Enhanced Fractal Image Coding by Combining IFS and VQ[1]

Jean-Luc Dugelay
Dept. of Multimedia Communications,
Institut EURECOM,
2229, Route des Crêtes
B.P. 193, 06904 Sophia Antipolis – France
dugelay@eurecom.fr

Allen Gersho
Signal Compression Lab
Dept. of Electrical & Computer Eng.
University of California,
Santa Barbara, CA 93106 U.S.A.
gersho@ece.ucsb.edu

Abstract

A novel paradigm for fractal coding selectively corrects the fractal code for selected domain blocks with an image-adaptive VQ codebook. The codebook is generated from the initial uncorrected fractal code and is therefore available at the decoder. An efficient trade-off results between incremental performance and bit rate.

1. Introduction

Since the landmark paper in 1992 by A. Jacquin [1] on image coding with iterated function systems (IFS), many authors have studied IFS (or "fractal") and proposed many improvements to Jacquin's algorithm [2]. While IFS has been viewed as a promising technique that might overtake more established coding methods, generally it has not so far lived up to these expectations.

One fundamental problem of IFS coding methods is the lack of direct control over the reconstruction error (between original and decoded images) since the encoder attempts only to minimize the collage error (between the original and the self-similar transformation of the original generated in the encoder). Although the well-known Collage Theorem gives an upper bound on the reconstruction error as a function of the collage error, minimizing collage error does not minimize the reconstruction error. In practice, the reconstruction error is larger and often much larger than the collage error.

This paper introduces a technique based on vector quantization (VQ) to reduce the collage error in such a way that the difference between the collage and reconstruction errors will also be substantially reduced. This is done in a selective manner without significantly increasing the bit rate.

1.1 Review of fractal coding

Basic concepts. Let, μ_{orig} be the image to be compressed, μ_{init} an arbitrary initial image, μ and ν two generic images, and $d(\mu,\nu)$, a distortion measure which measures the dissimilarity between images. A transformation τ which maps an image into another image is said to be *contractive* if: $d(\tau(\mu),\tau(\nu)) < \sigma . d(\mu,\nu)$ with $0 < \sigma < 1$, where σ is the *contractivity* of τ. Then $\tau^n(\mu_{init})$ converges to an *attractor* μ_a as n approaches infinity, where μa is independent of μ_{init}. The Collage Theorem states: that if there exists a transformation τ such that, $d(\mu_{orig}, \tau(\mu_{orig})) < \varepsilon$ and τ is contractive with contractivity σ, then $d(\mu_{orig}, \mu_a) < \varepsilon / (1-\sigma)$.

The task of the encoder is to determine a transformation τ (an "IFS code") for which $\tau(\mu_{orig})$ is as similar as possible to μ_{orig} subject to a limitation on the number of bits needed to specify τ. The IFS code τ is transmitted to the decoder which then computes the attractor μ_a as the reconstructed image. The reconstruction error is upper-bounded by the Collage Theorem.

Encoding stage. In Jacquin's approach, the encoder finds the transformation τ from the original image μ_{orig} as a sum of affine transformations (τ_i), one for each range block, R_i, each of which maps a particular domain block into the corresponding range block R_i where the domain and range are partitioned at different resolutions; typically with square range blocks of size *BxB* and domain blocks of size *2Bx2B*. (Generally *B* is 8 pixels.) For each range block, the encoder searches for the best collage match from a suitably transformed and selected domain block. For this search, candidate domain blocks are transformed in three steps, by performing sub-sampling, isometry, and scale and shift operations on the block luminance values.

Decoding Stage. To decode an image, the received IFS code, τ, is applied to an arbitrary initial image μ_{init} to form an image μ_1. The process is repeated to obtain μ_2 from μ_1, and so on, until it reaches μ_a. Typically, less than ten iterations are needed for convergence.

Weakness of Fractal Coding. The collage theorem gives only an upper bound on the reconstruction error as a function of the collage error. There is no encoding method

[1] This work was supported in part by the National Science Foundation under grant no. NCR-9314335, the UC MICRO program, ACT Networks, Advanced Computer Communications, Cisco Systems, DSP Group, DSP Software Engineering, Fujitsu Labs, General Electric, Hughes Electronics, Intel, Nokia Mobile Phones, Qualcomm, Rockwell International, and Texas Instruments.

known that directly minimizes reconstruction error. Here, we take a first step in this direction by modifying the IFS code τ in order to simultaneously make both the reconstruction error and collage error as low as possible.

