A List Directed Approach to Fractal Image Coding in the Wavelet Transform Domain

I Messing and D Malah
Department of Electrical Engineering*
Technion, Haifa 32000, Israel
imessing@siglab.technion.ac.il

ABSTRACT

This paper describes a technique for Fractal Image Coding in the Discrete Wavelet Transform (DWT) domain employing variable size range and domain subtrees. The DWT domain is partitioned by means of a top-down quad-tree algorithm. The splitting decision function is implemented by a list ordered with respect to a rate-distortion based partitioning gain. The partitioning and encoding are done at the same pass. Each range subtree has three optional coding modes: zero-tree, fractal prediction by a best fit domain subtree from a pool of the same size and "Intra" scalar quantization of the subtree coefficients. The choice between fractal coding and Intra coding is rate-distortion based.

1 Introduction

In recent years the Fractal Transform range-domain coding algorithm of Jacquin [1] in the spatial domain has been extended into the DWT domain [2, 3, 4, 5]. In addition to the spatial self-similarity present in the image, the wavelet domain is characterized by self-similarity across subbands at the same spatial location. These two attributes can be utilized by the fractal transform scheme for encoding sets of DWT coefficients. As in most wavelet coding schemes, after a multi-resolution (pyramid) DWT the coefficients are clustered into sets of hierarchical spatial orientation trees which have the form of quad-trees and are rooted at the lowest resolution level. These trees serve as range trees for fractal coding in the DWT domain. The corresponding domain trees are formed by taking the pyramid decomposition one step further and excluding the finest resolution. The relationship between the range and domain trees in the wavelet domain and the corresponding relationship in the image domain are shown in Fig. 1.

Each range tree is predicted by the best domain tree with respect to a selected matching criterion. The fractal transform parameters: coordinates of domain tree root coefficient, the scaling factor multiplying the coefficients of the domain tree and, if employed, the isometry index, are the compressed information to be transmitted. The best matched domain tree is obtained by a search (either extended or reduced) over the

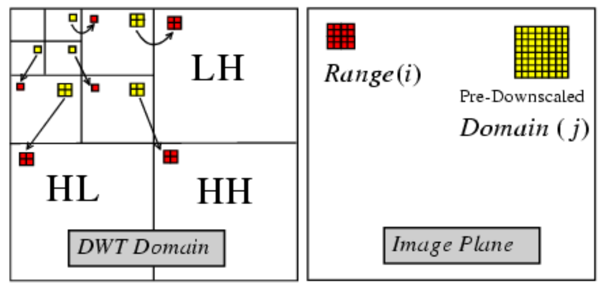

Figure 1: **Range and Domain subtrees in the DWT domain and corresponding range and domain blocks in the image domain**

pool of candidate domain trees, to minimize a certain distortion function such as MSE. This DWT-Fractal coding scheme was adopted by Krupnik et al. [2] (following the hierarchical image domain fractal coding approach of Baharav et al. [6]), Davis [3], Li et al. [4] and Zhang et al. [5]. In [2] the range trees were classified with respect to a threshold criterion into two classes: Zero-trees which do not have to be transmitted and fractal-predicted trees. Zero-trees are range trees whose energy is lower than a threshold value predefined in terms of the required PSNR value. Its introduction is inspired by the high efficiency of the Embedded Zero-trees Wavelet coding algorithm (EZW, [9]). The coefficients of the lowest resolution level which consists of a LL ("DC") subband and three directional sub bands forming the roots of the domain pool of trees are scalar quantized and transmitted together with the fractal transform parameters. The decoder initially reconstructs the lowest pyramid level coefficients. The higher resolution coefficients are hierarchically reconstructed from the lower coefficients and the fractal parameters. The performance of the above mentioned DWT-Fractal coding algorithms exceeds that of image domain fractal coders.

The algorithm of Li et al. [4] is an hybrid DWT-fractal algorithm in the sense that other options beside fractal prediction are available for encoding a range tree. It can be encoded as either "Inter" (using fractal prediction and encoding of the residual) or "Intra" (direct encoding of the tree coefficients). Both residual and "Intra" trees are encoded by an efficient bitplane wavelet coder and mode selection is based on a rate-

*This work was performed at the Signal and Image Processing Lab (SIPL) and supported by the Ministry of Science, Grant No. 8635196

distortion criterion. Its performance exceeds EZW.
Recently, Zhang et al. [5] have extended the DWT-fractal coding algorithm to enable variable range and domain sizes, employing a top-down quad-tree partitioning method for the wavelet domain. Starting from the coarsest wavelet pyramid level, each range tree is predicted by a domain tree from a pool of corresponding size if the collage error does not exceed a predefined threshold value. Otherwise, the tree root is split off and coded independently and a fractal prediction for the root-less tree is attempted by means of an appropriately modified domain pool. If the threshold value is exceeded, this range tree is split into four children trees to be coded separately and the process is repeated with the smaller size trees until either a minimum size tree is reached or the predefined bit budget is exhausted.
In this paper we describe a variable range size DWT-fractal coding algorithm which employs a simpler quad-tree approach by omitting the root-less tree prediction step. Also, rather than the simple predefined threshold method we propose a novel partitioning algorithm termed list directed top-down approach. This approach is described in section 3 following a brief discussion of DWT-Fractal coding in section 2. The simulations results are described in section 4 and are followed by concluding remarks in section 5.

