Wavelet-Based Fractal Transforms for Image Coding with no Search

Saeed Asgari
Western Digital Corporation
1599 N. Broadway
Rochester, MN 55906
asgari@wdroc.wdc.com

*T. Q. Nguyen**
Boston University
Electrical & Computer Engineering
44 Cummington St.
Boston, MA 02215
nguyent@bu.edu

William A.Sethares
University of Wisconsin-Madison
Electrical & Computer Engineering
1415 Engr. Dr.
Madison, WI 53705
sethares@ece.wisc.edu

ABSTRACT

The compression performance of fractal image coding is considered using the wavelet-based fractal coder with no search or classification of the domain blocks. A new partitioning scheme is introduced as a variant of earlier schemes which further improves the compression performance. The Wavelet-Based Fractal Transform (WBFT) links the theory of multiresolution analysis (MRA) with iterated function systems (IFS). This not only provides a local time-frequency analysis on (the partitions of) the image using multiresolution representation but also an iterative construction of the same (partitions of the) image using IFS and fixed point theory. A set of experiments and simulations show the potentials of using the WBFT for image coding after uniform quantization and entropy coding of the coefficients of the transform. Possibilities for further improvements are discussed.

Keywords: fractals, nonlinear contractive mapping, wavelet decomposition, image coding

1. Introduction

The basic idea in data compression is to reduce the number of bits required to represent stored data, by choosing an encoding scheme that can remove redundancy while keeping the useful information intact. There must also be a robust decoding scheme to reconstruct the original data. Many approaches have been suggested for image compression including JPEG, fractal techniques (see [4], [7], and [9]), and wavelet based compression (see [8]). This paper proposes a technique that incorporates many of the best features of both the fractal and the wavelet based approaches. The method of fractal image compression identifies an image as the unique fixed point of a finite collection of contractive maps on the space of fractals which is often referred to as Hyperbolic Iterated Function System (IFS) [4]. The attractor of an IFS is the "fixed point" of the collection of contractive maps whose important characterizing property is its self-similarity.

The objective is that by subdividing the image into non-overlapping partitions, referred to as range blocks, and bigger (overlapping) partitions, refered to as domain blocks, and searching for the best set of (affine) contractive transformations among all possible domain block candidates that match the selected range block (in mean square sense). The domain blocks are taken from a decimated version of the same image. Information will inevitably be lost since the fixed point or the attractor of the IFS is only an approximation to the original image. The closer the Collage (or union) of obtained maps to the original image the better the approximation [4].

Although one can rarely come up with a perfect reconstruction system using fractal image compression, very high compression ratios can often be achieved using IFS techniques such that the decoded image is reconstructed with little loss in resolution and fidelity [6]. This suggests very promising potential applications for fractal image compression using IFS.

The Wavelet-Based Fractal Transform (WBFT) as introduced in [1] and [3] links the theory of multiresolution analysis (MRA) with Iterated Function System (IFS). It is shown in [1] that the WBFT may be used as a (near) perfect reconstruction system where the standard (affine) fractal transforms (used by others), in general, may not. The hope is that by using nonlinear (rather than affine) operators, together with the theory of multiresolution analysis, better image fidelity can be achieved while retaining significant compression. Since the mathematical foundation of IFS theory is based on the Contraction Mapping Theorem on the space of fractals, the non-linear maps must be contractive. These non-linear contractive maps will then be used in place of the affine contractive maps (used by others) as building blocks for an IFS image compression scheme.

The paper is organized as follows. The wavelet-based fractal transform is introduced and its properties are discussed in section 2. Experimental results and future extensions are described in section 3.

*This work is supported in part by the National Science Foundation under grant MIP-9501589 and by the Wisconsin Alumni Research Foundation.

2. Wavelet-Based Fractal Transforms

The WBFT is introduced in [1] where its properties and condition for (strict) contractivity are investigated. In the paper, even though only the 2D case is considered, the theory holds true for the 1D case as well. Define $\mathcal{F} = \mathcal{L}^2(I^2)$ to be the space of all images (i.e. functions) $f : I^2 \to \Re$ where I^2 is the unit square and $\mathcal{L}^2(I^2) = \{f : \int_{(x,y)\in I^2} | f |^2 \, dxdy < \infty\}$. The definition above assumes infinite resolution for a given image where in fact the "pixels" of a digitized image, $\mu(f)$, can be regarded as the samples of an image (or a function) $f \in \mathcal{F}$. Any given image $f \in \mathcal{F}$ is assumed to have a range in the unit interval. Therefore, the range of the image is first mapped to the interval $[0, 1]$ using an affine transformation. The transformation has to be invertible so that the image can be mapped to its original range after the image is reconstructed at the decoder (see [1] for more detail).

