Compression of multitemporal remote sensing images through Bayesian segmentation

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

Dipartimento di Ingegneria Elettronica e della Telecomunicazioni

Università Federico II di Napoli, via Claudio, 21 – 80125 Napoli, Italy

{cagnazzo, poggi, giscarpa, verdoliv}@unina.it

Abstract— **Multitemporal remote sensing images are useful tools for many applications in natural resource management. Compression of this kind of data is an issue of interest, yet, only a few paper address it specifically, while general-purpose compression algorithms are not well suited to the problem, as they do not exploit the strong correlation among images of a multitemporal set of data.**

Here we propose a coding architecture for multitemporal images, which takes advantage of segmentation in order to compress data. Segmentation subdivides images into homogeneous regions, which can be efficiently and independently encoded. Moreover this architecture provides the user with a great flexibility in transmitting and retrieving only data of interest.

I. INTRODUCTION

Multitemporal images allow one to follow the evolution in time of a given region of interest by means of change detection techniques [1-3], and therefore represent valuable tools for natural resource management. For many advanced applications the basic data-unit of interest becomes a set of multispectral or hyperspectral images acquired at different times. The transmission and archival of such a huge amount of data is a very demanding task, despite the constant improvement of storage and communication media, and some form of data compression is often desirable or necessary. General-purpose image compression techniques, such as JPEG2000, are not suitable in this case, as they neglect important information about the source, especially the strong functional dependence among the various spectral bands and temporal images.

Despite the relevance of this problem, only a few papers address the compression of multitemporal images in the literature [4,5], with approaches not really tailored to the task. More relevant to the problem is the literature on the compression of multispectral images, where techniques for the joint analysis and compression of images are currently under investigation [6,7].

Segmentation-based compression techniques, in particular, seems very promising for several reasons. First of all, segmentation can single out regions where significant changes have occurred; this information is embedded in the encoded stream and represents a very valuable information for many applications. Turning to compression efficiency, segmentation allows one to encode each region independently (object-based coding) and therefore to devote more encoding resources to regions of interest (*e.g.* those where changes have occurred), adapt encoding parameters to local features of the region

(adaptive coding), or even select different encoding techniques for different regions (dynamic coding). Finally, for multispectral multitemporal images, the additional cost of encoding the segmentation map is shared by all bands, and hence it is typically negligible.

Given all these considerations, we propose here a coding scheme for multitemporal images based on the following major steps:

1) the images collected at times $t_1, t_2, \ldots, t_N$, are jointly segmented;
2) based on the analysis of the segmentation maps, changed and unchanged regions are detected;
3) the segmentation maps are jointly encoded without loss of information;
4) the region textures are lossy coded independently, with rate allocation based on the results of change detection.

Segmentation is carried out by means of a Bayesian technique, based on a TS-MRF (tree-structured Markov random field) image model, recently developed and assessed in [8,9]. Texture compression is carried out by means of shape-adaptive wavelet transform [10] followed by a shape-adaptive version of the SPIHT algorithm [11] to quantize transform coefficient.

The paper has the following structure: in Section 2 we provide more detail on the processing tools used for segmentation and coding, while in Section 3 we describe the experimental setting and discuss numerical results; finally Section 4 draws conclusions.

II. SEGMENTATION AND CODING TOOLS

Our compression algorithm comprises four major steps

1) supervised classification of the multitemporal image;
2) extraction of a simple change detection map;
3) lossless encoding of the map;
4) lossy coding of the image regions.

that will be briefly described in the following.

To make description more concrete and easy to understand, we refer already to the data that will be used in the experiments. We work on two Landsat TM images (only optical bands) of an agricultural area near the river Po in Italy taken in April and May 1994 and accurately registered. The images are 494x882 pixel but we take a smaller square region of 512x512 pixel (some white lines added at the margin) to speed up processing. Fig. 1 and Fig. 2 show band 3 of the selected area in April and in May.
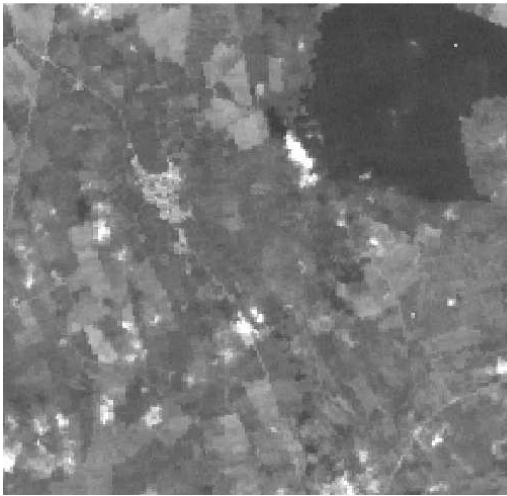
Fig. 1. Band 3 in April

Fig. 2. Band 3 in May

Together with the image, we also have information about which classes were present at both times (wet rice fields, cereals, wood, and bare soil in April, both wet and dry rice fields, cereals, wood, and corn in May) and some patches where ground truth data, involving all classes of interest, were collected at both times.

