Adaptive Fractal Image Coding in the Frequency Domain

KAI UWE BARTHEL AND THOMAS VOYÉ
Institut für Fernmeldetechnik, Technische Universität Berlin
Einsteinufer 25, D-10587 Berlin, Germany
Phone: ++ 49 30 31423838 , E-mail: barthel@ftsu00.ee.tu-berlin.de

ABSTRACT

Fractal image coding has been used successfully to encode digital grey level images. Especially at very low bitrates fractal coders perform better than cosine-transform-based JPEG coders.

A block-based fractal image coder is able to exploit the redundancy of grey level images by describing image blocks through contractively transformed blocks of the same image. Previous fractal coders used affine linear transformations in combination with 1st order luminance transformations that change the brightness and scale the luminance values of image blocks.

We propose an extension to high order luminance transformations that operate in the frequency domain. With this transformation and an adaptive coding scheme a better approximation of image blocks can be achieved. Bitrate reductions are higher than those achieved by "spatial-domain" fractal coding schemes. An additional effect of this new transformation is a better convergence at the decoder.

I. INTRODUCTION AND OVERVIEW

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 (*self-transformability*).

Fractal image coding is based on the mathematical theory of iterated function systems (IFS) developed by Barnsley [1]. Jacquin [2] was the first to propose a block-based fractal coding scheme for grey level images. In [3] we have shown that the coding performance can be greatly improved by applying a vector quantization to the optimal luminance transformation and using a better geometrical search scheme.

In this paper we describe a new luminance transformation in the frequency domain. With this transformation the coding efficiency can be further enhanced. At the decoder fewer iterations are needed to reconstruct the image.

In section II, we briefly present the principle of a block based fractal image coder. An improved codebook design and an adaptive geometrical search scheme are described in section III. The proposed new luminance transformation is presented in section IV. The description of the new coder can be found in section V. Finally, in section VI, we present some results and discuss the merits of the new coding scheme.

II. THE PRINCIPLE OF A FRACTAL BLOCK-CODER

The image to be encoded is partitioned into non-overlapping square blocks. $R_{i,j}$ is the image block at the position (i, j) and is called a *range block*.

The task of a fractal coder is to find a good approximation for all range blocks. Each range block is approximated by a transformed larger block $D_{l,k}$ of the same image (*domain block*) as shown in figure 1.

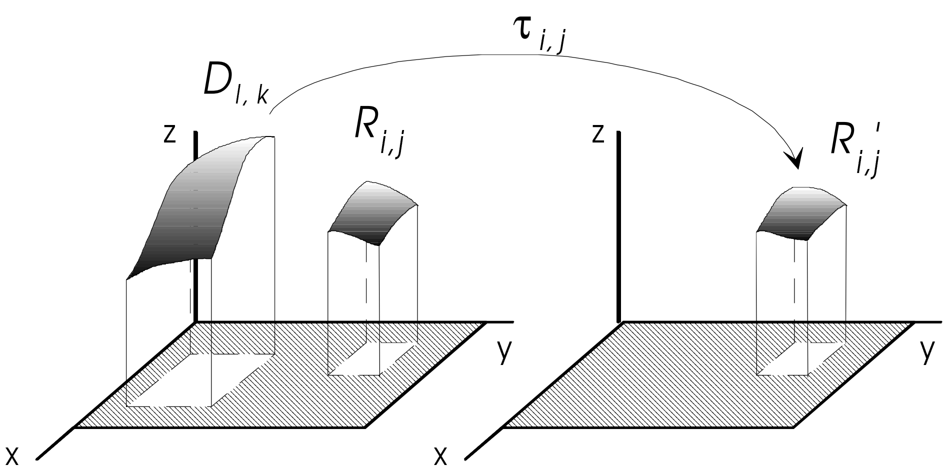

Fig. 1. Approximation of a range block through a transformed domain block

The transformation $\tau_{i,j}$ combines a geometrical transformation and a luminance transformation. The geometrical transformation is an affine linear transformation that consists of a spatial contraction and a position shift that maps the domain block to the position of the range block. The domain block that has been scaled down to the size of the range block is referred to as *codebook block*.

Jacquin proposed a 1st order luminance transformation that scales the dynamic range and changes the brightness of the pixel values of a codebook block.

