*Journal of University of Science and Technology Beijing*
*Volume 10, Number 6, December 2003, Page 67*



# Fractal image encoding based on adaptive search

*Kya Berthe, Yang Yang, and Huifang Bi*

Information Engineering School, University of Science and Technology Beijing, Beijing 100083, China
(Received 2002-08-20)

**Abstract:** Finding the optimal algorithm between an efficient encoding process and the rate distortion is the main research in fractal image compression theory. A new method has been proposed based on the optimization of the Least-Square Error and the orthogonal projection. A large number of domain blocks can be eliminated in order to speed-up fractal image compression. Moreover, since the rate-distortion performance of most fractal image coders is not satisfactory, an efficient bit allocation algorithm to improve the rate distortion is also proposed. The implementation and comparison have been done with the feature extraction method to prove the efficiency of the proposed method.

**Key words:** fractal optimization; bit allocation; adaptive search; image compression

## 1 Introduction

Since the introduction of fractal theory in image compression [1-4], several research works have been made to improve the efficiency of this method [5-9]. The motivation for fractal image compression is that images contain not only the spatial redundancy incorporated into the transformation coder image model, but also redundancy in scale [4, 10]. Fractal compression takes advantage of this redundancy in scale by using coarse-scale image features to quantize fine-scale features [9, 11]. While the transformation coder takes advantage of every simple structure in images, the fractal compression on the contrary considers all the image features, including straight edges and constant regions. When compared with the Vector Quantization (VQ) [5, 12, 13], the vector codebook in fractal compression is constructed from locally averaged and sub-sampled isometries of larger blocks of the image; this codebook is effective for coding constant regions and straight edges due to the scale invariance of these features.

The recent research works on fractal compression theory [5-7, 14, 15] allow us to conclude that (1) the potential of transformations in the geometric domain is huge and probably poorly exploited by a linear model of only 8 isometries; (2) the extension of the number of transformations does not necessarily result in an increased coding complexity. If a good prediction mechanism is in place, the best matching transformation can be found without involving a complicated search. This approach has been focused on in this paper using an adaptive search based on the orthogonal projection and the least-square error optimization.

One key problem with any extended model would be to provide a good control mechanism for solving the local inverse problem, *i.e.* selecting the best transformation between two blocks without performing a complete search through a large pool of transformations [16].

The conditions imposed by the optimization and the orthogonal projection can disqualify many domain block candidates during the encoding process [6, 17, 18,]. The well-known classification methods also allow significant reduction of the encoding complexity. The problem with a classification method is that it reduces only the search of the domain pool by a factor at most equal to the number of classes [19, 5]. Similarly research which have been done in this field, such as the nearest neighbor method [20, 13] and the fast search approach, has failed in the quality of compression [7, 9].

In this paper, we use the feature extraction method to extract the feature of each range and domain block at first, then we make the pixel-by-pixel comparison to match the best domain block, a method for bit allocation is also introduced to improve the rate distortion and the conclusion is given finally.

## 2 Review of fractal image compression

The mathematical principle of fractal image com-

**Corresponding author:** Yang Yang, **E-mail:** yyang@ustb.edu.cn

pression consists of approximating an image $f$ seen as an element of a complete metric space $(X,\delta)$, by the unique fixed point $f_w = W(f_w)$ of a contractive mapping $W: X \to X$ [1]. The encoding process is as following, the sampled square grey-scale digital image is partitioned into disjoint $2^n \times 2^n$ blocks $r = \{r^i, i = 1,...,n_R\}$ called range blocks and $d = \{d^j, j = 1,...,n_D\}$ of square domain blocks called domain blocks, which are taken from the same image as the range blocks. The domain block consists of all $2^{n+1} \times 2^{n+1}$ square blocks whose top-left pixels are situated on a lattice with a fixed vertical and horizontal spacing. For each range block $r^i$, a transformation $W_i$ is found, which approximates the range block $r^i$ by the block

