REGION BASED COMPRESSION OF MULTISPECTRAL IMAGES BY CLASSIFIED KLT

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,gaetano,poggi,verdoliv}@unina.it

ABSTRACT

A new region-based algorithm is proposed for the compression of multispectral images. The image is segmented in homogeneous regions, each of which is subject to spectral KLT, spatial shape-adaptive DWT, and SPIHT encoding. We propose to use a dedicated KLT for each region or for each class rather than a single global KLT. Experiments show that the classified KLT guarantees a significant increase in energy compaction, and hence, despite the need to transmit more side information, it provides a valuable performance gain over reference techniques.

1. INTRODUCTION

Remote sensing of the Earth produces such a huge amount of data that compression is essential in one or more phases of the data life (before transmission to the ground station, in the archival facilities, or during dissemination to the end users) provided that image quality is preserved.

Region-based (a.k.a. object-based) coding is emerging as a very convenient way to of accomplishing this task [1, 2, 3]. The idea is to single out the elementary regions of the image through a preliminary segmentation step, and then encoding them as individual entities. In the end, this amounts to more than just compression, and could be more properly considered as an alternative representation of the image, with data saving as a by-product.

This approach guarantees some obvious advantages, related to the opportunity to deal with the semantic components of an image. First of all, the user is provided with a segmentation map (obtained on the original data) which gives the basic structure of the image, and might even be the only product of interest if image classification is pursued. In addition, the user can single out just one or a few regions of the image, considered of interest because of their shape or their synthetic statistics or else, and download just those regions, with huge bandwidth saving. More in general, this form of compression can be seen as more fit to data mining applications, which draw much interest lately.

Besides these high-level, or "semantic" advantages, there are also advantages related to low-level signal processing tasks. For example, if regions are accurately selected, one can easily and effectively denoise them, without the risk of smearing the important object contours. Our goal, here, is to show that region-based coding can be effective also in a rate-distortion sense, despite its need to transmit some side information. In fact, given a reliable region map, one can tailor the data compression algorithm and its parameters to each specific region, and allocate the coding resources among all regions so as to optimize performance and be competitive with state-of-the-art conventional techniques.

In [3] we implemented and tested a number of conventional (or "flat") and region-based transform coding techniques, to conclude that the best region-based technique had approximately the same performance of the best flat one. From that analysis it also emerged clearly that the best techniques had in common the use of the Karhunen-Loeve transform (KLT) in the spectral dimension to decorrelate the image bands. Such a result, supported by similar findings in the literature, stresses the key role of an effective spectral decorrelation for the overall coding performance. Therefore, we decided to move one step forward and turn to *adaptive* KLT to further improve performance, that is to use a dedicated KLT matrix for each region, adapted to its specific statistics. Adaptive KLT is not new by itself, a simplified form was already used in [4], and then in [5] in the framework of a class-based coder for multispectral images. Its application to region-based coding, however, deserves some caution because there are significant differences between class-based and region-based coding, which we summarize, for the time being, by saying that the first relies on a small number of highly homogeneous but sparse classes, while the latter is based on a larger number of spatially-compact but less homogeneous regions.

In this work we test the use of classified KLT (CKLT) in the context of region-based coding and compare results with the natural references, that is region-based coding with *global* KLT, and class-based coding with classified KLT. Section 2 describes in more detail the various coding schemes considered, while Section 3 analyzes the effects of CKLT on region-based coding and draws 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.1. There is first a segmentation of the image which singles out N objects (regions or classes), then each object O_i is subject to a spectral KLT, a spatial transform (DCT or DWT) and finally quantization and coding, to produce an output stream s_i. Of course, the segmentation map must be itself coded and sent as a side information, together with the object statistics which allow to reproduce the KLT at the receiver.

The coder proposed in [5] uses point-wise segmentation and singles out N_C classes, that is N_C types of land cover that share similar statistical properties. All other coders considered here, instead, use a contextual segmentation, which takes into account both spatial and spectral properties, and identify N_R regions rather than classes, with N_R typically much larger than N_C.

