IEEE International Conference on Image Processing (ICIP' 94), 13-16 November 1994, Austin Texas

A NEW IMAGE CODING TECHNIQUE UNIFYING FRACTAL AND TRANSFORM CODING

Kai Uwe Barthel, Jörg Schüttemeyer, Thomas Voyé and Peter Noll

Institut für Fernmeldetechnik
Technische Universität Berlin
Einsteinufer 25, D-10587 Berlin, Germany
email: barthel@ftsu00.ee.tu-berlin.de

ABSTRACT

We present a new image coding scheme based on an unification of fractal and transform coding. We introduce a generalization of the luminance transformation used by fractal coding schemes. By extending the luminance transformation to the frequency domain, fractal and transform coding become subsets of the proposed transformation. Our new coding scheme FTC (fractal based transform coding) combines the advantages of both techniques. Compared to JPEG a coding gain of 1.5 - 2.5 dB [PSNR] is obtained. The encoding time is reduced compared to conventional fractal coding schemes and a better convergence at the decoder is attained. At equal error rates the subjective quality of images coded with the new scheme is superior compared to transform coded images.

1. INTRODUCTION

The principle of fractal image coding consists in finding a construction rule that produces a fractal image which approximates the original image. Redundancy reduction is achieved by describing the original image through contracted parts of the same image.

The principle of fractal image coding is based on the mathematical theory of *iterated function systems* (IFS) developed by Barnsley [1]. Jacquin was the first to propose a block-based fractal coding scheme for grey level images [2]. We have shown that the coding performance can be greatly improved by applying an aliasing free codebook design, an enhanced luminance transformation combined with a vector quantization, and an adaptive geometrical search scheme [3].

In this paper we describe a generalization of the luminance transformation in the frequency domain. Fractal and transform coding are subsets of this transformation. Our proposed method called *Fractal based Transform Coding* (FTC) is derived from this universal transformation. With FTC the coding efficiency can be further enhanced.

2. THE PRINCIPLE OF A FRACTAL BLOCK-CODER

The image to be encoded is partitioned into non-overlapping *range blocks* $\boldsymbol{R}$. The task of a fractal coder is to find a larger block of the same image (a *domain block* $\boldsymbol{D}$) for every range block such that a transformation of this block $\tau(\boldsymbol{D})$ is a good approximation of the range block (figure 1).

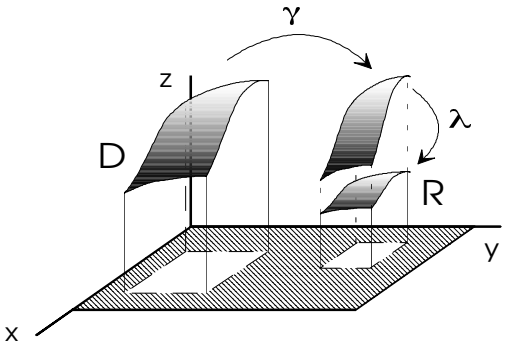

Fig. 1. Approximation of a range block through a transformed domain block.

The transformation τ is a combination of a *geometrical transformation* γ and a *luminance transformation* λ. In matrix form τ can be expressed as follows:

$$\tau\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & 0 \\ k_{21} & k_{22} & 0 \\ 0 & 0 & a \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ b \end{bmatrix} \qquad (1)$$

z denotes the pixel intensity at the position x, y. (a, b, $k_{m,n} \in \mathbb{R}$)

Only the transformations of each range block are transmitted to the decoder. This code, iteratively applied to any initial image, generates the reconstructed image. To ensure the convergence at the decoder the transformations τ have to be *contractive*.

The process of fractal encoding is lossy. At the decoder the reconstruction error increases, since the domain blocks are generated from the decoded image.

3. THE PRINCIPLE OF FRACTAL BASED TRANSFORM CODING

3.1. Conventional Transform and Fractal Image Coders

Both transform and fractal image coders are block based. Transform coders (like JPEG) use an orthogonal set of basis functions to transform each image block to the frequency domain. The coding gain results from the compacted energy distribution in the frequency domain. The quantized spectral coefficients are transmitted to the decoder. At very low bitrates the image quality is very unsatisfactory because the quantization levels are very high and many spectral coeffi-

cients are not transmitted. This leads to very annoying blocking artifacts.

