Zerotree Wavelet Coding Using Fractal Prediction

Kai Uwe Barthel, Sven Brandau, Wolfgang Hermesmeier, and Guido Heising*

Institut für Nachrichtentechnik und Theoretische Elektrotechnik
Technische Universität Berlin
*now with NTEC MEDIA Babelsberg
e-mail: kai@ntec-media.de

ABSTRACT

Fractal image coding can be seen as spectral prediction of coefficients in the wavelet decomposition of an image. Fine scale coefficients are described by shifted and scaled coefficients from the next coarser frequency scale. We analyzed the ability of fractal coders to predict wavelet coefficients and found that the prediction of entire wavelet subtrees - as done by conventional spatial domain fractal coders - leads to poor coding results. This paper discusses how to design an improved fractal coder in the wavelet domain and proposes a new hybrid fractal-wavelet zerotree coding scheme. Although the coding results of the new coder outperform conventional fractal coders, they are unfortunately inferior compared with a pure wavelet version of this zerotree coder.

1. INTRODUCTION

The *discrete wavelet transform* of an image provides a set of wavelet coefficients which represent the image at multiple scales [3]. The wavelet representation can be seen as a multilevel pyramid tree-structure with the coarsest scale coefficients at the top and the finest scale coefficients at the bottom of the pyramid. Typically, coefficients at different levels of the pyramid are assumed to be independent and they are quantized according to the signal statistics of the corresponding pyramid level. As an extension *zerotree wavelet coding* schemes exploit the correlation between different pyramid levels by setting entire insignificant subtrees to zero, so their coefficients do not have to be transmitted [4]. For high coding efficiency a careful selection of these zerotrees is crucial. A quantization of the wavelet coefficients followed by a zerotree encoding of the remaining indices does not lead to optimal coding performance.

Fractal image coding achieves redundancy reduction by describing the original image through contracted parts of the same image, thus taking advantage of *scale-redundancy* or *self similarity* of images. The image to be encoded is partitioned into non-overlapping *range blocks*. For each of these range blocks a larger *domain block* of the same image has to be determined such that a *contractive* (*geometrical* and *luminance*) *transformation* of this block is a good approximation of the range block. At the decoder all transformations are iteratively applied to an arbitrary initial image which then converges to the fractal approximation of the image [1]. Fractal coding relies on the existence of self similarities within the image. As "fractal self similarity" can only be found in certain regions of natural images, fractal coding fails for non-fractal image contents. In its original form fractal coding suffered from low coding efficiency, difficulties to obtain high quality encodings of images, and blocking artifacts at low bitrates. Adaptive and hybrid schemes as proposed by the authors in [2] helped to improve coding efficiency. Blocking artifacts can be avoided if fractal coding is performed in the wavelet domain [5].

The original intention of this paper was to show how an "optimal" fractal coder in the wavelet domain should be designed. Fractal coding assumes that all wavelet-subtrees that make up an image can be described by other wavelet-subtrees from coarser scales. We analyzed this supposed ability of the fractal prediction of wavelet coefficients. As we found that the prediction of entire wavelet subtrees - as done by conventional spatial domain fractal coders - leads to poor coding results, we propose a new coding scheme predicting only parts of the wavelet subtree. We either allow a direct quantization or a predictive fractal approximation of the wavelet coefficients. For the selection of the coding method we employ an *entropy constrained selection* using a rate-distortion based *gradient method.*

With the new coding scheme an improved coding efficiency can be achieved compared to conventional fractal coders. However, we unfortunately found, that the best coding results are always obtained if no fractal prediction at all is performed. This is an unpleasant result for "admirers" of fractal coding. The rest of the paper is organized as follows: Section 2 describes and analyzes fractal coding in the wavelet domain. Section 3 explains the proposed new hybrid fractal-wavelet coding scheme. In section 4 some results and a conclusion is given.