A. Classification

As said before, to carry out segmentation we resorted to a Bayesian approach, so as to take into account spatial dependencies through the suitable definition of the image prior. In particular, we use the TS-MRF model proposed in [12] which provides several advantages over non-structured models, and especially allows us to organize a priori the classes in a meaningful way through a binary tree. Based on the available information, we chose the tree of classes shown in Fig. 3. If we consider the tree only down to the second level, we find
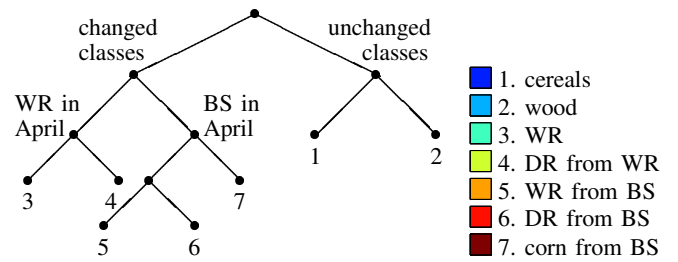

Fig. 3. Tree structure used for data fitting (left) and classes (right). The classes 4-7 represent land changes from April to May (WR = wet rice fields; DR = dry rice fields; BS = bare soil).
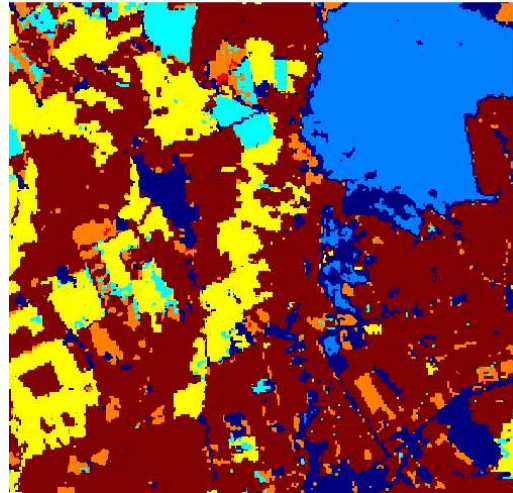

Fig. 4. Segmentation map provided by TSMRF algorithm

the four classes present in April. Then new nodes are added to account for changed classes, corresponding to regions that change label from April to May. The right part of the tree does not grow further, as the cereals and wood areas do not change in May. The left part, on the contrary, changes radically: the wet-rice node is split in two, wet-rice (unchanged) and dry-rice, and the bare soil class disappears altogether, giving way to wet or dry rice and to corn. All in all, combining information from April and May, seven classes are singled out, five of them changed and the rest unchanged.

Given this semantic tree and the class statistics (mean vectors and covariance matrices), the classification algorithm works, for each pixel, by descending the classification tree down to a leaf, and thereby refining progressively its label, taking into account both observed data and the labels of neighboring pixels. A detailed description goes beyond the scope of this work (the reader is referred to [8] and [9] for details); the final result on our image is shown in Fig. 4.

B. Change detection map

Given the classification of Fig. 4, the creation of a change detection map is immediate. Without any further processing, however, such a map would contain a very large number
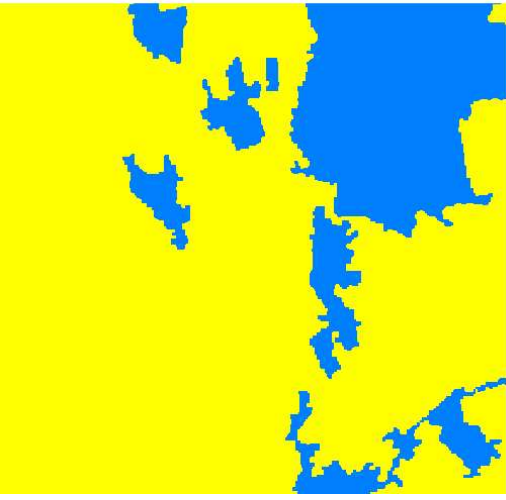
Fig. 5. Change detection map: dark areas are unchanged

Fig. 6. Performance comparison with 3D-SPIHT

of objects, sometimes composed of just a few points, which would not add information to the map's semantic but would compromise the efficiency of the subsequent texture encoding step. Therefore we carry out two simple morphological operations, the elimination of small *unchanged* regions, and the opening/closure of the regions to smooth their contours. The final result is shown in Fig. 5.

