IEEE International Conference on Image Processing (ICIP), Thessaloniki, Greece, 2001

ON THE COMPRESSION OF TWO-DIMENSIONAL PIECEWISE SMOOTH FUNCTIONS

Minh N. Do[†], *Pier Luigi Dragotti*[†], *Rahul Shukla*[†], *and Martin Vetterli*[†§]

[†]Audio-Visual Communications Laboratory
Swiss Federal Institute of Technology Lausanne (EPFL), 1015 Lausanne, Switzerland
[§]Department of EECS, University of California at Berkeley, Berkeley CA 94720, USA
Web: `lcavwww.epfl.ch`

ABSTRACT

It is well known that wavelets provide good non-linear approximation of one-dimensional (1-D) piecewise smooth functions. However, it has been shown that the use of a basis with good approximation properties does not necessarily lead to a good compression algorithm. The situation in 2-D is much more complicated since wavelets are not good for modeling piecewise smooth signals (where discontinuities are along smooth curves). The purpose of this work is to analyze the performance of compression algorithms for 2-D piecewise smooth functions directly in a rate distortion context. We consider some simple image models and compute rate distortion bounds achievable using oracle based methods. We then present a practical compression algorithm based on optimal quadtree decomposition that, in some cases, achieve the oracle performance.

1. INTRODUCTION

The interaction of approximation theory and compression is an active area of research, especially in the context of wavelet based signal processing (see [1, 2] for reviews). The set up usually considers classes of signals (piecewise smooth signals, specified by an appropriate norm) and classes of bases (Fourier series, wavelet bases). Then, the intuition is that if a certain expansion has good approximation properties (e.g. its N-term linear or non-linear approximation converges fast) then a compression scheme based on such an expansion will be efficient. The analysis usually works for the limit of very large N, or what is called the high rate limit in compression. An example of such a result is the following [3]. Assume a piecewise smooth function with pieces that are in C^p and a finite number of discontinuities. The distortion of a wavelet based scheme can be shown to be

$$D(R) \sim c_1 R^{-2p} + c_2 \sqrt{R} 2^{-c_3 \sqrt{R}} \qquad (1)$$

where the first term corresponds to the smooth pieces, while the second corresponds to the discontinuities. At high rates, the first term clearly dominates.

While good approximation properties are necessary for good compression, it might not be enough. In particular, in non-linear approximation, the indexing and individual compression of expansion coefficients might be inefficient. This was shown for a simpler class of signals, namely piecewise polynomials, in [4]. In that case, it is possible to achieve

$$D(R) \sim c_4 2^{-c_5 R} \qquad (2)$$

This work was supported the Swiss National Science Foundation grant numbers 20-52347.97 and 20-061493.

using an oracle method, and this can be realized with a reasonable (polynomial) computational cost using dynamic programming. At low rates, such an algorithm works well for piecewise smooth signals also. The basic ingredient is to precisely model discontinuities, which can be done also in a wavelet scheme using footprints [5].

In two dimensions, the situation is much more open. First, wavelets are not good for modeling piecewise smooth signals (where discontinuities are along smooth curves). The N-term non-linear approximation is far from optimal, so new bases are needed. The search for such new bases has been led by Candes, Donoho and co-workers [6, 7]), and interesting results on optimal or quasi-optimal non-linear approximation have been obtained.

But just as in the one-dimensional case, these results may or may not give the right answer for compression. It is the purpose of the present paper to investigate the interplay of good representations and compression. In order to do so, we consider some very simple models, like for example piecewise polynomials with piecewise polynomial curves as edges. While simplistic, such a model has two advantages. It leads to precise analysis and the results may be relevant for low rate behavior of more general signals. Intuitively, if an object has a finite number of free parameters (as is the case for this simple model) we expect an exponential decay of the rate-distortion function. We verify this using an oracle based method. In two dimensions, oracle based method are difficult to realize with polynomial cost, and thus, we consider a quadtree method and its performance. We show the performance of this method on simple instances of the image model. A comparison with a wavelet based coder is also done.

