AN UNSUPERVISED SEGMENTATION-BASED CODER FOR MULTISPECTRAL IMAGES

M.Cagnazzo, L.Cicala, G.Poggi, L.Verdoliva

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

ABSTRACT

To fully exploit the capabilities of satellite-borne multi/ hyperspectral sensors, some form of image compression is required. The Gelli-Poggi coder [1], based on segmentation and class-based transform coding, has a very competitive performance, but requires some a-priori knowledge which is not available on-board. In this paper we propose a new version of the Gelli-Poggi coder which presents about the same performance than the original but is fully unsupervised, and is therefore suited for use on-board a satellite. Numerical experiments on test multispectral images validate the proposed technique.

Key-words: Multispectral image coding, region-based coding, on-board implementation.

1. INTRODUCTION

The performance of satellite-borne sensors increases ever more in terms of spatial resolution, radiometric accuracy, and number of spectral bands. All these aspects, and especially the latter, contribute to increase the data volume that such sensors must transmit to the ground station to the point that the required data rate largely exceeds the available channel capacity and large chunks of data must be simply discarded. To avoid this loss one can resort to data compression which allows one to reduce the data volume by one/two orders of magnitude without serious effects on the image quality and on their diagnostic value for subsequent automatic processing. To this end, however, one cannot resort to general purpose techniques as they do not exploit the peculiar features of multispectral remote-sensing images, and in fact several ad hoc coding schemes have been proposed in recent years, e.g., [1, 2, 3, 4].

One of the most promising such schemes, based on classified transform coding, is the Gelli-Poggi coder originally proposed in [1]. The image is first segmented, so that each pixel is associated with one of a given number of classes based on its spectral response vector. Then, all vectors of the same class are grouped together and compressed by means of transform coding techniques. This way, transform coding operates on stationary homogeneous sources, thereby maximizing its efficiency, and leading to an excellent overall rate-distortion performance, which is in fact superior to that of other state-of-the-art coders.

The Gelli-Poggi coder, however, relies heavily on a-priori information which is hardly available to both encoder and decoder, and makes the coder unsuited for compression on-board a satellite before transmission to the ground station. In this paper we address this problem, by suitably modifying the various steps of the original coder in order to obtain more practical coding schemes suited for on-board operations. Next Section describes the Gelli-Poggi coder in detail, highlighting its weak points. Section 3 presents the various improvements proposed and Section 4 assesses the performance of the various alternative schemes by means of numerical experiments on test multispectral images. Finally Section 5 draws conclusions.

2. THE GELLI-POGGI CODER

The coding scheme is articulated in three main steps as shown in Fig.1:

1. image segmentation;
2. lossless coding of the segmentation map;
3. lossy coding of the radiometric information.

The original scheme is fully supervised, meaning that all statistical parameters are computed in advance from a training set. In the following, these three steps are described in more detail.

Segmentation amounts to a simple spectral clustering. Specifically, each pixel is classified by computing the Euclidean distance between its spectral vector and a set of template vectors, one for each class, and assigning the pixel to the minimum distance class. The set of template vectors can be viewed as a VQ codebook, computed off-line on a suitable training set, and the segmentation itself as a vector quantization. In particular, to limit computation complexity, the VQ codebook is tree-structured so that only a few binary comparisons are needed.

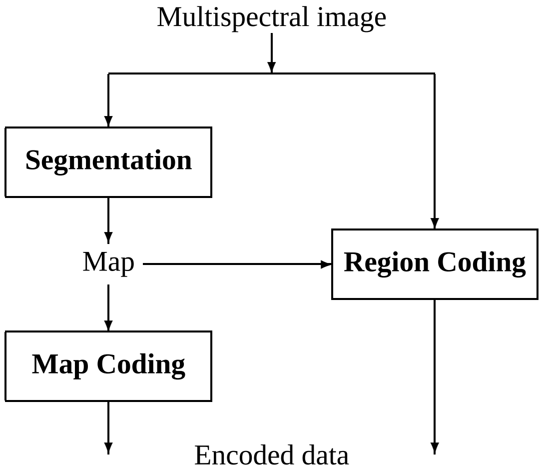

Fig. 1. Block diagram of the coding scheme.

