A NEW IMPROVED COLLAGE THEOREM WITH APPLICATIONS TO MULTIRESOLUTION FRACTAL IMAGE CODING

G. E. Øien, Rogaland University Centre * *Z. Baharav, Technion – Israel Institute of Technology* †

S. Lepsøy, Consensus Analysis A/S ‡ *E. Karnin, IBM Israel Science and Technology*

D. Malah, Technion – Israel Institute of Technology

ABSTRACT

The paper presents an improved collage theorem, valid for a class of signal mappings called *Affine Blockwise Averaging (ABA) Mappings.* The ABA structure is exploited to form a bound depending on norms of collage error signals at several resolutions. Compared to previously published collage theorems, the new theorem provides a much tighter bound on the maximum distance between the original signal (real world image) and the decoded attractor, given the distance between the original and the "collage" optimized by the encoder. This is achieved without contractivity constraints on the ABA mapping parameters. Finally, an encoding method in which one attempts to minimize the upper bound directly is suggested.

1. INTRODUCTION

In recent years, the interest in fractal-based signal compression has been steadily growing. The first automatic fractal-based coding algorithm that was practically applicable to real world images was suggested by Jacquin in 1989 [1, 2]. Jacquin's block-oriented method has been the basis for almost all research in the field since it was first published, and has been analyzed and refined by among others Øien, Lepsøy *et al.* [3, 4, 5], Fisher *et al.* [6], and Baharav *et al.* [7]. A forthcoming book [8] will collect many of the most recent contributions to fractal compression research.

2. FRACTAL CODING PRINCIPLES

The encoder in a fractal coder optimizes a nonlinear signal mapping T such that the distance between the original signal x. and the *collage* Tx of x, is minimized. According to the *Fixed Point Theorem*, the decoder may then generate the *attractor* x_T of T (given certain constraints on T), as

$$x_T = \lim_{k\to\infty} T^k x_0 \qquad \text{for an arbitrary } x_0 \tag{1}$$

x_T can also be uniquely defined in terms of its *fixed point property*, i.e. its invariance with respect to T: $x_T = Tx_T$.

The decoder only needs the description of T to synthesize x_T, since x_0 is arbitrary. If x_T is an approximation of the signal x to be coded, the description of T can therefore be used as a *lossy code* for x.

x_T will generally not be exactly equal to Tx, so the code computed by the encoder will not necessarily produce the globally optimal attractor. However, we are able to ensure that x_T will lie *almost* as close to x as Tx does, by suitably constraining the mapping T. This can be seen from the *collage theorem* which traditionally has been the justification behind the fractal coder structure. This theorem provides a bound on the distance between x and x_T which is given by [9]

$$d(x, x_T) \le \frac{1}{1-s_K} \cdot \frac{1-s_1^K}{1-s_1} \cdot d(x, Tx) \tag{2}$$

where s_1 is the Lipschitz factor of T, and K is the smallest integer such that the following holds for all signals x, y:

$$d(T^K x, T^K y) \le s_K \cdot d(x,y) \text{ for some } s_K < 1 \tag{3}$$

Note that the existence of such a K is necessary for the decoding process to converge. Mappings for which such a K exists are called *contractive.*

If $s_K \ll 1$, and s_1 and K are not too large, it is seen from 2 that the original and the attractor are guaranteed to be almost as close to each other as the original and the collage. Hence collage optimization is in that case justified as a means of finding a good signal code. However, for the types of mappings used in practice it seems that even better coding results can be achieved by allowing for larger s_1's and K's. This was observed both by Fisher [6] and by Øien *et al.* [3, 10], but existing collage theorems do not give bounds even remotely close to the actual distances observed. This has made it seem somewhat "dangerous" to perform an unconstrained optimization of the collage in the encoder - one has been afraid that there *might* exist signals for which such an optimization leads to a mapping T having an attractor lying far away from the original. The solution most often used has been to put rather severe constraints on important parameters in the allowed mappings, with a corresponding decline in signal quality as a result. "Strictly" contractive mappings — with $s_1 < 1$ — have been most common.

