# Adaptive Fractal Coding of Still Pictures

*Bernd Hürtgen, Frank Müller and Christoph Stiller*

Institut für Elektrische Nachrichtentechnik
Rheinisch-Westfälische Technische Hochschule (RWTH) Aachen, 52056 Aachen, Germany
Phone: #49–241–807683, Fax: #49–241–8888196, Email: huertgen@ient.rwth-aachen.de

## Abstract

This paper presents a method for lossy coding of still pictures using the theory of iterated function systems (IFS) which is a standard description for deterministic fractals. The coding concept is based on a blockwise fractal approximation of the original image by contraction mappings of itself using affine transformations which are a special case of IFSs. Data reduction is achieved by exploiting the self-similarity within natural or artificial pictures. The proposed algorithm accounts for the local image characteristics by an adaptive quadtree segmentation. In contrast to existing fractal coding concepts an overall constant reconstruction quality is achieved.

## 1. Introduction

Picture coding has attracted great interest since the early days of image processing. Most common block-oriented techniques essentially revolve around *transform coding*, e.g. [1] and *vector-quantization*, e.g. [2]

One of the latest fields of research for coding of still pictures is *fractal coding* proposed by Barnsley [3] and at first implemented by Jacquin [4]. For a given image the encoding process consists of finding among a class of *a priori* defined contractive transformations one which leaves this image approximately invariant. Instead of the original image the optimal transformation is encoded. Reconstruction then starts with decoding this transformation and applying it iteratively to any arbitrary image. The collage theorem [5] states that the sequence of reconstructed images converges to the approximation of the original one if one can find a set of transformations which approximately maps the image onto itself.

For practical purposes the entire image is segmented into rectangular nonoverlapping blocks of same size and the optimal transformation is searched for every image block independently. Because of using a fixed blocksize most existing coding concepts are not adaptable to different image structures and therefore reconstruction quality is highly instationary.

The algorithm presented in this paper overcomes those problems by using a variable quadtree-segmentation with differently sized blocks adapted to the structure of the image which results in an approximately stationary reconstruction quality.

The paper is organized as follows: The second section describes the theory of encoding and decoding still images using a fractal approach. After this an algorithm for optimal block-splitting is introduced which accounts for the local structure of the image. At the end some coding results are presented and compared to other concepts.

## 2. Theory

Let $\mathbf{X} = (\mathbf{X}, d)$ be a complete metric space with metric $d$ and $W = \{w_k : k = 1, \ldots, N\}$ be a finite set of $N$ transformations $w_k : \mathbf{X} \to \mathbf{X}$. $W$ is called a *contraction mapping* if there exists a constant $0 \le \lambda_k < 1 \forall w_k \in W$ such that

$$d(w_k(\mathbf{x}), w_k(\mathbf{y})) \le \lambda_k d(\mathbf{x}, \mathbf{y}) \qquad (1)$$

$\forall \mathbf{x}, \mathbf{y} \in \mathbf{X}$ and $k \in \{1, \ldots, N\}$. In this case there exists a unique attractor

$$A = W(A) = \bigcup_{k=1}^{N} w_k(A) \tag{2}$$

which is invariant with respect to $W$. It turns out that $A$ is determined by $W$. This means that $A$ is the limit of approximating sequences of sets which can be constructed from $W$. In this paper we restrict the transformations to be affine mappings from the two dimensional space $\mathbb{R}^2$ into itself which can be written as

$$w_k : \mathbf{x} \rightarrow \mathbf{T}_k \mathbf{x} + \mathbf{b}_k, \tag{3}$$

where

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \mathbf{b}_k = \begin{pmatrix} b_x^{(k)} \\ b_y^{(k)} \end{pmatrix}$$
$$\mathbf{T}_k = \begin{pmatrix} M_x^{(k)} a_{11}^{(k)} & M_y^{(k)} a_{12}^{(k)} \\ M_x^{(k)} a_{21}^{(k)} & M_y^{(k)} a_{22}^{(k)} \end{pmatrix}.$$

$W$ is a *contractive* linear mapping, if $\|\mathbf{T}_k\| < 1 \forall k \in \{1, \ldots, N\}$ holds. To represent grey-scale values on an attractor $A$ we define an affine mapping extended to three dimensions where $z(x, y)$ denotes the grey-value at the position $(x, y)$. In addition to (3) the mapping for $z$ is defined as follows:

$$z(x, y) \rightarrow p_1^{(k)} x + p_2^{(k)} y + p_3^{(k)} z(x, y) + p_4^{(k)}. \tag{4}$$

Our task is now to determine for a given attractor $A$ - representing the original picture - a set $W$ of transformations $w_k$ so that $A$ is invariant with respect to $W$. In order to find an appropriate set of mappings $W$, a search in the $n$-dimensional parameter-space has to be performed where $n$ is the number of independent parameters.

