A COMPARISON OF FLAT AND OBJECT-BASED TRANSFORM CODING TECHNIQUES FOR THE COMPRESSION OF MULTISPECTRAL IMAGES

M. Cagnazzo, G. Poggi, L. Verdoliva

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

ABSTRACT

In this work we implement and compare several state-of-the-art transform coding schemes for the compression of multispectral images, in order to better understand which elements have a deeper impact on the overall performance, and which tools guarantee the best results. All schemes are based on Karhunen-Löeve transform and/or Wavelet Transform, in various combinations, and use SPIHT as the coding engine. Moreover, besides the ordinary techniques, their object-based counterparts are also examined, so as to study the viability of such approach [1] for these images. Whenever possible, an optimal rate allocation strategy is applied. The experiments, performed on images acquired by two different sensors, highlight the superiority of KLT as spectral transform; the rough equivalence between object-based and ordinary techniques in terms of rate-distortion performance; and the importance of the optimal allocation.

1. INTRODUCTION

A Multispectral (MS) image is composed by several bi-dimensional images or *bands*, representing the same area of the Earth surface in different spectral intervals. Each MS image thus corresponds to a huge amount of data, to be transmitted to the ground station, stored and disseminated to the final users, and is therefore evident the need for efficient compression techniques. One could resort to general-purpose encoding techniques but, since they do not take into account the peculiar characteristics of MS images, they can end up providing far sub-optimal performances. As a consequence, many algorithms have been specifically designed for MS images in latest years, e.g., [2, 3, 4]. Many of them are based on a three-dimensional (3D) transform in order to exploit not only the *spatial* correlation among data, but also the inter-band or *spectral* correlation. In the spectral domain, both Discrete Wavelet Transform (DWT) and Karhunen-Löeve Transform (KLT) have been used, while DWT and the Discrete Cosine Transform (DCT) are the most popular choices in the spatial domain, KLT being too complex in this case. After the transform, the coefficients have to be quantized and encoded, and again several techniques have been considered to this end, like JPEG2000 or SPIHT and its variations. Finally, in all possible coding schemes, a pivotal role is played by the rate allocation strategy which allows one to spend the encoding bits in the most effective way.

More recently, object-based (OB) methods have begun to appear [1], where the image is segmented in a number of regions (or objects) which are then encoded individually. This approach arises naturally considering that these images are often destined to data mining or classification or other region-oriented processing tools for which a native object-based image description is highly convenient. Of course, the OB approach requires Shape-Adaptive (SA) spatial transforms and SA encoding techniques, which can be less effective than their non-SA or *flat* counterparts. Nonetheless, the region-based coding paradigm can be appealing even in a rate-distortion sense. In fact, when the image is composed of a limited number of well defined, homogeneous regions, it can be advantageous to segment it beforehand, so as to remove and encode separately the high frequency components related to the boundaries, and encode the region textures more efficiently in a following pass. Finally, OB coding lends itself also to unequal resource assignment strategies, in which some regions (of greater interest for the user) are assigned more resources than the others.

In this work we implement and compare several state-of-the-art transform coding schemes for the compression of multispectral images, in order to better understand which elements have a deeper impact on the overall performance, and which tools guarantee the best results. All schemes are based on KLT and/or DWT, in various combinations, and use SPIHT as the coding engine, given its simplicity and high efficiency. Moreover, besides the flat techniques, their OB counterparts are also examined in the same operating conditions, so as to study the viability of the approach for these images. Whenever possible, an optimal rate allocation

This work was partially supported by the *Centro Regionale di Competenze ITC* of Benevento, Italy

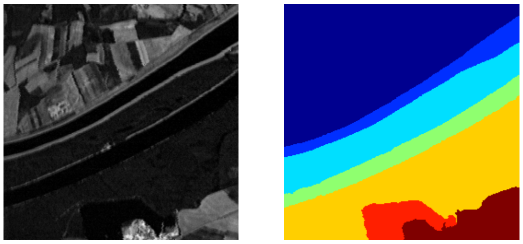

Fig. 1. One band of a MS image and its segmentation map

strategy is applied.

In next Section we briefly analyze our coding tools to point out some relevant properties, Section 3 presents the implemented coding schemes, and Section 4 reports on the experimental results and draws conclusions.