2. MODELING AND RATE-DISTORTION ANALYSIS

2.1. Image Models

We consider two simple models for piecewise smooth images where the edge is also a smooth curve. Our first model is the "Horizon" model [8] of an image $f(x_1, x_2)$ defined on the unit square $[0, 1]^2$:

$$f(x_1, x_2) = 1_{x_2 \geq b(x_1)} \qquad 0 \leq x_1, x_2 \leq 1,$$

where $b(x_1) \in C^p$ - the class of functions that are p-times continuously differentiable, and $b(x_1)$ has finite length inside the unit square.

When coding such an image, a possible optimal (or oracle-based) method could simply spend all the available bit rate to code the smooth edge. In other words, our 2-D object has the complexity of a 1-D function. Given a bit budget R, coding the edge b produces $\hat{b}$ and the corresponding 2-D function $\hat{f}$. The distortion can be written as

$$\begin{aligned} D(R,f) &= \int_{[0,1]^2}(f-\hat{f})^2 \leq \int_{[0,1]}|b-\hat{b}| \\ &\leq \left(\int_{[0,1]}(b-\hat{b})^2\right)^{1/2} = D(R,b)^{1/2}. \end{aligned}$$

where the second inequality comes from the Schwartz's inequality. With the edge $b(x_1) \in C^p$, coding it using a wavelet basis with enough vanishing moment gives $D(R,b) \sim R^{-2p}$ [3]. Thus the oracle R-D behavior for the "Horizon" model is

$$D(R,f) \sim R^{-p}. \tag{3}$$

In the second model, 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 P vertices of the polygon. With R/P bits for each vertex, a regular grid on the unit square provides quantized points within a distance $\Delta = 2^{-R/2P}$ 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 $D(R,f) \leq L\Delta$. Therefore for the "Polygon" model, the oracle R-D function decay exponentially as

$$D(R,f) \sim 2^{-CR}. \tag{4}$$

For a practical algorithm to code shapes in discrete space, see [9].

2.2. Wavelet Performance

Let us now demonstrate how a wavelet coder performs for such image models. Assume that wavelet transform with separable Haar wavelet is employed. At level j, wavelet basis functions have support on dyadic square of size 2^{-j}. Call n_j the number of dyadic squares at level j that intersect with the edge curve on the unit square. In both "Horizon" and "Polygon" models, the edge has finite length so we have

$$n_j \sim 2^j \tag{5}$$

Therefore, there are typically $O(2^j)$ nonzero wavelet coefficients at the scale 2^{-j} . This is the failure point of the separable wavelet transform for 2-D piecewise smooth functions: comparing with the similar 1-D case where the number of nonzero coefficients is a constant at each scale; here this number grows exponentially as scale gets finer. As a result, the total number of nonzero wavelet coefficients up to level J is

$$N_J \sim \sum_{j=0}^{J} 2^j \sim 2^J.$$

Along the boundary, we can verify that these nonzero wavelet coefficients decay like $|c_{j,\boldsymbol{k}}| \sim 2^{-j}$ at the j-th level [10]. Now assume that we code up to J levels of wavelet coefficients with quantization step Δ. Then $\Delta \sim 2^{-J}$ or we need about J bits for each nonzero coefficient. If these nonzero coefficients are organized in a quadtree data structure, then for each coded coefficient we need at most 4 more bits to specify if its children are coded or not. For large J this indexing cost is insignificant compared to the quantization cost. Thus the total bit rate is

$$R \sim N_J J \sim J 2^J. \tag{6}$$

The total distortion is the sum of quantization errors and wavelet series truncation error

$$\begin{aligned} D(R,f) &\sim N_J \Delta^2 + \sum_{j=J+1}^{\infty} 2^j (2^{-j})^2 \\ &\sim 2^{-J}. \end{aligned} \tag{7}$$

Combining (6) and (7) we obtain the following R-D behavior of the wavelet coder for the "Horizon" model

$$D(R,f) \sim \frac{\log R}{R}. \tag{8}$$

So typically, the wavelet coder is far less effective compared with the oracle based method. It is important to stress that the regularity of the edge curve is irrelevant to the performance of the wavelet coder. In the next section we propose an efficient adaptive coding scheme that can achieve (or nearly achieve) the oracle R-D behaviors.