2 Fractal Coding in the DWT Domain

The translation of the fractal transform into the wavelet domain can be formulated by means of the simplest orthogonal wavelet transform- the Haar-DWT. We apply the Haar transform U_L with $L = \log_2(B)$ subbands to a range block R_i of size B^2. The wavelet coefficients ${r^l}_i, l = 1, ..., B^2$, of the range block R_i are given by: $[r_i] = {U_L}^T \cdot R_i \cdot U_L$. The wavelet coefficients can be arranged in a L level tree structure [2]. Thus, each tree has $B^2 - 1$ high-pass coefficients, and r^1 is the only low-pass coefficient which corresponds to the DC of the block denoted by $\bar{R}$. The high-pass coefficients are distributed in L subband levels, each consisting of 3 directional subbands: LH, HL and HH (Fig. 1). A similar expression can be obtained for the downscaled domain block D_j. The relationship between the coefficients of the range and domain subtrees is: ${r_i}^k = a_i \cdot d^k_{j(i)}$, $2 \leq k \leq B^2$, where ${d_j}^k$ are the coefficients of domain subtree D_j and a_i are scaling factors. The relationship between wavelet trees and blocks in the image domain exists also for other wavelet transforms employing filters which are longer and smoother than the Haar pair. The 9 taps filter originally suggested in [7] is adopted for the present work.
In fixed-size range and domain trees fractal coding algorithms [6], [3] and [4], the wavelet domain is partitioned into $(N/B)^2$ range trees (for an image of size NxN) of equal size rooted at the coarset subband. For each tree the best domain tree is matched by exhaustive search over $(N/2B)^2$ trees comprising the domain pool. The matching error used is MSE, defined in the wavelet domain by:

$$d(R_i, \hat{R}_i) = \sum_{k=2}^{B^2} \left(r_i^k - a_i \cdot d^k_{j(i)}\right)^2 \qquad (1)$$

where $\hat{R}_i$ is the fractal predicted value of range block R_i. The fractal transform parameters are obtained by minimizing this error function with respect to the scaling parameter a_i and domain location $j(i)$. (If isometry is applied, we have to minimize over its 8 possible operations too). In order to exploit the sparseness and self similarity properties of the wavelet domain representation of images, a threshold value is defined in [6] and all subtrees whose energy values $\sum_{k=2}^{B^2} (r^k)^2$ are below it are classified as zero-trees and their coefficient values are set to zero. In the variable-size case [5], several domain pools are prepared, corresponding the different size range subtrees but the prediction procedure is the same as for the fixed-size range case.

3 A list-directed approach to DWT-fractal image coding

Our basic idea is to sort the subtrees (of variable size) with respect to a partitioning gain and prepare an ordered list of subtrees or quad-tree nodes. A pyramid DWT decomposes the image into an hierarchy of Kmax+1 levels, Kmax being the coarsest (upper) quad-tree stage. Starting from the nodes rooted at Kmax, we calculate a partitioning gain for each subtree. This gain expresses the benefit in the rate-distortion (R-D) sense of splitting it into four children subtrees and encoding them together with the parent root coefficient as compared to encoding the subtree itself. The list is sorted with respect to the partitioning gain which is obtained by computing $\frac{\Delta \mathbf{D}}{\Delta \mathbf{R}}$, $\Delta \mathbf{D} =$ (distortion of subtree) - (sum of distortions of 4 children subtrees) and $\Delta \mathbf{R} =$ (sum of bit-rates of 4 children subtrees) - (bit-rate of subtree). The node with the highest gain value is split into four children nodes and the list is than updated by adding its four children subtrees, deleting the parent subtree and re-sorting. If subtrees are not to be split from R-D considerations or because they are rooted at level Kmin+1, Kmin being the finest (lowest) quad-tree stage, their partitioning gain is set to a negative value and they are entered into a second list of leaf nodes.
The proposed algorithm steps are as follows:

1. Pyramid DWT decomposition into Kmax+1 levels.
2. Encoding and reconstructing the four lowest resolution subbnands at Kmax+1 using a uniform scalar quantizer.
3. Initiating the list of range-tree nodes by calculating the partitioning gain of all trees rooted at level Kmax. The leaf list is empty.
4. Scanning the tree nodes list. If the partitioning gain is negative the tree is moved into the leaf list. Nodes list is sorted with respect to gain.
5. Checking termination condition. If met, all remaining trees in node list are moved to the leaf list, terminating the quad-tree splitting procedure. If not met, the node with highest gain value is split into four children nodes, their partitioning gain is calculated and the parent node is removed from the list. If level=Kmin+1, the children are added to the leaf list, else they are added to the node list then returning to step 4.