$$W_i(r^i) = s_i \tau_p d^j + o_i \tag{1}$$

where $\tau_p, p = 1,\cdots,8$ is an isometry which transforms the domain block. $s = \{s_1,\cdots,s_{n_s}\} \subset \Re$ is a scaling factor (contract), and $o = \{o_1,\cdots,o_{n_o}\}$ is an offset (brightness), calculated by the below formulae:

$$s_i = \frac{\left\langle r^i - \bar{r}^i I, d^j - \bar{d}^j I \right\rangle^2}{\left\| d^j - \bar{d}^j I \right\|^2}, \qquad o_i = \bar{r}_i - s_i \bar{d}_j \tag{2}$$

where $\langle \cdot,\cdot \rangle$ is inner product and $\bar{r}^i = \sum_{\substack{l \le i \le n \\ k \le j \le n}} r^i_{l,k}/n^2$, $\bar{d}^j = \sum_{\substack{l \le i \le n \\ k \le j \le n}} d^j_{l,k}/n^2$ and $I$ are the average intensities of $r^i$ and $d^j$, and the identity matrix respectively.

The mapping $W$ is found by minimizing the distance $\Delta(\cdot,\cdot)$ (local collage error) between $f$ and $W(f)$ in a set of many candidate contractive mappings. In practice the encoding process is as following, for a given range block $r^i$ each approximation of the local collage error is computed using the formula

$$\Delta(d^j, r^i) = \sum_l \sum_n \left( s\tau_p d^j_{l,n} + o - r^i_{l,n} \right)^2 \tag{3}$$

And the rate distortion is determinate by

$$\begin{aligned} E(r^i, d^j) &= \left\| r^i - \bar{r}^i I \right\|^2 - \frac{\left\langle r^i - \bar{r}^i I, d^j - \bar{d}^j \right\rangle^2}{\left\| d^j - \bar{d}^j I \right\|^2} \\ &= \left\| r^i - \bar{r}^i I \right\|^2 - s^2 \left\| d^j - \bar{d}^j I \right\|^2 \end{aligned} \tag{4}$$

where

$$s = \max_k s_k, \quad o = \max_k o_k \tag{5}$$

and $\| \ \|$ is the usual two-norm.

The compressed image is obtained by storing the coefficients $(r^i, d^j, s_i, o_i, \tau_p)$. The decoding process starts with an arbitrary image, then after a certain number (depending on the encoding process) of iterations we obtain the original is obtained [4, 5, 14].

The main drawback of this method is the long encoding time [4, 16, 17]. Therefore, this broad principle encompasses a very wide variety of coding schemes which have been introduced to minimize this encoding time and improve the quality of compression [5, 6, 10].

## 3 Optimization of the collage error

### 3.1 Static approach

The first decision to be made when we use fractal theory to compress an image is the choice of the type of a partition of range block [3-5, 21]. This is because domain blocks must be transformed to cover range blocks, and restrict the possible sizes and shapes of the domain blocks.

Let $\Delta(d^j, r^i)$ be the collage distance error between range block $r^j$ and domain block $d^j$. By combining formula (3) and formula (1), we have:

$$\Delta(d^j, r^i) = \sum_l \sum_n \left[ s\tau_p (d^j_{l,n} - \bar{d}^j_{l,n}) - (r^i_{l,n} - \bar{r}^i_{l,n}) \right]^2 \tag{6}$$

Formula (4) can be written as

$$\Delta(d^j, r^i) = \sum_l \sum_n (s\tau_p \tilde{d}^j_{l,n} - \tilde{r}^i_{l,n})^2 \tag{7}$$

where

$$\tilde{r}^i = r^i_{l,n} - \bar{r}^i \text{ and } \tilde{d}^j = d^j_{l,n} - \bar{d}^j \tag{8}$$

The optimization of the collage distance error can be formulated as follows, for each isometric $\tau_p$ find the most similar domain block, which can be written as the follow:

$$W_i(\tilde{r}^i) \approx s\,\tau_p \tilde{d}^j \tag{9}$$

### 3.2 Orthogonal projection

According to the orthogonal projection described by Dietmar Saupe and the theorem 6.1 (see reference [4], [5]) the minimization of the square error can be written as follows:

$$\Delta(d^j, r^i) = \langle r^i, \phi(r^i) \rangle \sqrt{1 - \langle \phi(d^j), \phi(r^i) \rangle^2} \tag{10}$$

where $\phi(\cdot)$ is the normalized operator.