Define the saturation function $h(x)$ to be

$$h(x) = \begin{cases} x & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } 1 < x \\ 0 & \text{otherwise.} \end{cases}$$

The saturation function may be used (at the decoder) to make sure that the range of the image belongs to the interval $[0, 1]$ where $f(x, y) = 0$ corresponds to black and $f(x, y) = 1$ to white. Any given image $f \in \mathcal{F}$ is assumed to have a range in the unit interval. Therefore, the image is clipped if necessary. Moreover, define the d_{rms} metric between two images f and g to be $d_{\text{rms}}(f, g) = \left(\int_A (f - g)^2 dL\right)^{\frac{1}{2}}$ where L is the Lebesgue measure on $\Re^2$. It is shown [1] that the saturation function h is a contractive function with contractivity factor of one under the d_{rms} metric.

Since the domain of the image (I^2) is partitioned into rectangular tile cells, the geometric maps, $\mathbf{w}_k$, (which take a parent cell to its k'th tile cell) are represented with affine transformations, i.e., $\mathbf{w}_k(x, y) = \mathbf{w}_k(\mathbf{x}) = \overline{\mathbf{A}}_k\mathbf{x} + \overline{\mathbf{B}}_k$. The $\mathbf{w}_k$'s for $k = 1, \ldots, N$ are assumed to be invertible (i.e. $\det(\overline{\mathbf{A}}_k) \neq 0$ and det stands for determinant) where N is the number of nonoverlapping tile cells. Let $\mathcal{A}_k = \mathbf{w}_k(I^2)$ for $k = 1, \ldots, N$ be the support of each tile cell then it is assumed that $\bigcup_{k=1}^{N} \mathcal{A}_k = I^2$ and $\mathcal{A}_i \bigcap \mathcal{A}_j = \mathcal{A}_{i,j}$ for all $i \neq j$ where $\mathcal{A}_{i,j}$ is the set whose elements are the edge of the intersection between $\mathcal{A}_i$ and $\mathcal{A}_j$. Observe that $\mathcal{A}_{i,j}$ are sets of (Lesbegue) measure zero for all $i \neq j$.

The Multiresolution Analysis (MRA) suggests a new operator in the space of all images, $\mathcal{F}$, which is no longer an affine transformation. This operator will be used in place of the affine "massic" map of Jacquin [7] and others, and is a generalization of the Monro's operator [9]. Using separable scaling functions (see [8] and [5]) and given any $g \in \mathcal{F}$, the

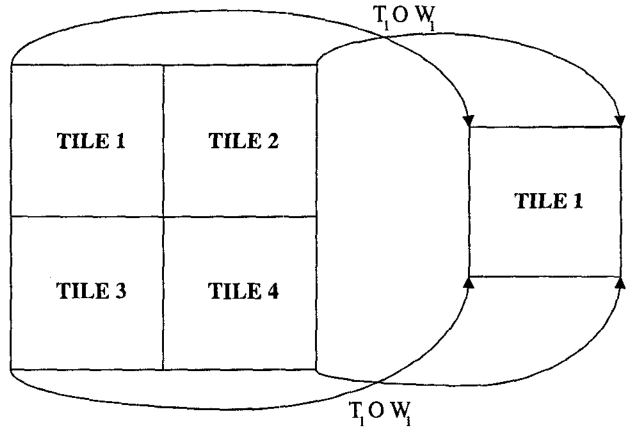

Figure 1. The mapping $T_1 \circ \mathbf{w}_1$ in action.