2. FRACTAL CODING IN THE WAVELET DOMAIN

Under the partitioning constraint that every domain block is made up of an integer number of range blocks conventional fractal coding can be described in the Haar-wavelet domain [5]. The approximation of a range block through a

0-8186-8183-7/97 $10.00 © 1997 IEEE

contracted domain block in the spatial domain also can be described as the prediction of fine scale coefficients from coarse scale coefficients in the wavelet domain. The spatial contraction (lowpass average filtering and subsampling) corresponds to moving coefficients to the next higher scale (fig 1). Now the decoding can be performed in a non-iterative way by consecutively extrapolating higher frequency coefficients from lower frequencies. Fractal coding in the wavelet domain is not limited to the Haar-wavelet. The usage of smooth basis wavelets corresponds to fractal coding with overlapping range blocks in the spatial domain, thus avoiding blocking artifacts.

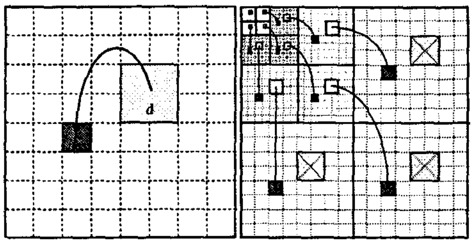

Fig. 1: Interpretation of fractal coding in the wavelet domain. The approximation of a range block f through a spatial contracted domain block d (left) can be done in the wavelet domain (right).

Design of a fractal coder in the wavelet domain

For the design of a fractal coder in the wavelet domain the wavelet basis functions and the partitioning of the domain blocks have to be chosen:

If every domain block is made up of an integer number of range blocks, no contractivity constrains on the luminance transformation have to be imposed, as this corresponds to a fractal prediction in the wavelet domain which then is causal. However, this leads to a smaller geometrical code-book, which does not contain the domain blocks that surround the range blocks. Earlier investigations on spatial domain fractal coders showed the importance of a domain pool with a finer position resolution than the size of the range blocks [8]. A finer domain block partitioning scheme including the surrounding domain blocks becomes possible by generating additional wavelet decompositions of shifted versions of the image. However, as this corresponds to a non-causal prediction, the luminance transformation therefore has to be contractive and the decoding must be done iteratively. We compared different wavelet functions and partitioning schemes. Best fractal coding results were obtained using the classical 24 tab QMF-wavelets from Johnston [6] and the 9/7-tab biorthogonal wavelets from [7]. The usage of these smooth wavelets leads to an improvement of about 1 dB compared to fractal coding in the spatial domain (= Haar-wavelet domain). While code-books with a fine domain block shift are useful in the spatial domain, in the smooth wavelet domain an unconstrained luminance transformation (non iterative dec.) is more important (fig 2).

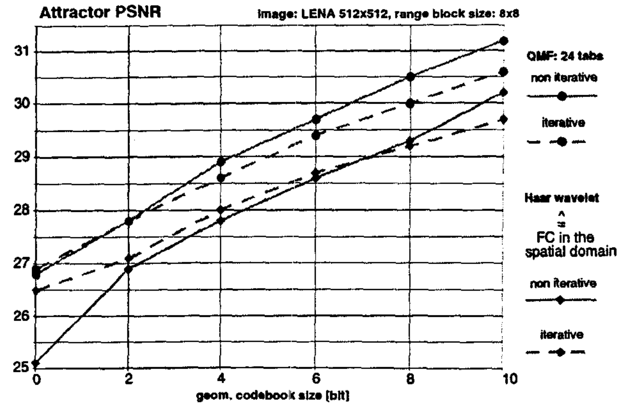

Fig. 2: Comparison of fractal coding using a Haar-wavelet and a QMF-wavelet (24 tabs) with iterative and non-iterative decoders (constrained and unconstrained luminance transformation). (For each range block its geometrical codebook of size n corresponds to the set of the 2^n closest domain blocks.)

Another important investigation was the analysis of the capability of fractal coders to predict wavelet coefficients of higher scales. Two aspects have to be considered:

- During the decoding process prediction errors will propagate from lower to higher scales.
- Doing no prediction at all may be better than a bad prediction, as a bad prediction may increase the error.