Thus, the minimization of the collage error $\Delta(d^j, r^i)$ among domain codebook blocks $d^j$ can be achieved by an angle criterion: the minimum of $\Delta(d^j, r^i)$ occurs when the squared inner product

$\langle\phi(d^j),\phi(r^i)\rangle^2$ is maximal. Since $\langle\phi(d^j),\phi(r^i)\rangle^2 = \cos^2\angle(\phi(d^j),\phi(r^i))$, the maximization of $\Delta(d^j,r^i)$ is made at the minimal angle.

In practice we compute the normal vector of the range and the domain block, then we determine the angle between two vectors.

### 3.3 Adaptive search

Formulae (7) and (9) do not guarantee the best estimation of the collage error, because only part of the upper bound for $\Delta(d^j,r^i)$ is minimized, and it is assumed that the contracting factor for $s$ would not significantly affect the bound. Furthermore, if $s$ is only eventually contractive, it is also assumed that $s$ would minimize $\Delta(d^j,r^i)$ for suitably large number of domain blocks, which is not necessarily true.

So in order to make a best optimization of the collage error, we need to constrain the range of the scaling parameter (to ensure a contractivity) as well as quantizing the parameters for compression. So, it is better to work with a set of quantized fractal parameters $\{s_{i,i}\}$. And formulas (7) and (9) can be written as

$$W_i(\tilde{r}^i) \approx (s - s_i)\tau_p \tilde{d}^j \tag{11}$$

$$\Delta(d^j,r^i) = \sum_l \sum_n \left((s - s_{i,j})\tau_p \tilde{d}^j_{l,n} - \tilde{r}^i_{l,n}\right)^2 \tag{12}$$

**Figure 1** shows the comparison of the number of range and domain blocks between the proposed method and the feature extraction method.

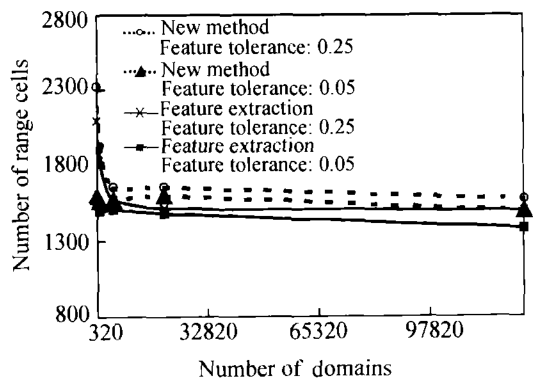


**Figure 1 Range-domain block simulation result.**

Figure 1 is obtained using the Lena image with 256×256 size. We can see that, for the same number of domain blocks and feature tolerance, the number of the range blocks using the new method is greater than the number of range blocks of the feature extraction method (at the same feature value). This means that the new method increases the number of domain blocks (because the number of range blocks is fixed). In addition the new method is more flexible than the feature extraction method.

**Figure 2** shows the compression ratio. Peak Signal Noise Ratio (PSNR) curves for the proposed method and the feature extraction method applying the same image Lena with size 256×256. The new method improve significantly the compression ratio.