In matrix form $\tau_{i,j}$ can be expressed as follows:

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

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

Only the transformations of each range block have to be transmitted to the decoder. The set of all transformations can be seen as the *fractal code* for the original image. This code, iteratively applied to any initial image, generates the reconstructed image. To ensure the convergence at the decoder the transformations $\tau_{i,j}$ have to be contractive. This means:

$$\left|\det\begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix}\right| < 1 \quad \text{and} \quad |a| < 1 \qquad (2)$$

The process of fractal encoding is lossy. The approximation error ε, that is determined at the coder increases during the decoding process since the codebook blocks are generated at the decoder from the fractal reconstruction image which is not free of errors. If the scaling factor a is assumed constant, the upper bound for the approximation error after the decoding is given by $\varepsilon / (1-|a|)$.

As the total number of transformations has to be kept low, hierarchical coders with variable range block sizes are used. If the approximation error for a large range block exceeds a given level, this block is split into up to four smaller range blocks for which additional transformations are determined.

For high coding efficiency well chosen coding parameters in combination with efficient coding of the fractal transformation parameters are necessary.

III. GEOMETRICAL TRANSFORMATION

The search for a geometrical transformation can be seen as a search in a codebook that contains the set of contracted domain blocks. Coding efficiency strongly depends on the construction of this codebook. Another important aspect is the order in which this codebook is searched.

When constructing the codebook, the set of all possible affine-linear transformations (equation 1: $k_{m,n}$, Δx, and Δy) has to be reduced to a suitable subset. As digital images are sampled images with a given spatial resolution not all affine linear transformations are possible. The size ratio of range to domain block is usually chosen to 1:2 in *x*- and *y*-directions. A smaller contraction ratio allows a better approximation of range blocks but results in a higher error propagation at the decoder. Using higher contraction ratios leads to decreasing similarities between range and codebook blocks.

To assure contractivity the codebook blocks are generated from the filtered and sub-sampled original image. Jacquin proposed a simple averaging filter. We obtained better coding results using a 10-tap anti-aliasing filter with a cut-off frequency below $\pi/2$.

We determined an efficient search path by examining 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. This fact can be used for an optimized adaptive search scheme. The codebook blocks being the most probable are examined first. The search path has the form of a spiral and starts with the codebook block directly above the range block (Fig. 3). By introducing *search regions* a variable length of the search path is possible. The search is aborted at the end of each search region if the approximation error is below a threshold value. This search scheme reduces the encoding time and the average search index. Figure 2 shows the probability density function of the search indices. In the given example an image was encoded with an 8 bit geometrical codebook as shown in figure 3. The entropy of the search indices is reduced if additional smaller search regions with error thresholds are introduced. Note that we use a relative addressing of the codebook blocks and a variable domain block shift.

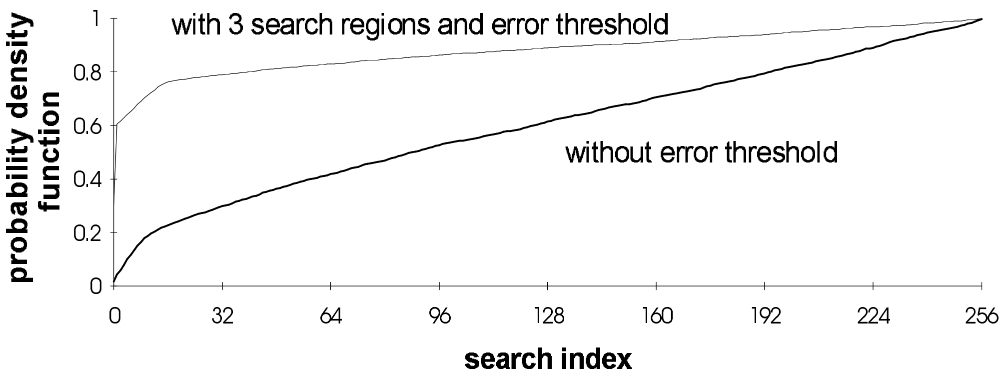

Fig. 2. Probability density function of the geometrical codebook indices using a search scheme as shown in fig. 3. The maximum search region was 8 bit. The additional search regions used a search width of 0 and 4 bits.