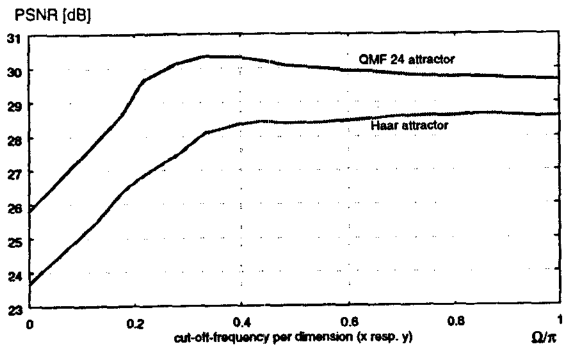

Fig. 3: Comparison of fractal coding using a Haar-wavelet and a QMF-wavelet (24-tabs) using varying prediction depths.

Figure 3 compares the results of a fractal coding where the prediction depth was varied (only the lower frequency part of the wavelet subtree was predicted, the high frequency part was set to zero). For fractal coding in the spatial domain (corresponding to the Haar-wavelet) the prediction of all range block wavelet coefficients is useful. Using smooth wavelets however, the coding performance will decrease if too many coefficients are predicted.

3. PROPOSED CODING SCHEME

The previous results can be resumed as follows: Fractal coding should be done in the wavelet domain using biorthogonal or QMF wavelets. Domain blocks should be made up of an integer number of range blocks to allow an unconstrained luminance transformation. These two aspects lead to a coding gain of 1 dB compared to spatial domain fractal coding. The best prediction of wavelet coefficients can be achieved if only the lower frequency coefficients of a wavelet subtree are predicted. These results lead to our proposed coding scheme shown in fig 4.

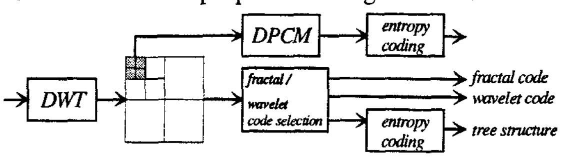

Fig. 4: Proposed coding scheme

The image to be encoded is decomposed with a discrete wavelet transform (DWT). The DC-coefficients are encoded using a DPCM followed by an entropy coder. In the beginning all wavelet subtrees are set to zero. During the encoding these trees are extended using an entropy constrained selection criterion for every node of the wavelet decomposition tree. Each node of the tree has three possible states: its four child nodes may be set to zero (which is the initial state), they may be set to their quantized values, or they may be described through a fractal prediction using four scaled coefficients from the next coarser frequency level (see fig. 5). For the encoding of the quantized wavelet coefficients for each level an adaptive arithmetic coder is used.

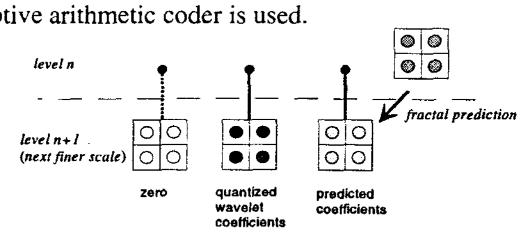

Fig. 5: Each node of a wavelet subtree may be followed by a zerotree (left), four quantized coefficients (middle) or four fractally predicted coefficients (right) which are obtained from four scaled coefficients of the next coarser frequency scale.

The fractal codes consist of two parts: The quantized scaling factor which again is entropy coded and the index of the codebook, which describes the position in the next coarser frequency level. For each set of four coefficients the codebook is centered at the corresponding position of the node in the next coarser frequency level, which is the position with maximum correlation.

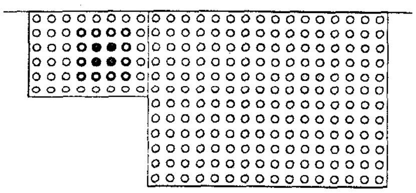

Fig. 6: The codebook for a set of four coefficients is located in the next coarser subband, centered at the corresponding position of the next higher node in the subtree. The starting positions of a 4 bit codebook are shown left.

