SEGMENTATION BASED IMAGE CODING WITH l_∞ NORM ERROR CONTROL

Marco Dalai and Riccardo Leonardi

Universitá degli Studi di Brescia
Dipartimento di Elettronica per l'Automazione
via Branze, 38 - I-25123 Brescia, Italy
Email: {marco.dalai, riccardo.leonardi}@ing.unibs.it

ABSTRACT

In this paper we present a segmentation based technique for lossy image coding with minmax error control. The technique is an improvement of a scheme presented in [1]. In the base scheme, the image is recursively and adaptively divided into rectangular regions with a binary tree decomposition. Every region is approximated by mean of a bilinear surface, whose values on the four corners are encoded. It is well known that binary tree decompositions can generate many adjacent regions that contain the same texture, but have been separated higher up in the decomposition of the hierarchy. Inspired by the work of Sukla et al. ([2]), we here investigate the possibility of jointly coding such similar adjacent regions, so as to reduce the bit rate.

1. INTRODUCTION

Lossy image coding has received a lot of attention in the last decades and second generation techniques have revealed very good performance when high compression ratio is required. The drop in quality due to compression is usually evaluated by means of the mean square error or the peak signal-to-noise ratio (PSNR). In some cases, nevertheless, the l_∞ norm of the error provides a more attractive evaluation method, for example in application used for medical purposes. Up to now, most of the interest in this sense is dedicated to the so called "near-lossless" coding, i.e. with small l_∞ error norm (typically at most 4 out of 256 levels). For this task, predictive and transform based methods seem to give very good results ([3, 4]). On the other hand, when the more general problem of wider range l_∞ error control is considered, one may expect that second generation techniques can give better performance. This question seems to have received very little attention up to now; the aim of this work is to propose an image segmentation technique for l_∞ lossy coding, which gives an improvement with respect to the first idea exposed in [1].

2. SEGMENTATION WITH l_∞ NORM

In [1] an algorithm for the segmentation of images in rectangular domains has been presented in which bilinear approximation are performed within each domain. The image, supposed to be rectangular, is initially approximated with a bilinear surface; if the l_∞ error is larger than a given threshold δ, the image is divided in two rectangular regions, and the procedure is applied recursively on every one of them, thus leading to a binary tree decomposition. The bilinear approximations are suboptimal, and are constructed with a separable application of the method explained in [5]. Basically, the procedure is the following (see [1] for details). Suppose the rectangular domain $D = \{(x, y) : x_0 \leq x \leq x_1, y_0 \leq y \leq y_1\}$ has R rows and C columns. For every $r = 1 \dots R$ we compute the l_∞ optimal linear approximation of the r-th row of the signal s. We thus obtain a segment $l(r)$ which has an associated error $e(r)$; this segment intersects the planes $x = x_0$ and $x = x_1$ in two points that we call respectively $q_0(r)$ and $q_1(r)$. Consider the $q_0(r)$ points (the same argument holds for $q_1(r)$), $r = 1 \dots R$; in general these are not aligned, and, in order to have a bilinear surface, they should all lie on a straight line. If g is a straight line approximation of these points, call $\varepsilon_0(g; r) = |g(r) - q_0(r)|$ the error over the point $q_0(r)$. This error should be added to $e(r)$ to obtain a bound on the actual error introduced on the points of the r-th row. The best linear approximation of the $q_0(r)$'s should be redefined as

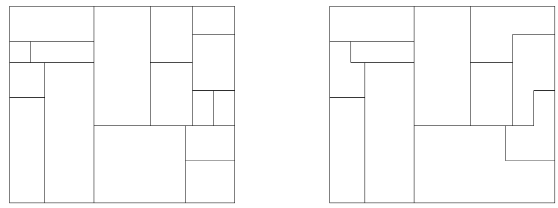
$$h = \arg\min_g \left(\max_r (e(r) + \varepsilon_0(g; r)) \right). \qquad (1)$$