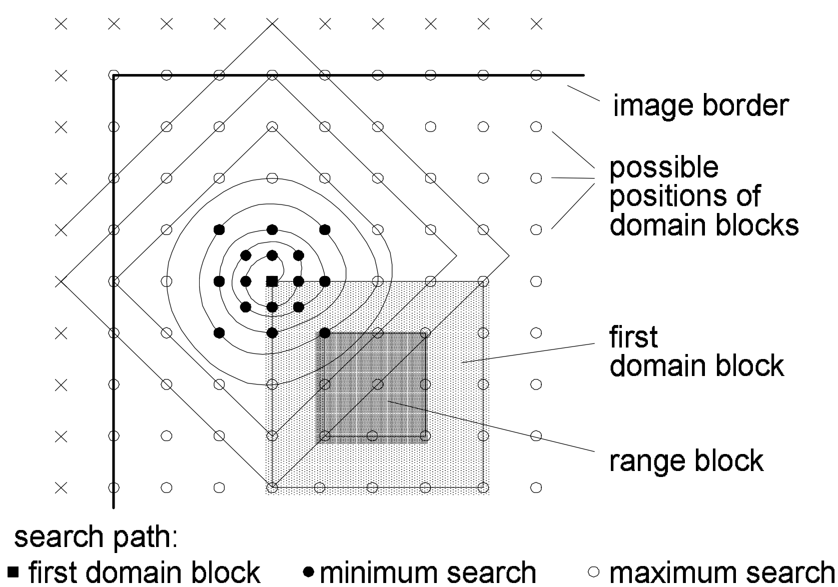

Fig. 3. Adaptive search scheme using a minimum search region

IV. LUMINANCE TRANSFORMATION

Problems of 1st order luminance transformations

The 1st order luminance transformation λ_1 proposed by Jacquin scales the dynamic range (a) and changes the brightness of the pixel values (b) of a codebook block $\boldsymbol{g}$:

$$\lambda_1(\boldsymbol{g}) = a \cdot \boldsymbol{g} + b \tag{3}$$

This 1st order transformation has two disadvantages:

- Only small and 'simple structured' range blocks can be approximated well.
- The convergence at the decoder is poor. In particular if a high approximation error is tolerated at the coder, the error propagation at the decoder is very high. In this case the number of iterations necessary to decode the reconstruction image will rise.

Modified 1st order luminance transformation

Figure 4 shows the distribution of optimal non-quantized a/b-values obtained from a fractal coder using the conventional 1st order luminance transformation.

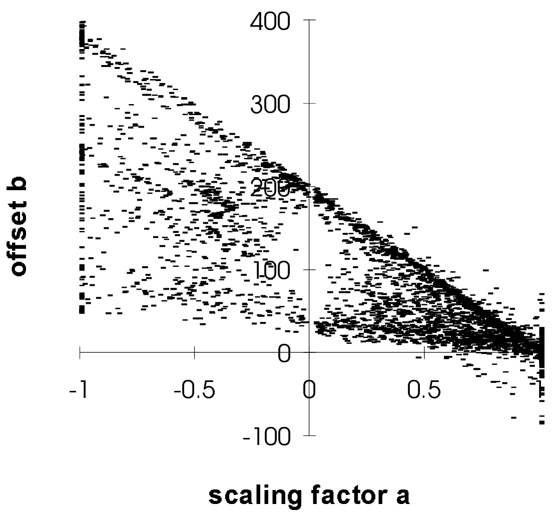

Fig. 4. Distribution of the optimal a/b-coefficients using the conventional 1st order luminance transformation

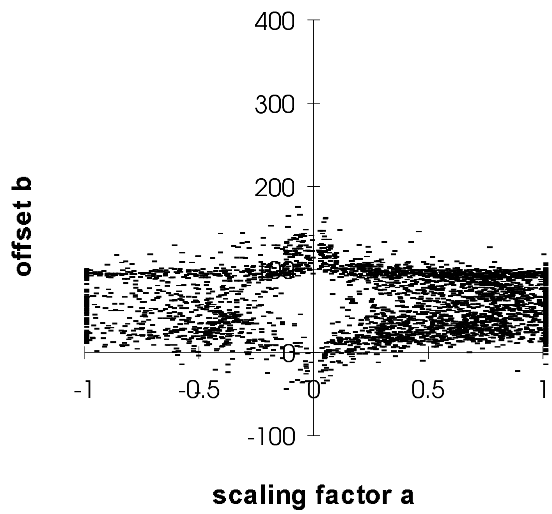

Fig. 5. Distribution of the optimal a/b-coefficients using the modified 1st order luminance transformation

From eq. (3) it can be seen that the b-offset serves to adjust the scaled means of the codebook blocks and is dependent on a. Decreasing a-values generally require increasing b-values. This leads to a triangular shaped a/b-distribution (Fig. 4).