The map of class indexes, *i.e.* the segmentation map resulting from VQ, must be sent to the decoder as a side information. Since neighbouring pixels are highly correlated the map is significantly compressed, without loss of information, by resorting to a predictive scheme followed by Huffman coding, with the code computed on the training set as well.

Using the selected template vectors for every pixel instead of the original spectral vectors, we have a first VQ approximation of the multispectral image. The difference between the original image and the VQ approximation (called *residual image*) is compressed by means of transform coding. First, a classified Karhunen-Löeve Transform (KLT) is performed along the spectral dimension. In order to account for class information, a different transformation matrix for each class is derived off-line from the training-set. Then, a Discrete Cosine Transform (DCT) is used to decorrelate the spatial information within each transformed band. Finally, each transform coefficient is sorted by spectral class, KLT band and DCT frequency, and is included in a quantization set which is quantized by a specific tree-structured Lloyd-Max quantizer designed off-line on the training set. Rate allocation is decided on-line with a greedy bit allocation algorithm.

3. THE UNSUPERVISED VERSION

The obvious weakness of the original Gelli-Poggi coder in view of on-board implementation is that several pieces of information are supposed to be known in advance, that is

- the VQ classifier;
- the class-adapted KLT matrices;
- the set-adaptive Lloyd-Max quantizers.

We will therefore abandon these hypotheses and consider an alternative coding scheme in which all needed parameters are designed on-line based on the very same data to be encoded. Of course, this entails an increase in both the computational complexity and the side information to be transmitted along with the quantized coefficients. We will examine the new steps in turn under these two points of view.

3.1. The VQ classifier

The design of a VQ codebook can be extremely demanding in terms of CPU power, but since only a limited number of land covers are typically present in a given image we are interested in a very small codebook (e.g., 4 to 20 classes [5], which largely reduces computation time. In addition, our codebook is tree-structured, which further reduces both design and segmentation complexity. Finally, the design need not be carried out on all the data to be encoded, but only on a sample subset, which can be as small as a few thousands of spectral vectors, although extreme subsampling can produce some performance losses. All in all, computational complexity is probably not an issue for the VQ classifier.

As for the side information, a tree-structured VQ codebook for C classes is composed of 2C-1 vectors, with B components each if B is the number of bands in the image. For images in the order of 1 Mpixel, and coding rates not unreasonably small, this cost is always negligible, even if 16 bits were spent to encode each vector.

3.2. The class-adapted KLT matrices

To compute a KL transform matrix, we must first estimate the $B \times B$ correlation matrix of the data, and then compute its eigenvectors. Since we use class-adaptive KLT, we need C such matrices, one for each class.

The estimation part is not extremely demanding, especially if we resort again (with due care) to some subsampling of the training data. Computing the eigenvectors, instead, can require a significant time, which grows as the third power of the number of bands. This can become a problem if B is very large. On the other hand, if B is large, almost all of the image energy is compacted in the first few transform coefficients, to the point that the less significant coefficients are assigned no bits at all. This suggests us to resort to low-complexity iterative techniques, such as the power method, to compute the B' most relevant eigenvectors which comprise almost all the energy (say 99.9%). This condition can be tested on-the-fly, and helps limiting complexity in critical cases.

Concerning the side information, for each KLT matrix we must send $B \times (B+1)/2$ parameters in the conventional case, and approximately $B' \times B$ coefficient in the

Fig. 2. Band 24 of the test image.