2. COMPRESSION TOOLS

2.1. Segmentation and Map Coding

When OB techniques are considered, the quality of segmentation plays an important role. Here, we resort to a recently proposed algorithm [5] based on the Tree-Structured Markov Random Field model, which proved to be both fast and accurate. In addition, in order to remove small regions produced by segmentation and to smooth object contours, a further morphological processing of the map is performed. One band of a MS image used in the experiments is shown in Fig.1. together with the segmentation map obtained with the proposed algorithm.

It must be pointed out that the segmentation map must be transmitted as a side information in order for the decoder to work properly. Of course this represents an overhead which, however, is usually quite limited, all the more considering that a single map is used for many bands. We resort to a simple context-based arithmetic encoder, achieving a coding cost of roughly 0.04 bit per pixel for the map of Fig.1.

2.2. Transforms and Encoding Techniques

Both the Karhunen-Löeve Transform (KLT) and the Discrete Wavelet Transform (DWT) have been employed for the spectral decorrelation of MS images. The KLT is theoretically optimal, but it is data-dependent, and hence requires some computation and the transmission of the transform matrix, namely, an additional coding cost. Both problems are negligible if only a few bands are considered while can become significant for larger groups. It must be also pointed out that the optimality of KLT is somewhat reduced by the fact that data are typically non-stationary, and hence the transform matrix is a weighted average over different areas. In these conditions, the DWT, which is simpler and data independent, might well be a reasonable alternative.

As for the spatial transform, KLT becomes prohibitively complex and is no longer a viable alternative, so we consider only the DWT (which has proved to be very effective in image coding) neglecting the DCT which does not seem to be competitive. In the case of OB schemes, the transform must be able to operate on regions of arbitrary shape, so we use the Shape-Adaptive (SA) version of the DWT proposed by Li and Li [6],

For the quantization and encoding of the coefficients we use SPIHT [7], a popular zerotree-based compression technique, which is simple, efficient, and scalable. In addition, we have already implemented and tested [1] a SA version of SPIHT which therefore allows us to conduct a meaningful comparison between flat and object-based techniques.

2.3. Rate Allocation

Resources must be allocated to different *objects* constituting the MS image. The definition of object here is extended to encompass several cases. A *2D-region* is a set of pixels belonging to a single frequency band and having the same label in the segmentation map; all the 2D-regions of the MS image corresponding to the same label form a *3D-region*; we will consider also all the 2D-regions in the same band as an object. Therefore, in the following, we will generically refer to objects which can actually be bands, 2D-regions or 3D-regions.

Rate allocation (RA) could well be driven by the application or even by the user itself in order to privilege certain areas over the others. In this work, however, we take as our goal only the minimization of the distortion, and hence resort to the optimal Lagrangian approach, progressively allocating resources to the various objects.

The optimal allocation is obtained at rates corresponding to the same slope on the rate-distortion curves of each object. This result holds independently from the nature, dimensionality and spatial shape of the objects. Using the SPIHT algorithm, which provides a naturally embedded coding, it is quite easy to develop an algorithm which achieves the optimal allocation.

3. IMPLEMENTED CODING TECHNIQUES

With the tools described in previous Sections, we can assembly several encoding techniques. They differ for the spectral transform, the spatial transform, the encoding algorithm, the dimensionality of objects, the rate allocation (RA) strategy. We considered the techniques listed in Tab. 1 whose main characteristics we summarize in the following:

1. **3D WT + 3D SPIHT.** This is the ordinary 3D SPIHT, and is considered as the reference technique. A three-

Spatial Transform	Spectral Transform	Object Dimension	Encoding Technique	Rate Allocation
2D WT	WT	3d	SPIHT	N/A
	KLT	3d	SPIHT	N/A
		2d	SPIHT	band
2D SA WT	WT	3d	SA-SPIHT	3D-region
	KLT	3d	SA-SPIHT	3D-region
		2d	SA-SPIHT	2D-region

Table 1. Summary of encoding techniques

dimensional WT is performed, using Daubechies 9/7 filters in the spatial domain and Haar filters in the spectral domain. The WT coefficients are encoded with the 3D version of SPIHT.