There is a strong accumulation of the a/b-values in the region of a-values near 1. This indicates that the dynamic range of most codebook blocks is kept almost constant. Scaling values close to 1 have the disadvantage that they result in a high error propagation at the decoder. The a/b-distribution and the fact that the means represent the largest energy component of the codebook blocks result in a high upper error bound. Avoiding the scaling of the codebook means by large a-values reduces the error propagation at the decoder.

We propose a simple modification of the luminance transformation. We decorrelate the a/b-values by only scaling the dynamic part of the codebook blocks. With this modified transformation a similar approximation of the range blocks is possible. With a well-chosen factor a_0 a lower upper error bound at the decoder can be achieved.

$$\lambda_{1\,\text{mod.}}(\boldsymbol{g}) = a \cdot (\boldsymbol{g} - \mu_g) + a_0 \cdot \mu_g + b \tag{4}$$

The constant factor a_0 can be chosen from 0 to 1 and is found as a compromise: For $a_0 = 1$ the variance of the a/b-coefficients reaches its minimum, but the luminance transformation is not contractive anymore. If a_0 is set to 0 we obtain a minimal error propagation and a minimal decoding time at the decoder. In this case the variance of the *a*/*b*-distribution is maximum. Our studies have shown that with quantized *a*/*b*-coefficients the best coding results are reached for $a_0 = 0.5$. Figure 5 shows the distribution of optimal non-quantized *a*/*b*-values of the *modified 1st order luminance transformation*.

Figure 6 compares the convergence at the decoder for a critical part of the "Clown"-image. We compare the conventional and the modified luminance transformation with quantized and non-quantized coefficients. It can be seen that the modified luminance transformation outperforms the conventional transformation in the reconstruction error and the number of iterations needed. Decoding examples are shown in figure 7. The significance of the artifacts as shown in figure 7b is image dependent and they only occur with quantized parameters. By using the modified luminance transformation these artifacts can always be avoided.

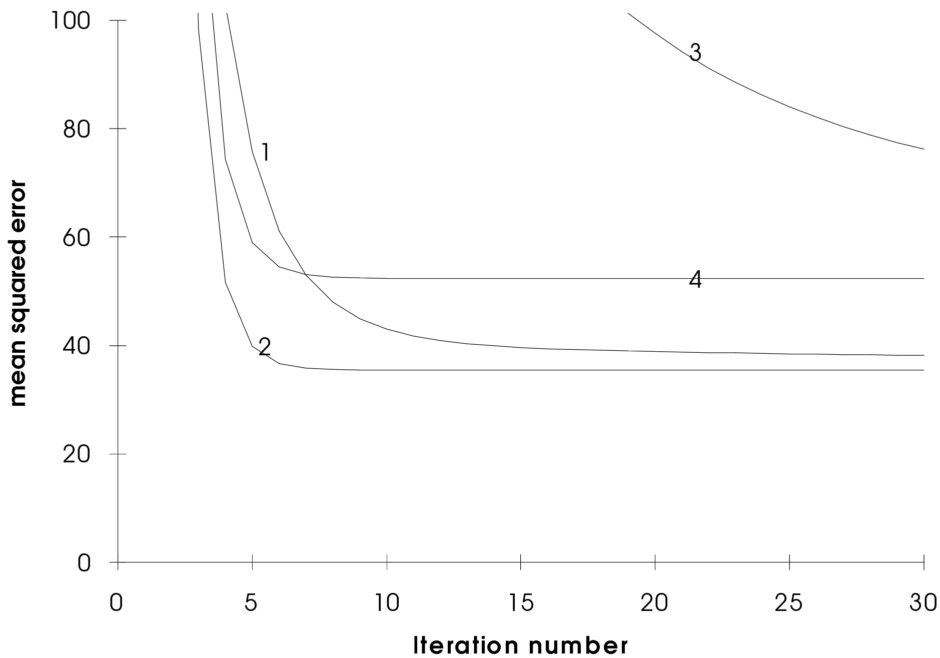

Fig. 6. Comparison of the conventional and the modified 1st order luminance transformation: conventional: 1 non-quantized, 3 quantized, modified: 2 non-quantized, 4 quantized

a) original

b) conventional 1st order luminance transformation (10th iteration)

c) modified 1st order luminance transformation (10th iteration)

Fig.7. Original and decoded images using the conventional and the modified 1st order luminance transformation with quantized a/b-values.