Conventional fractal image coders use a fixed geometrical scaling factor of 0.5 in *x*- and *y*-directions. The domain blocks that have been scaled down to the size of the range blocks are referred to as *codebook blocks*.

Jacquin proposed a 1st order luminance transformation that scales the dynamic range and changes the brightness of a codebook block $\boldsymbol{g}$ to get an approximation of the range block $\boldsymbol{f}$ [2].

$$\hat{\boldsymbol{f}} = a \cdot \boldsymbol{g} + b \qquad (2)$$

Fractal coders using this 1st order transformation have several disadvantages:

- Only small and 'simple structured' range blocks can be approximated well.
- Coding time (for good quality approximations) is high, as the search for the best combination of geometrical and luminance transformation is exhaustive.
- Coarse quantization of the *a*/*b*-luminance parameters can lead to artifacts in the decoded image.
- The convergence at the decoder is poor. High approximation errors at the coder result in a high error propagation at the decoder.

We introduced a modified 1st order luminance transformation [3]. With decorrelated *a*/*b*-parameters (3) artifacts in the decoded image can always be avoided. The mean of a codebook block is always scaled with $a_0 = 0.5$ instead of a as done by conventional fractal coders. This results in a lower error propagation at the decoder and a better convergence is obtained.

$$\hat{\boldsymbol{f}} = a \cdot \left(\boldsymbol{g} - \mu_g\right) + a_0 \cdot \mu_g + b \qquad a_0 = 0.5 \qquad (3)$$

In literature many proposals have been made to improve the approximation of range blocks. One approach is to use additional 'basic codebook blocks', such as simple polynomial blocks [4]. Another possibility is the use of squared and cubic scalings of the pixel intensities of the codebook blocks.

3.2. An Universal Luminance Transformation in the Frequency Domain

We propose a fractal coding scheme using a high order luminance transformation in the frequency domain. This transformation uses orthogonal basis vectors and assures a linearly independent approximation and an individual contractivity control for all parameters.

An universal approach to a high order luminance transformation can be expressed in the following way:

All range and codebook blocks (of size $N \cdot N$) are transformed via the discrete cosine transform (DCT). In the frequency domain we obtain the energy compacted spectra $\boldsymbol{F}$ and $\boldsymbol{G}$ of the range and codebook blocks $\boldsymbol{f}$ and $\boldsymbol{g}$. The spectral coefficients of a block are zigzag-scanned and are referred to as F_i and G_i . By individually setting or scaling the spectral coefficients of the codebook block we can approximate the spectrum $\hat{\boldsymbol{F}}$ of the range block :

$$\hat{\boldsymbol{F}} = \begin{bmatrix} a_0 & & 0 \\ & a_1 & \\ & \ddots & \\ 0 & & a_{N^2-1} \end{bmatrix} \cdot \boldsymbol{G} + \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{N^2-1} \end{bmatrix}$$

$$\hat{\boldsymbol{F}} = \left[\hat{F}_0 \;\; \hat{F}_1 \;\; \hat{F}_2 \;\; \ldots \;\; \hat{F}_{N^2-1}\right]^T = DCT(\hat{\boldsymbol{f}})$$
$$\boldsymbol{G} = \left[G_0 \;\; G_1 \;\; G_2 \;\; \ldots \;\; G_{N^2-1}\right]^T = DCT(\boldsymbol{g}) \qquad (4)$$

Conventional fractal coding and transform coding are special cases of this universal transformation. For a fractal coder all spectral scaling factors a_i are equal for a given codebook block. b_0 serves to adjust the mean. To assure contractivity $|a_i|$ must not exceed 1. Transform coding does no scaling of any spectral codebook component ($a_0, \ldots, a_{N^2-1} = 0$). The values of b_0 to b_{N^2-1} are set to the spectral weights of the 'range block'.

