

IMPROVED QUADTREE ALGORITHM BASED ON JOINT CODING FOR PIECEWISE SMOOTH IMAGE COMPRESSION

Rahul Shukla[†], Pier Luigi Dragotti[†], Minh Do[‡], and Martin Vetterli^{†§}

[†]Audio-Visual Communications Laboratory

Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland

[‡]Department of ECE, University of Illinois at Urbana-Champaign, USA

[§]Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA

ABSTRACT

We are presenting a novel coding algorithm based on the tree structured segmentation, which achieves the oracle like exponentially decaying rate-distortion (R-D) behavior for a simple class of signals, namely piecewise polynomials in the high bit rate regime. We are considering R-D optimization framework, which employs optimal bit allocation strategy among different signal segments to achieve the best tradeoff between description complexity and approximation quality. First, we describe the basic idea of the algorithm for the 1-D case. It can be shown that the proposed compression algorithm based on an optimal binary tree segmentation achieves the oracle like R-D behavior ($D(R) \sim c_0 2^{-c_1 R}$) with the computational cost of the order $O(N \log N)$. We then show the extension of the scheme to the 2-D case with the similar R-D behavior without sacrificing the computational ease. Finally, we will conclude with some experimental results.

1. INTRODUCTION

Image coding for stripping the redundancy from the typical highly correlated image waveforms is an active area of research. The field of image compression offers diverse source coding schemes ranging from classical lossless techniques and popular transform approaches to the segmentation based coding methods [3, 4, 8]. For low bit rate compression applications, segmentation based coding methods generally provide high compression ratios when compared with traditional transform, vector quantization (VQ) and subband (SB) coding schemes.

Quadtree based image compression, which recursively divides the image into simple geometric regions has been one of the most popular segmentation based coding schemes investigated by researchers [5, 6, 13, 14]. The tree structure based segmentation for image coding has been considered by several researchers [12, 14, 15], however none have incorporated a scheme to find truly optimal tree structures, given some additive distortion measure such as squared error or absolute difference. Few methods, which are partially capable of searching such optimal structures, were presented by Chou *et al.*'s (G-BFOS)[1], Sullivan and Baker [13], Ramchandran and Vetterli [9] and Radha *et al.* [8]. All of these schemes are very elegant and practical but result in suboptimal R-D behavior because these schemes fail to exploit the dependency among the neighboring nodes with different parents. As we

will figure out in this paper, the independent coding of dependent nodes leads to the suboptimal R-D behavior.

Since geometrical regions, image semantics like edges, and image signal within a small region can be well described by simple polynomial models, recently there has been a growing interest in the study of piecewise polynomial functions as an approximation to piecewise smooth functions. Wavelets have long been considered ideal candidates for piecewise smooth function due to their vanishing moment properties. It was shown in [2, 7] that for piecewise polynomial signals, the distortion of wavelet based coder decays as $D(R) \sim d_0 \sqrt{R} 2^{-d_1 \sqrt{R}}$. But in the same case, it is possible to achieve

$$D(R) \sim c_0 2^{-c_1 R} \quad (1)$$

using an oracle method, and this can be realized with a reasonable (polynomial) computational cost using dynamic programming [7]. At low rates, such an algorithm works well for piecewise smooth signals too. The basic ingredient is to precisely model singularities. The dynamic segmentation algorithm in [7] achieves the right R-D characteristic, but its computational cost is high. Moreover, this algorithm cannot be generalized to the 2-D case.

In this paper, we are interested in a coding scheme, which achieves oracle like asymptotic R-D behavior with polynomial complexity for 1-D as well as 2-D signals. In particular, for the 1-D case we present a coding scheme which utilizes binary tree segmentation with optimal bit allocation among different segments. Investigation of the algorithm reveals the inherent weakness in the initial coding scheme, leading to a suboptimal performance ($D(R) \sim c_2 \sqrt{R} 2^{-c_3 \sqrt{R}}$). Reinvestigation of the algorithm also reveals the path to an optimal scheme, which achieves the desired R-D behavior as indicated in eqn.(1).

Most importantly, the optimal binary tree scheme in the 1-D case can be easily extended to the 2-D case as an optimal quadtree scheme with the similar computational cost. The proposed optimal quadtree scheme also achieves the oracle like R-D performance for some simple classes of images whereas for the 2-D case there is no known algorithm, which achieves the right R-D behavior with a reasonable computational cost. This paper is organized as follows: first, we will briefly outline the coding scheme for the 1-D case in the form of the binary tree segmentation algorithm. Then, we show the extension of 1-D scheme to 2-D using the quadtree based scheme along with the R-D behavior of the algorithm for the simple image class. Finally, we present the performance of this coding scheme in comparison to the wavelet based coder.