The algorithm is a single pass procedure: The quad-tree partitioning and subtree encoding are carried out at the same pass. The splitting process is stopped when either of the two conditions is first met: A preset termination condition is met, or the list of splitting candidates nodes is empty. The termination condition can be a predefined distortion (quality) value or a target bit-rate. In most simulations we employed the bit-rate (or total number of bits) target. At each step of the algorithm which involves encoding of data, the additional bits needed are calculated and their total number updated. The range and domain subtrees were constructed by combining the three directional subtrees LH, HL and HH into one compound subtree. This is the simplest and least bit consuming option for subtree formation. ([6] employs also a mode where these 3 subbands are fractal-predicted independently. We have found that the rate-distortion gain of such a mode does not justify its inclusion in our variable range size algorithm). Thus, each subtree has three root coefficients instead of one. A splitting decision involves the encoding of the three directional root coefficients of the subtree by means of a uniform scalar quantizer.

Leaf subtrees can be encoded by either of the following modes: zero-tree, fractal predicted or "Intra". Their mode classification is performed as part of the partitioning gain determination procedure. A zero-tree subtree is one whose energy does not exceed a predefined threshold value. The threshold value can be related to a required PSNR by means of the expression [6]: $Th = 255^2 \cdot 10^{-\frac{PSNR(dB)}{10}}$. We have found in our simulations that a different threshold value for each subtree size, usually a decrement of PSNR by 1-3 dB per level, enhances performance. The initial threshold and decrement values are the only adjustable parameters in our algorithm. A non zero-tree range tree is classified as either fractal predictable or "Intra", depending on which of the options has the lowest R-D based gain (the ratio of block distortion error to the number of assigned bits).

No attempt was made to encode the fractal prediction error (residual coding). Since the actual distortion incurred by fractal prediction is not available beforehand it is substituted by the collage error . We have attempted to estimate the expected difference between the fractal prediction error and the collage error by analyzing the results for a given image. An average correction factor was computed and served for future simulations. As expected, small subtrees which comprise of coefficients from finer subbands have bigger correction factors (for the Lena image: 1.0 for the coarsest level Kmax=5 and 1.4 for Kmin=2). The overall effect of this correction on the image quality or PSNR is rather small. If the "Intra" mode is chosen, the wavelet coefficients of the subtree are independently scalar quantized and transmitted.

The transmitted code consists of the following data:

- Parameters essential for decoding the data, i.e. Kmax and Kmin.
- Indices of quantized coefficients of the four coarsest subbands.
- The quad-tree structure which maps the leaf nodes.
- Coding mode index for each leaf node.
- Index of domain location and quantized scaling factor for fractal encoded subtrees.
- Indices of the three quantized root coefficients in the case of split subtrees.
- Indices of quantized subtree coefficients in the case of Intra coded subtrees.

4 Experimental results

The proposed algorithm was tested on the image Lena (512x512). A 6 level DWT pyramid of the image was constructed by means of the 9 taps orthogonal filter of [7]. The quad-tree partitioning is initialized by range subtrees rooted at level Kmax=5 of the pyramid. The smallest allowed subtrees are rooted at level Kmin=2 of the pyramid, corresponding to a four levels quad-tree. The coarsest four sub-bands (level 6) of the pyramid are quantized by means of a 7 bits linear quantizer. Since an extended search over the multiple pools of domain trees is time consuming, we have employed a reduced search mode whereby only the small group of 8x8 nearest domain trees about the position of the range tree is actually searched. For bit rates lower than 0.5 bpp there is no noticeable degradation in performance and there is a 10 fold processing-time saving. The scaling factors of the fractal transform are linearly quantized by a 6 bits quantizer.

The quantization of the wavelet coefficients (3 root coefficients of split subtrees and the coefficients of Intra subtrees), is done according to a bit allocation computed on the basis of the variances of the corresponding subbands. The method is a simplified version of the scheme employed in [8] for the FBI fingerprint compression standard:

$$r_k = r_0 - log_2 Q'_k + log_2(\frac{\Delta_k}{\Delta_0}) \qquad (2)$$

r_i is bit rate for subband i, Δ_i is the input range for the quantizer of subband i (twice the maximum absolute value of the coefficients) and Q'_k is the quantizer step size ratio $\frac{Q_k}{Q_0}$. subscript 0 refers to the coarsest (Kmax+1) subband.