1.2 IFS versus VQ

In fractal encoding, the set of domain blocks of μ_{orig} may be viewed as a basic VQ codebook. The affine transformations of the sub-sampled code vectors in this basic codebook form an extended codebook, corresponding to mean-shape-gain VQ [3], providing a large pool of code vectors for obtaining a collage approximation to μ_{orig}. With conventional VQ, either a fixed or image-adaptive codebook may be used, but in the latter case, the code vectors must be transmitted. The main advantage of the IFS approach is that an image-adaptive codebook is obtained without the need to explicitly transmit it to the decoder. The main drawback is the lack of control over the reconstruction error since IFS coding is lossy and the "virtual" codebook (set of domain blocks of μ_a) obtained at the decoder is not exactly the same as the codebook which had been used for computing the IFS code during the coding stage.

2. New fractal codec

2.1 Introduction

The proposed scheme can be divided into two stages. The first is the conventional IFS code estimation following Jacquin. The second stage directly reduces the collage error by adding a quantized vectorial shift to selected local block IFS codes. In this step, the local blocks are chosen according to their contribution to the gap between the collage and reconstruction errors.
In this paper, the dissimilarity measure between images is the Euclidean norm of the difference image. We define the collage error: $E_c = \left\| \mu_{orig} - \tau(\mu_{orig}) \right\|$, the reconstruction error: $E_r = \left\| \mu_{orig} - \mu_a \right\|$ and the gap between the collage and reconstruction errors, later referred to simply as the «gap»: $E_f = \left\| \mu_a - \tau(\mu_{orig}) \right\|$.

2.2 Reducing the collage error

Considering that the image μ_{orig} has been partitioned into a set of N range blocks, R_i, E_c can be expressed as:

$$E_c = \left\| \sum_{0 \le i \le N} \left(\left(\mu_{orig}\right)_{|R_i} - \left(\tau_i\left(\mu_{orig}\right)\right)_{|R_i} \right) \right\|$$

where the notation of Jacquin is followed. The collage error can be reduced by correcting one range block, as follows:

(a) Select the range block R_k to maximize the error component $\left\| \left(\mu_{orig}\right)_{|R_k} - \left(\tau_k\left(\mu_{orig}\right)\right)_{|R_k} \right\|$.

(b) Consider the domain block $D_{h(k)}$ where h(k) is the index of the domain block that is mapped to R_k. The local IFS map τ_k can be expressed as a sequence of four operations: a sub-sampling, r, an isometry, i_k, a scalar multiplication, s_k, and a shift o_k to each component. Thus:

$$\left(\tau_k\left(\mu_{orig}\right)\right)_{|R_k} = s_k\left(i_k \circ \left(r\left(\mu_{orig}\right)_{|D_{h(k)}} \right)\right) + o_k.1$$

Then, by modifying τ_k by τ'_k as follows, the error of collage for this range block will be reduced:

$$\left(\tau'_k\left(\mu_{orig}\right)\right)_{|R_k} = \left(\tau_k\left(\mu_{orig}\right)\right)_{|R_k} + O'_k$$

where O'_k is a quantized version of the vector $O_k = \left(\mu_{orig}\right)_{|R_k} - \left(\tau_k\left(\mu_{orig}\right)\right)_{|R_k}$ using a VQ correction codebook, described later.

2.3 Selection of Blocks for Correction

The difference between errors of collage and reconstruction E_f can be expressed as,

$$E_f = \left\| \sum_{0 \le i \le N} \left(\left(\tau_i\left(\mu_a\right)\right)_{|R_i} - \left(\tau_i\left(\mu_{orig}\right)\right)_{|R_i} \right) \right\|$$

where we have used the attractor property: $\mu_a = \tau(\mu_a)$.
To decrease the error between collage and reconstruction, we then select the domain D_l for correction which provides the maximum value according to the following criterion:

$$\left\| \sum_{0 \le i \le N} \left(s_i \left(r\left(\mu_{orig}\right)_{|D_l} - r\left(\mu_a\right)_{|D_l} \right) \right) \right\|$$

Thus, we consider the composite effect that this domain block has on all range blocks. In this way, we can determine the global impact that the error between the original and virtual values of a domain block will have on the entire image. Note that the shift and isometry operations in τ_i do not affect the gap, however, an element of the virtual domain pool is responsible for part of the gap if the difference between it and its corresponding element in μ_{orig} is large after (only) sub-sampling, and if the IFS code associates it with many range blocks R_i, which are themselves associated with high scale factors (s_i).