High order luminance transformations in the frequency domain

Any improvement in the approximation of range blocks will improve the image quality and can reduce the total number of transformations needed to describe the fractal approximation of the image to be encoded.

One approach to do this is to use additional 'basic codebook blocks' [4], such as simple polynominal blocks. We feel that such an approach is not very promising because these simple blocks are easy to encode with the fractal coder itself.

Another possibility is the use of squared and cubic scaling of the pixel intensities of the codebook blocks. The optimal scaling parameters are difficult to determine because of the dependency of the parameters on each other. A further problem is to guarantee the contractivity of such a transformation.

A high order luminance transformation has to fulfill the following conditions:

- To enable their individual adaptation the transformation coefficients should be independent of each other.
- To assure a control of the contractivity, the requirements for the contractivity should be controllable independently by the transformation coefficients.

Our proposal for a high order luminance transformation is an extension of our modified 1st order luminance transformation:

First we transform all range and codebook blocks via the discrete cosine transform (DCT). In the frequency domain we obtain the energy compacted spectra of range and codebook blocks. Then by individually setting or scaling the spectral values of the codebook block $G(u,v)$ we can approximate the spectrum of the range block $F(u,v)$.

$$\lambda(\boldsymbol{g}) = \mathrm{IDCT}\left(\bigcup_{u=0}^{N-1}\bigcup_{v=0}^{N-1} a(u,v)\cdot G(u,v) + b(u,v)\right) \qquad G(u,v) = \mathrm{DCT}(\boldsymbol{g}), \quad F(u,v) = \mathrm{DCT}(\boldsymbol{f}) \tag{5}$$

whereby N denotes the size of the blocks, the IDCT is the inverse DCT

Many coding schemes are possible using subsets of this *general luminance transformation* (5). If all spectral values were set or scaled, the number of transformation parameters to be transmitted would increase drastically. However, many range blocks can be approximated with low order luminance transformations.

In this paper we propose a coding scheme using one or more scaling factors for the dynamic part of the codebook spectrum. For a luminance transformation of order K we merge subsets of the spectral values to non-overlapping regions $\mathbf{R}_1$ to $\mathbf{R}_K$. The mean is approximated the same way as with the modified 1st order luminance transformation.

If using a 1st order luminance transformation, all dynamic coefficients are scaled with only one scaling factor (a_1). A 2nd order luminance transformation has got three regions, so the dynamic part of the spectrum is scaled with two coefficients (a_1, a_2). For luminance transformations of order 2 and higher various *frequency domain partitions* are possible. Figure 8 shows some examples of partitions for 2, 3, and 4 regions.
These high order luminance transformations can be expressed the following way:

$$\lambda_K(\boldsymbol{g}) = \mathrm{IDCT}\left(\bigcup_{u=0}^{N-1}\bigcup_{v=0}^{N-1}\begin{cases} a_0\cdot G(u,v)+b & \text{if } \; u=0, v=0 \\ a(u,v)\cdot G(u,v) & \text{else}\end{cases}\right) \qquad a(u,v) = a_i \quad \mathit{if} \quad (u,v)\in \mathbf{R}_i \quad i=1,..,K \tag{6}$$

For every region $\mathbf{R}_i$ the optimal scaling factor $a_{i\,opt}$ can be evaluated :

$$a_{i\,opt} = \frac{\sum\limits_{(u,v)\in R_i}\{G(u,v)\cdot F(u,v)\}}{\sum\limits_{(u,v)\in R_i} G(u,v)^2} \qquad b_{opt} = F(0,0) - a_0\cdot G(0,0) \tag{7}$$

To assure contractivity $a_{i\,opt}$ must not exceed a value of 1. a_0 again is set to 0.5.