This work was supported by the Swiss National Science Foundation grant numbers 20-63664.00, 20-52347.97 and 21-52439.97.

2. 1-D CASE: BINARY TREE ALGORITHM

Our goal is to implement a compression algorithm based on the modeling assumption that signals are piecewise smooth functions. In this case, if we segment the signal into smaller pieces, then each sub-signal can be well represented by a simpler signal (e.g. polynomial) model. For instance, we can choose a very simple polynomial model of degree one.

Classical binary tree algorithm: Our algorithm employs a binary tree segmentation followed by a coding algorithm on each signal block in an operational R-D optimal sense. We employ an operational rate-distortion optimization that is similar to the approach used in [9] in finding the best wavelet packet bases. A decision strategy based on optimizing R-D performance for each signal block is designed so that the coder can decide if a signal block is worth to be further divided and coded with some appropriate quantization level. The algorithm can be summarized as follows [9]:

1. Segmentation of the input signal: a binary tree segmentation scheme is employed.
2. Optimal representation of each sub-signal by the polynomial model, in a R-D sense.
3. Optimization of this representation to achieve the best reconstructed signal for a given bit rate constraint and distortion measure (MSE). Optimization is performed in R-D sense using Lagrangian cost functional for an optimal quality factor(λ) [10].

Since the binary tree segmentation acts like a singularity detector, so when the tree grows to capture the singularity in a region, it divides the same polynomial piece repeatedly. The classical coding scheme encodes these subdivided pieces of the same polynomial separately (see Fig. 1(b)). This independent coding of the dependent nodes results in a suboptimal R-D behavior.

Improved binary tree algorithm: Investigation in the earlier paragraph suggests that the inter node dependency should be exploited. For instance, if neighboring leaves have the same model (polynomial) parameters, then they should be encoded jointly. For ensuring that, we have introduced a binary variable which indicates whether neighbors are coded jointly or separately. Fig. 1 illustrates, how the improved coding scheme outperforms the balanced as well as the classical binary tree algorithms for the piecewise polynomial signal class. The improved coding scheme employs the same previous classical binary tree coding scheme followed by the neighbor joint coding algorithm which, in fact, decides whether incoming information should be transmitted or not. Thus the neighbor joint coding algorithm is essentially introduced to account for the already transmitted information. We have shown in our earlier work [11] that this dependent coding leads to the desired exponentially decaying R-D behavior with the computational complexity of the order $O(N \log N)$. It is interesting to note that the improved tree segmentation map is *shift invariant* if the tree is decomposed up to the full depth, e.g. up to the pixel level for the discrete case.

3. EXTENSION TO 2-D

Although the situation is much more open and complex in two dimensions (2-D), yet it is not hard to visualize the extension of the proposed 1-D coding scheme to the 2-D coding scheme. Since all the sub-algorithms developed for the 1-D scheme have a very nice equivalent in the 2-D world, e.g. the binary tree segmentation can be replaced by the quadtree segmentation, polynomial model can

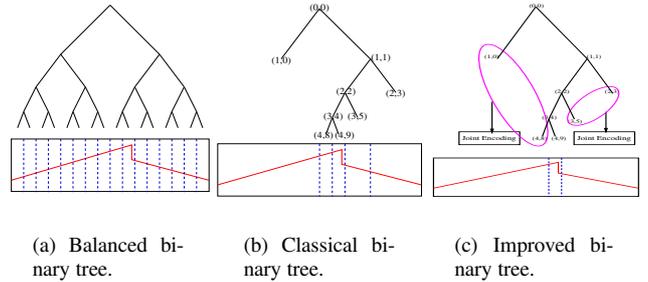


Fig. 1. Comparative study of different algorithms.

be replaced by the 2-D geometrical model consisting of two 2-D polynomials separated by a polynomial boundary. The Lagrangian optimization algorithm remains the same. The neighbor joint coding algorithm is a bit more involved but can be implemented efficiently. Therefore we can have an efficient quadtree based coding scheme for 2-D signals with a reasonable computational cost.