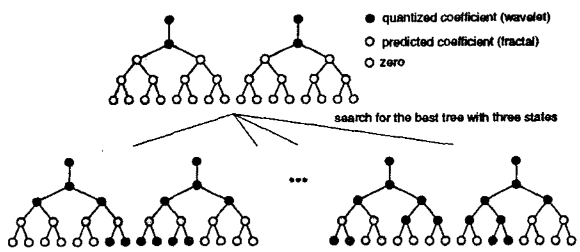

Fig. 7: A search for the best three-valued zerotree has to be performed to achieve best coding efficiency (here shown with a binary tree). Two possible trees are shown.

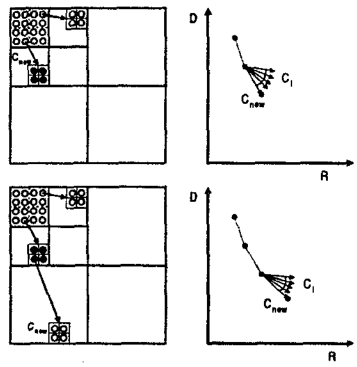

Fig. 8: Successive approximation of wavelet subtrees

The search for the best three-valued zerotree is done the following way: The entire image is described by k nodes which each describe four coefficients that are approximated by a code c_k. The Rate R and the approximation error D of the entire image are given by:

$$R = \sum_{\forall k} COSTS(c_k) \quad D = \sum_{\forall k} MSE(c_k)$$

Our coding algorithm starts to encode the image with a code with minimal R and corresponding D. From this point we improve the approximation of the image by increasing the rate step by step. At each step we change that code of a node describing four previous non coded coefficients from c_{zero} to c_{new} (see fig. 8) which yields the maximum gradient

$$\frac{\Delta D}{\Delta R} = \frac{MSE(\mathbf{c}_{zero}) - MSE(\mathbf{c}_{new})}{COSTS(\mathbf{c}_{new})} = \frac{\sum_{i=1}^{4} \hat{c}_i \cdot (2 \cdot c_i - \hat{c}_i)}{COSTS(\mathbf{c}_{new})}$$

where $\hat{c}_i$ is the approximation, while c_i is the original coefficient. The code for the entire image is changed until the desired rate or distortion is reached. At the end of the coding the set of three-valued subtrees describing the approximation, either fractal, wavelet or zero is used, are entropy coded as well. The costs for the different kind of nodes depend on the level in the tree and the approximation used. As these costs are not know in the beginning we start with an estimate of these costs and use an iterative coding scheme by encoding the image two to three times in order to determine the real costs.

The selection of the optimal zerotree consists of three steps. First all nodes are sorted by the magnitude of the gradients. In a second step the list of these nodes is resorted starting a the bottom of the tree. Now step by step the list is resorted by putting the most inefficient nodes back in the list as long the tree structure is preserved (each used node must be connected to the root of the tree). During this resorting it may happen that the approximation has to be switched from fractal to wavelet or vice versa. Fig 9 shows the four possible cases that may occur. In a final step starting at the top of the list the most efficient nodes are selected from this resorted list until the desired rate or distortion is achieved.

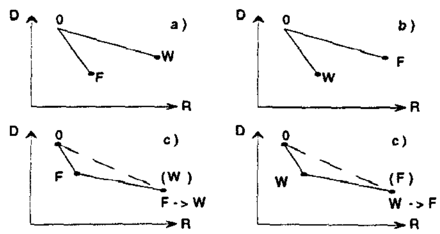

Fig. 9: The choice of the better approximation scheme is evident in cases a) and b). Switching of the approximation scheme may become necessary in the cases c) and d).