2. **1D KLT + 2D WT + 3D SPIHT.** With respect to the previous technique, we change the spectral transform, resorting to the KLT, but keep using 3D SPIHT for encoding. Of course, SPIHT was originally developed for tree-structured data such as those provided by the DWT, so it is likely not really suited in this case.

3. **1D KLT + 2D WT + 2D SPIHT + band RA.** This scheme arises from the considerations made above: since DWT is not used in the spectral domain, we renounce using SPIHT along the bands. After the KLT, each transformed band, or eigenimage, is spatially transformed by DWT and encoded by 2D SPIHT. The encoding rate of each “band” is then decided by the RA algorithm.

4. **1D WT + 2D SA-WT + 3D SA-SPIHT + 3D-region RA.** This is the first OB scheme. The difference with respect to the reference scheme is that both 3D WT and 3D SPIHT take place on 3D regions rather than on the whole image and therefore their SA versions must be used. The RA algorithm of course will work on 3D regions.

5. **1D KLT + 2D SA-WT + 3D SA-SPIHT + 3D-region RA.** The difference with respect to scheme 4 is only in the spectral transform.

6. **1D KLT + 2D SA-WT + 2D SA-SPIHT + 2D-region RA.** In this scheme, the objects are very finely defined, since they are 2D regions, and therefore RA must handle a very large number of objects and plays a central role.

4. EXPERIMENTAL RESULTS

The first experiment is carried out on a MS image, acquired by the GER sensor, composed by 16 bands of 256 × 256

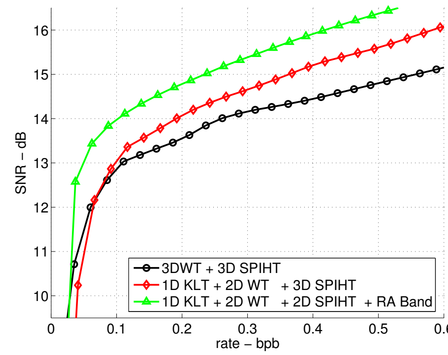

Fig. 2. Performance of flat coding schemes on GER images

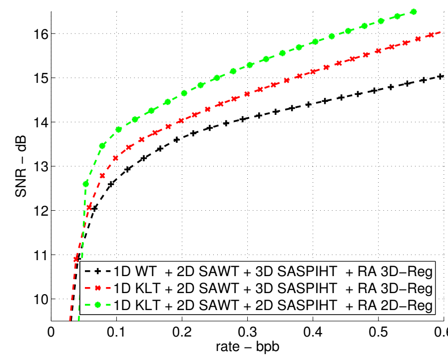

Fig. 3. Performance of OB coding schemes on GER images

pixels. A sample band and the segmentation map are shown in Fig.1. To obtain a good encoding efficiency, we use 5 levels of WT decomposition in the spatial domain (Daubechies 9/7 filters), and 4 in the spectral domain (Haar filters). After encoding all objects, when applicable, the optimal rate allocation procedure is carried out.

In Fig. 2 we report the rate-distortion results for the flat (non-OB) techniques 1, 2 and 3, using SNR as a quality measure of decoded images.

From this first experiment, we see that the spectral transform accounts for a remarkable difference in performance. Using the KLT in the spectral direction instead of the DWT provides a gain of up to 1 dB (observed between techniques 1 and 2 which only differ in the spectral transform). We observe also that coding bi-dimensional bands with an optimal resource allocation among them provides a further advantage (about 0.8 dB between techniques 2 and 3), but this advantage seems to fade out at higher bit-rates and, as a matter of fact, further experiments prove that at high bit-rate, the 3D technique performs better than the 2D version. It is worth underlining again that KLT is only applicable when a small number of spectral bands have to be encoded, as in our case, otherwise its computational cost and side information become prohibitive. From this first experiment