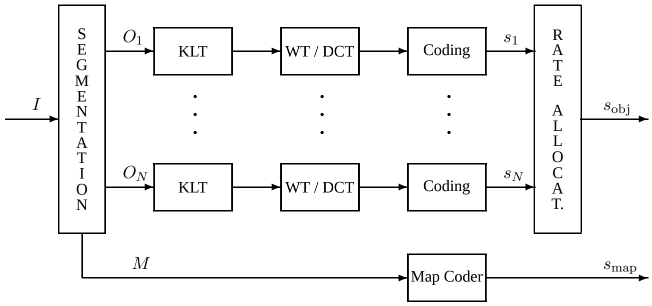


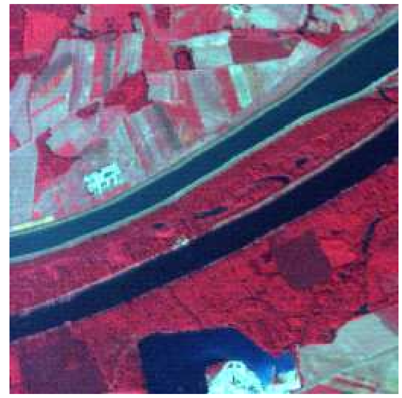
Figure 1: Generic segmentation-based coding scheme

To gain insight about this difference consider Fig.2, which shows an example image in false colors (a), its pixel-wise segmentation (b), contextual segmentation (c), and a further processing of this latter map (d). The segmentation of Fig.2(b) is just a clustering of the spectral response vectors associated with each pixel. The output is a set of 5 classes, corresponding to the colors, each one composed by several disjoint regions as well as by many isolated points or small clusters of points that can be hardly regarded as regions. Within each class there is a strong statistical homogeneity, and therefore an adaptive KLT, with a dedicated transform matrix for each class, is very effective in compacting energy, as verified experimentally in [5]. Unfortunately, the subsequent spatial transform turns out to be quite inefficient since it has to operate on objects (the classes here) that are very fragmented in space. The best option in these conditions turns out to be one-dimensional rather than two-dimensional DCT [5], with an obvious performance impairment.

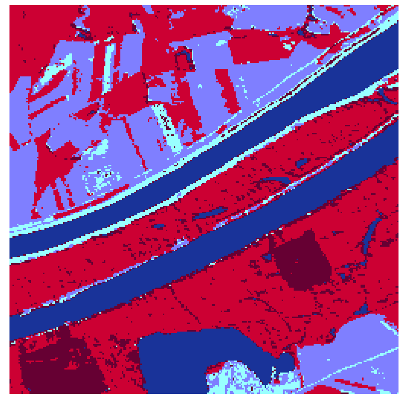
The segmentation of Fig.2(c) is instead obtained by means of a bayesian contextual technique [6] which penalizes isolated points and provides therefore a smaller number of more compact regions. If one further regularizes such a map, by erasing small regions and smoothing boundaries between different regions, one obtains a map like that of Fig.2(d), with a limited number of compact regions. For example, in this map there are just 8 disjoint regions belonging to the same "water basins" class (dark blue).

The region-based coder proposed in [3] uses this kind of segmentation. With reference again to Fig.1, the objects $O_1, \dots O_n$ are now the individual regions, each region is subject to spectral KLT, with *the same transform matrix* for all the regions, then undergoes a two-dimensional transform, the shape-adaptive WT proposed in [7], which makes sense for these compact regions, and is finally encoded them by means of shape adaptive SPIHT [8, 2].

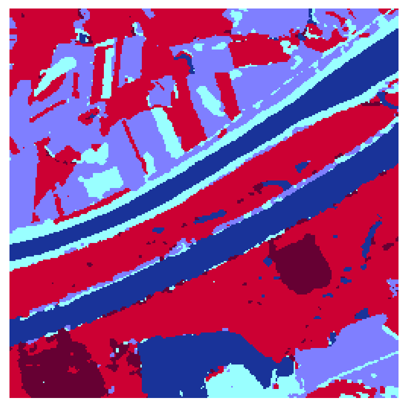
We now want to analyze a coding scheme based on compact regions, like in [3], but with a dedicated KLT for each region, so as to better compact energy in a few bands. It is not obvious in advance that such dedicated transform will be really effective, because the regions they operate on are not statistically homogeneous anymore. In fact, the contextual segmentation algorithm pools together pixels with different statistics just because of their spatial proximity. Even worse, the subsequent regularization is based on a purely geometric processing. A small region can be absorbed in a larger one no matter what their statistics are, and boundary pixels can change region in a similar way. In addition, sending a differ-



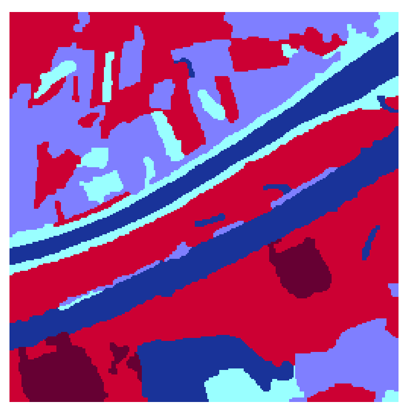
(a)