(piecewise) wavelet-based fractal operator, T_k, has the form

$$T_k(g)(x, y) = 2^{\frac{M_1}{2}} \sum_n c_n^k \phi(2^{M_1} g(x, y) + n) + 2^{M_2} \sum_{l,m} b_{l,m}^k \phi(2^{M_2} x + l)\phi(2^{M_2} y + m) \quad (1)$$

for all $k = 1, \ldots, N$. The transformation T_k is a (piecewise) nonaffine transformation on $\mathcal{F}$. Observe that given the coefficients c_n^k and $b_{l,m}^k$, $\tilde{g}_k(x, y) = T_k(g)(x, y) = \nu_k(x, y, g)$ is a fractal function approximation to tile k, $g(\mathbf{w}_k(x, y))$, for all $(x, y) \in I^2$ and $k = 1, \ldots, N$. The parameters M_1 and M_2 are integers (≥ 1) representing the number of resolution levels needed, $\phi(.)$ is the (orthonormal) scaling function (e.g. of Battle-Lemarie kind), and the coefficients c_n^k and $b_{l,m}^k$ are the inner product coefficients in the sense that is explained in [8] and [5]. Observe that by the resolution levels M_1 and M_2 we actually mean the resolutions 2^{M_1} and 2^{M_2}. The index n runs from $-K_1 - L_1, \ldots, K_1 - L_1$ where $K_1 = \frac{2^{M_1} - p_1 + 1}{2}$ and $L_1 = 2^{M_1 - 1}$. Similarly, m and l run from $-K_i - L_i, \ldots, K_i - L_i$ where $K_i = \frac{2^{M_2} - p_i + 1}{2}$ and $L_i = 2^{M_2 - 1}$ for $i = 2, 3$. The coefficients p_i for $i = 1, \ldots, 3$ are integers (≥ 1) which act as safeguard parameters analogous to the ones used in the wavelet network of [11]. Figure 1 shows a parent cell along with its four tile cells and the operator $T_1 \circ \mathbf{w}_1$ in action.

Now let $\mathbf{1}_{\mathcal{A}}(x, y) = 1$ if $(x, y) \in \mathcal{A}$ and zero otherwise be the indicator function associated with the set $\mathcal{A}$. Furthermore, let $\bar{g}_k(x, y) = g(\mathbf{w}_k^{-1}(x, y))\mathbf{1}_{\mathcal{A}_k}(x, y)$ and the image (or function) $g_k(x, y) = g(x, y)\mathbf{1}_{\mathcal{A}_k}(x, y) = g(\mathbf{w}_k(x, y))$ for all $(x, y) \in I^2$ and $k = 1, \ldots, N$. Then the operator T defined as $T = \sum_{k=1}^{N} T_k$ or more precisely

$$T(g)(x, y) = \nu(x, y, g) = \tilde{g}(x, y) = \sum_{k=1}^{N} T_k(\bar{g}_k)(x, y) \quad (2)$$

is a self-transformation which maps $\mathcal{F}$ to itself for all $(x, y) \in I^2$ where $\tilde{g}(x, y)$ is a fractal function approximation

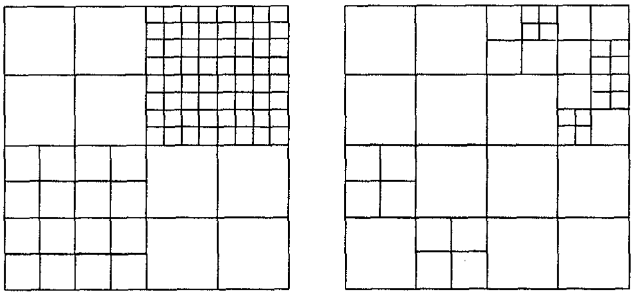

Figure 2. a). The partitioning method 1 and b). The partitioning method 2.

to a given parent cell $g \in \mathcal{F}$ (for a given set of coefficients of the contractive operator T_k).

The coefficients of the operator T_k can be found by minimizing $d(g_k, T_k(\bar{g}_k))$ with respect to the corresponding parameters under the constraint that the operator T_k is (strictly) contractive where d is some suitable metric such as the d_{rms} metric. The conditions for (strict) contractivity of the operator T_k and the WBFT, T, under the d_{rms} metric are derived in [1]. Observe that the same conditions hold for the operators $T_k \circ h$ and $T \circ h$ since h has contractivity factor one under the d_{rms} metric. Furthermore, it is shown in [1] that this constrained optimization problem can be solved via any interior point method of optimization such as the LMI toolbox of MATLAB. Since solving the constrained optimization problem may slow the coder down, one can first find the least square solution without considering the constraints for (strict) contractivity of the WBFT. If the constraints are satisfied save the coefficients or else force the coefficients c_n^k to zero and solve for the least square solution again where in this case the contractivity factor of the operator T_k is known to be zero.