If the a_i/b-parameters are limited or quantized, we get an approximation error plus a quantization error:

$$e(\boldsymbol{F},\boldsymbol{G}) = \Delta b^2 + \sum_{i=1}^{K}\left[\sum_{(u,v)\in R_i} F(u,v)^2 - \frac{\left(\sum_{(u,v)\in R_i} F(u,v)G(u,v)\right)^2}{\sum_{(u,v)\in R_i} G(u,v)^2} + \Delta a_i^2 \cdot \sum_{(u,v)\in R_i} G(u,v)^2\right]$$

$$\Delta a_i = a_{i\,opt} - a_i \quad \Delta b = b_{opt} - b \qquad (8)$$

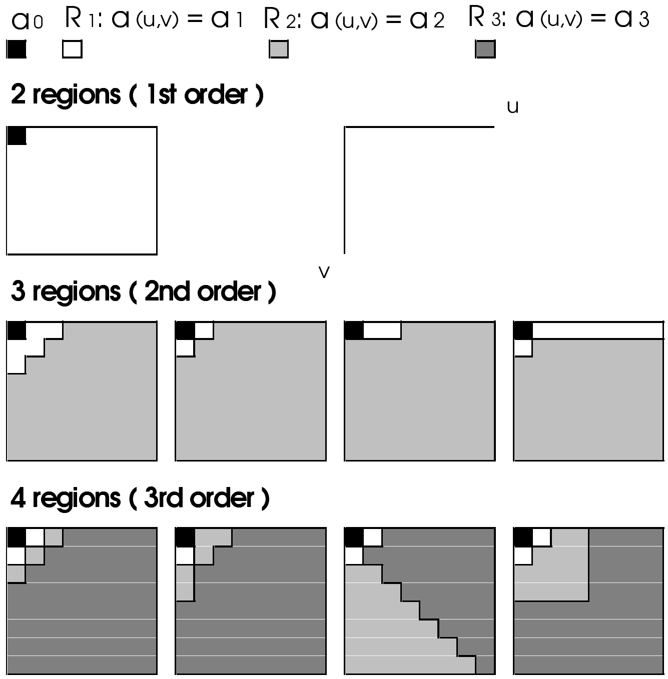

Fig. 8. Examples of different frequency domain partitions for luminance transformations of 1st, 2nd, and 3rd order (block size 8x8)

V. DESCRIPTION OF THE CODER

Before describing our new coding scheme, we suggest some important modifications of the conventional fractal coding scheme:

Improvements of the fractal coding scheme

Partial approximation:

To reduce the total number of transformations generally hierarchical coding schemes with variable range block sizes are used. In a first step, transformations for the largest range blocks of the highest hierarchy level are determined. If the approximation error is too high for any of the four range blocks of the next hierarchy level, the large range block containing these smaller blocks is split into sub blocks. For these sub blocks additional transformations are determined. The transformation for the large range block is kept if the number of additional transformations does not exceed two.

The total number of transformations can be significantly reduced if new transformation parameters are determined for the remaining part of the large range block. The coding procedure for large range blocks can then be described as follows: The transformation of a large range block and the resulting errors in the sub blocks are determined. The sub block that is responsible for the highest error component is excluded and a new transformation for the remaining *¾-block* is searched. If necessary, this procedure is repeated for the ¾-block and leads to a *½-block*. Many large range blocks that were totally split using the conventional scheme can now be coded as ½- or ¾-blocks.

Codebook-update and *a*/*b*-update:

One problem of fractal image coding is the error propagation at the decoder. It results from the fact that at the decoder the codebook is generated from the reconstructed image whereas at the coder the codebook is generated from the original image. The error propagation at the decoder can be reduced if the coder codebook is updated with the coded versions of the range blocks.

The *a/b-update* is comparable to the codebook-update. At the end of the coding process the best possible approximation of the original image is known. Now the coder could start coding the image again and again getting a better and better approximation of the decoder codebook. As this increases the coding time we propose to keep the geometrical transformations, but to redetermine the best *a*/*b*-values. The *a*/*b*-update can be repeated. We found that 1 to 2 *a*/*b*-updates are useful.

Using the modified luminance transformation and the described update procedures, the error propagation can be reduced and a slightly higher coding efficiency is obtained. The increase of the decoding error can be reduced to approximately 1 - 4 % of the coding error compared to more than 10 % using the conventional scheme.

Coder description

We use a coder with a three level hierarchy with range block sizes of 16x16, 8x8 and 4x4 pixels. For the quantization of the luminance transformation we apply a vector quantization (VQ) technique.

We use an adaptive search algorithm to determine the order of the luminance transformation and the search region that is used for the geometrical transformation.

For each hierarchy level we define a set of *search classes*. A search class contains a fixed search region and a luminance transformation with fixed order and VQ-codebook size. These search classes are searched successively. If the approximation error after searching one class fulfills a given *search stop criterion* (error threshold) the search is aborted, otherwise the next search class is examined. To obtain good coding efficiency the bit costs are increased during the search. This assures to encode a range block with the lowest necessary rate. Simulations have shown that it is useful to increase both the VQ-codebook size and the search width.

