PERFORMANCE BOUNDS FOR FRACTAL CODING

Bernd Hürtgen

Institut für Elektrische Nachrichtentechnik
Rheinisch-Westfälische Technische Hochschule, RWTH Aachen
52056 Aachen, Germany

ABSTRACT

This paper reports on investigations concerning the performance of *fractal transforms*. Emerging from the structural constraints of fractal coding schemes, lower bounds for the reconstruction error are given without regarding quantization noise. This implies finding an at least locally optimal transformation matrix. A full search approach is by definition optimal but also intractable for practical implementations. In order to simplify the calculation of some appropriate encoding parameter, the *collage theorem* and other fast but also suboptimal approaches are applied. For a memoryless Gaussian source and some real world images the optimal encoding parameters in view of the structural constraints are determined together with the minimal reachable distortion. This allows to quantify the performance of the suboptimal encoding procedures.

1. INTRODUCTION

Recently so called *fractal* schemes gained some degree of interest in the (image) coding community. Basic ideas for encoding and modeling of signals by use of fractal techniques go back to Barnsley et al., e.g. [1] and a first implementation for automatic encoding of images has been proposed by Jacquin, e.g. [2]. In contrast to common linear transformations, e.g. the DCT, whose coding gain mainly emerges from the bindings between neighboring samples, the non-linear fractal coding schemes also exploit some sort of long-range correlations within the signal. In this context those correlations are termed global and/or local self-similarities which arises from the fact that many parts of a natural signal are in some sense similar to the entire signal or at least a part of it.

Coding schemes are termed *fractal* if a given input vector is approximated by an unique *fixed point* of a contractive transformation. Since not the signal itself but the approximating fixed point - sometimes also termed *attractor* - is encoded, the term *attractor coding* would be more appropriate but did not gain acceptance in the recent literature.

Let $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n$ be the entire signal, then most common fractal coding schemes employ a non-linear affine transformation

$$W : \mathbf{x} \to W(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}. \qquad (1)$$

consisting of a linear part $\mathbf{A}$ and a non-linear offset $\mathbf{b}$. For the given input signal $\mathbf{x}$ the goal of the encoding process is determining the matrix $\mathbf{A}$ and the offset vector $\mathbf{b}$ such that

1. the distance $d(\mathbf{x}, \mathbf{x}_f)$ between the original signal $\mathbf{x}$ and the fixed point $\mathbf{x}_f = W(\mathbf{x}_f)$ of the transformation W is minimal,
2. the transformation W obeys a contractivity constraint, and
3. the representation of $\mathbf{A}$ and $\mathbf{b}$ is simple.

In this case the parameters $\mathbf{A}, \mathbf{b}$ of the transformation W serve as fractal code. A (quantized) description of the fractal code is transmitted to the decoder. Since data compression is the aim of fractal coding, some constraints are imposed on the structure of the transformation matrix and the offset vector in order to keep the representation as simple as possible.

Due to the contractivity constraint the fixed point or attractor $\mathbf{x}_f = W(\mathbf{x}_f)$ exists and is uniquely determined by the transformation itself. According to the *contraction mapping theorem* the decoder reconstructs the fixed point from the fractal code by solving the equation $W(\mathbf{x}_f) = \mathbf{x}_f$ in an iterative way. Starting from any arbitrary initial signal $\mathbf{x}_0 \in \mathbb{R}^n$ the contraction mapping theorem states that the sequence of iterates

$$\mathbf{x}_k = W^{\circ k}(\mathbf{x}_0) = \mathbf{A}^k\mathbf{x}_0 + \sum_{i=0}^{k-1} \mathbf{A}^i\mathbf{b} \qquad (2)$$

converges to the unique fixed point

$$\mathbf{x}_f = \lim_{k\to\infty} \mathbf{x}_k = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \qquad (3)$$

which is completely independent from the initial signal $\mathbf{x}_0$.