This paper considers two different partitioning schemes to study compression performance which requires no search for the best match beteen a given tile cell and a parent cell candidate. In each case, for a given tile cell, the parent cell will be the larger (square) block which contains it (refer to Figure 1). The first partitioning scheme is a modification of level zero search of Monro [9]. We call this partitioning method 1. The performance of fractal image coding based on partitioning method 1 is shown in [3] without entropy coding of the coefficients of the blockwise transforms. The second partitioning scheme is analogous to quadtree partitioning of Fisher et al. [6] et al. with the modification that a search is not performed over the whole image. First, we begin with a large sized parent cell. Only the tiles that do not match (in mean square sense) with the parent cell will be broken to smaller tile cells and this is iterated until a match is found. We call this partitioning method 2. Each parent cell is assumed to be, a square, partitioned into (four) nonoverlapping tile cells with equal sizes [9] (refer to Figure 1). Figure 2a-b compares the two partitioning methods.

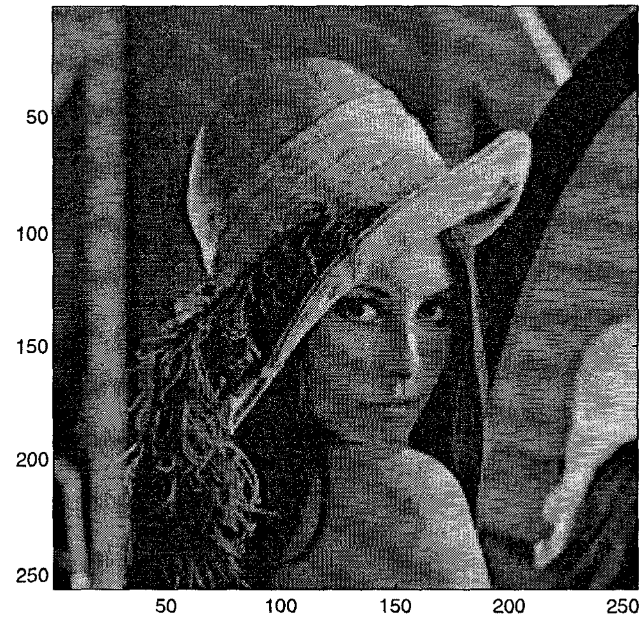

Figure 3. The reconstruction of Lenna.

3. Experimental Results and Conclusion

This section discribes the results of applying the WBFT for image coding. For compression, we first evaluate the WBFT by using either partitioning scheme 1 or 2 applied on the components of a given test image (e.g. Lenna). Next, we scan the coefficients of each block transformation independently at each resolution M_2 in different block sequences and subtract their mean. This will give rise to zero mean block coefficients at each resolution level. We then assess the decrease in the rms error of approximation which occur by uniformly quantizing each zero mean block with a different stepsize at each resolution. Finally, Huffman coding with runlength is applied to the quantized (zero mean) block sequences to achieve further compression without loss. Figure 3 shows the reconstructed image for a 256 x 256 Lenna starting with a black initial image after three iterations of the IFS algorithm. Further iterations did not result in much improvement. The unquantized reconstructed image has a rms error of 3.204 or a PSNR of 38 dB with respect to the original image. The reconstructed image after quantization and entropy coding has a compression ratio of 4.6:1 and PSNR of 35.33 dB.

From these results we select combinations to achieve high, meduim, low and poor fidelity images. Observe that image coding via WBFT does not require either search or parent block classifications. Moreover, it has the power of a (near) perfect reconstruction system and most of the other fractal approaches, in general, do not.

Table 1 summarizes the simulation results using the partitioning method 2 and Figure 4 compares the wavelet-based coder with other fractal methods sited in UWaterloo BragZone public domain (at http://links.uwaterloo.ca) tested on

Table 1. Compression of Lenna at various bit rate using no search.

Fidelity	High	Medium	Low	Poor
PSNR in dB (quantized)	35.58	31.18	28	25.06
rms (quantized)	4.24	7.04	10.2	14.24
rms (unquantized)	3.204	6.4	9.703	13.49
Max Error (quantized)	42	47	80	104
CR (unquantized)	2.92:1	4.3:1	7.6:1	16:1
CR (entropy)	4.6:1	8:4	15:1	32:1