For practical reasons the values for the parameters $M_x, M_y, a_{ij}, i, j \in \{1, 2\}$ and $p_l, l \in \{1, \ldots, 4\}$ cannot be chosen arbitrarily, but have to be calculated from an *a priori* given set of allowed values. Therefore we are in general not able to determine the set $W$ for a given attractor $A$ exactly, but the collage theorem guarantees that if we can find an approximation $\widetilde{W}$ to $W$ the attractor $\widetilde{A} = \cup_{k=1}^{N} \tilde{w}_k \widetilde{A}$ will be an approximation to $A$ [5].

The decoding procedure is as follows. Starting with an arbitrary image $V^0$ the transform $\widetilde{W}$ is applied iteratively. Note that in accordance with (4) the mapping $\widetilde{W}$ is extended to three dimensions.

$$V^n = \widetilde{W}^{\circ n}\left(V^0\right) \tag{5}$$

$\widetilde{W}^{\circ n}\left(V^0\right)$ denotes that $\widetilde{W}$ is iteratively applied $n$-times to $V^0$. Because the mapping $\widetilde{W}$ is contractive, the sequence of reconstructed images $V^0, V^1, \ldots, V^n$ converges to the approximation $\widetilde{A}$ of the original image $A$

$$\lim_{n \rightarrow \infty} V^n = \widetilde{A} \approx A. \tag{6}$$

## 3. Image Segmentation

The coding concept proposed by Monro and Dudbridge [6] partitions the image into rectangular blocks of same size and sets the parameters $M_x^{(k)} = M_y^{(k)} = \frac{1}{2}, a_{11}^{(k)} = a_{22}^{(k)} = 1$ and $a_{21}^{(k)} = a_{12}^{(k)} = 0$. The $b_x^{(k)}, b_y^{(k)}$ are such that the considered block is mapped to four tiles of equal size. The encoder determines the $p_l, l = \{1, \ldots, 4\}$ for each block using least square approximation.

Jacquin's coder [4] fixes $M_x^{(k)} = M_y^{(k)} = \frac{1}{2}$ too, but allows the $a_{ij}, i, j \in \{1, 2\}$ to be in a set of eight parameter-vectors which he calls *isometries*. The $p_1^{(k)}, p_2^{(k)}$ are set to zero which means that no planar facets are used and $p_3^{(k)}, p_4^{(k)}$ are determined using the least square method. The $a_{ij}$ and the offset-vector $\left(b_x^{(k)}, b_y^{(k)}\right)^T$ is calculated by applying a computationally very expensive search method which is the major disadvantage of Jacquin's coding concept.

As a fixed image partitioning results in a fixed number of blocks per image, the codelength is the same for each image of same size if no entropy coding is applied to the parameters of the transform. Hence there is no possibility to adapt the coder to different image material, compression-ratio or quality demands. One step towards adaptivity has been made by Jacquin [4] by distinguishing between parent- and child blocks in size and smooth, mid-range and range blocks in shape.

In order to overcome the above mentioned constraints a block-based image segmentation scheme is introduced which is not restricted to two different blocksizes only, but permits the blocksizes and therefore the number of blocks to vary within a wide range.

The algorithm presented in this paper is mainly based upon Jacquin's coding concept. Initially the entire image is segmented into non-overlapping squared blocks of same size $B \times B$ pixels yielding the initial segmentation. In accordance with Jacquin we call these blocks *range blocks*. We now seek to exploit similarities between those range blocks and other blocks within the image called *domain blocks*. The encoding process is carried out as follows: In a first step a *library* $L$ is created from 'all' possible domain-blocks which are allowed to be transformed by the *a priori* defined mappings. Each entry in the library therefore represents one possible set of parameters for the affine mapping.

We presuppose that if we can find a good match between a range and a library block the resulting reconstruction error is small. Determining the optimal mapping-parameters for a range block $\mu$ is therefore performed by finding an approximating library block $\tilde{\mu}$ which minimizes the distance

$$\epsilon(\mu, \tilde{\mu}) \to \min . \tag{7}$$

One common distance measure is the *mean-square-error (MSE)* defined for blocks of size $B \times B$ pixels as follows:

$$\mathrm{MSE}(\mu) = \frac{1}{B^2} \sum_{i=0}^{B-1} \sum_{j=0}^{B-1} \left(\mu(i,j) - \tilde{\mu}(i,j)\right)^2 \tag{8}$$

Let $\mu_k$ be a range block and $w_k$ its appropriate mapping minimizing $\mathrm{MSE}(\mu_k)$. We now split the block $\mu_k$ into four tiles of equal size and independently search the optimal mapping $w_k^i, i \in \{1, \ldots, 4\}$ for each subblock $\mu_k^i$. We define an improvement $I(\mu_k)$ in terms of MSE for each block $\mu_k$

$$I(\mu_k) = \mathrm{MSE}(\mu_k) - \sum_{j=1}^{4} \mathrm{MSE}\left(\mu_k^j\right) \tag{9}$$

