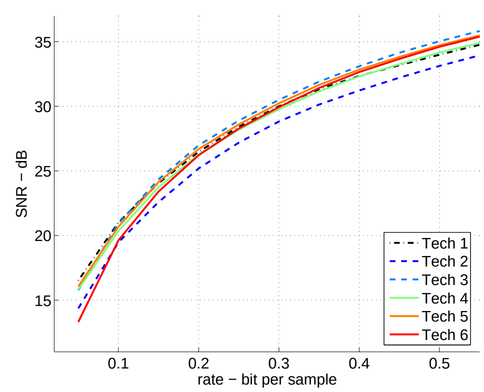


Fig. 6. Performance of test techniques on the AVIRIS image

Fig. 5 and Fig. 6 report the rate-distortion curves of all these schemes for the Landsat and, respectively, the AVIRIS test image. We observe that at low bit-rates, flat and class-based schemes have the best performances, thanks to their low side-information requirements. However #2 is always worse than #3, meaning that classified KLT is worth its cost with respect to global KLT. This is still true for the region-based techniques, where #4 is worse than #5 and #6 (but for low rates). At medium rates, the flat scheme is overcome by the class-based and region-based techniques; moreover, we see that region-based techniques #5 and #6 surpass class based techniques #2 and approach performances of #3, which however can be considered the best technique overall. We can conclude that the more complex description of image objects carried by class-based and region-based techniques causes a global improvement of performances with respect to a completely flat technique as #1. The best compromise between cost and effectiveness of shape information seems to be the class-based technique #3, but the region-based techniques (namely #5 and #6) have quite close performance, except for very low bit-rates.

Similar considerations can be made for performances on the AVIRIS image, see Fig. 6. The main difference with respect to the previous case is that now, region-based techniques performance is even closer to the best class-based technique #3. We ascribe this behavior to the smaller number of objects, that reduce side-information, and the higher number of bands, that makes more important the effectiveness of a precise spectral transform.

Finally, in Fig. 7 we report a false color representation of the original and decoded (at 0.3 bit per sample) AVIRIS image: no visible artifact is present and the quality is good.

4. CONCLUSIONS

In this work we introduce several region-based coding technique for the compression of multispectral image. The main interest of this approach is that it allows a semantics-based access to image content, which in turn can be very useful for sophisticated techniques of image processing and retrieval. Nevertheless, we are interested to the effect that the region-based paradigm has on rate-distortion performance. Our experiments lead to the conclusion that the region-based approach does not penalize significantly RD performance with respect to classical, non object-based schemes; on the contrary, their performance is usually better than that of completely flat approaches, and always quite close to the class-based one, which optimize the trade-off among cost and effectiveness of object-based transforms.

Of course, a central role in a region-based encoding scheme is played by the segmentation stage. At the moment, we use a segmentation technique which is independent from the subsequent compression stage. Further investigations will be devoted to the development of a segmenter which performs its task with the aim of optimizing the global rate-distortion performance.

Fig. 7. Original and reconstructed (0.3 bit per sample) AVIRIS image, false colors

5. REFERENCES

[1] G.Gelli, G.Poggi, “Compression of multispectral images by spectral classification and transform coding”, *IEEE Transactions on Image Processing*, pp.476-489, Apr. 1999.

[2] M.Cagnazzo, G.Poggi, L.Verdoliva, A.Zinicola, “Region-oriented compression of multispectral images by shape-adaptive wavelet transform and SPIHT”, *Proc. IEEE ICIP*, pp.2459-2462, Singapore, Oct. 2004.

[3] M.Cagnazzo, G.Poggi, L.Verdoliva, “A Comparison of flat and object-based transform coding techniques for the compression of multispectral images”, *Proc. IEEE ICIP*, pp.657-660, Genova (I), Sept. 2005.

[4] C.D’Elia, G.Poggi, G.Scarpa, “A Tree-Structured Markov random field model for Bayesian image segmentation”, *IEEE Transactions on Image Processing*, pp.1259-1273, Oct. 2003.

[5] M.Cagnazzo, R.Gaetano, S.Parrilli, L.Verdoliva, “Region based compression of multispectral images by classified KLT”, *Proc. EUSIPCO*, Firenze (I), Sept. 2006.

[6] V.Ratnakar, “RAPP: Lossless image compression with runs of adaptive pixel patterns” *Proc. 32nd Asilomar Conf. Signals, System and Computers*, pp.1251-1255, Nov. 1998.

[7] S.Li, W.Li, “Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding”, *IEEE Trans. Circuits Systems Video Technology*, pp.725-743, Aug. 2000.

[8] A.Said, W.A.Pearlman, “A new fast and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Trans. Circuits and Systems for Video Technology*, vol.6, pp.243-250, June 1996.

[9] M.Cagnazzo, G.Poggi, L.Verdoliva: “Costs and advantages of shape-adaptive wavelet transform for region-based image coding”, *Proc. IEEE ICIP*, pp.197-200, Genova (I), Sept. 2005.