Many different coding schemes are possible using any subset of the universal approach to a high order luminance transformation (4). However if all spectral coefficients were scaled or set individually, the number of transform parameters to be transmitted to the decoder would increase drastically.

3.3. Fractal Based Transform Coding (FTC)

We propose FTC as an unification of fractal and transform coding. Transform coding reduces intra-block redundancy. Fractal coding exploits inter-block similarities between different scales of the image. As a consequence FTC uses a combined approximation method.

The major part of the spectrum of a range block is approximated using the fractal transform. Any spectral coefficients that cannot be approximated by the fractal transform are individually coded as with transform coding. These spectral coefficients are excluded from fractal approximation. An approximation with n individually coded dynamic spectral coefficients can be seen as a luminance transformation of order n+1.

The following table shows examples of typical approximations, one for each of the discussed methods:

$\hat{\boldsymbol{F}} = \begin{bmatrix} a \cdot G_0 \\ a \cdot G_1 \\ a \cdot G_2 \\ \vdots \\ a \cdot G_{N^2-1} \end{bmatrix} + \begin{bmatrix} b_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\hat{\boldsymbol{F}} = \begin{bmatrix} 0.5 \cdot G_0 \\ a \cdot G_1 \\ 0 \\ 0 \\ a \cdot G_i \\ \vdots \\ a \cdot G_{N^2-1} \end{bmatrix} + \begin{bmatrix} b_0 \\ 0 \\ b_2 \\ b_3 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$	$\hat{\boldsymbol{F}} = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$
fractal coding (1st order)	fractal based transform coding (3rd order)	transform coding

Table 1. Comparison of the different approximation methods (examples).

The optimal scaling factor a_{opt} for the rest of the spectrum is found as:

$$a_{opt} = \frac{\sum_{i \notin S_{TC}} \{G_i \cdot F_i\}}{\sum_{i \notin S_{TC}} G_i^2} \qquad (5)$$

S_{TC} is the set of individually coded spectral coefficients of the block.

A good choice of the individually coded spectral coefficients is very important for high coding efficiency. As each variation of the set S_{TC} leads to a variation of the fractal code, a good solution is to be determined iteratively.

The FTC coding scheme is superior to conventional fractal coding combined with transform coding of the approximation error, because the best combination of a fractal approximation and individually coded spectral coefficients is determined.

4. FAST CODEBOOK SEARCH SCHEME

The search for a fractal transform can be seen as a search in a codebook that contains the set of contracted domain blocks. Coding efficiency strongly depends on the way in which this codebook is searched. We determined the distribution of codebook block positions that yield the best approximation for a given range block. Very often the best codebook block corresponds to the domain block directly above or close to the position of the range block to be encoded. Nevertheless for some blocks a *full search* covering the entire image is useful. Increasing the search width generally leads to a better approximation quality. To profit from the distribution of suitable codebook blocks we introduce four *search regions* relative to the position of the range block as shown in figure 2. Variable *costs* (amount of bits) for the geometrical transformation can be achieved using search regions of different sizes. For each search region the codebook block with the lowest approximation error is determined and stored. In a later step of the encoding process the most cost efficient block is selected.

To find the best codebook block of the large and full search region a huge number of codebook blocks has to be examined. Usually the correlation for every range-codebook block pair has to be evaluated. The pair with the highest correlation coefficient $|\rho|$ represents the best λ-γ-pair. Applying a full search leads to very high coding times. To cope with this problem Saupe proposes a *fast search* method for fractal image coding [5]. Range and codebook blocks are seen as vectors in a multidimensional space. The method exploits the fact that similarities between blocks are described by the orientation of these vectors in space. By applying a normalization *key-vectors* are generated. Now the search can be reduced to a nearest neighbourhood search. Friedman et al. propose an algorithm to find the *d* best neighbors in a *k*-dimensional space in logarithmic time [6].

The search time for this fast search scheme strongly depends on the dimension of the space to be searched. To reduce the dimension of the key vectors Saupe uses a sub-sampling of the blocks. We obtained reduced coding time by generating lower dimensional key vectors from the low-frequency spectral coefficients of the blocks.