Image model and Oracle R-D performance: we consider the case where the edge is a piecewise polynomial curve. In particular, with piecewise linear edge, we have the “Polygon” model where there is a polygon-shaped object against a uniform background. In such case, a possible oracle method would simply code the position of the V vertices of the polygon. With R/V bits for each vertex, a regular grid on the unit square provides quantized points within a distance $\Delta = 2^{-R/2V}$ from the original vertices. Let L be the finite length of the edge (or the boundary of the polygon) then the distortion for the 2-D object is upper bounded by $D(R) \leq L\Delta$. Therefore for the “Polygon” model, the oracle R-D function decays exponentially as

$$D(R) \sim L2^{-R/2V}. \quad (2)$$

4. CLASSICAL QUADTREE ALGORITHM

Our algorithm employs a quadtree segmentation followed by a coding algorithm on each image block in an operational R-D optimal sense. We employ an operational R-D optimization that is similar to the approach used in [9] in finding the best wavelet packet bases. A decision strategy based on optimizing R-D performance for each image block is designed so that the coder can decide if an image block is worth to be further divided and coded with some appropriate quantization level.

4.1. R-D analysis of the classical quadtree compression algorithm for the polygonal image class

In this section, we will consider the “Polygon” image model, where the edge is a piecewise linear curve with a *finite* number of discontinuity points and derive an asymptotic R-D behavior of the quadtree based coding scheme. In an optimal quadtree, at each level, the only dyadic blocks that need to be divided further are the ones containing a singularity point (vertex) of the edge in the given image. Other dyadic blocks contain either no edge or a straight edge, which can be efficiently represented by an edge tile.

An optimal tree is obtained via recursively dividing nodes with transition (vertex) points. Essentially quad tree grows only in the region where algorithm observes transition points (see Fig. 4(a)).

That means quad tree segmentation algorithm acts like the singularity (vertex) detector. Assume that the polygonal image has V vertices, then at each level there are at most V splitting nodes. Thus they will generate no more than $3V$ leaves with straight edge at the next level. Therefore the total number of the terminal-nodes (leaves) in an optimal quad tree are bounded as follows:

$$\begin{aligned} N_J &\leq N_{J-1} + 3V \\ \Rightarrow N_J &\leq N_0 + J3V = 1 + 3JV \sim 3JV. \end{aligned} \quad (3)$$

In the similar fashion, the total number of nodes in an optimal tree can be derived: $N_{T,J} \sim 4JV$. We are investigating the performance of the algorithm in the high bit rate regime and interested in the asymptotic R-D behavior. Suppose that the quadtree is decomposed up to depth J . Let us simply allocate J bits to each of the quantized points to ensure that the maximum distance between the true vertices and their quantized version is bounded by $O(2^{-J})$. The total bit rate can be seen as the sum of the costs of coding the quad tree itself and the quantized model parameters of the leaves. Hence the total description complexity can be written as follows:

$$\begin{aligned} R_T &= R_{\text{TreeInfo}} + R_{\text{Leaves}} \sim 8JV + 6J^2V \\ \Rightarrow J &\sim \sqrt{\frac{R_T}{6V}} \end{aligned} \quad (4)$$

For the computation of the associated distortion, firstly we have to observe that the edge length in the given original polygonal image is finite, assume total edge length is L . Secondly, since the maximum distance between the true vertices and their quantized version is bounded by $O(2^{-J})$, the net distortion will be

$$\begin{aligned} D_T &\sim L2^{-J} \\ \Rightarrow D_T &\sim L2^{-\sqrt{\frac{R_T}{6V}}}; \text{ Combining eqn.(4) and (5).} \end{aligned} \quad (5)$$

Therefore $\sqrt{R_T}$ term is present in the exponent, which indicates a suboptimal decay in comparison to the desired exponential decay given by eqn.(2). Hence, the classical quadtree algorithm also exhibits the suboptimal R-D performance due to the independent coding of the dependent nodes (see Fig. 4(a)). For correcting the suboptimal behavior, we propose an improved quadtree algorithm.

5. IMPROVED QUADTREE ALGORITHM

The improved coding scheme employs the same previous classical quadtree coding scheme followed by the neighbor joint coding algorithm which, in fact, decides whether neighbors should be coded jointly or independently (see Fig. 2).