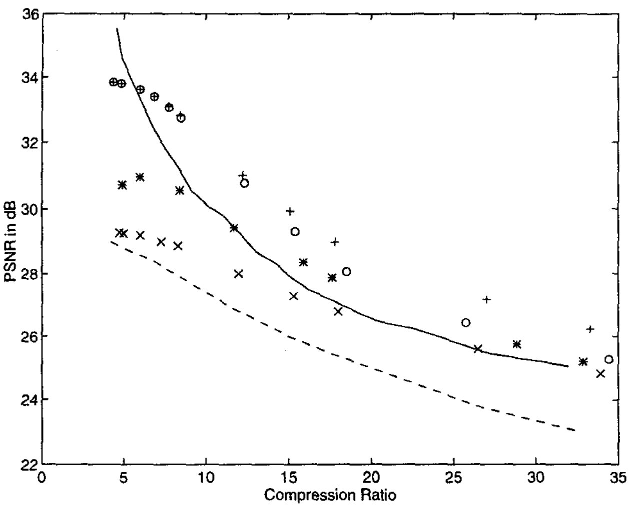

Figure 4. PSNR vs. Compression ratio.

a 256 x 256 standard Lenna image.

The other fractal methods used for comparison are TRNA (Fisher, 3-level quadtree) marked with '+', TRNB (Fisher, 4-level quadtree) marked with 'o', FIFC (Images Incorporated) marked with '*', FIFA (Images Incorporated, version 3.1) marked with 'x', and finally FIFB (Images Incorporated, version 4.0) marked with the dashed line in Figure 4. The result of simulations for the proposed wavelet-based coder is marked with the solid line in the same graph.

As the results show, the performance of the wavelet-based coder is very comparable to the best fractal method sited in UWaterloo BragZone which is by Fisher et. al (i.e. TRNA and TRNB).

Future studies are in progress to further improve the current wavelet-based fractal coder. One consideration is to incorporate search into the proposed algorithms. Although the encoding time will incearse, it is easy to see that performing a search can only improve the coder/decoder. Also different choice of the scaling functions and in general basis functions for the blockwise image (function) approximation and representation can be investigated. Another important future extension of this work is to look into better ways of entropy coding the wavelet-based fractal coder which optimally takes advantage of the existing correlation among the obtained coefficients of the fractal transform. To this end, one immediate approach may be to incorporate the use of Shapiro's zerotree embedded scheme [10] into the current proposed entropy coder. Therefore, look for a parent/child relationship and further exploit redundancy due to inherent self-similarity in the coefficient space between the parents and their associated childs. Furthermore, the notion of bit plane mapping used in Shapiro [10] zerotree algorithm, which relies on the fact that larger coefficients correspond to the areas with the most energy concentration and hence important to be encoded with more accuracy, holds true in this case as well and for the same reason.

References

[1] Saeed Asgari, *Constrained Networks for Fractal Image Compression.*, PhD thesis, University of Wisconsin-Madison, Madison, Wisconsin, May 1997.

[2] Saeed Asgari, T.Q. Nguyen, William A. Sethares, *Wavelet Networks with Constraints for Fractal Image Compression*, submitted for publication to IEEE Transactions on Image Processing.

[3] Saeed Asgari, T.Q. Nguyen, William A. Sethares, *A Wavelet-Based Approach to Fractal Image Compression, Proceeding of NORSIG*, pp. 323-326, 1996.

[4] M. Barnsley, *Fractals Everywhere,* Acedemic Press, Boston, 1988.

[5] I. Daubechies, *Ten Lectures on Wavelets,* Philadelphia, PA: SIAM Press, 1992.

[6] Y. Fisher, D. Rogovin, T.P. Shen, "A Comparison of fractal methods with DCT and wavelets," *Proceedings of the SPIE,* Vol. 2304-16, *Neural and Stochastic Methods in Image and Signal Processing III,* 1994.

[7] A. Jacquin,"Image coding based on a fractal theory of iterated Contractive image transformation," *IEEE Trans. Image Processing,* Vol. 1, pp. 18-30, 1992.

[8] S.G. Mallat, "A theory for multiresolution signal decomposition the wavelet representation," *IEEE Trans. PAMI,* Vol. 11, No. 7, July 1989.

[9] M. D. Monro, "A hybrid fractal transform," *Proceeding of ICASSP,* Vol. 5, pp. 169-172, 1993.

[10] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients, " *IEEE Trans. Signal Processing,* vol. 41, pp. 3445-3462, Dec. 1993.

[11] Jun Zhang, Gilbert G. Walter, Yubo Miao, and Wan Ngai Wayne Lee, " Wavelet neural networks for function learning," *IEEE Trans. on Signal Processing,* Vol. 43, No. 6, June 1995.