If even with the maximum search class no good approximation can be found, then this transformation is rejected and additional transformations are determined using the partial block approximation. For smaller range blocks this scheme is repeated until the highest search class of the lowest hierarchy level is reached. As the splitting criterion we check all errors of the smallest range block size.

The advantage of this coding scheme is that we can locally adapt the bitrate to the image contents. No classification of the range blocks is done before the coding process. Figure 9 shows the search classes of the lowest hierarchy level (block size 4x4 pixels). A complete set of coding parameters is shown in table 1.

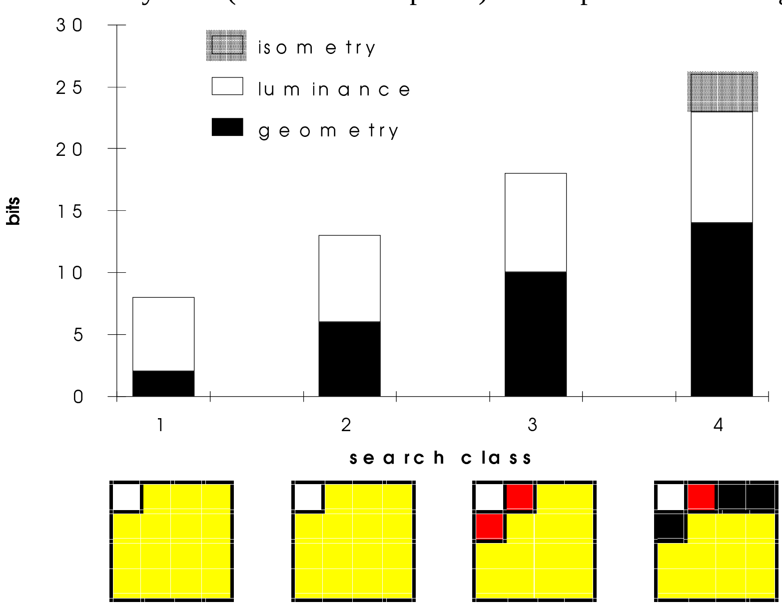

Fig. 9. Adaptive search scheme combining geometrical and luminance transformation (search classes 1 to 4) The applied frequency domain partitions of the luminance transformations are shown.

Transformation parameters to be transmitted to the decoder are:

- splitting partition of range blocks of the higher hierarchy levels,
- search class,
- geometrical index of the codebook block and the isometry (if used) and
- codebook index of the luminance transformation VQ (scaling factors a_1 to a_K and the offset b).
  (The search class of a block and the splitting partition are entropy-coded.)

The image quality respectively the bitrate can be controlled over a large range by only adjusting the error thresholds. For very low bitrates however, the block sizes have to be enlarged to 32x32, 16x16, and 8x8 pixels.

hierarchy level	**block size**	**search class**	**split error threshold** [msq]	**search stop error threshold** [msq]	***luminance transformation*** order	***luminance transformation*** VQ codebook size [**bits**]	***geometrical transformation*** codebook size (search) [**bits**]
3	16	*1*	60	15	1^{st}	**8**	**3**
		2			1^{st}	**8**	**7**
2	8	*1*	90	30	1^{st}	**8**	**3**
		2			1^{st}	**8**	**7**
1	4	*1*	-	120	1^{st}	**6**	**2**
		2			1^{st}	**6**	**7**
		3			2^{nd}	**8**	**10**
		4			3^{rd}	**9**	**14 + 3** (isom.)

Table 1. : Coding parameters used for the coding results shown in table 2.

level and search class	**classification [bits]**	**partition [bits]**	**geometry [bits]**	**luminance [bits]**	**sum [bits]**	**number of blocks**	**product [bits]**
3__x	**2**	**-**	**-**	**-**	**2**	350	**700**
3_1	**2**	**-**	**3**	**8**	**13**	313	**4069**
3_1x	**2**	**3.13**	**3**	**8**	**16.13**	148	**2387**
3_2	**3**	**-**	**7**	**8**	**18**	68	**1224**
3_2x	**3**	**3.13**	**7**	**8**	**21.13**	145	**3063**
2__x	**2**	**-**	**-**	**-**	**2**	378	**756**
2_1	**2**	**-**	**3**	**8**	**13**	517	**6721**
2_1x	**3**	**3.22**	**3**	**8**	**17.22**	236	**4063**
2_2	**2**	**-**	**7**	**8**	**17**	394	**6698**
2_2x	**3**	**3.22**	**7**	**8**	**21.22**	340	**7214**
1_1	**1**	**-**	**2**	**6**	**9**	1301	**11709**
1_2	**2**	**-**	**7**	**6**	**15**	461	**6915**
1_3	**3**	**-**	**10**	**8**	**21**	371	**7791**
1_4	**3**	**-**	**17**	**9**	**29**	225	**6525**
						total :	**69835**
				PSNR =	33.45 dB	0.266	bpp