5.1. Neighbor joint coding algorithm

The neighbor joint coding algorithm can be described as follows:

1. As soon as neighbor joint coding algorithm receives the leaf information, the algorithm looks for the leaf's neighbor, which has already been transmitted or encountered by the algorithm.
2. If the algorithm finds a transmitted neighbor, then it compares the model parameters for both the leaves using a suitable comparison metric. One possible choice of comparison metric could be l_2 -error. That means if the l_2 -error between the model parameters of the leaf and its neighbor is within some predefined distortion, then the incoming leaf information will not be transmitted and the neighbor joint coding variable will be set to one followed by the

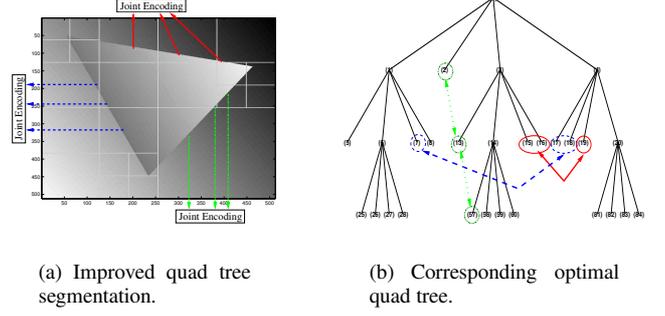


Fig. 2. Illustration of the joint encoding by the improved coding scheme.

two bits for neighbor index (considering only 4 neighbors), otherwise leaf information will be transmitted and the neighbor joint coding variable will be set to zero.

3. If the neighbor joint coding algorithm does not find any transmitted neighbor, then the neighbor joint coding variable will be set to zero and the leaf information will be sent.

5.2. Rate-distortion analysis

We are presenting a simplified asymptotic R-D analysis of the proposed coding scheme. The R-D function can be computed in two steps. First, we will compute the rate and the distortion as a function of the decomposition depth J , then combine the rate and the distortion by eliminating J . Optimal tree leaves counting will be done in the similar fashion as it was done for the previous scheme that means $N_J \sim 3JV$ (eqn.3).

In the improved coding scheme, the side information consists of two parts: 1) Bits required to encode the quad tree. 2) Bits required to encode leaf joint coding tree. The quad tree split-merge decision variable will consume bits proportional to the total number of the nodes ($N_{T,J} \sim 4JV$) in the optimal quad tree whereas joint coding decision variable will consume bits proportional to the total number of leaves ($N_J \sim 3JV$) in the optimal quad tree. Hence coding of the side information will cost the following bit rates: $R_{\text{TreeInfo}} \sim 8JV$ and $R_{\text{Leaf-Joint}} \sim 9JV$.

Let us naively allocate $4J$ bits to every linear piece of the polygonal edge to ensure that the distortion for any leaf is around $O(2^{-J})$. The total bit rate can be seen as the sum of the costs of coding the quad tree, leaf joint coding tree and the quantized model parameters of the leaves. The total bit budget can be related to the decomposition level (J) in the following way:

$$\begin{aligned} R_T &= R_{\text{TreeInfo}} + R_{\text{Leaf-Joint}} + R_{\text{Leaves}} \\ \Rightarrow R_T &\sim 8JV + 9JV + 4JV \sim 21JV \\ \Rightarrow J &\sim \frac{R_T}{21V} \end{aligned} \quad (7)$$

The net distortion computation is similar to the analysis done for the previous version of the coding scheme and it follows that

$$\begin{aligned} D_T &\sim L2^{-J}; \text{ from eqn. (5)} \\ D_T &\sim L2^{-\frac{R_T}{21V}}; \text{ from eqn. (7)}. \end{aligned} \quad (8)$$

Therefore the introduction of the joint coding binary variable itself consumes bits but provide the desired rate-distortion behaviour, which is exponentially decaying. Thus the joint encoding

scheme achieves the right asymptotic rate-distortion characteristic.

6. SIMULATION RESULTS

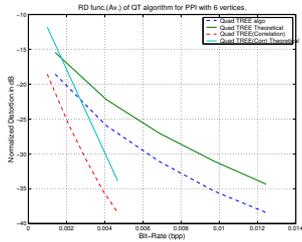
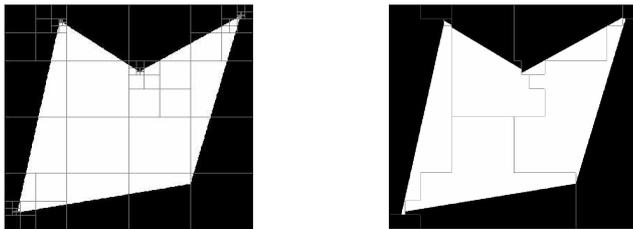


Fig. 3. Theoretical(solid) and numerical(dotted) R-D curves for the classical and improved quadtree algorithms for the B/W Polygonal image class.

Numerical experiments are performed for the two type of image classes: 1) B/W images with polygonal singularity. Polygon's vertices are generated randomly using uniform distribution on the space $[0, 1]^2$. For experimentation, polygons with 5 and 6 vertices are considered as singularities. 2) Real life images like Lena. For both cases we have observed a better R-D behavior of the improved quadtree algorithm with respect to JPEG-2000.