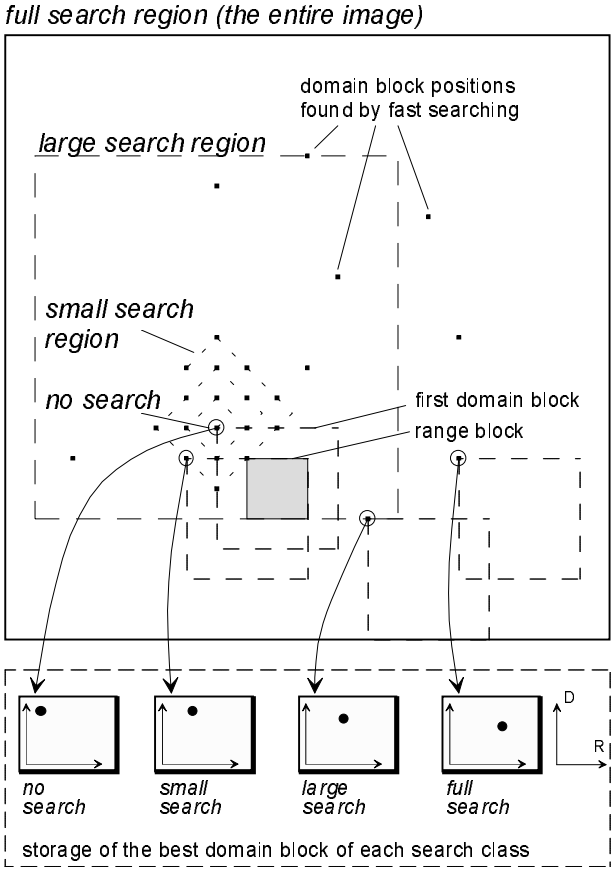

Fig. 2. Geometrical search scheme. For each search region the best codebook block is determined. Approximation errors and rates are stored.

5. CODER DESCRIPTION

Encoding an image with the fractal based transform coding scheme can be seen as an optimization process. For each range block the fractal code and the set of individually coded spectral coefficients has to be determined such that a good approximation is obtained at lowest bitrate costs.

The approximation quality can be enhanced either by enlarging the search width or by increasing the order of the luminance transformation. Both leads to higher coding costs. Conventional fractal coding schemes code all range blocks independently of each other. In order to obtain high coding efficiency we use an adaptive coding scheme that operates globally on the whole image. Our iterative coding algorithm successively improves an initial code by minimizing the rate-distortion function.

For a range block at position *i,k* a block code $\boldsymbol{c}_{i,k}$ can be defined. This block code consists of:

geometrical transformation parameters γ :
- classification (search region)
- geometrical index, and the

luminance transformation parameters λ :
- 1st order luminance transformation parameters (a/b_0)
- positions and values of TC coefficients.

The costs for describing one block code depend on the number of TC-coded coefficients and the search region used. They are referred to as $COSTS(\boldsymbol{c}_{i,k})$. The approximation error for a range block called $MSE(\boldsymbol{c}_{i,k})$ is dependent on $\boldsymbol{c}_{i,k}$ as well.

The union of all block codes forms the code $\boldsymbol{C}$ for the entire image.

$$\boldsymbol{C} = \bigcup_{\forall i} \bigcup_{\forall k} \boldsymbol{c}_{i,k} \tag{6}$$

The total amount of bits R and the average approximation error D are given by:

$$R = \sum_{\forall i} \sum_{\forall k} COSTS(\boldsymbol{c}_{i,k}) \quad D = \frac{1}{I \cdot K} \sum_{\forall i} \sum_{\forall k} MSE(\boldsymbol{c}_{i,k}) \tag{7}$$

It is desirable to choose a rate R and obtain the best possible approximation error D (or vice versa). Our adaptive coding algorithm starts to encode the image with a code with minimal R and corresponding D. This means that all block codes are initialized to 'no search' and a 1st order luminance transformation. From this point we improve the approximation of the image by increasing the rate step by step. With each step only one block code is changed. The change of a block code with luminance transformation of order n is limited to enlarge the search region or to set the luminance transformation to order n+1.