We point out explicitly that no *changed* regions are eliminated and therefore, in case a coding modality will be chosen that privileges these regions, they will be all faithfully represented.

C. Map coding

The lossless coding of a binary map is a well-understood problem, for which a large number of techniques exist. Efficiency here is not really an issue because the map is very simple, so we resort to standard context-based arithmetic coding, for a coding cost of just 0.033 bit/pixel, to be further divided among all the image bands. Note that this map is sent immediately as a side information and is itself a useful product for many applications.

D. Texture coding

Classification provides us with a set of homogeneous regions, or objects, which can be efficiently encoded with a technique based on three-dimensional Shape-Adaptive Wavelet Transform (SAWT) and three-dimensional Shape-Adaptive SPIHT (SA-SPIHT), as it has been done in [13] for multispectral images. SA-SPIHT enables an efficient encoding of arbitrarily shaped objects. The first step is the three dimensional shape adaptive transform, which allows the representation of any object with a set of wavelet coefficients, with similar properties to regular wavelet transform's: energy is concentrated at lower frequency subbands, and there is a strong self-similarity among subbands. These properties can be exploited by a zero-tree encoding scheme, as SPIHT, with just some
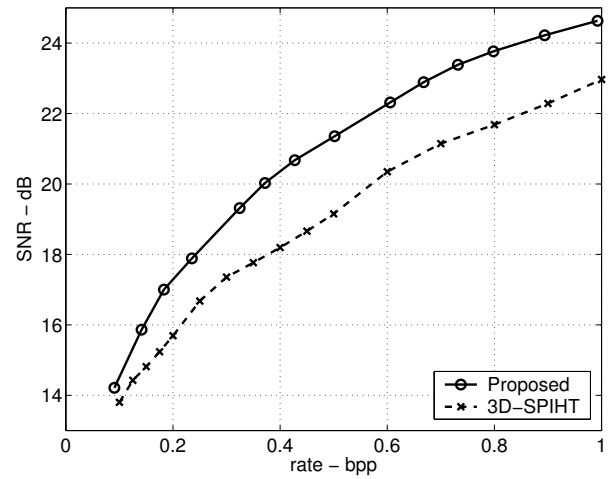
adjustment in order to manage the non rectangular shape of objects. Moreover this algorithm produces an embedded bitstream, which can be truncated at any point, still providing an efficient representation of original data.

This object-based coding scheme achieves good compression performances and provides the user with a great flexibility in encoding strategy. The good encoding capabilities of this approach depend mainly from two elements. The first one is that, thanks to segmentation, each object is quite smooth and, therefore, easy to encode. Sharp transition between objects, which usually requires many coding resources, no longer concern texture coding. Actually they are losslessly and efficiently encoded along with segmentation map. The second element is the possibility to employ an optimal rate allocation strategy among objects, which allows to minimize reconstruction distortion for a given total rate.

Another interesting feature of the proposed object-based technique, is that any arbitrary rate allocation strategy can be chosen. Here we consider a few application-driven scenarios: once fixed the overall bit budget, the user can

1) decode completely both sets of images (April and May);
2) decode completely April images, but only regions which change in May;
3) decode only changed regions of both sets of images.

Going from scenario 1 to 3, the user subtracts resources to unchanged regions to improve the quality of changed ones. Of course, one can also envision a situation in which changed regions are only prioritized in time, but transmitted anyway, as well as many others made possible by this flexible data structure.

III. NUMERICAL RESULTS

In a first experiment, we encoded the complete set of multitemporal images with the proposed algorithm, using a rate allocation strategy intended to minimize distortion of decoded

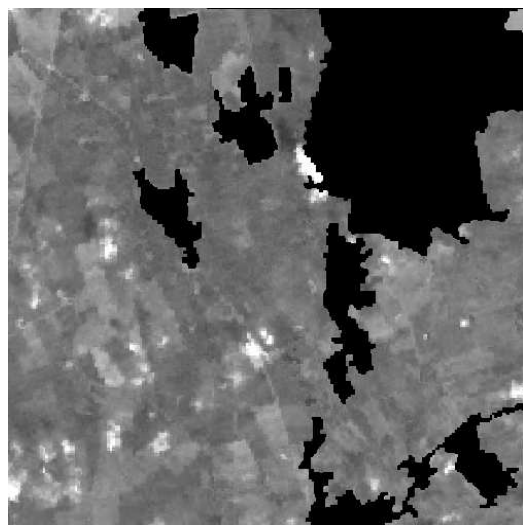Fig. 7. Decoding only changed regions for band 3, April