The experimental results shown for the Polygonal image class in Fig. 3 also confirm the theoretical R-D behaviors. Fig. 4 presents example of reconstructed images by the classical and improved quadtree algorithm.



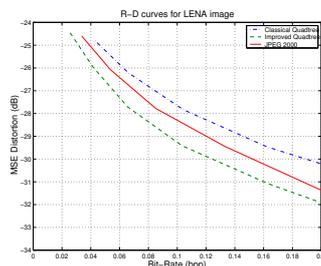
(a) Classical quadtree segmentation.

(b) Improved quadtree segmentation.

Fig. 4. Examples of the quadtree representation.



(a) Improved quad tree (0.12 bpp).



(b) R-D performance.

Fig. 5. Comparison of operational R-D curves using a wavelet coder (JPEG-2000) and quadtree coders for Lena image.

7. DISCUSSION AND FURTHER WORK

In this work we have analyzed the asymptotic R-D behavior of the binarytree and quadtree decomposition algorithms and we have proposed an improved version of these algorithms which have the optimal R-D behavior for the case of 1-D and 2-D piecewise polynomial signals.

Numerical simulations (Fig. 3) confirm that these algorithms achieve optimal performance if the input image fits the model exactly. In addition, preliminary simulations show that our quad-tree algorithm outperforms JPEG2000 also in the case of compression of real life images such as 'Lena'. Our aim is to improve such a performance further by better compressing the texture present in images. Another issue is to understand whether the proposed quad-tree scheme provides an efficient and fast tool to detect and compress contours in object oriented multimedia applications.

8. REFERENCES

- [1] P.A. Chou, T. Lookabaugh and R.M. Gray, "Optimal pruning with applications to tree structured source coding and modeling," *IEEE Trans. Inform. Th.*, vol. IT-35, pp. 299-315, Mar. 1989.
- [2] A. Cohen, I. Daubechies, O.G. Guleryuz and M.T. Orchard, "On the importance of combining wavelet-based nonlinear approximation with coding strategies," *IEEE Trans. Inform. Th.*, 2000, submitted.
- [3] A.K. Jain, "Image data compression: A review," *Proc IEEE*, vol. 69, pp. 349-389, Mar. 1981.
- [4] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second generation image coding techniques," in *Proc. IEEE*, vol. 73, pp. 549-574, Apr. 1985.
- [5] M. Kunt, M. Bernard and R. Leonardi, "Recent results in high compression image coding," *IEEE Trans. Circuits Syst.*, vol. CAS-34, no. 11, pp. 1306-1336, nov. 1987.
- [6] R. Leonardi and M. Kunt, "Adaptive split and merge for image analysis and coding," *Proc. SPIE*, vol.594, 1985.
- [7] P.Prandoni and M.Vetterli, "Approximation and compression of piecewise smooth functions," *Phil. Trans. Royal Society London*, August 1999.
- [8] H. Radha, M. Vetterli and R. Leonardi, "Image compression using binary space partitioning trees," *IEEE Trans. Image Proc.*, vol. 5, no. 12, pp. 1610-1624, Dec. 1996.
- [9] K.Ramchandran and M.Vetterli, "Best wavelet packet bases in a rate distortion sense," *IEEE Trans. Image Proc.*, vol. 2, no.2, pp. 160-175, April 1993.
- [10] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust. Speech Signal Proc.*, vol. 36, pp. 1445-1453, Sept. 1988.
- [11] R. Shukla, M.N. Do, P.L. Dragotti and M. Vetterli, "Rate-distortion optimal tree based coding algorithms for piecewise polynomials," accepted to the *IEEE International Symposium on Information Theory*, Switzerland, June 2002.
- [12] P. Strobach, "Quadtree structured recursive plane decomposition coding of images," *IEEE Trans. Signal Proc.*, vol. 39, pp. 1380-1397, June 1991.
- [13] G.J. Sullivan and R.L. Baker, "Efficient quadtree coding of images and video," *IEEE Trans. Image Proc.*, vol. 3, no. 3, pp. 327-331, May 1994.
- [14] D.J. Vaisey and A. Gersho, "Image coding with variable block size segmentation," *IEEE Trans. Signal Proc.*, vol. SP-40, pp. 2040-2060, Aug.1992.
- [15] X. Wu, "Image coding by adaptive tree structured segmentation," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1755-1767, Nov. 1992.