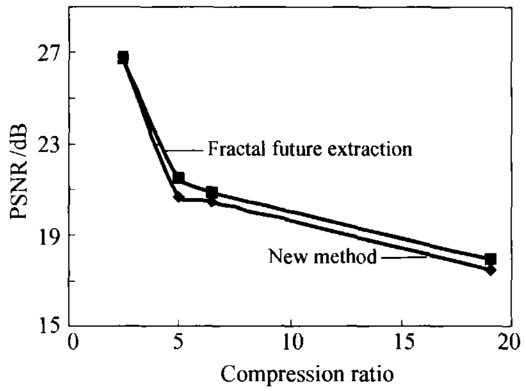


**Figure 2 Implementation of the proposed method.**

## 4 Bit allocation optimization

The Rate-Distortion theory is often used for solving the problem of allocating bits to a set of classes, or for bit rate control in general. The theory aims at reducing the distortion for a given target bit rate by optimally allocating bits to the various classes of data. In this section we tackle the problem of rate-distortion by introducing an algorithm to accelerate the domain block search and improve both the rate-distortion performance and the decoding speed.

Let us denote by $b_{\text{rate}}$ the total number of bits for a quadtree encoding in which the minimum range size is $2^2 \times 2^2$ and the maximum range size is $2^{n_{\max}} \times 2^{n_{\max}}$. Our objective is to find a quadtree encoding that minimizes $\Delta(d^j,r^i)$ subject to a constraint on $b_{\text{rate}}$. The optimal solution for the $b_{\text{total}}$ can be found by solving independently for each $i \in \{i,\cdots,m_R\}$ of the unconstrained problems

$$\min \Delta(d^j,r^i)d_i + \lambda b^i_{\text{rate}} \tag{13}$$

where

$$\Delta(d^j,r^i) = \begin{cases} \sum_l \sum_n \left((s - s_{i,j})\tau_p \tilde{d}^j_{l,n} - \tilde{r}^i_{l,n}\right)^2, & \text{if } s \neq s_{i,j} \\ \sum_l \sum_n \left(s\tau_p \tilde{d}^j_{l,n} - \tilde{r}^i_{l,n}\right)^2, & \text{otherwise} \end{cases} \tag{14}$$

and

$$b_{\text{rate}}(d^j,r^i) = \begin{cases} \lceil \log_2 n_{s_{i,j}} \rceil + \lceil \log_2 n_D \rceil + 5, & \text{if } s \neq s_{i,j} \\ \lceil \log_2 n_s \rceil & \text{otherwise} \end{cases} \tag{15}$$

$\lambda \geq 0$ is a Lagrange multiplier. The optimal value of $\lambda$ is not a known priori and depends on the particular bit-budget. The optimal $\lambda$ can be found by using the bisection search [11] (where recursions with different $\lambda$ values are used to approach the bit budget) which is computed intensively. Using the approximation described in [22], the relationship between $\lambda$ and $s$ could be reasonably approximated as

$$\lambda = 0.85s^2 \tag{16}$$

The uniform quantization and fixed-length codes for the scaling factors and the offsets, where $\Delta(d^j, r^i)$ and $b^i_{\text{rate}}$ denote the total collage error and the total number of bits corresponding to a quadtree structure in a region of maximum size.

### 4.2 Extension to quadtree partition

In this section, we will generalize formula (16) to the range $r^i$ where the least square error is greater than the estimated tolerance.

Let $n \in \{2, \cdots, n_{\max}\}$ and let $\Re^n$ be a $2^n \times 2^n$ block in a region of maximum size. We denote the number of bits rate and the collage error by $b^0_{\text{rate}}(r^i, d^j)$ and $\Delta(d^j, r^n)$, respectively corresponding to the bit ratio and the collage distance error. The optimal encoding as a leaf node is found using formula (16).