The aforementioned direct determination of the optimal encoding parameters $\mathbf{A}$ and $\mathbf{b}$, sometimes termed the *inverse problem,* is in general computational infeasible especially if very high dimensional input sources (e.g. images) are concerned. This is due to the non-linear dependency between the fixed point $\mathbf{x}_f$ and the encoding parameters. A way out of this dilemma is the *collage theorem* originally introduced by Barnsley in this context [3, 1]. It greatly simplifies the determination of some suited parameters $\mathbf{A}, \mathbf{b}$ during the encoding process since it demands to minimize the distance $d(\mathbf{x}, \hat{\mathbf{x}})$ between the original signal $\mathbf{x}$ and a so called *collage* $\hat{\mathbf{x}} = W(\mathbf{x})$ rather than between the original signal $\mathbf{x}$ and the fixed point. By this way the time consuming calculation of the fixed point during the encoding process can be avoided. Unfortunately collage coding is suboptimal which means that in general the closest attractor is not found. But as is shown below, the collage theorem at least provides a good initial guess which can be modified in a subsequent optimization process. So minimizing the "collage error" as suggested by the collage theorem still makes sense.

Some improvements can be achieved if the collage theorem is modified. An interesting approach is presented in [4]. Another way is simulating the reconstruction also at the encoder and successively modifying the transformation in a way that a closer attractor is achieved as proposed in [5].

But due to the structural limitations of the fractal transformation even without any quantization of the fractal code no exact reconstruction of the input signal will be possible. Hence the reconstruction error $d(\mathbf{x}, \mathbf{x}_f)$ consists of one part due to the structural limitations of the affine mapping and another part which is due to the quantized description of the fractal code.

Despite the growing interest and available literature concerning fractal coding there are still many open questions demanding for answers, e.g.:

- Given the inherent structural constraints of fractal coding schemes, how good can they perform for a specific source model? How small is the minimal reachable distortion for this model?
- How good performs collage coding compared with an optimal scheme? Which improvement can be achieved by a modified collage theorem or other proposals?
- How can a (nearly) optimal fractal coding scheme be developed with tractable computational effort in the encoding phase?

This paper is concerned with the aforementioned questions and is organized as follows: Section 2 describes a simple fractal coding scheme which serves as basis for the presented investigations. Section 3 then deals with the performance of this scheme and especially considers the effects of the non-optimal encoding process. Some simulation results and a brief discussion in section 4 conclude the paper.

2. A BASIC FRACTAL CODING SCHEME

Our investigations emerge from a simplified version of Jacquin's scheme [6]. Instead of searching for similarities within the entire signal, each block of the signal is treated independently from all others. Let $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$ be a vector of dimension n representing any arbitrary block of the signal to be encoded. This source vector is partitioned into $m = n/n_R$ consecutive non-overlapping parts denoted *range blocks* each consisting of n_R samples. Further let the structure of the transformation matrix be given in a way that the first sample of each range block is formed by the weighted average of the first m samples of the signal, the second sample of each range block by the weighted average of the second m samples and so on. The structure of the transformation matrix can then be outlined as follows:

$$\mathbf{A} = \frac{1}{m}\left(\begin{array}{cccc} \multicolumn{4}{c}{\overbrace{\hspace{8em}}^{n_R \text{block columns}}} \\ \mathbf{a}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{a}_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{a}_1 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}_m & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{a}_m & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{a}_m \end{array}\right); \mathbf{a}_i = \left(\overbrace{a_i, a_i, \ldots, a_i}^{m \text{ times}}\right). \tag{4}$$

The offset vector

$$\mathbf{b} = (\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_m})^T;\ \mathbf{b}_i = \left(\overbrace{b_i, b_i, \ldots, b_i}^{n_R \text{ times}}\right)^T \tag{5}$$

consists of m constant parts $\mathbf{b}_i$ each belonging to one of the non-overlapping range blocks of the original signal.

In order to apply the contraction mapping theorem at the decoder, the weights or scaling parameters a_i are constrained by the contractivity condition to ensure the convergence of the iterative reconstruction. This can be guaranteed if all eigenvalues of the transformation matrix $\mathbf{A}$ lie within the unit circle. In this case the contractivity constrains the scaling coefficients to fulfill $\sum_{i=1}^{m} |a_i| < m$. For n and n_R being integral powers of two the above constraint may be released to $|\sum_{i=1}^{m} a_i| < m$. For a more detailed description of the convergence properties of fractal transforms the reader is referred to [7, 8].