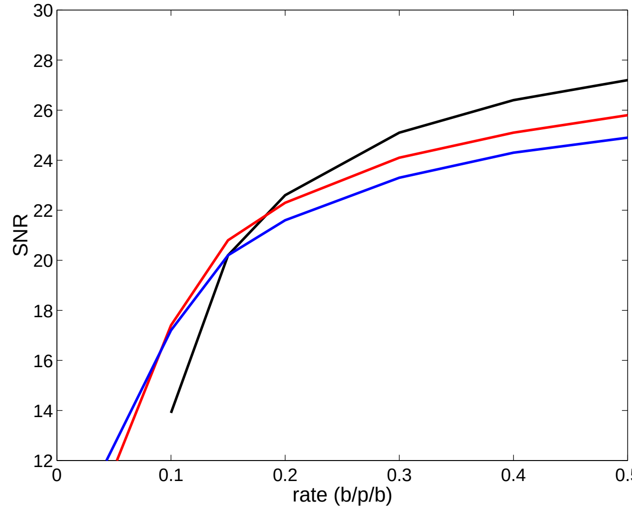

Fig. 3. Gelli-Poggi vs. wavelet-based coders.

reduced dimensionality version. In some non-typical conditions (small images, very low coding rates, many classes, many bands) this could become significant and some care must be therefore taken to encode all parameters with as few bits as possible, provided no significant performance loss is provoked.

3.3. The set-adaptive quantizers

The problem, here, is that a very large number of quantizers are needed, $C \times B \times K$ in the most general case, with K the DCT vector length. In fact, an ad hoc quantizer is used for the first DCT coefficient of the first KLT band of the first class, another one for the second DCT coefficient of the first KLT band of the first class, and so on. Even considering that most of these sets of coefficients will be assigned no encoding bits, and no information need be transmitted for them, so many quantizers remain to be designed and transmitted that this approach becomes clearly unreasonable. We resort therefore to parametric quantizers: each set of coefficients is modeled as a zero-mean generalized Gaussian, characterized by its variance and shape parameters, which univocally identify the optimal quantizer. To preserve the scalability of the original scheme however, we designed embedded quantizers and we made them midtread to increase robustness, so that the tree structure has, at every depth level, one ternary node besides the binary ones. Exception is made only for the (1,1) coefficients of each class (*i.e.* the DCT low frequency coefficients of the first KLT band), for which the Lloyd-Max algorithm keeps being used. The coefficient variances are then used to perform rate allocation by means of the Huang-Schultheiss algorithm [5]. The information to be sent is therefore composed of three sets: the active/inactive bits for each set, the pdf parameters for the active sets, and the Lloyd-Max quantizers for the baseband coefficients.

Of course, parametric quantizers cannot guarantee the same performance of the optimal Lloyd-Max, but complexity and side information are now fully manageable.

4. EXPERIMENTAL ANALYSIS

All experiments presented here are performed on a hyperspectral image acquired by the GER airborne sensor which portrays an agricultural area in Germany near the river Rhein. In particular, we use a square region of 512x512 pixels, and select only 8/16 bands that have a constant spectral resolution of 25.4 nm, and 9 bit of meaningful information. A sample band of the test image is shown in Fig. 2.

In the first experiment we analyze the absolute performance of the Gelli-Poggi coder by comparing it with two state-of-the-art coders, one based on 3d-wavelet transform followed by 3d-SPIHT [3], and the other based on spectral KLT followed by JPEG-2000 on the transform bands, with optimal rate allocation. For the Gelli-Poggi coder, all needed parameters are supposed to be known a-priori and are actually evaluated on another 512x512 section of the same GER image. The rate-distortion curves are reported in Fig.3, and show that the Gelli-Poggi coder (black) is fully competitive with the two wavelet-based coders, outperforming them by more than 1 dB at rates beyond 0.3 b/p/b (bit/pixel/band).

In the second experiment we address the effect of designing on-line the tree-structured VQ codebook. We use the well-known splitting algorithm, and build an unbalanced binary tree by splitting the vector that contributes most to distortion. Each split is carried out by binary GLA, which convergence after a few iteration. The training set is obtained by regular sampling of the test image, and its size ranges from 1024 to 16384 vectors. In all cases the computation time turns out to be affordable but, what is more important, the overall performance is very mildly affected by the codebook size, with a loss within 0.1 dB in the worst case (for this reason curves are not reported), and therefore we will proceed with the training set of size 1024[1].