In this paper we shall take a different approach: We shall develop a collage theorem especially suited for the class of mappings introduced by Jacquin [1, 2], and subsequently modified by Øien *et al.* [8, 10]. This is a class of *affine* mappings, operating on the space of discrete signals of a given size $N = 2^n$ samples. We term them *Affine Blockwise*

*P.B. 2557 Ullandhaug, 4003 Stavanger, NORWAY. E-mail: geir@elektro.tih.no

†Dept. of Electrical Engineering, Haifa 32000, ISRAEL. E-mail: zachi@techunix.technion.ac.il

‡P.B. 1391,1401 Ski, NORWAY

Averaging (ABA) mappings. We shall use Øien's modified (orthogonalized) versions of Jacquin's mappings, but our main result holds also without this modification. The modified mappings are used because they give us several advantages, including fast signal-independent decoder convergence and a simpler proof.

The special structure of the ABA mappings allows us to construct a potentially much stricter collage theorem bound than that given by 2, thus ensuring that no dramatic discrepancy can occur between attractor and original even when an unconstrained collage optimization is used. Hence we are assured that such optimization can be used without fearing the consequences for the decoding. The new theorem also suggests a way of improving the encoder by optimizing the code with respect to the upper bound of the theorem instead of with respect to the collage only.

3. AFFINE BLOCKWISE AVERAGING (ABA) MAPPINGS

By an *Affine Blockwise Averaging (ABA) Mapping* we mean a mapping $T : \mathbf{R}^N \to \mathbf{R}^N$ fulfilling:

1. $Tx = Lx + t$ where $L : \mathbf{R}^N \to \mathbf{R}^N$ is a linear operator and $t \in \mathbf{R}^N$ is a constant vector.
2. t is constant-valued over consecutive blocks of length 2^b samples.
3. L replaces consecutive 2^b-blocks (*range* blocks) in x by inserting in their place 2^d-blocks (*domain* blocks, $d > b$) — taken from another part of x — which
   (a) are *decimated* to size 2^b by averaging over $2^{(d-b)}$ samples at a time, i.e. for a domain block $[l(1) \cdots l(2^d)]$, each subvector $[l(1 + i \cdot 2^{(d-b)}) \cdots l((i+1) \cdot 2^{(d-b)})]$ is replaced by its average value $2^{-(d-b)} \cdot \sum_{j=1}^{2^{(d-b)}} l(j + i \cdot 2^{(d-b)})$, for $i = 0, \ldots, 2^b - 1$.
   (b) have their DC value removed (are orthogonalized with respect to t)
   (c) are scaled by real multipliers α
4. Each domain block is made up of an integer number of range blocks.

For each range block, the output of an encoder optimizing an ABA mapping is a domain block address, the value of the best scalar multiplier, and the value of the constant vector t in the position of the range block. The former can be found by a systematic search while the latter two can be directly optimized in a least squares sense [3, 8, 10]. t will then equal the mean value of the samples within each range block of x.

In [8, 10] an important property of ABA mappings of the form given above is shown: The attractor (at the same resolution as the original signal) can be written as

$$x_T = T^{K-1}t = \sum_{k=0}^{K-1} L^k t \quad (4)$$

where

$$K = \lceil \frac{d}{d-b} \rceil \quad (5)$$

i.e. convergence towards the attractor always occurs in a finite number of steps. (The result can also be generalized to the perhaps more common case of several range and domain block sizes used at once in a quadtree structure [2, 6].) For practical parameter choices, typically $K = 2 - 4$.

Baharav *et al.* have given an interpretation of the fractal decoding process in a *multiresolution* context [7, 8]. From their work it is clear that the expression $T^M t = \sum_{k=0}^{M} L^k t$, $0 \le M \le K-1$, is equivalent to the attractor x_T after decimation by averaging over $2^{b-M(d-b)}$ samples at a time, followed by sample duplication back to original size N. We write this as

$$T^M t = x_T^{(K-1-M)} \text{ for } 0 \le M \le K-1 \quad (6)$$

where $x_T^{(k)}$ means x_T after averaging over $2^{k(d-b)}$ samples followed by sample duplication back to original size. It follows from this that

$$x_T^{(K-1)} = t \quad (7)$$

Note also that

$$x^{(K-1)} = t \quad (8)$$

since t contains the mean value over each range block. Thus the decoding consists of first constructing the attractor at resolution 2^b samples, then, with each new term in the sum, adding correct details at resolution $2^{2b-d}, 2^{3b-2d}$, and so on down to single-sample resolution.