For $n > 2$ let the four subblocks of $r^i$ be denoted by $r^{i,n}, n = 1, \cdots, 4$. Suppose that the optimal quadtree structures for these four subblocks are known. Let the total number of bits and the total error of the subtrees corresponding to these optimal quadtree structures be denoted by $b^{\bullet}_{\text{rate}}(d^j, r^{n,i})$ and $\Delta^{\bullet}(d^j, r^{n,i})$, then the four subtrees are combined into a leaf [5]. Let $\Delta_{\text{dif}}(d^j, r^n)$ and $b_{\text{dif}}(d^j, r^i)$ be respective the error of the collage distance and the bits rate, which can be used for the range block matching. The relationship between them can be written as:

$$\Delta_{\text{dif}}(d^j, r^n) \le b_{\text{dif}}(d^j, r^i) \tag{17}$$

where

$$\Delta_{\text{dif}}(d^j, r^i) = \Delta^0(d^j, r^i) - \sum_{n=1}^{4} \Delta^{\bullet}(d^j, r^{i,n}) \tag{18}$$

and

$$b_{\text{dif}}(d^j, r^i) = \sum_{n=1}^{4} b^{\bullet}_{\text{rate}}(d^j, r^{i,n}) - b^0_{\text{rate}}(d^j, r^i) \tag{19}$$

The number of bits needed to represent the subtree for block $r^n$ is

$$b^{\bullet}_{\text{rate}}(d^j, r^i) = \begin{cases} 1 + b^0_{\text{rate}}(d^j, r^i), & \text{if } \Delta_{\text{dif}} \le \lambda b_{\text{dif}} \\ 1 + \sum_{n=1}^{4} b^{\bullet}_{\text{rate}}(d^j, r^{i,n}), & \text{otherwise} \end{cases} \tag{20}$$

and the resulting error is

$$\Delta^{\bullet}(d^j, r^i) = \begin{cases} \Delta^0(d^j, r^i), & \text{if } d_{\text{dif}} \le \lambda b_{\text{dif}} r \\ \sum_{n=1}^{4} \Delta^{\bullet}(d^j, r^{i,n,i}), & \text{otherwise} \end{cases} \tag{21}$$

### 4.3 Proposed algorithm

There is no optimal method for bit allocation; for example: the method mentioned in [1, 2, 5] starts with zero bit allocated for all classes, and to find the class which is most 'benefited' by getting an additional bit. The 'benefit' of a class is defined as the decrease in distortion for that class. In our approach, we keep on reducing one bit at a time until we achieve optimality, either in distortion or target rate or both. The following steps describe our approach:

1) Initially, all range block are allocated a predefined maximum number of bits.

2) For each range and domain block, one bit is reduced from its quota of allocated bits, and the distortion due to the reduction of that 1 bit is calculated.

3) Of all the selected domain blocks, the domain block with minimum distortion for a reduction of 1 bit is noted, and 1 bit is reduced from its quota of bits.

4) The total distortion for the domain candidate blocks is calculated using formula (16).

5) The total rate for all range blocks is calculated as in (16).

6) Compare the target rate and distortion specifications with the values obtained above. If not optimal, go to step 2.

**Figure 3** shows the effective compression ratio at each level of bit plane decoding.

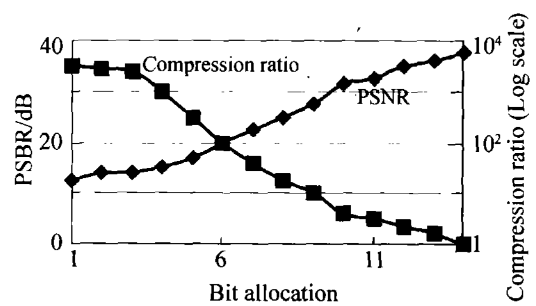


**Figure 3 relationship between the Bit Allocation, the Compression Ratio and the PSNR.**

Figure 3 represents error and compression as a function of bit plane number. The number here corresponds to the 256×256 "Lena" image encoded with the Adaptive search method.

## 5 Conclusions

A coding scheme that efficiently combines fractal image compression based on the optimization of the collage error and the orthogonal projection has been presented. The results prove the efficiency of this method. This method can also be extended to other compression methods, such as the nearest neighbor search method, the classification domain to improve their efficiency, and the fast encoding method.

## References


[1] Stephen Welstead, *Fractal and Wavelet Image Compression Techniques* [M], Published by the SPIE Optical between two consecutive range Engineering Press, 1999.