This line is actually the optimal l_∞ approximation of the set of points $\{q_0(r) \pm e(r)\}_{r=1\dots R}$, and this justifies the "separable" characteristics of the suboptimal approximation. It is obvious that the same scheme can be used by interchanging the role of rows and columns, i.e. linear approximating the columns and then adjusting the extremities of the straight lines in the planes $y = y_0$ and $y = y_1$. In this way we find two generally different solutions, and obviously we choose the one with smallest error. This suboptimal solution, not only has the advantage of being computational efficient, but also gives knowledge of the local nonlinearities of the signal along rows and columns. Thus, when the bilinear surface has an approximation error that is over the threshold δ, the segmentation task can be performed by considering the errors $e(r)$ and $e(c)$. We identify the row or the column such that the error committed by its minmax straight line approximation ($e(r)$ or $e(c)$) is the largest with respect to all values of r and c. We set then the partition line to be orthogonal to the selected row or column close to the point with maximum approximation error.

This procedure leads to a decomposition of the image as shown, for example, in figure 1(a) which we call Binary Rect-Tree (BRT). The image is then coded by specifying the tree structure and the approximating bilinear surfaces of every leaf. As motivated in [1], a good way of coding the bilinear surfaces within every rectangle is to code their values on the four corners, quantizing these values with the same quantization step of the original image. Furthermore, the corner values of every rectangle are predicted from the previously encoded grey levels of the neighboring pixels. This technique has shown to perform quite well in coding piecewise smooth images for large ranges of the δ parameter, even if for small values of δ state of the art techniques such as JPEG-LS still outperform the present approach.

3. JOINT CODING OF NEIGHBORING LEAFS

Concerning the coding method explained in the previous section we can find that the main disadvantage of the binary-tree decomposition is the "oversegmentation", that is many regions are unnecessary split as a consequence of the cross-partitioning straight lines. In fact, many leaves of the obtained tree contain adjacent similar regions that have been separated higher up in the decomposition. This cause a penalty in terms of bit per pixel used in coding the image, as many leaves attributes must be coded even in some regions where only few coefficients would be sufficient. Following the idea presented in [2] we think that a study of the possible joint coding of these similar adjacent regions would give an interesting improvement with respect to the simple



(a) Binary tree decomposition.

(b) After Joint Coding of neighboring leaves.

Fig. 1. Example of decomposition of the domain into rectangular regions and successive joining of neighbors.

BRT coding scheme. Here we only consider the problem of grouping pairs of rectangles, verifying that even this simple strategy gives good results. An example of equivalent segmentation of the joint coding technique is shown in figure 1(b).

Given the BRT decomposition we scan the leaves in the same order as they should be coded, for example from top left to right bottom. For the generic leaf (say A), if it is not already flagged as joined with a previously coded leaf, we search for a neighboring leaf (say B) which should be coded in future steps; thus, we compute the optimal[1] approximation of the signal over $A \cup B$ and, if the error is smaller that the threshold δ, we flag B as already coded and we jointly code it with A. The procedure is repeated for every leaf of the tree and, for every fixed A-leaf all the possible B-leaves are checked. This coding scheme (which we call JBRT) is not optimal, but it performs well, has relatively small computational complexity, and is rasonably simple to implement.

An efficient way of handling the tree structure and the neighbors searches is the following. The tree is analyzed in a recursive fashion; at every parent node the recursion is called on "left-child" or "up-child" first (depending on the partition line direction) and on "right-child" or "bottom-child" successively. This means that the top left leaf of the tree is the first analyzed leaf. Furthermore, this scanning order implies that for every A-leaf we can only search neighbors on the right and to the bottom, as top left leaves are already coded. Finally, for searching the neighbors of a given A-leaf, it is important to use the constructed tree structure, by moving from parent nodes to children and viceversa, so as to improve the efficiency. We show for example, the case where a right-neighbor of a leaf $A = \{(x, y) : x_0 \leq x \leq x_1, y_0 \leq y \leq y_1\}$ is searched:

[1] In this step it is advisable to compute optimal approximations instead of suboptimal ones, as the segmentation is already performed, and optimal approximations are preferable when working close the δ threshold.

1. Starting from the A leaf, go up on the tree until a “left-child” node i reached, and then move to its “right-child” brother.