Before we prove our main result, we note another important property of ABA mappings: Due to their decimation-by-averaging structure, the signal to be mapped may be decimated by averaging over 2^{d-b} samples and then brought back to original size by sample duplication without changing the result of the mapping. I.e. the mapping is "blind" to details smaller than 2^{d-b} samples.

4. AN IMPROVED COLLAGE THEOREM FOR ABA MAPPINGS

In this section we state an improved collage theorem, originally found by Baharav [11], which holds for ABA mappings (with *or* without orthogonalization).

Theorem 1 (Collage Theorem for ABA Mappings) *Let $T : \mathbf{R}^N \to \mathbf{R}^N$ be an ABA mapping with range blocks of size 2^b and domain blocks of size 2^d. Assume that T^k has Lipschitz factor s_k for $k = 0, 1, \ldots$[1]. Let $x^{(k)}$ denote an arbitrary vector $x \in \mathbf{R}^N$, after decimation by averaging over $2^{k(d-b)}$ samples at a time followed by sample duplication back to original size N. Let $K = \lceil \frac{d}{d-b} \rceil$. Then the following holds:*

$$d(x, x_T) \le \sum_{k=0}^{K-2} s_k \cdot d(x^{(k)}, Tx^{(k)}) \quad (9)$$

Proof: By the triangle inequality and the fixed point property of x_T,

$$\begin{aligned} d(x, x_T) &= d(x^{(0)}, x_T^{(0)}) \\ &\le d(x^{(0)}, Tx^{(0)}) + d(Tx^{(0)}, Tx_T^{(0)}) \quad (10) \end{aligned}$$

[1] $T^0 = I$, with $s_0 = 1$.

Now, remember that T only "sees" the average value of the samples of any vector over consecutive 2^{d-b} blocks. Thus, for any $y \in \mathbf{R}^N$,

$$Ty^{(k)} = Ty^{(k+1)} \qquad (11)$$

In addition, for the attractor x_T of T, it follows from equation 6 that

$$T^k x_T^{(k)} = T^{k-1} x_T^{(k-1)} \text{ for all } k = 1, 2, \ldots, K-1 \qquad (12)$$

Equations 11 and 12 and repeated use of the triangle inequality implies that

$$\begin{aligned} d(x^{(0)}, x_T^{(0)}) &\leq d(x^{(0)}, Tx^{(1)}) + d(Tx^{(1)}, Tx_T^{(1)}) \\ &\leq d(x^{(0)}, Tx^{(0)}) + d(Tx^{(1)}, T^2 x^{(1)}) \\ &\quad + d(T^2 x^{(2)}, T^2 x_T^{(2)}) \\ &\leq \sum_{k=0}^{K-2} d(T^k x^{(k)}, T^{k+1} x^{(k)}) \\ &\quad + d(T^{K-1} \underbrace{x^{(K-1)}}_{=t (by\ 8)}, T^{K-1} \underbrace{x_T^{(K-1)}}_{=t (by\ 7)}) \quad (13) \end{aligned}$$

The last term in equation 13 is seen to be zero. Thus,

$$\begin{aligned} d(x^{(0)}, x_T^{(0)}) &\leq \sum_{k=0}^{K-2} d(T^k x^{(k)}, T^{k+1} x^{(k)}) \\ &\leq \sum_{k=0}^{K-2} s_k \cdot d(x^{(k)}, Tx^{(k)}) \qquad (14) \end{aligned}$$

□

A comment is in order on how to compute the factors $s_1, s_2, \ldots$. Remember that L^k is an $N \times N$ matrix for $k = 0, 1, \ldots, K-1$. In the case of the l_2 metric, $s_k = \|L^k\|$, the norm of the linear operator L, which is the square root of the largest eigenvalue of $(L^k)^T(L^k)$. For the special case of an ABA mapping and $k = 1$, it has been shown by Lundheim [8] that

$$s_1 = 2^{d-b} \cdot \max_{\text{all } l} \sum_{\text{all } r \text{ using } l} \alpha_r^2 \qquad (15)$$

where r denotes range block, l denotes domain block and α_r are the scalar multipliers.