Table 1 displays the PSNR results for the image Lena at three bit-rate values employing the reduced search algorithm and the collage error correction factor. Fig. 2 show the original image Lena and reconstructed images at 0.25 bpp and 0.5 bpp.

These results were obtained without entropy coding of the encoder bit stream. Recently it was demonstrated [10] that applying arithmetic coding to the bit stream of the proposed coder results in significant improvement in PSNR performance. For the Lena image, for example, the same PSNR values cited in Table 1 are achieved at a 16-20% lower bit rate.

Figure 2: **Lena Image. From left to right: Original, reconstructed at 0.25 bpp and reconstructed at 0.5 bpp**

bit rate bits per pel	PSNR dB	ZT thres. dB	ZT dec. dB	Collage error dB
0.25	30.86	32.0	2	31.36
0.35	32.42	38.0	3	32.51
0.5	33.59	38.0	1	33.67

Table 1: PSNR values for image Lena. ZT tres is the threshold for zero-tree classification. ZT dec is the decrement of the threshold per quad-tree level, starting from Kmax. Results shown are without entropy coding

In order to estimate the effectiveness of the list directed partitioning approach, we have done some simulations on a simpler quad-tree partitioning algorithm employing a predefined threshold value as a splitting criterion. Comparing the results for the image Lena at 0.35 bpp, for example, the list directed approach shows a 1.5dB improvement in PSNR with significantly better visual quality.

5 Conclusion

In this paper a modified version of an hybrid fractal-wavelet coding algorithm is presented. We described a list directed approach to quad-tree splitting in the DWT domain. The splitting criterion is based on rate-distortion considerations instead of a threshold parameter commonly employed and results in improved performance. The procedure is a single pass, top-down, thus compromising optimality for simplicity. The performance demonstrated is not as good as current best wavelet coders but can still be improved, e.g., by using better wavelet filters, more efficient coding of the “Intra” subtrees, encoding the residual in the case of fractal predicted subtrees and by the introduction of an additional coding mode of “in situ” prediction of a range subtree from a domain subtree at the same position.

References

[1] A.E. Jacquin, ”Fractal Image Coding: A Review”, Proc. IEEE, Vol.81, No.10, pp. 1451-1465, October 1993.

[2] H. Krupnik, D. Malah, and E. Karnin, ”Fractal Representation of Images via the Discrete Wavelet Transform”, IEEE 18th National Convention of EE in Israel, March 1995, pp 2.2.2-1 - 2.2.2-5.

[3] G. M. Davis, “A Wavelet Based Analysis of Fractal Image Compression”, IEEE Trans. Image Proc. , Vol. 7(2), pp. 141-154, Feb. 1998,.

[4] J. Li and C.-C. Jay Kuo, ”Image Compression with a hybrid Wavelet-Fractal Coder” IEEE Trans. Image Proc., Vol. 8(6), pp. 868-874, June 1999.

[5] Ying Zhang and Lai-Man Po, ”Variable Tree Size Fractal Compression for Wavelet Pyramid Image Coding”, Psignal Processing: Image Communication, 14, pp. 209-227, 1999.

[6] Z. Baharav, H. Krupnik, D. Malah, and E. Karnin, ”A Multi-Resolution Frame Work for Fractal Image Representation and its Applications”. *This work was presented at the Fractal Image Encoding and Analysis Advanced Study Institute held in Trondheim, Norway, July 1995, and was erroneously omitted from the book (edited by Y. Fisher) containing the proceedings. It can be downloaded from: http://www-sipl.technion.ac.il/Sipl/NatoAsiPaper.ps.*

[7] E.H. Adelson, E. simoncelli and R. Hingorani, ”Orthogonal pyramid Transforms for Image Coding, SPIE Proc., Vol. 845, pp. 50-58, 1987.

[8] C.M. Brislawn,J.N. Bradley, R.J. Onyshczak and T. Hopper, “The FBI Compression Standard for Digitized Fingerprint Images”, Proc. SPIE Vol. 2847, (Denver), August 1996.

[9] J.M. Shapiro ”Embedded Image Codec Using Zerotrees of Wavelet Coefficients”, IEEE Trans. Signal Proc., Vol. 41, No. 12, pp. 3345-3462, Dec. 1993.

[10] A. Kubrik and D. Malah “Fractal Image Coding Final Report”, SIPL internal report, Technion 1999.