2. If the node is not a leaf, go down recursively in the tree following the rule
   - if the partition line is vertical, search only on “left-child” node,
   - if the partition line is horizontal, say $y = y_p$, search on the “up-child” if $y_p \leq y_0$, on the “bottom-child” if $y_p \geq y_0$ and on both if $y_0 < y_p < y_1$

   until a leaf is reached. An example of this procedure is shown in fig. 2.

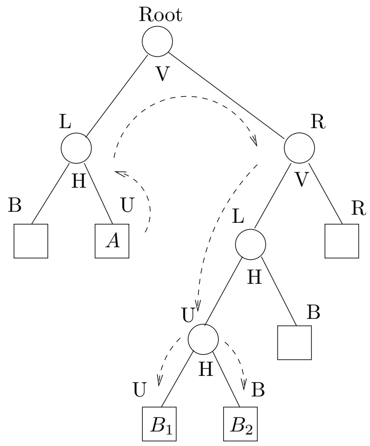
With this method we can easily find all leaf neighbors in the tree; when a neighbor B is found the optimal approximation is computed. If the two leaves can be joined, a joint coding flag is set and a reference from A to B is specified. The approximating surface is then coded by using its values on the four corners of the bounding box of $A \cup B$, as shown in fig. 3, rounded with the same precision as the original image. As explained in [1], this quantization method is very effective when we are interested in coding an integer valued signal s with an allowed maximum integer error δ. In this case in fact, when the explained coding method is used, it is only necessary to have a bilinear approximation with error smaller than δ in order to ensure a reconstructed signal with error less than or equal to δ, and thus quantization effects are negligible.

4. EXPERIMENTS

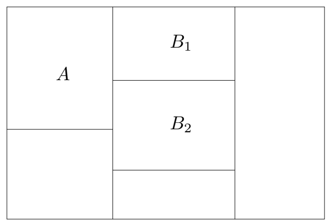
We have implemented the JBRT scheme and we have compared the compression performance with that of the base BRT scheme. We have tested the two methods on the images Lena and Bird obtaining the compression results, in bit per pixel, shown in table 1. As we can see, the simple joining of pair of neighbor leaves, without optimal pairing strategies and with very simple prediction techniques, can reduce the bpp values of a quantity between 15% and 20%. Figure 4 shows the obtained segmentation of the Lena and Bird images when an error threshold $\delta = 16$ is used in the BRT and JBRT schemes.

5. CONCLUSION AND FUTURE WORK

We have presented a segmentation based coding scheme for image compression with l_∞ norm error control. The method is an extension of the base scheme proposed in [1] where the image was coded without joint coding



(a) Decomposition tree



(b) Domain partition

Fig. 2. Example of right hand neighbor search for the leaf A. The circles are nodes and rectangles are leaves. Letters V and H indicate vertical or horizontal partitions respectively, while L, R, U, B represent left, right, up or bottom children.