Table 2. : Coding results for the Lena image (512 x 512 pixels).
k__x : level *k*, block is totally split;
k_j : level *k*, search class *j*, block is not split;
k_jx : level *k*, search class *j*, block is partially split (the splitting partition is additionally coded)

VI. SIMULATION RESULTS AND CONCLUSION

We have proposed a new block-oriented fractal coding scheme using an adaptive search scheme with an extended luminance transformation in the frequency domain. This transformation is able to better approximate codebook blocks to range blocks and has a better convergence at the decoder.

The bitrate is reduced because fewer transformations are needed to describe the fractal approximation of the image to be encoded. The subjective quality of images coded with our new scheme is superior compared to conventional fractal coded images. Blocking artifacts are reduced and detailed structures are better preserved.

In our simulations we used a hierarchical fractal coder with variable block sizes. Our results show that the 'Lena'-image (512 x 512 pixels) can be coded at the rate of 0.1 bpp to yield a peak-to-peak SNR of 30 dB. Figure 10 shows the coder performance compared to JPEG.

Due to the high number of parameters detailed investigations are needed to achieve optimal coding efficiency. With optimized parameters and better codebooks for the luminance transformation further improvements are to be expected.

Many different coding schemes are possible using the general luminance transformation expressed in equation 5. A new efficient coding scheme unifying fractal and transform coding will be presented in a further publication.

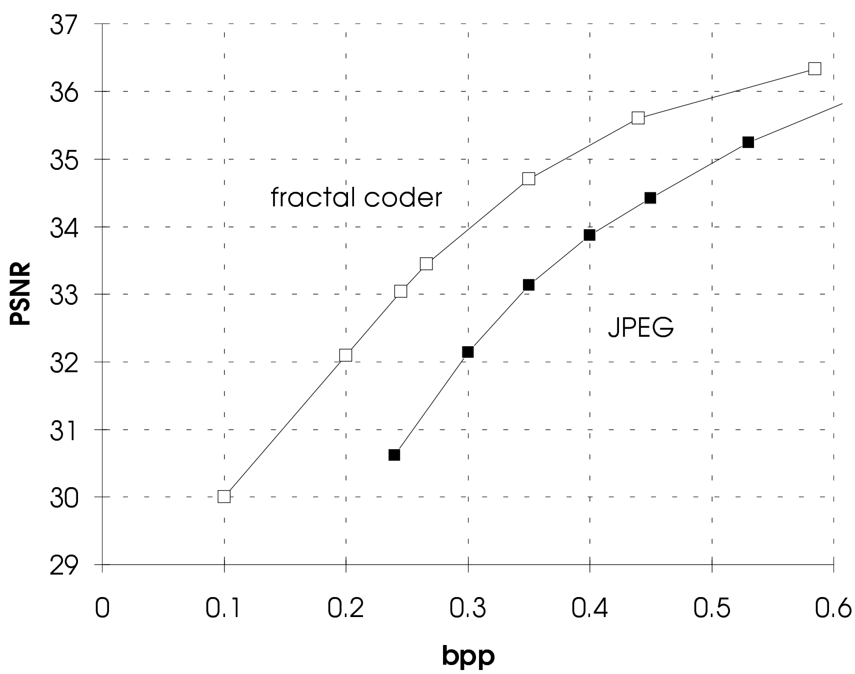

Fig. 10. Coding results of the new fractal coding scheme compared to JPEG. (Image: Lena 512x512 pixels)

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, T. Voyé, P. Noll, *Improved Fractal Image Coding*, Proceedings of PCS ' 93 , section 1.5

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

[5] E. W. Jacobs, Y. Fisher and R. D. Boss, *Image Processing. A study of the iterated transform method, Signal Processing* 29 (1992) 251-263