5. DISCUSSION

The new collage theorem bound, equation 9, suggests several amendments to established fractal coding theory. Most importantly, for the first time the *frequency content* of the signals in question is taken into account. Since the bound now includes norms of collage error signals at successively coarser resolutions, which amounts to low-pass filtered versions of the collage error $x - Tx$, it will work differently for signals with differing frequency content. If $x - Tx$ were a low pass signal, the new bound would therefore not necessarily be much stricter than the classical one. However, an error or noise signal, such as $x - Tx$, is fundamentally a high pass signal. Thus, we expect the low pass filtering implied by the decimation and sample duplication to remove a significant amount of the collage error energy, which in turn improves the bound.

Another thing to note is that for the special case of $d = 2b$, the new bound shows that $d(x, x_T) = d(x, Tx)$. In fact, Lepsøy has shown that in this case we always have $x_T = Tx$ (and the decoding is *noniterative*) [5, 4]. Our new bound reflects this fact.

5.1. Modified encoding procedure

Since the new bound has so few terms (typically 1-3) we have a possibility of *minimizing it directly* instead of concentrating on the collage only. By doing this we expect to obtain an even more "direct" optimization of the attractor. However, since the terms of the bound become successively less important the higher k is (as k increases we observe a significant decrease both in s_k and $d(x^{(k)}, Tx^{(k)})$), we suggest simplifying the problem by keeping maximally the first two terms. Of course, we do not know s_1 before we have found the mapping T, but this problem may be bypassed by initially optimizing T to minimize sums of the form $d(x, Tx) + s \cdot d(x^{(1)}, Tx^{(1)})$, where s is viewed as a free parameter which is to be varied within a suitable interval. By doing this on a training set of representative signals we might find the s giving the minimal sum on the average. This s value, which we expect to be close to the average Lipschitz factor for the resulting mappings, can then be used instead of s_1 in subsequent encodings.

6. EXPERIMENTS

6.1. Example: Modeling of synthetic vector

In this section we will illustrate the efficiency of the new collage theorem bound by means of modeling of a synthetic data vector. We consider the vector

$$x = \begin{bmatrix} 60 & 40 & 24 & 20 & 24 & 22 & 20 & 14 \\ 51 & 49 & 33 & 27 & 14 & 8 & 8 & 2 \end{bmatrix}^T \qquad (16)$$

and divide it into four range blocks of size 4 (no. 1 - 4), and two domain blocks of size 8 (no. 1 and 2). This yields $K = 3$. We performed ABA modeling of this vector with a code summarized in Table 1.

r no.	α_r	l no.	t
1	0.75	2	36
2	0.5	1	20
3	0.75	2	40
4	0.5	1	8

Table 1. ABA parameters

Applying this code, we find the attractor vector

$$x_T = \begin{bmatrix} 57 & 39 & 27 & 21 & 30 & 18 & 18 & 14 \\ 61 & 43 & 31 & 25 & 18 & 6 & 6 & 2 \end{bmatrix}^T \qquad (17)$$

By equation 15, we find $s_1 = 0.75$, i.e. T is "strictly" contractive in this case. We summarize the computed original-attractor l_2-distance and corresponding collage theorem bounds in Table 2. ("Old bound" refers to equation 2 and "new bound" refers to equation 9.)

$d(x, Tx)$	$d(x, x_T)$	Old bound	New bound
3.9627	3.9051	15.8509	5.1486

Table 2. Errors and error bounds for Example 5.1

From the table we see that the bound is drastically improved. Whereas the classical collage theorem predicts an original-attractor rms error that is over four times as high as the original-collage error, the improved theorem predicts an error being only 30 % higher. As seen from the table, the actual $d(x, x_T)$ is almost identical to $d(x, Tx)$ (in fact slightly *lower* in this case), so the bound is still too pessimistic, but the improvement is vast.

6.2. Image coding experiments

In our real world image experiments, the 256×256 8 bpp image "Lena" was used. Two ABA models were derived: One (ABA1) where the mapping T was restricted to be strictly contractive in the l_1 norm (all multipliers of absolute value less than 1) and a maximal vertical/horisontal distance between range and domain block of 96 pixels was allowed, and one (ABA2) where unconstrained optimization of multipliers and exhaustive search for domain blocks was used. We restricted ourselves to 8 × 8 range blocks and nonoverlapping 16×16 domain blocks. The scalar quantizers were kept as real values, so no bit rate is available. The metric used was the rms distance. The main results are summarized in table 3.