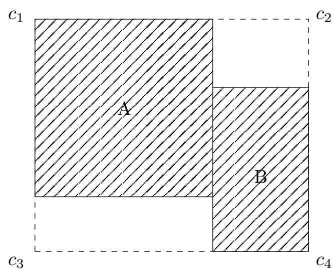
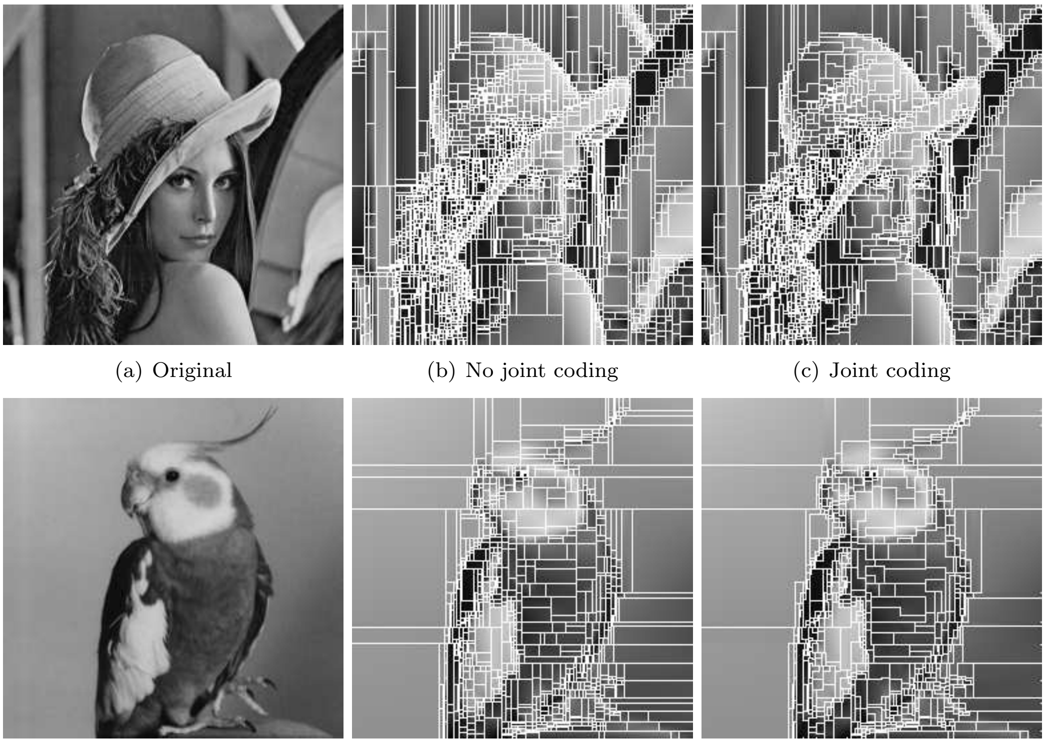


Fig. 3. Coding the approximation for a pair of joined rectangular leaves. The bilinear surface values on the four corners c_1, c_2, c_3 and c_4 are coded.

of neighboring leaves. In this paper we have studied the performance gain when only pairs of neighboring rectangular domains are coded jointly, obtaining a gain between 15% and 20% in terms of bpp for the images Lena and Bird. This suggests that a more sophisticated study on the possible grouping of more than two leaves, together with prediction and coding optimizations, could give very interesting results in term of com-



(a) Original (b) No joint coding (c) Joint coding

(d) Original (e) No joint coding (f) Joint coding

Fig. 4. Examples of segmentation obtained with the BRT and JBRT schemes. Here $\delta = 16$.

	Lena		Bird	
δ	BRT	JBRT	BRT	JBRT
4	3.57	3.18	1.68	1.41
6	2.73	2.40	1.12	0.92
8	2.26	1.97	0.89	0.71
10	1.95	1.68	0.75	0.63
12	1.72	1.48	0.65	0.51
14	1.54	1.30	0.57	0.44
16	1.39	1.19	0.50	0.39

Table 1. Occupation in bpp of the images Lena and Bird coded respectively with BRT and JBRT methods for different values of δ.

pression performance. Future work will thus concern the study of a scheme where many leaves are grouped and jointly coded.

Furthermore, it is clear that in our coding scheme most of the rate is spent in coding non-smooth regions for which, obviously, bilinear approximation are not useful. Future work will be dedicated thus to the study of efficient strategies for coding texture regions which can be efficiently isolated in our segmentation scheme.

6. REFERENCES

[1] M. Dalai and R. Leonardi, "l^∞ norm based second generation image coding," *To appear in Proc. of ICIP'04*.

[2] R Shukla, P.L. Dragotti, M. Do, and M. Vetterli, "Improved quadtree algorithm based on joint coding for piecewise smooth image compression," *Proc. of IEEE Int. Conf. on Multimedia*, August 2002, Swizzerlad.

[3] L. Karry, P. Duhamel, and O. Rioul, "Image coding with an L^∞ norm and confidence interval criteria," *IEEE Trans. Image Proc.*, vol. 7, no. 5, pp. 621–631, 1998.

[4] A. Alecu, A. Munteanu, P. Schelkens, J. Cornelis, and S. Dewitte, "Wavelet-based fixed and embedded L-infinite-constrained coding," *Proc. IEEE Int. Conf. Digital Signal Processing*, pp. 509–512, 2002, Santorini, Greece.

[5] M. Dalai and R. Leonardi, "Efficient (piecewise) linear minmax approximation of digital signals," *Proc. of ICASSP'04*, pp. 585–588, 2004.