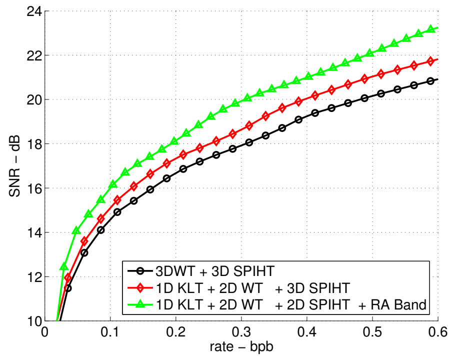

Fig. 4. Performance of flat coding schemes on TM images

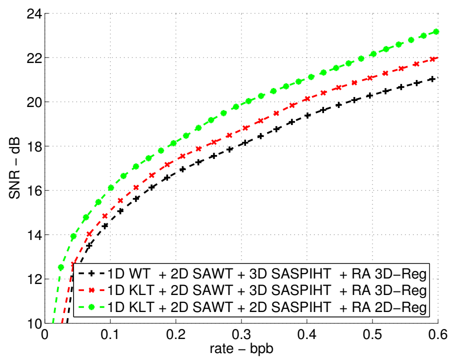

Fig. 5. Performance of OB coding schemes on TM images

we observe that just a few changes, as the use of KLT and optimal resource allocation, allow us to clearly outperform a state-of-the-art technique like 3D-SPIHT.

Let us now turn to the OB techniques, 4–6, whose performances are shown in Fig.3. Here the cost of encoding the segmentation map must be taken into account, but it just 0.040 bit/pixel, which amounts to a negligible 0.0025 bit/pixel on a per-band basis. In this case as well we see that the the spectral transform heavily affects performances, with KLT being clearly better than DWT. Indeed there is a difference of about 1 dB between techniques 4 and 5 which differ only under this point of view. As far as rate allocation is concerned, we note, once again, that the coding performance improves when it is possible to determine each band bit-rate (*i.e.* using 2d-regions as objects): technique 6 outperforms technique 5. However, in this case the gap is smaller, about 0.7 dB.

A comparison between the best flat and the best OB techniques (both using spectral KLT and per-band rate allocation) shows a slightly advantage for the former at all rates, but only about 0.2 dB.

A similar experiment was repeated on a 4-band Landsat TM image of 512 × 512 pixels. Results are reported in Fig. 4 and Fig. 5 and, apart from the obvious difference of scale, they are qualitatively the same as before, except that, in this case, the best OB technique is slightly better than the best flat one.

To summarize this analysis, it seems clear that, whenever possible, KLT should be used as a spectral decorrelating transform. It is also confirmed that a suitable rate allocation is able to provide non-negligible advantages even when compared with finely tuned and tested algorithms such as SPIHT and its SA version. Finally, the OB coding schemes, even neglecting their added value in terms of further processing steps, are already comparable with their flat counterparts, and can be probably further improved by working, for example, on segmentation or on the use of adaptive coding.

5. REFERENCES

[1] M. Cagnazzo, G. Poggi, L. Verdoliva, and A. Zinicola, “Region-oriented compression of multispectral images by shape-adaptive wavelet transform and SPIHT,” in *Proceed. of IEEE Intern. Conf. on Image Proc.*, Singapore, Oct. 2004, pp. 2459–2462.

[2] J.A. Saghri, A.G. Tescher, and J.T. Reagan, “Practical transform coding of multispectral imagery,” *IEEE Signal Processing Mag.*, pp. 32–43, Jan. 1995.

[3] P.L. Dragotti, G. Poggi, and A.R.P. Ragozini, “Compression of multispectral images by three-dimensional spiht algorithm,” *IEEE Trans. Geosci. Remote Sensing*, pp. 416–428, Jan. 2000.

[4] X. Tang, W.A. Pearlman, and J.W. Modestino, “Hyperspectral image compression using three-dimensional wavelet coding: A lossy-to-lossless solution,” *IEEE Trans. Geosci. Remote Sensing*, 2004, Submitted.

[5] C. D’Elia, G. Poggi, and G. Scarpa, “A tree-structured Markov random field model for bayesian image segmentation,” *IEEE Trans. Image Processing*, pp. 1259–1273, Oct. 2003.

[6] S. Li and W. Li, “Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding,” *IEEE Trans. Circuits Syst. Video Technol.*, pp. 725–743, Aug. 2000.

[7] A. Said and W. A. Pearlman, “A new, fast and efficient image codec based on set partitioning in hierarchical trees,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243–250, June 1996.