This procedure is performed for each block $\mu_k$ and the maximum improvement

$$I_{\max} = \max\left(0, I(\mu_k) \forall k \in \{1, \ldots, N\}\right) \tag{10}$$

is determined. If $I_{\max} = 0$ then the segmentation algorithm stops, because no improvement can be achieved by further blocksplitting. Otherwise the block $\mu_k$ for which $I_{\max} = I(\mu_k)$ holds is split into four subblocks and the segmentation is modified so that these subblocks are considered as range blocks in the next iteration step. This procedure continues until the number of range blocks exceeds a former defined threshold or until the reconstruction error remains under a predefined limit. In addition to the parameters of the affine mapping for each range block, the segmentation information has to be coded which is performed by a quadtree-code.

## 4. Simulation Results and Conclusion

We have described the implementation of a block oriented segmentation scheme based on a fractal coding system for digital images. As can be seen in fig. 2 our segmentation algorithm significantly increases the reconstruction quality in terms of *signal-to-noise-ratio* for a given total number of coded blocks by about 1-3 dB. Results have been compared with the same coder using a fixed blocksize for the entire image.

In our experimental studies we used images of size 512x512 pixels and allowed the largest block to be 32x32 pixel and the smallest one 4x4 pixel. In order to avoid distortions due to quantization no one has been applied to the parameters of the mapping. The overhead for the quadtree segmentation has been discarded because its below 0.02 bit/pixel.

Simulations have been carried out employing the above described segmentation scheme with various types of test images. Fig. 3 shows the reconstructed image 'lena' and fig. 4 the appropriate segmentation consisting of 4096 blocks.

Further investigations are directed to a fast library-search increasing mainly the encoding speed, to a suitable quantization of the transform parameters and to the extension of the fractal block-coding technique to video sequences of images.

Figure 1 Original 'lena' 512x512 pixel

Figure 3 Coded 'lena' 4096 blocks

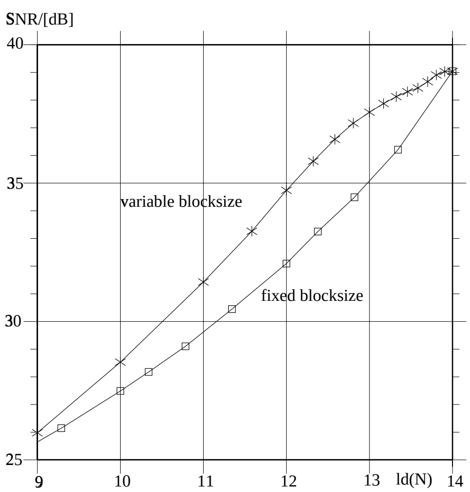


Figure 2 Variable vs. fixed blocksize
(N denotes the total number of blocks)

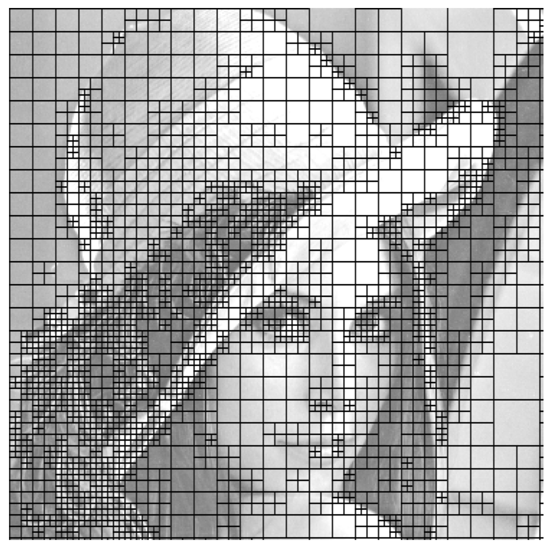

Figure 4 Segmentation result 4096 blocks

## 5. References

[1] R. J. Clarke, *Transform coding of images*. London: Academic Press, 1985.

[2] N. M. Nasrabadi and R. A. King, "Image coding using vector quantization: A review," *IEEE Transactions on Communication*, vol. COM–36, pp. 957–971, Aug. 1988.

[3] M. F. Barnsley, *Fractals Everywhere*. London: Academic Press Inc., 1988.

[4] A. E. Jacquin, "Fractal image coding based on a theory of iterated contractive image transformations," in *Proceedings SPIE Visual Communications and Image Processing '90*, vol. 1360, pp. 227–239, 1990.

[5] M. F. Barnsley, V. Ervin, D. Hardin, and J. Lancaster, "Solution of an inverse problem for fractals and other sets," in *Proc. Natl. Acad. Sci. USA*, vol. 83, pp. 1975–1977, Apr. 1986.

[6] D. M. Monro and F. Dudbridge, "Fractal block coding of images," *Electronics Letters*, vol. 28, no. 11, pp. 1053–1054, 1992.