3. OPTIMAL QUADTREE BASED COMPRESSION

3.1. Algorithm

Our objective is to implement a compression algorithm based on the modeling assumption that images are 2-D piecewise smooth functions. In this case, if we segment the image into smaller pieces, then each sub-image can be well represented by a simpler geometrical model. For instance, we can chose a very simple model which consists of two smooth regions separated by a line.

The quadtree data structure is applied widely in digital image processing and computer graphics for modeling spatial segmentation of images. Our algorithm employs a quadtree segmentation followed by a coding algorithm on each image block in an operational R-D optimal sense. We attempt to capture the smooth discontinuities as concatenations of linear segments. We employ an operational rate-distortion optimization that is similar to the approach used in [11] in finding 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. The algorithm can be summarized as follows:

1. Segmentation of the input image: a quadtree segmentation scheme is employed as it provides a very efficient and tractable data structure.
2. Optimal representation of each sub-image by the geometrical model, in a rate distortion sense.
3. Optimization of this representation to achieve the best reconstructed image for a given bit rate constraint and distortion measure. Optimization is performed in rate distortion sense using Lagrangian cost functional for optimal quality factor.

In short we have a compression scheme which provides rate distortion optimal representation of an image with a computationally efficient algorithm.

To code a block with a separation segment, we quantize its boundary using a certain quantization step depending on a given rate and assign the best quantized values to the end points of the separation line. This algorithm will provide us an approximation of the separation curve as a concatenation of linear segments as shown in Fig. 1. It is constructive to note that the underlying dictionary of "tiles" in the above algorithm is precisely the wedgelet dictionary [8].

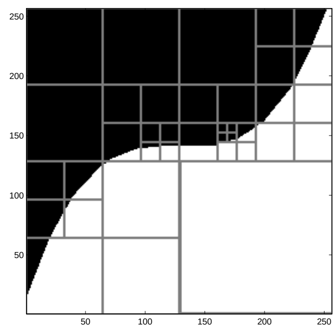

Fig. 1. Example of quadtree representation.

Notice that one can apply the ridgelet transform [6, 12] on each of these image block separately and this will result in very few transform coefficients as ridgelet transform is well adapted for line-like singularities.

3.2. R-D Behaviors

Let us first consider the "Horizon" model. The following analysis is based on [8], but we translate the non-linear approximation results of [8] to rate-distortion performance. Assume that $b(x_1) \in C^2$. The above coding algorithm employs a quadtree segmentation scheme. Each node of this quadtree corresponds to one dyadic square that is one of these four types: *intermediate, black, white, or edge* (i.e. where the boundary goes across). Thus we need 2 bits to code each node type. In addition, for an edge square, we spend $2K$ more bits for the quantized versions of the two vertices of the edge line on the boundary of that square. The encoding order of these two vertices can be used to specify the value of these two regions: for example when one traverses from the first vertices to the second one, the white region is on the left.

Let N_a and N_e be the number of all nodes and edge nodes, respectively, then the total bit rate for the quadtree is

$$R = 2N_a + 2KN_e.$$

If J is the depth of the quadtree then the number of nodes can be upper bounded as follows. N_a can be counted as the number of children of all dyadic squares up to level $J-1$ that intersect with the boundary $b(x_1)$ or $N_a \leq 4\sum_{j=0}^{J-1} n_j$. Using (5) we get $N_a \sim 2^J$. For the edge nodes, it is easy to see that $N_e \leq n_J\ 2^J$. Thus we have

$$R \sim K2^J. \tag{9}$$

In terms of distortion, consider the situation where we divide the interval $[0,1]$ of the x_1-axis into 2^J dyadic segments of length $\Delta = 2^{-J}$. Since $b(x_1) \in C^2$, it can be written by Taylor series as

$$b(x+h) = b(x) + b'(x)h + O(h^2).$$

Thus there exist a concatenation of line segments $b_J(x_1)$ of $b(x_1)$ where the vertices are in the perimeter of dyadic squares of size length $\Delta = 2^{-J}$ such that

$$\max_{x_1 \in [0,1]} |b(x_1) - b_J(x_1)| \leq C_1\Delta^2 = C_1 2^{-2J}. \tag{10}$$