4. RESULTS AND CONCLUSION

A new scheme using a combined approximation allowing a fractal prediction and a direct encoding of the wavelet coefficients was presented. The idea of this scheme was to exploit both intra block and inter scale redundancy of an image. A rate-distortion based gradient-method was used to determine the best selection of wavelet subtrees to be encoded. Figure 10 compares the coding results of the new scheme with other coding techniques. The results of two versions of the proposed coder are given. A five step wavelet decomposition of the image was used. For the fractal prediction the scaling factor was clamped at +/- 2 and was quantized with 4 bits. The codebook sizes were set to 1 bit for the highest frequency level up to 4 bits for the lowest level. In addition to this zerotree coder with fractal prediction (W+Fr.Pr.) a version using no prediction (W) was examined as well. It can be seen that the new scheme outperforms the JPEG standard and the hybrid fractal /DCT coder from [2]. The results of the proposed coder using the fractal prediction are inferior to those using no prediction. This seems astonishing at first because an optimal selection of the coding method fractal / wavelet was used. However, the gain resulting from the fractal prediction turned out to be rather small (0.1 - 0.2 dB) but in addition a more complexly structured three-valued tree has to be transmitted which actually diminishes the gain. It can be seen that the results of the new rather complex coder are still inferior compared to the embedded zerotree wavelet (EZW) coder from [9], which uses implicit knowledge of the statistics of the wavelet spectrum. This paper was intended to present an improved fractal coder which after optimizations turned out to be a pure wavelet coder. To be able to better address the self similarities within the wavelet spectrum future schemes should rather employ context based modeling and context based entropy coding. Another possibility could be a predictive coding in the sense that the fractal part makes an estimate while the wavelet part of the coder encodes the residuum.

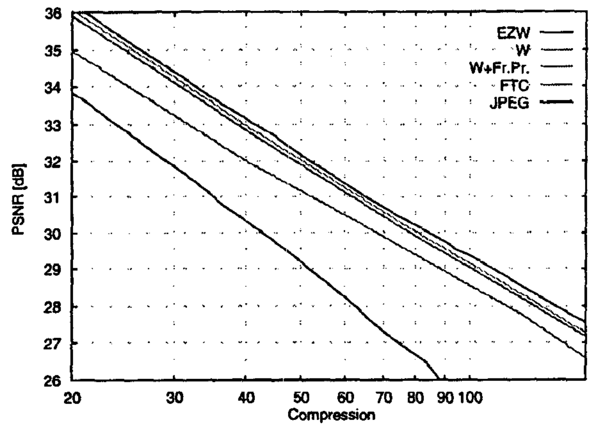

Fig. 10: Comparison of coding results for Lena (512x512). EZW: Embedded Zerotree Wavelet Coder, W: the proposed coder using no prediction, W-Fr.Pr.: with fractal prediction, FTC: Fractal Transform Coding, and the JPEG standard.

REFERENCES

[1] A. Jacquin, *Image Coding Based on a Fractal Theory of Iterated Contractive Image Transforms.* SPIE Vol. 1360 Visual Communications and Image Processing '90.

[2] K. U. Barthel, J. Schüttemeyer, T. Voyé and P. Noll, *A New Image Coding Technique Unifying Fractal and Transform Coding*, ICIP 94, Austin Texas, Nov. 94, Vol. III, pp. 112 -116.

[3] S. G. Mallat, *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 11,#7, 674 -693, Jul 1989.

[4] J. M. Shapiro, *Embedded Image Coding Using Zerotrees of Wavelet Coefficients*, IEEE Transactions on Signal Processing, vol. 41, pp. 3445 - 3462, Dec 1993.

[5] G. Davis, *Adaptive Self-Quantization of Wavelet Subtrees: A Wavelet-based Theory of Fractal Image Compression*, SPIE Conf. on Mathematical Imaging: Wavelet applications in Signal and Image Processing, San Diego, July 1995.

[6] J. D. Johnston, *A Filter Family Designed for Use in Quadrature Mirror Filter Banks*, Proc. 1980 IEEE Int. Conf. Acoust. Speech Signal Process., pp. 291 - 294., April 1990

[7] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, *Image coding using wavelet transform*, IEEE Trans. Image Processing, vol. 1 pp. 205-205, Apr. 1992

[8] K. U. Barthel, *Festbildcodierung bei niedrigen Bitraten unter Verwendung fraktaler Methoden im Orts- und Frequnzbereich*, Ph.D. Thesis, Technische Universität Berlin

[9] A. Said, W. Pearlman: *A New Fast and Efficient Image Codec Based on Set Partitioning in HierarchicalTrees*, IEEE Trans. on Circuits and Systems for Video Techn., 1995.