Fig. 8. Decoding only changed regions for band 3, May

data. We compared our technique's performances with the popular 3D-SPIHT algorithm's [14], by computing the Signal-to-Noise Ratio (SNR) achieved for several coding bit-rate. The results are shown in Fig. 6, where we can verify that the favorable properties of object-based compression algorithms actually lead to a remarkable performance improvement, up to over 2 dB. The subjective quality improvement is even more relevant, as borders sharpness, which is a subjectively valuable feature, is completely preserved through compression.

We also tried some other different encoding strategies, in order to test the flexibility of our architecture. As an example, we only decoded changed region of both sets of images, obtaining the results in Fig. 7 for April and in Fig. 8 for May (band 3). Here, changed regions are encoded at 0.5 bit/pixel/band. With respect to the case where the whole image is encoded (with the same available bit budget), the MSE on the region of interest decreases from 3.06 to 2.06, which corresponds to an increment of 1.72 dB in SNR.

IV. Conclusion

In this paper, a new architecture for segmentation-based compression of multitemporal images is presented. The Bayesian segmentation approach allows to easily find some valuable information, as object classification and change detection. The proposed compression scheme exploits an object-based algorithm, which on one hand can take advantage of segmentation information, while on the other allows a remarkable flexibility in allocation of coding resources. Moreover, the proposed algorithm is efficient, as its performance are superior to a state-of-the-art technique as 3D-SPIHT, and produces an embedded bit-stream.

Future work will be directed toward a better investigation of filtering along spectral and temporal axis, where an optimal transform as the Karhunen-Löeve Transform could be employed. Moreover, a deeper analysis should be performed on the effects of joint segmentation and compression.

References

[1] P.Lombardo, C.J.Oliver: "Maximum likelihood approach to the detection of changes between multitemporal SAR images", *IEE Proceedings Radar, Sonar and Navigation*, vol.148, pp.200-210 Aug. 2001.

[2] G.G.Hazel: "Object-level change detection in spectral imagery", *IEEE Transactions on Geoscience and Remote Sensing*, vol.39, pp.553-561, March 2001.

[3] L.Bruzzone, D.F.Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images", *IEEE Transactions on Image Processing*, pp.452-466, April 2002.

[4] G.Mercier, M.C.Mouchot, G.Cazaguel: "Multitemporal SAR image compression", *1998 International Conference on Image Processing*, vol.1, pp.760-763, 1998.

[5] H.S.Chae, S.J.Kim, J.A.Ryu: "A classification of multitemporal Landsat TM data using principal component analysis and artificial neural network", *1997 International Conference on Geoscience and Remote Sensing* vol.1, pp.517-520, 1997.

[6] G.Gelli, G.Poggi: "Compression of multispectral images by spectral classification and transform coding", *IEEE Transactions on Image Processing*, vol.8 pp.476-489, April 1999.

[7] 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, Mar. 2001.

[8] 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.

[9] G.Poggi, G.Scarpa, J.Zerubia: "Supervised segmentation of remote-sensing images based on a tree-structured MRF model", submitted to *IEEE Transactions on Geoscience and Remote Sensing*, May 2004.

[10] S.Li, W.Li, "Shape-adaptive discrete wavelet transforms for arbitrarily shaped visual object coding", *Transactions on Circuits and Systems for Video Technology*, pp.725-743, Aug. 2000.

[11] A.Said, W.A.Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", *Transactions on Circuits and Systems for Video Technology*, pp.243-250, June 1996.

[12] L.Cicala, G.Poggi, G.Scarpa, "Supervised segmentation of remote-sensing multitemporal images based on the tree-structured Markov random field model", *2004 International Conference on Geoscience and Remote Sensing*

[13] M.Cagnazzo, G.Poggi, L.Verdoliva, A.Zinicola, "Region-oriented compression of multispectral images by shape-adaptive wavelet transform and SPIHT", *2004 International Conference on Image Processing*

[14] Xiaoli Tang; Sungdae Cho; Pearlman, W.A.; "Comparison of 3D set partitioning methods in hyperspectral image compression featuring an improved 3D-SPIHT" , *2003 Data Compression Conference*