3. OPTIMIZED FRACTAL ENCODING

The aim of optimal encoding is to determine the scaling and offset parameter a_i, b_i such that the distance between the original signal $\mathbf{x}$ and the fixed point $\mathbf{x}_f$ is minimized. This involves in a first step the calculation of an optimal m-dimensional subspace $\mathcal{F}$ defined by $\mathcal{F} = \left\{(\mathbf{I} - \mathbf{A})^{-1}\mathbf{b},\ \mathbf{b} \in \mathrm{I\!R}^m\right\}$ of all possible fixed points $\mathbf{x}_f \in \mathcal{F}$. In a second step the offset vector $\mathbf{b}$ is determined by the orthogonal projection of the original signal $\mathbf{x} \in \mathrm{I\!R}^n$ onto the subspace of fixed points $\mathcal{F}$. Collage coding works suboptimal in both steps. It neither finds the optimal subspace nor the orthogonal projection onto this space.

Therefore an optimization procedure can be carried out in both steps:

1. Given the non-optimal subspace $\mathcal{F}^*$ provided by collage coding, determine a new offset vector $\mathbf{b}$ such that $(\mathbf{x} - \mathbf{x}_f) \perp \mathcal{F}^*$.
2. Modify the subspace $\mathcal{F}^*$ such that it becomes optimal or at least locally optimal and determine new offset vector (described in 1).

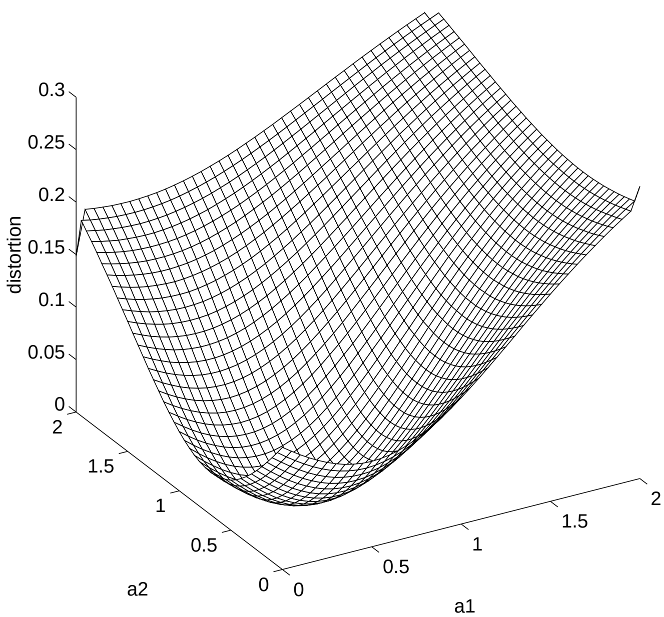

Figure 1. Minimal reconstruction error for given source vector and various choices of two scaling parameters a_1, a_2 and least square optimal offset vector $\mathbf{b}$

Optimizing the offset vector $\mathbf{b}$ for a given subspace $\mathcal{F}^*$ is rather simple. A detailed description of this process may be found in [9]. But determining an optimal subspace with reasonable effort is incomparably more difficult. In order to provide a lower bound for the minimal achievable reconstruction error a "full search" in the parameter space would be necessary. Fortunately the form of the "error surface" is continuous which simplifies the construction of a non-full-search encoding scheme. For a typical source vector the resulting reconstruction error $d(\mathbf{x}, \mathbf{x}_f)$ is displayed in fig. 1 for various choices of two scaling parameters a_1, a_2 and least square optimal offset vector $\mathbf{b}$.

By calculating the minimal reconstruction error for a large number of random input vectors according to a given source model a lower bound for the expected value $E(d(\mathbf{x}, \mathbf{x}_f))$ of the mean squared error for a simple unit variance memoryless Gaussian source can be determined. Tab. 1 shows the results for "optimal" encoding (optimal subspace selection and optimal offset) versus collage coding with and without offset optimization. One can see that

$E(d(\mathbf{x}, \mathbf{x}_f))$ collage coding	$E(d(\mathbf{x}, \mathbf{x}_f))$ opt. collage coding	$E(d(\mathbf{x}, \mathbf{x}_f))$ "optimal" encoding
0.414	0.386	0.295

Table 1. Expected value of reconstruction error for unit variance memoryless Gaussian sources. The source vector of length $n = 6$ is partitioned into $m = 2$ non-overlapping parts.

collage coding results in a significantly larger error compared with optimized collage coding or even with "optimal" encoding. The results obtained by "optimal" encoding constitute a lower bound for the given source model and the presumed structural constraints of fractal coding without quantization and cannot be outperformed for this scheme by any other fractal encoding procedure.