[2] M.F. Barnsley, *Fractal everywhere* [M], Academic Press, New York, 1988.

[3] A.E. Jacquin, Image coding based on a fractal theory of iterated contractive image transformations [J], *IEEE Trans. Image Process.*, 81(1993), No.10, p.1451.

[4] Y. Fisher, *Fractal Image Compression: Theory and Application* [M], Berlin, Germany: Springer-Verlag, 1995.

[5] Yuval Fisher, *Fractal Image Encoding and Analysis* [M], NATO ASI series F Springer, 1998.

[6] Chong Sze Tong and Minghong Pi, *Fast fractal Image Encoding Based on Adaptive Search* [J], *IEEE Tran. Image Process.*, 10(2001), No.9, p. 1269.

[7] E. Amram and J. B.T., Fractal image compression: an efficient acceleration technique, [in] *Proceeding of International Conference on Imaging Science, Systems and Technology* [C], CSST, Las Vegas 1997, p.372.

[8] H. Lin and A.N. Venet santopoulos, Fast fractal image coding using pyramids, [in] *Processing ICIAP'95* [C], San Ramon, USA, 1995, p.649.

[9] B. Bani-Eqbal, Fractal image compression, [in] *SPIE Proceedings* [C], San Jose, CA, USA, 2418(1995), p.67.

[10] B. Wohlberg and G. Jager, A review of the fractal image coding literature, *IEEE Transaction Image Processing* [C], 8(1999), p.1716.

[11] K. Berthe and Y. Yang, Motion video for Multimedia application in engineering teaching, [in] *5th UICEE Annual Conference on Engineering Education*, Chenej India, February 2002, p.156.

[12] Raouf Hamzaoui and Dietmar saupe, Combining Fractal Image Compression and Vector Quantization [J], *IEEE Trans. Image Process.*, 9(2000), No.2, p.197.

[13] S. Arya and D.M. Mount, Approximate nearest neighbor queries in fixed dimensions [J], [in] *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms* [C], Austin, TX, USA, 1993, p.271.

[14] Trieu-Kien Truong, Jyh-Horng Jeng, Irving S. Reed, P.C. Lee, and Alan Q. Li, A fast encoding algorithm for fractal image compression using the DCT inner product [J], *IEEE Trans. Image Process.*, 9(2000), No.4, p.529.

[15] K. Berthe and Y. Yang, Optimization of Fractal Iterated Function System (IFS) with probability and fracta image generation [J], *J. Univ. Sci. Technol. Beijing*, 8(2001), p. 152.

[16] D.M. Monro and F. Dudbrindge, Fractal block approximation of images [J], *Electron. Lett.*, 28(1992), No.11, p.870.

[17] S. Lepsoy, G.E. Oien, and T.A. Ramstad, Reducing the complexity of a fractal-based coder, [in] *Signal Processing VI: Theories and Applications* [C]. J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlink (eds.), Elsevier Science Publishers, 1998.

[18] D.J. Bone, Orthonormal fractal image encoding using overlapping blocks [J], *J. Fractals*, 28(1996), p. 52.

[19] K.U. Barthel, J. Schuttemeyer, T. Voye, and P. Noll, A new image coding technique unifying fractal and transform coding, [in] *Proc ICIP-94 IEEE Int. Conf. Image Processing* [C], 3(1994), Austin, TX, p.112.

[20] Chou-Chen wang and Chaur-Heh Hsieh, An efficient Farctal image-coding method using interblock correlation search [J], *IEEE Trans. Circuits Syst. Video Technol.*, 11(2001), No.1, p.257.

[21] Z. Barahav, D. Malah, and E. Karnin, Hierarchical interpretation of fractal image coding and its application to fast decoding, [in] *Proceedings of the IEEE International Conference on Digital Signal Processing* [C], Nicosia, Cyprus, 1993, p.190.

[22] E.A. Riskin, Optimum bit allocation via the generalized BFOS algorithm [J], *IEEE Trans. Inf. Theory*, 37(1991), p.400.