At each step we change that block code which yields

$$max\left(-\frac{\Delta MSE}{\Delta COSTS}\right) = max\left(-\frac{MSE(\boldsymbol{c}) - MSE(\boldsymbol{c}_{new})}{COSTS(\boldsymbol{c}) - COSTS(\boldsymbol{c}_{new})}\right). \tag{8}$$

The code for the entire image is changed until the desired rate or approximation error is reached.

The coding algorithm can be expressed the following way:

0. *Initialization*:
   For each range block determine the best codebook blocks of all search classes for the luminance transformation of order 1 and 2.
   Set all codes to "no search" and 1st order luminance transformation.
1. *Search* the best new block code over the whole image (figure 3). If the luminance transformation of the block code was changed to n+1 then determine the best codebook blocks of all search classes for the luminance transformation of order n+2.
2. If the rate or the approximation error fulfills the desired conditions: then stop.
   else go to 1.

When increasing the order of the luminance transformation, one problem is to determine the most suitable spectral coefficient to be coded independently. An easy approach is to select that coefficient which is responsible for the highest error component in the approximation error. This method is not optimal as it does not take into account the change of the approximation error from the fractal transform which is redetermined for the rest of the spectrum. An optimal choice has to consider a possible alteration of the geometrical and the luminance transformation. If the optimal choice is used coding time increases.

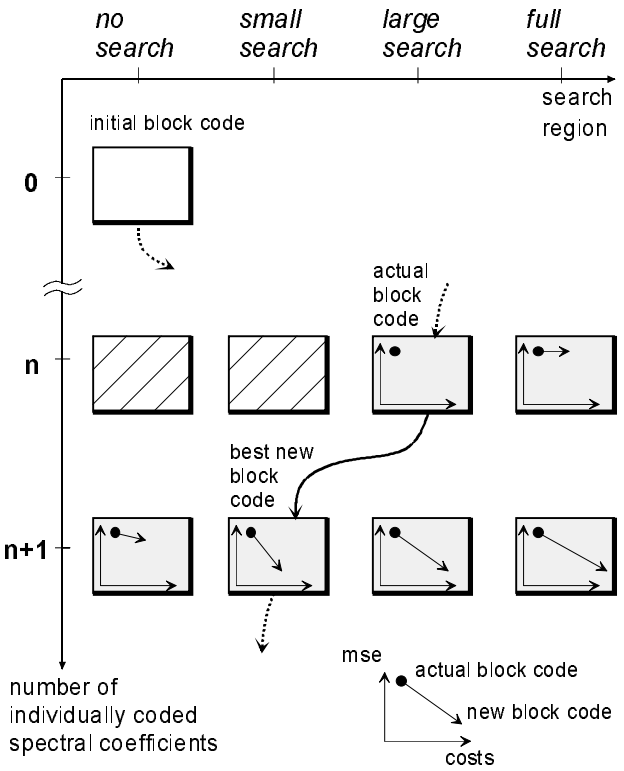

Fig. 3. Determination of the best new block code to replace the actual block code for one range block. This new block code is chosen, if $-\frac{\Delta MSE}{\Delta COSTS}$ is maximum for the whole image.

At the end of the coding process all block codes are classified to eight *classes* according to their search region and their luminance transformation (1st order or higher). These eight classes are entropy-coded.

A VQ-technique with fixed codebook size is used for the 1st order luminance transformation parameters. The high order luminance coefficients are linearly quantized. Their positions and values are transmitted using an UVLC *scheme* (*universal variable length coding*) [7].

For the coding results as shown in figure 6 and 7 a block-hierarchical coding scheme with block sizes of 16x16 and 8x8 pixels was used.

6. SIMULATION RESULTS AND CONCLUSION

We have proposed a new block-oriented fractal coding scheme using an approximation in the frequency domain. Compared to conventional fractal coders FTC better approximates larger codebook blocks to range blocks and has an improved convergence at the decoder.