For a given domain location, the reconstructed error is generally greater than but of the same order of magnitude as the collage error, i.e., $\left(\mu_a\right)_{|D_l} \approx \left(\tau\left(\mu_{orig}\right)\right)_{|D_l}$.
Furthermore, if several corrections are made and the IFS iterative process is re-used, only the difference between $\left(\mu_{orig}\right)_{|D_l}$ and $\left(\tau\left(\mu_{orig}\right)\right)_{|D_l}$ need be considered. The difference between $\left(\mu_a\right)_{|D_l}$ and $\left(\tau\left(\mu_{orig}\right)\right)_{|D_l}$ will continue to be reduced, approaching zero, by correcting other parts of the picture.

Hence, the "optimal" vectorial shift to add to $r(\mu_a)|_{D_l}$ in order to reduce the gap will be: $\left(r(\mu_{orig})|_{D_l} - r(\tau(\mu_{orig}))|_{D_l} \right)$.

For range blocks of size BxB and domain blocks of size 2Bx2B (with a step of 2B between 2 consecutive domain blocks), the location of a domain block also corresponds to a set of 4 range blocks. Hence, the above expression can also be viewed as a simultaneous correction, according to the collage error, of a set of 4 range blocks, up to a scale factor of 2, thus motivating the method of Section 2.2).

In other words, in order to ensure that the resulting attractor image will be nearly equal to the collage image, the IFS code τ is modified to obtain τ', by adding only vectorial shifts (without altering the correspondence between range and domain blocks or the contractivity of the code). This is done in such a way that: $\tau'(\tau'(\mu_{orig})) = \tau'(\mu_{orig})$. To achieve this, we must partially reduce the collage error, i.e., $\left\| \tau'(\mu_{orig}) - \mu_{orig} \right\|$, to make τ' invariant to the collage image.

2.4 Algorithm

Initialization.

a) Compute τ_0 and the associated attractor according to Jacquin's method.
b) Define the number of correction n.
c) Set the counter *loop* equal to 0

Correction stage.

a) Stop if *loop* count is equal to n;
b) Select a domain block (which has not yet been corrected) according to the criterion defined in 2.3.
c) Select the pixel region of size 2Bx2B which corresponds to the selected domain block location
d) Sub-sample the selected area using the operator r to compute the correction vector O of size BxB.
e) Quantize O by selecting the closest error vector O' in the correction codebook (see next section for the design of this codebook);
f) Update the IFS code to obtaining τ_{loop+1}, the next improvement to the IFS code by adding the VQ shift code vector O', after having over-sampled it, by pixel duplication from size BxB to 2Bx2B.
g) Iterate the upgraded IFS code to spread the correction and update the domain pool.
h) Increment *loop* and go to (a).

Decoding. With the new IFS code, the associated attractor can be found in the same manner as in classical IFS decoding, except there is an additional step of vectorial shifting for some local transformations.

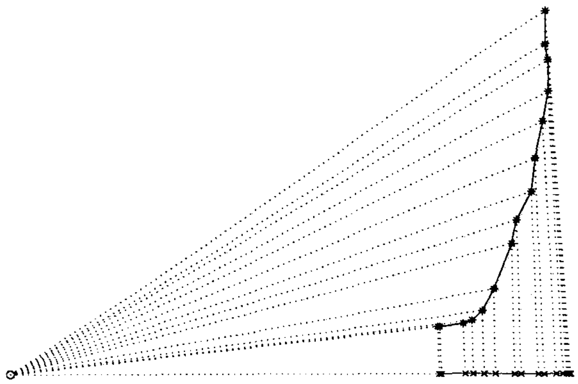

Figure 1. Visual representation of distances
This figure shows how the three distances E_c, E_r and E_f decrease as a function of the number of corrected domain blocks according to the algorithm described in this section without VQ, i.e., step (e) of stage 2) is omitted. The quantitites μ_{orig}, $\tau_{loop}(\mu_{orig})$, and $\tau^n_{loop}(\mu_{init})$ are respectively associated with symbols 'o', 'x' and '*'. Results are for the 256x256 Lena image. The number '*loop*' of corrected domain blocks vary from 0 (at right) to 256 (at left), with intermediate steps of 1, 2, 5, 10, 20, 30, 40, 50, 75, 100, 128, 150, 200 and 208. The initial values of E_c and E_r (i.e. with *loop* = 0), in terms of RMSE, are respectively 10.92 and 11.21; the final ones are both 8.39.

2.5 Correction codebook design

In order to preserve an efficient compression rate, the vectorial corrections have to quantized. Several design methods can lead to an efficient codebook [3]. Here, we take the following approach:

1. The codebook is initialized to the range pool (blocks of size BxB) extracted from the attractor associated with τ_0.
2. To each sub-block of size B/2xB/2 (i.e. corresponding to the size of a subsampled range block), its average value is subtracted (based on the remark that the matching between blocks for IFS code estimation is realized in such a way the resulting average value of collage errors is zero for all range blocks),
3. The number of vectors is reduced from M to M' (with M' << M) using a modified version of the Lloyd algorithm: that is to say, final vectors of the codebook must be elements of the training set (the initial codebook.).
4. The size of the codebook is enlarged firstly from M' to 8.M' by using isometries (isometry-VQ), and secondly to 8.g.M' by using a scale factor (gain-VQ). Moreover, if needed, the codebook can be further enlarged by constraining one quarter of the correction block's pixels to be zero, so that only ¾ of the block is corrected.