(b)



(c)



(d)

Figure 2: An example image and its segmentation maps

Row	Transform	C	N	O	Side Info	Energy	E(1)	E(2-6)	%E(1)/E	σ^2_{GM}
1	flat	1	0	1	0.000	5035	4597	438	91.3	50.33
2	block-adap	16	0	16	0.004	3163	2866	297	90.6	33.85
3	class-adap	3	0	3	0.026	1048	773	274	73.8	26.88
4	region-adap	3	0	546	0.175	836	657	179	78.6	17.59
5	”	3	30	167	0.062	935	740	195	79.1	19.45
6	”	3	100	97	0.040	1004	801	203	79.8	20.33
7	”	3	300	59	0.028	1153	938	214	81.4	22.37
8	region/flat	3	300	59	0.015	1153	876	277	76.0	29.83
9	class-adap	5	0	5	0.050	700	459	242	65.5	21.99
10	region-adap	5	0	2041	0.609	417	310	107	74.3	10.13
11	”	5	30	366	0.126	654	507	147	77.5	14.39
12	”	5	100	200	0.076	767	610	156	79.6	15.88
13	”	5	300	92	0.041	1039	848	191	81.6	20.30
14	region/flat	5	300	92	0.021	1039	787	253	75.7	27.12

Table 1: Performance indicators for classified KLT.

ent KLT matrix for each region can have a significant cost in terms of side information which could cancel out the benefits of better energy compaction[1]. It is therefore left to the experiments to establish whether, and in which conditions, classified KLT is an interesting option under the rate-distortion point of view for region-based coding.

3. EXPERIMENTAL RESULTS

All experiments presented here concern a 6-band 512x512 pixel Landsat TM image of a rural area near Lisbon shown in false color in Fig.3. Similar results have been obtained with other images but are not reported for lack of space.

First of all we want to compare some different approaches to the spectral transform, with special attention to the region adaptive KLT proposed here. To this end, in Tab.1 we report some indicators of efficiency, for flat KLT, where the same transform is used for the whole image, block-adaptive KLT, proposed in [4], where the image is partitioned in large blocks and a different transform is used for all blocks, and class-adaptive KLT, proposed in [5], where the image is segmented point-wise in a few classes, with a different transform for each class. As for region-based coding, we show results for the region-adaptive KLT proposed here, where a dedicated transform is used for each region, and for the technique considered in [3], where the image is segmented and later coded in regions, but the same KLT is used for all regions.

Let us begin by comparing the flat and block-adaptive approaches: we see that even a simple segmentation in squares (16 blocks of 128x128 pixels here) allows one to reduce the average pixel energy from 5035 to 3163, just because local block means are used rather than the global one. In both cases, most of this energy (more than 90%) is then compacted in the first KLT band. It must be underlined, however, that such result is more remarkable for the block-adaptive KLT where much of the correlation has been already removed by using the local means. As a figure of KLT efficiency, we should then look at the residual energy left in the other bands E(2-6), which goes down from 438 to 297. This provides a rough prediction of low bit-rate performance, when resources are so scarce that little more than the first band can be encoded, and the E(2-6) energy goes all in the error energy. For a prediction (equally rough) of high bit-rate performance we can look instead at the σ^2_{GM} parameter, that is the geometric mean of the band variances, which also goes from 50 to 34. Of course such improvements are obtained at the price of increased side information, but there are only 16 objects (O) in this case, and the overall cost remains negligible[2]

Turning to the class-adaptive KLT, with 3 classes and hence 3 objects (row 3 in the table) we see that, thanks to a meaningful segmentation, total pixel energy is much reduced, to little more than 1000, and the σ^2_{GM} parameter decreases for 34 to 27. Due to the reduced correlation among residuals, only 74% of such energy is compacted in the first band, leading to a relatively large E(2-6) value. Even though the side information now includes the cost for encoding the segmentation map[3], it is still relatively small, about 0.025 bit/sample. Notice that segmentation quality becomes very important here; for example, with 5 classes (row 9 of the table) performance indicators improve significantly even though the cost of side information doubles.