The bitrate compared to TC is reduced because with FTC the number and the entropy of the individually coded spectral coefficients can be considerably reduced. This is because many spectral coefficients are approximated well enough by the fractal transform.

Fig. 4. Fractal based transform coded image (MSE = 50).

Fig. 5. Transform coded image with the same error (MSE = 50).

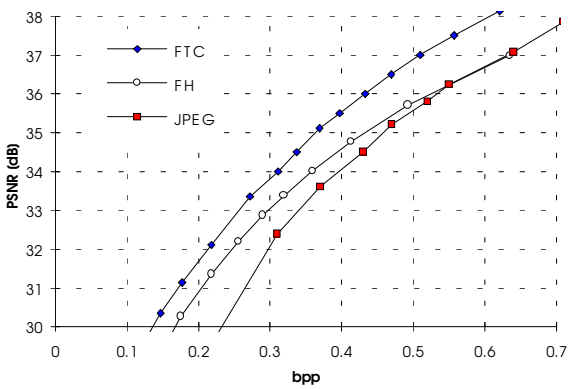

Fig. 6. Coding results for the "boats" image (512x512 pixels).
(FH: hierarchical fractal coder; FTC: fractal based transform coding)

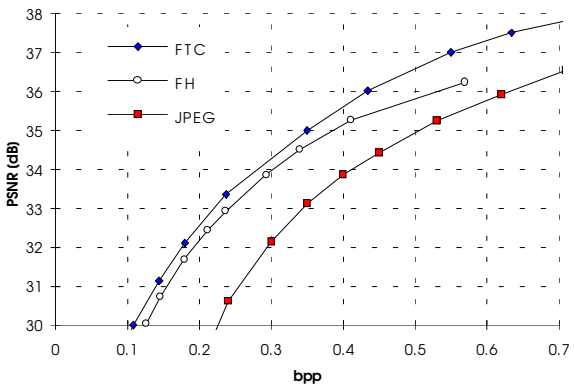

Fig. 7. Coding results for the "Lena" image (512x512 pixels).
(FH: hierarchical fractal coder; FTC: fractal based transform coding)

Figures 6 and 7 compare the coding efficiency to JPEG and our hierarchical fractal block coder using 1st order luminance transformations [3]. It can be seen that the new FTC coding scheme outperforms both other coders. Especially for higher PSNR-values the coding gain is improved compared to the conventional fractal coder.

At equal error rates the subjective quality of images coded with our new scheme is superior compared to transform coded images. Figures 4 and 5 show an example of an FTC- and TC-coded image. FTC shows less blocking artifacts, edges are better preserved. With TC many high-frequency components are set to zero, this is not the case with FTC. Our coder uses no psycho-visual weighting of the spectrum. This can easily be implemented in our scheme.

We think that this new approach is very promising as the coding results are excellent and the subjective quality of FTC coded images is very good.

REFERENCES

[1] M. F. Barnsley, *Fractals Everywhere*. New York: Academic Press, 1988.

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

[3] K. U. Barthel and T. Voyé, *Adaptive Fractal Image Coding in the Frequency Domain*, Proceedings of 'International Workshop on Image Processing: Theory, Methodology, Systems, and Applications', Budapest 1994, in Journal on Communications, Volume XLV, May-June 1994, pp. 33-37.

[4] M. Gharavi-Alkhansari and T. S. Huang, *A Fractal-Based Image Block-Coding Algorithm* , Proceedings ICASSP 93, V pp.345-348.

[5] D. Saupe, *Breaking the Time Complexity of Fractal Image Compressing*, Technical Report No. 53, Institut für Informatik, Universität Freiburg, May 1994.

[6] J. H. Friedman, J. L. Bentley, and R. A. Finkel. *An Algorithm for Finding Best Matches in Logarithmic Expected Time,* ACM Trans. Math. Software 3.3, 1977, pp. 209-226.

[7] P. Delogne, B. Macq, *Universal variable length coding for an integrated approach to image coding,* Ann. Telecommun., 46, no. 7-8, 1991, pp. 452-459.