Model	$d(x, Tx)$	$d(x, x_T)$	Old bound	New bound
ABA1	12.66	13.11	353.00	42.30
ABA2	12.13	12.60	543.71	38.81

Table 3. Errors and error bounds for "Lena"

This table shows several things: 1) Use of an unconstrained ABA mapping improves the modeling results. The original-to-attractor PSNR increases by about 0.35 dB. 2) Using an unconstrained ABA mapping does not result in a greater original-attractor distance (less than 0.5 dB in both cases). 3) The new collage theorem bound is vastly superior to the "classical" one. The predicted rms error is a decade lower with the new bound. 4) The new bound performs as well, or even slightly better, for the unconstrained mapping as for the constrained one. 5) There is still a gap between the prediction of the bound and the actual coding results.

We have not done extensive experiments with the "weighted" method suggested in Subsection5.1, but we have tried it on the "Lena" image. The results are given in Table 4.

s	0	1.2	2.4	3.6	4.8
PSNR(x, x_T)	26.85	26.91	26.93	26.91	26.88

Table 4. "Lena" encoding results with "weighted" method.

As is seen from the table, there is a (very) slight PSNR improvement. The optimum occurs for $s = 2.4$, while the Lipschitz factor for this mapping was computed as $s_1 = 2.36$. The visual improvement is also small, but is mostly concentrated in the notoriously "difficult" sections of the image, such as the eyes and other highly detailed regions.

7. CONCLUSION

A new collage theorem holding for ABA mappings has been presented. The theorem uses *a priori* knowledge about the structure of the mappings usually employed in fractal coding, which results in a much stricter upper bound on the distance between original and attractor than what has previously been published. Collage error signals over several resolutions are taken into account, and the amount of improvement over the old collage bound is hence dependent on the frequency content of the collage error signal; it is higher the less low frequencies this error contains. For real world image examples, the bound predicts an error which is a decade lower than that given by previous theorems. A modified encoding procedure in which we attempt to minimize the upper bound directly is shown to yield slightly improved attractors. However, more extensive experiments with this method should be performed before its performance can be accurately assessed.

REFERENCES

[1] A. E. Jacquin, *A Fractal Theory of Iterated Markov Operators with Applications to Digital Image Coding.* PhD thesis, Georgia Institute of Technology, 1989.

[2] A. E. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Trans. Image Processing*, vol. 1, pp. 18–30, Jan. 1992.

[3] G. E. Øien, S. Lepsøy, and T. A. Ramstad, "An inner product space approach to image coding by contractive transformations," in *Proc. ICASSP*, pp. 2773–2776, 1991.

[4] S. Lepsøy, G. E. Øien, and T. A. Ramstad, "Attractor image compression with a fast non-iterative decoding algorithm," in *Proc. ICASSP*, pp. V–337 – V–340, 1993.

[5] S. Lepsøy, *Attractor Image Compression: Fast Algorithms and Comparisons to Related Techniques.* PhD thesis, The Norwegian Institute of Technology, June 1993.

[6] Y. Fisher, E. W. Jacobs, and R. D. Boss, "Fractal image compression using iterated transforms," 1992. In J. Storer (ed.): *Image and Text Compression*, Kluwer Academic Publishers, Norwell MA, pp. 35–61.

[7] Z. Baharav, D. Malah, and E. Karnin, "Hierarchical interpretation of fractal image coding and its application to fast decoding," in *Proc. Digital Signal Processing Conference, Cyprus*, July 1993.

[8] Y. Fisher, ed., *Fractal Compression: Theory and Applications to Digital Images.* Springer Verlag, to be published 1994.

[9] L. M. Lundheim, *Fractal Signal Modelling for Source Coding.* PhD thesis, The Norwegian Institute of Technology, November 1992.

[10] G. E. Øien, *L_2-Optimal Attractor Image Coding with Fast Decoder Convergence.* PhD thesis, The Norwegian Institute of Technology, June 1993.

[11] Z. Baharav, "Hierarchical approach for fractal representation of signals," January 1994. M. Sc. Thesis (in Hebrew), Technion – Israel Institute of Technology, Department of Electrical Engineering, Haifa, Israel.