Now we can analyze the region-adaptive approach (rows 4-7) as a function of two parameters, the number of classes C, and the minimum number of pixels in a region N, which in turn determine the number of objects O, that is regions, in the image. The first case we consider, $C = 3$ and $N = 0$ is not realistic, because regions of any size are allowed here, even single isolated points, but is of interest because it provides limiting results. With respect to class-adaptive transform, there is a limited decrease in total energy, a better compaction of energy in the first band, and especially a much lower σ^2_{GM}. All this, however, cannot make up for the explosion of side information, which reaches almost 0.2 bit/sample, basically because of the many KLT matrices to be sent. Turning to more realistic cases, where regions below 30 (respectively 100 or 300) pixels are absorbed in larger ones, all energy compaction indicators worsen progressively, as expected, but the side information cost is also much reduced. Comparing the case $C = 3, N = 300$ with the reference class-adaptive ap-

[1]Another option, in order to limit the cost of side information, is to consider a single KLT for all regions of the same class. Experimental studies are under way to assess the potential of this approach.

[2]Note that we have used 16 bits for each free parameter to be encoded, but we expect that fewer bits could be used with little or no effects on performance.

[3]Computed as the bit-rate needed by the RAPP algorithm [9].

Figure 3: Test image in false colors

proach, we see that the side information is about the same (there are more KLT matrices but map coding cost decreases) while energy compaction indicator, as a whole, speak in favor of the region-adaptive solution. In more detail, total energy increases, but most of it is compacted in the first band (81.4% as opposed to 73.8%) and the σ^2_{GM} parameter is significantly lower. Finally, within the framework of region-based coding, let us compare the region-adaptive KLT with flat KLT approach, considering for the sake of simplicity only the case C=3, and N=300 (rows 7 and 8). The use of a single KLT matrix allows one to reduce the side information for 0.028 to 0.015, but entails a much lower energy compaction. 24% of the energy spills in the higher bands instead of 19%, and the σ^2_{GM} parameter grows from 22 to almost 30.

Similar considerations, with minor changes, apply to the $C = 5$ cases, reported in rows 9-14 of the table.

To gain a more intuitive insight about the increased compaction ability of the region-adaptive approach we compare, in Fig.4, the eigenimages obtained by flat KLT and region-adaptive KLT (case C=5, N=100). It is obvious that the latter images contain much less energy and that it is much concentrated in the first bands.

This preliminary analysis provides insight, and in general seems to support the use of region-adaptive KLT, but it remain to be proved our indicators correlate well with overall performance. Therefore, to complete this analysis, we report and discuss some actual rate-distortion curves.

In Fig.5, we consider region-based coding with region adaptive KLT for various choices of C and N (rows 5-7 and 11-13 of the table). Results are quite informative. First of all, performance depend clearly on the number of regions, improving when fewer regions are considered, with the case C=3, N=300, and only 59 regions as the best. This is certainly due to the increase in side information, both for the segmentation map and the KLT matrices, but also to the lower coding efficiency of SPIHT when small and irregular regions are considered. It must be also observed, however, that the performance gap all but disappears at higher bit-rates (except for the 366-region case) showing that this approach is viable even when a very accurate map is given, if resources are not too scarce.

Finally, in Fig.6 we compare the performance of the proposed region-adaptive coding algorithm (red), with that of