The line segments in $b_J(x_1)$ need to be "quantized" with K bits for each vertex along the perimeter of dyadic squares of size 2^{-J} to produce the new function $b_{J,K}(x_1)$ where

$$\max_{x_1 \in [0,1]} |b_J(x_1) - b_{J,K}(x_1)| \leq C_2 2^{-J} 2^{-K}. \tag{11}$$

The construction of $b_{J,K}(x_1)$ leads to a tiling coding scheme of the function $f(x_1, x_2)$ with J quadtree levels and K bits for the edge vertices. By choosing $J = K$, the distortion is bounded by

$$\begin{aligned} D(R,f) &= \int_{[0,1]^2} (f - \hat{f})^2 \\ &\leq \max_{x_1 \in [0,1]} |b(x_1) - b_{J,K}(x_1)| \\ &\leq C2^{-2J}. \end{aligned}$$

Combining (9) with (12) we will result into the following R-D upper bound for the optimal quadtree algorithm

$$D(R,f) \sim \frac{\log(R)}{R^2}. \tag{12}$$

Thus for $p = 2$, the optimal quadtree algorithm has the same decay rate for the R-D function as that of the oracle method. For general p, the same conclusion can be derived when we replace "linear" tile in the described algorithm with "polynomial" tile of degree $p-1$. Furthermore, using the generalization in [8], the result also hold when the discontinuity curve $b(x_1)$ is piecewise smooth or have more general geometry.

We now consider the "Polygon" model where the edge is a piecewise linear curve with a *finite* number of discontinuity points. In such a case, we will show that the R-D function of the quadtree algorithm decays exponentially with the rate.

In an optimal adaptive quadtree, at each level, the only dyadic squares that need to be divided further are the ones containing a discontinuity points of the edge in the original image. Other dyadic squares contain either no edge or a straight edge which can be efficiently represented by an edge tile. Let P be the number of discontinuity point along the original piecewise linear edge. Then for the quadtree described above, at each level there are P splitting nodes, and thus they generate no more than $3P$ dyadic squares with straight edge at the next level. With those edge squares, extra bits are needed to code the vertices of the line segment as the quantized points along square boundary.

With J the depth of the quadtree, we can spend J bits for each of the quantized points[1] to ensure that the maximum distance

[1]A variable bit rate applying to different size length of dyadic squares can be used, but it will not change the main behavior of the R-D function

between the true vertices and their quantized version is bounded by 2^{-J}. The total bit rate is

$$R \leq 8PJ + 6PJ^2, \tag{13}$$

where the first term is the cost of coding the quadtree itself, while the second term is the cost for the quantized vertices of the edge squares. Since the original edge has finite length L, the distortion is bounded by $L2^{-J}$. Together, the R-D function for coding a "Polygon" image using optimal quadtree method behaves like

$$D(R) \sim 2^{-C\sqrt{R}} \tag{14}$$

4. NUMERICAL EXPERIMENTS

Numerical experiments are performed for two type of classes of B/W Images: 1) Images with polygonal singularity. Polygon's vertices are generated randomly using uniform distribution on the space $[0, 1]^2$. For experimentation, polygons with 4 and 5 vertices are considered as singularities. 2) Images with polynomial singularity. Polynomial coefficients are generated randomly using uniform distribution on the range $[-1, 1]$. For experimentation purposes, polynomial singularities of degree 2 and 3 are considered. For both models we have observed a better rate distortion behavior of optimal quadtree algorithm with respect to JPEG2000. In Fig. 2, we show the two rate distortion behaviors for the case of images with polygonal singularity.