Let us now consider the KLT transforms. With 8 bands the computation time is still affordable, but it grows quite rapidly, and with 16 bands it entails already a 40% increase in the overall encoding time, indicating that with more bands one must resort to some more efficient computation strategy, like the power method mentioned above. As for the encoding cost of the matrices, with 8 bands, even in the worst case of 20 classes, and using 16 bits for each coefficients, it is less that 0.005 b/p/b. It then increases linearly with the number of bands, but it seems safe to say that it remains always negligible for the coding rates of interest in remote-sensing applications. In addition, since the matrices are designed on the same data they are used on, there is a significant gain in the compactation ability, which more than compensates the increased rate.

Let us finally turn to the quantizers. In this case a CPU-time comparison is meaningless, as the on-line design and especially the transmission of all Lloyd-Max quantizers is simply not affordable. On the contrary, the estimate of the variance of each set (we use Laplace quantizers and therefore do not estimate the shape parameters) increases the CPU time of a few percents, and is then fully acceptable.

The increase in side information is also very limited. With 20 classes, 8 bands, and 64-point DCT, we have a total of 10240 sets of coefficients, most of which however will be assigned 0 bit for quantization. Therefore we must send 10240 bit to signal activity, followed by, in a typical case, 1000 variances for the active quantizers, and 20 full-fledged Lloyd-Max quantizers for the baseband coefficients of each class. All in all, this amounts again to little more than 0.01 b/p/b, and does not increase with the number of bands. Under the rate-distortion point of view, the use of parametric quantizers turns out to be even advantageous. Fig.4 compares the rate-distortion performance of the supervised Gelli-Poggi coder (solid line) and of the new fully unsupervised version (dashed line). It can be seen that the two curves are almost coincident, but for very small rates, where the supervised version performs slightly better.

5. CONCLUSION

We set to implement an unsupervised version of the Gelli-Poggi coder for multispectral images with the goal of making it suitable for use on-board a satellite and thus reduce the problems encountered in the transmission to the ground stations. Although experiments have not been extensive thus far, they are very encouraging. Of course, the overall encoding time increases, but never more than 60% w.r.t. the original supervised coder in the cases considered. In addition, the distortion-rate performance is essentially unaffected since the increase in side information is balanced, at least at higher rates, by the improved quality of encoding. Under this point of view, it should be also considered that typical remote-sensing images are larger than the 512x512 section considered here, which goes in the direction of further reducing the cost of side information.

[1] Of course, if one is willing to use much larger codebooks, the training set size must be adjusted accordingly.

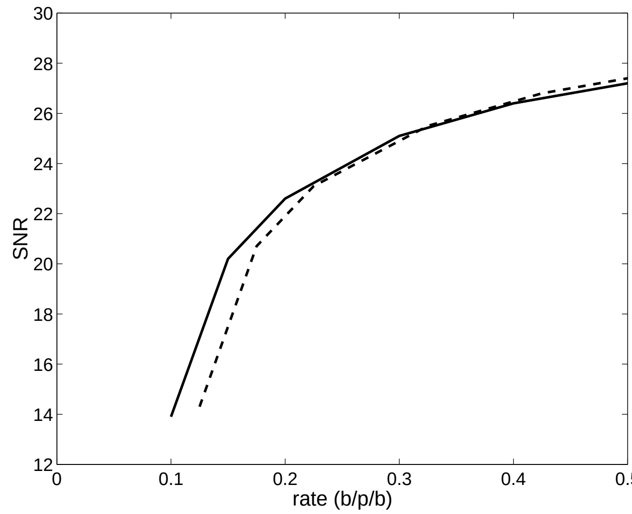

Fig. 4. Supervised vs. unsupervised coders.

6. REFERENCES

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

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

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

[4] H. S. Lee, N. H. Younan, R. L. King, “Hyperspectral image cube compression combining JPEG-2000 and spectral decorrelation”, *2002-IEEE IGARSS*, Toronto (Canada), June 2002, pp.3317-3319.

[5] A. Gersho and R. M. Gray. *Vector quantization and signal compression*. Kluwer Academic Publisher, Boston, 1992.