Since real world signals, e.g. images, are in accordance with such a simple source model only up to a certain degree, it is interesting to investigate the performance of collage coding in comparison with optimal encoding also for these sort of signals. For this purpose the former described basic fractal coding scheme has been applied to real world images. In a first step only simple collage coding has been carried out. Since neither the subspace selection nor the offset vector is optimal, in a second step a new optimized offset vector is calculated but the subspace is retained. Finally the third step consists in optimizing the subspace together with the offset. Since a full search in the parameter space of the scaling coefficients is infeasible for these sources, a modified gradient search has been employed. It consists in successively optimizing the single scaling coefficients until a locally optimal set has been found. Our experiments showed that in all cases the local optimum also equaled the global one. Tab. 2 shows the results in terms of signal to noise ratio for various test images. For this experiment the images have been segmented into single square blocks of size $n = 16$. Each of these blocks is independently encoded by use of the former described simple fractal coding scheme with parameter $m = 4$.

	collage coding	opt. collage coding	"optimal" encoding
lena	29.5	30.4	31.0
clown	28.4	29.2	29.9
camera	28.1	28.9	29.5

Table 2. Reconstruction error (signal to noise ratio in dB) for various test images and encoding schemes without quantization. Left/centre: Collage coding without/with offset optimization. Right: "Optimal" encoding scheme.

It can be seen that by employing an optimization step improvements of about 1.5 dB can be achieved. Unfortunately the encoding procedure of fractal schemes is computational very expensive, even if the collage theorem is applied. So one has to weigh up carefully the additional gain in reconstruction quality and the more complex optimization of the encoding parameters.

4. CONCLUSION

Due to the structural limitations in the choice of the transformation matrices and the offset vectors in general no exact reconstruction of the input signal can be obtained with fractal coding schemes even if no quantization is carried out. In this paper a simple fractal coding scheme is presented which serves as basis for calculating lower bounds for the minimal achievable reconstruction error without regarding the additional quantization noise.

Determining the optimal encoding parameters is a non-linear problem for which no simple and exact solution has been found up to now. An easy method known as collage coding greatly simplifies this task but achieves suboptimality only. This paper evaluates the results obtained by collage coding together with those obtained by an optimal encoding scheme employing a full search in the parameter space. The full search is of course optimal but also computational intractable and not suited for any practical implementation but in this case it provides the expected value for the lowest achievable reconstruction error ignoring any quantization.

In order to treat real world images the good-natured behavior of the reconstruction error has been exploited in order to construct a fast gradient search which approximately reaches full-search performance but with a fraction of its computational effort. The algorithm is based upon the collage theorem for which our experiments showed that it provides a good initial guess for the encoding parameters though it is suboptimal in nature. A subsequent modification of the encoding parameters finally leads to the optimal subspace and offset vector selection.

REFERENCES

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

[2] A. E. Jacquin, "Fractal image coding: A review," *Proceedings of the IEEE*, vol. 81, pp. 1451–1465, Oct. 1993.

[3] 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.

[4] G. E. Øien, Z. Baharav, S. Lepsøy, and E. Karnin, "A new improved collage theorem with applications to multiresolution fractal image coding," in *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing ICASSP'94*, vol. 5, pp. 565–568, 1994.

[5] J. Domaszewicz and V. A. Vaishampayan, "Iterative collage coding for fractal compression," in *Proceedings IEEE International Conference on Image Processing ICIP'94*, vol. 3, (Austin, Texas, USA), pp. 127–131, Nov. 1994.

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

[7] B. Hürtgen and T. Hain, "On the convergence of fractal transforms," in *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing ICASSP'94*, vol. 5, pp. 561–564, 1994.

[8] B. Hürtgen and S. F. Simon, "On the problem of convergence in fractal coding schemes," in *Proceedings of the IEEE International Conference on Image Processing ICIP'94*, vol. 3, (Austin, Texas, USA), pp. 103–106, Nov. 1994.

[9] J. Domaszewicz and V. A. Vaishampayan, "Structural limitations of self-affine and partially self-affine fractal compression," in *Proceedings SPIE Visual Communications and Image Processing '93*, vol. 2094, pp. 1498–1504, 1993.