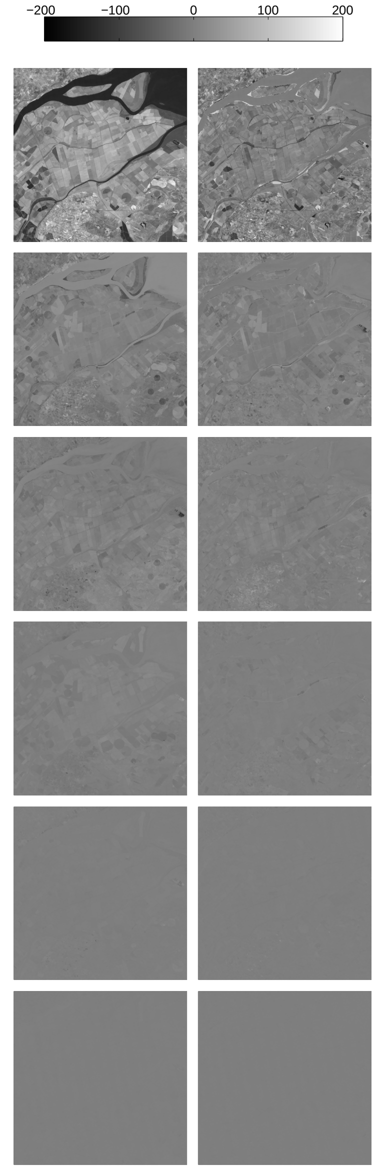


Figure 4: Eigenimages:left, KLT; right region adaptive KLT C=5, N=100

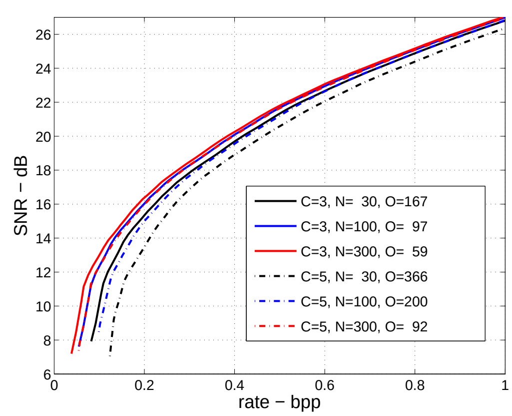


Figure 5: Rate-distortion performance for the region-adaptive coders

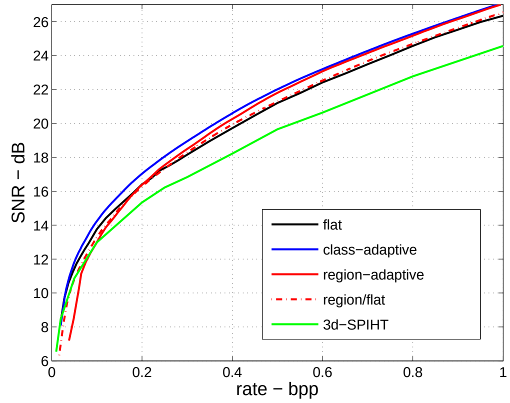


Figure 6: Rate-distortion comparison of several coders

Figure 7: Test image (false color) coded at 0.4 bit/sample

several reference coders, that is: the flat coder (black), the class-based coder (blue), the region-based coder with the same KLT for all regions (red, dashed), and finally the 3-d SPIHT flat coder available on the web [10] (green). All region or class based techniques make reference to the same segmentation map obtained with C=3 and $N = 300$ parameters. First of all, 3d-SPIHT loses 2-3 dB w.r.t. all other techniques, but this was expected since there is no explicit resource allocation among the transform bands, a major cause of performance impairment. On the other hand, it is by far the simplest technique. The class-adaptive coder is consistently the best, about 0.7 dB better than the flat coder at high-rates, since it exploits class information to adapt the KLT and improve energy compaction, but requires little side information. Region-based coders have intermediate performances: when a region-adaptive KLT is used, there is an initial loss at low rates because of the side information, but at high rates performance is very close to the best. A dual behavior is observed for the region-based coder with flat KLT which, all in all, is very close to the flat coder.

Fig.7 shows finally the test image encoded by the region-adaptive coder at 0.5 bit/sample, with the same false-color representation as Fig.3. Beyond the objective performance, it is reassuring to verify that no visible artefact is present, which is an important requirement in view of subsequent automatic elaborations of such images.

4. CONCLUSIONS

In summary, both the indicators of the compaction ability and the actual rate-distortion results seem to indicate that the adaptive KLT is a convenient option in the framework of region-based coding of multispectral images. More important, the region-based approach itself is definitely viable in terms of rate-distortion performance, as it guarantees performance very close or superior to that of the reference techniques. In addition, it allows one to describe and encode the image in terms of elementary regions, which can be useful for a number of applications. Although similar results have been obtained in other experiments not reported here, it must be underlined that performance depends significantly on the statistics of the image, like its spatial and spectral correlation, on its parameters, like size, number of bands, etc., and on the type of segmentation that is carried out, especially number and size of regions. In any case, there is certainly space for further improvements, which depend mainly, in our opinion, on the crucial step of segmentation.

REFERENCES

[1] M.Petrou, P.Hou, S.Kamata, C.I.Underwood, "Region-based image coding with multiple algorithms", *IEEE Transactions on Geoscience and Remote Sensing*, pp.562-570, March 2001.

[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] J.A.Saghri, A.G.Tescher, J.T.Reagan, "Practical transform coding of multispectral imagery", *IEEE Signal Processing Magazine*, pp.32-43, Jan. 1995.

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

[6] 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.

[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] X.Tang, W.A.Pearlman, J.W.Modestino, "Hyperspectral image compression using three-dimensional wavelet coding: a lossy-to-lossless solution", submitted to *IEEE Transactions on Geoscience and Remote Sensing*, available at http://www.cipr.rpi.edu/ pearlman/papers/tgrs04_tpm.pdf.

[9] 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.

[10] http://www.cipr.rpi.edu/research/SPIHT/