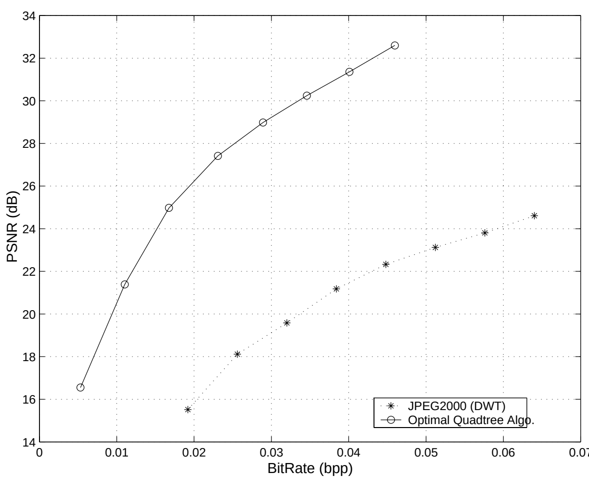

Fig. 2. Comparison of operational R-D curves using a wavelet coder (JPEG-2000) and optimal quadtree coder for B/W images with polygonal like discontinuities.

5. DISCUSSION

The optimal quadtree based compression algorithm that we have proposed in this work represents a possible way to model and approximate discontinuities in images. Several other efforts have been made to design algorithms or transforms that exploit the regularity of discontinuities in images [6, 7, 13].

Note that our analysis is completely done in a rate distortion framework. If discontinuities in our image are represented by smooth curves then the quadtree algorithm performs as well as an oracle method (the distortion rate curves have a polynomial decay in both cases). If discontinuities are piecewise polynomials (i.e. piecewise linear), the quadtree algorithm do not perform so well as an oracle method. Both distortions decay exponentially with the rate but the exponent is different ($2^{-\sqrt{R}}$ vs. 2^{-R}).

Our on-going work is to improve the quadtree based coder for this kind of image model so as to eliminate the gap with the oracle method. This will require methods similar to [4] and [5] so as to make sure that dependent coefficients are not coded independently. A second open issue is to understand if algorithms that perform correctly in a R-D sense for some restricted classes of functions can have an impact on more general classes of images.

6. REFERENCES

[1] D. L. Donoho, M. Vetterli, R. A. DeVore, and I. Daubechies, "Data compression and harmonic analysis," *IEEE Trans. Inform. Th.*, vol. 44, no. 6, pp. 2435–2476, October 1998.

[2] M. Vetterli, "Wavelets, approximation and compression," *IEEE SP Mag.*, 2001, to appear.

[3] 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. Info. Theory*, 2000, submitted.

[4] P. Prandoni and M. Vetterli, "Approximation and compression of piecewise smooth functions," *Phil. Trans. Royal Society London*, August 1999.

[5] P. L. Dragotti and M. Vetterli, "Wavelet transform footprints: Catching singualarities for compression and denoising," in *Proc. IEEE Int. Conf. on Image Proc.*, Vancouver, Canada, Sep. 2000.

[6] E. J. Candès and D. L. Donoho, "Ridgelets: a key to higher-dimensional intermittency?," *Phil. Trans. R. Soc. Lond. A.*, pp. 2495–2509, 1999.

[7] E. J. Candès and D. L. Donoho, "Curvelets – a suprisingly effective nonadaptive representation for objects with edges," in *Curve and Surface Fitting*, A. Cohen, C. Rabut, and L. L. Schumaker, Eds., Saint-Malo, 1999, Vanderbilt University Press.

[8] D. L. Donoho, "Wedgelets: nearly-minimax estimation of edges," Tech. Rep., Department of Statistics, Stanford University, 1997.

[9] G. M. Schuster, G. Melnikov, and A. K. Katsaggelos, "Operationally optimal vertex-based shape coding," *IEEE SP Mag.*, vol. 15, no. 6, pp. 91–108, Nov. 1998.

[10] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2nd edition, 1999.

[11] 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.

[12] M. N. Do and M. Vetterli, "Orthonormal finite ridgelet transform for image compression," in *Proc. IEEE Int. Conf. on Image Proc. (ICIP)*, Vancouver, Canada, Sep. 2000.

[13] E. Le Pennec and S. Mallat, "Image compression with geometric wavelets," in *Proc. IEEE Int. Conf. on Image Proc.*, Vancouver, Canada, Sep. 2000.