The main advantage of this approach is to preserve the features of fractal coding: the codebook of errors is included in the IFS code (i.e., the attractor) itself, and in this way:

(a) The codebook of corrections is adaptive and it is not necessary to transmit it to the decoder (as in the basic IFS coding approach where it is not necessary to transmit the domain pool) since the decoder can directly compute the codebook from the IFS code by omitting the corrections.

(b) The compressed data obtained by this scheme is still theoretically independent of the size of the image: if the attractor is decoded using a scale factor [4], the correction codebook will also be automatically created at the appropriate resolution.

Bit rate. The increase in the bit rate due to the correction stage will be given by, $s_{dom} + n.s_{cor}$ where s_{dom} is the size of the domain pool (1 bit per block indicates if there is a correction or not), n is the number of corrected domain blocks, and $2^{s_{cor}}$ is the size of the correction codebook.

Preliminary results. In spite of the performance penalty resulting from quantization of the corrections, the improvement offered by the new algorithm in terms of quality and control over the image reconstruction is very promising. Typically only 10 corrections are enough to reduce the reconstruction error to the point where it is equal to the original collage error.

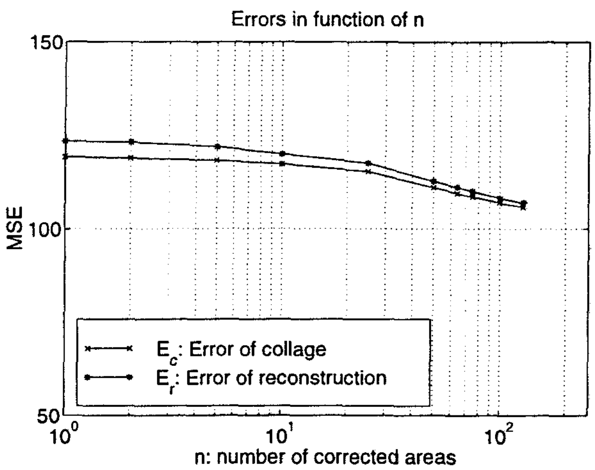

Illustration 2: Preliminary results
This illustration shows how the collage and reconstruction errors vary according to the number of corrected domain blocks using the entire algorithm (including VQ of corrections) for the 256x256 Lena image.

3. Concluding remarks

The proposed scheme retains the advantages of the classical IFS codec, i.e., an adaptive codebook is automatically available at the receiver, while mitigating the key weakness of the IFS codec, namely, a reconstruction error that can be substantially in excess of the collage error.

Each correction to the IFS code is based on adding a quantized vectorial shift to some local τ_i. This modification of the IFS code does not alter the contractivity and not only is the collage error reduced, but also the error of reconstruction is also controlled by the corrections.

We have shown that, in order to decrease the gap between the errors of collage and reconstruction, it suffices to correct, according to the collage errors, only some parts of the picture up to a scale factor of 2 (In our example, only one index is necessary to correct a region equivalent to 4 range blocks). Moreover, it is necessary to take into account only regions that include part of a domain block which itself creates a non-zero additional error, after having considered scale factors of its associated range blocks.

Initial simulations of the new algorithm show the correction of a single domain block propagates to improvements in other regions of the image after re-applying the iterative process. Thus, because the algorithm ties the collage error to the reconstruction error, each additional correction step further decreases both the collage error and, concurrently, the reconstruction error.

We are currently investigating several improvements to the proposed algorithm. Specifically, these include ways to increase the effective size of the correction codebook and the use of the scale factors to control the precision of correction (and hence, the bit rate).

References.

1. A.E. Jacquin, Image coding based on a fractal theory of iterated contractive image transformations, *IEEE Trans. on Image Proc.*, 2,1 (1992), pp. 18-30.
2. Y. Fisher (Ed.), Fractal Image Compression: Theory and Application to Digital Images, Springer Verlag, NY, 1994.
3. A. Gersho and R.M. Gray, Vector Quantization and Signal Compression, 1992 by Kluwer Academic Publishers.
4. E. Polidori and J.-L. Dugelay, Zooming using I.F.S., Fractals, Vol. 5, Supplementary Issue, (April 1997), pp. 111-123.