Available online at www.sciencedirect.com
SCIENCE DIRECT®
Signal Processing: *Image Communication* 18 (2003) 483–495
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Analysis of a hybrid fractal-predictive-coding compression scheme ☆

Chong Sze Tong[a,*], Minghong Pi[b]

[a] *Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong*
[b] *Department of Computing Science, University of Alberta, Edmonton AB, T6G 2E8, Canada*

Received 3 May 2002; received in revised form 4 December 2002; accepted 14 February 2003

Abstract

There has been tremendous progress in fractal compression since the pioneer work of Barnsley and Jacquin in the late 1980s. As the encoding time complexity issues are gradually being solved, there is a steady growth of applications of fractals, especially in hybrid systems. However, such fractal hybrid systems tend to be rather difficult to analyze, and part of that difficulty lies in the quantization of the scaling and luminance offset parameters adopted in most fractal compression schemes. In this paper, we present theoretical and empirical justification for a well-known but underused alternative parametrization for the fractal affine transform. In particular, we shall present a detailed analysis of a hybrid fractal-LPC (linear predictive coding) compression scheme using the aforementioned alternative affine transform parameters.

Keywords: Fractal image compression; Linear predictive coding; Fractal bit allocation

1. Introduction

Fractal image encoding has generated much interest due to its promise of high compression ratio at good decompression quality. Although it suffers from long encoding times, it has the advantage of very fast decompression. These properties made it a very attractive method for applications in multimedia: e.g. it was adopted by Microsoft for compressing thousands of images in its *Encarta* multimedia encyclopaedia [3]. Moreover, fractal encoding is much more than just another compression method since the encoding is capturing the self-similarity of the image and such fractal structures can be further exploited for value-added image processing. Thus fractal encoding has been widely used in, e.g. special image archive applications ranging from MR, ECG, to space images [21,22,30], feature extractions [25], image watermarking [23], image retrievals [1,6,16,26], texture segmentation [13] and many other image processing applications.

The basic implementation of Fractal Image Compression was pioneered by Jacquin [11,12] based on the theory of iterated functions developed by Hutchison [10] and Barnsley [4,5]. The

☆This work is partially supported by the Hong Kong Baptist University Faculty Research Grant Nos. FRG/99-00/II-16.
*Corresponding author. Tel.: +852-2339-7023; fax: +852-2339-5811.
E-mail address: cstong@hkbu.edu.hk (C.S. Tong).

0923-5965/03/$ - see front matter
doi:10.1016/S0923-5965(03)00037-7

essence of Jacquin's fractal compression is to exploit block-self-similarity. The image is first partitioned into $n \times n$ blocks called range blocks. A pool, Ω, of spatially contracted domain blocks D is generated from the image by sliding a $2n \times 2n$ size window through the image in fixed step size (domain block skip size), then spatially contract such $2n \times 2n$ domain blocks into the same size as the range blocks by spatial averaging of 2×2 sub-blocks. For each range block, R, we find the best matching domain block D such that $\|R - T(D)\|$ is minimized. The affine transform $T(D)$ is (usually) of the form:

$$T(D) = sD + o\mathbf{I}, \tag{1}$$

where the affine parameter s is the scaling coefficient, o is the luminance offset. D is the spatially contracted domain block, and $\mathbf{I}$ is a block of the same size as D, but with all entries equal to 1. In many fractal compression schemes, additional isometric operations are also applied after contraction. However, such isometries increase the encoding time with little gain in compression quality [24] and we shall not consider their use in this paper. Following Fisher [8], we shall adopt even value domain block skip size so that the contracted domain blocks can be directly obtained from the pre-decimated image for efficiency.

Although many enhancements to Jacquin's basic scheme have been proposed (see [29] for the latest review), virtually all of them use the same affine transform as in Eq. (1). There are of course attempts to generalize the transform, but most still adopt the same form while imposing extra structures on the affine parameters. For example, Øien et al. [20] suggested that the luminance offset assumes a linear function in the position indices, while Monro and Dudbridge [17] generalizes to third-degree polynomials.

Here we advocate replacing the luminance offset by an alternative affine parameter g. Specifically, the alternative affine transform takes the form:

$$T(D) = s(D - \bar{d}\mathbf{I}) + g\mathbf{I}, \tag{2}$$

where $\bar{d}$ is the mean intensity of the (contracted) domain block and $\mathbf{I}$ is an identity block where every entry is 1. In fact, other researchers have implicitly made use of this mean-removing transform. For example, Bani-Eqbal [2] carried out all the computations in an orthogonal complement sub-space, which is mathematically equivalent to using the transform in (2). Øien and Lepsøy also used orthogonal projection to improve decoder convergence [19]. Tong and Pi [27], and Tong and Wong [28] explicitly adopted this alternative parametrization, which considerably simplified the mathematical analysis. In this paper, we further discuss the properties of this parametrization and how it can facilitate analysis of a hybrid fractal-LPC compression scheme.

The idea of hybrid fractal compressions has been explored in a number of papers including [31,15,18]. The connections of fractal compression with wavelet transform was discussed by Davis [7], and successful hybrid system has been proposed by Kim et al. [14] that demonstrated superior rate-distortion performances over state of the art wavelet compression scheme. However, since such hybrids are essentially conventional encoders with enhancement provided by fractal codes, they are not fractal compression schemes as such and hence we would not be able to exploit the fractal (or self-similar) nature of the images for value-added image processing as discussed at the start of this section. In this paper, we shall discuss how a hybrid compression scheme that is built around the fractal encoder can be analyzed by adopting a non-conventional fractal parametrization. To demonstrate our approach, we shall construct a toy hybrid encoder and present a full analysis of the hybrid. The advantage of this approach is that we shall maintain the fractal encoder at the heart of the hybrid and hence will be able to exploit the fractal nature of the image as captured by the fractal codes.

In Section 2, we review our mathematical analysis of an alternative parametrization and show its superiority over the conventional affine transform. In Section 3, we describe how the new parametrization can be exploited to yield a hybrid compression scheme that combines fractal with linear predictive coding (LPC). A mathematical analysis is carried out to provide theoretical justification for preferring the alternative

transform to the conventional one. Experimental results are presented in Section 4. Finally, conclusions are discussed in Section 5.

2. Mathematical analysis of the alternative parametrization

2.1. Unconstrained fractal compression

We begin with no constraints imposed. That is, we wish to minimize the matching error for each pair of range and domain block:

$$\arg\min_{s,g} E(R,D) = \|R - s(D - \bar{d}\mathbf{I}) - g\mathbf{I}\|^2 = \sum (r_{ij} - g - s(d_{ij} - \bar{d}))^2. \quad (3)$$

The optimal solution is easily obtained by least squares:

$$g = \frac{\sum r_{ij}}{n^2} = \bar{r}, \quad s = \frac{\langle R - \bar{r}\mathbf{I}, D - \bar{d}\mathbf{I}\rangle}{\|D - \bar{d}\mathbf{I}\|^2}, \quad (4)$$

where $\bar{r}$ is the average pixel intensity of the range block. Here $\langle .,. \rangle$ is inner product, and $\|\ \|$ is the usual 2-norm. In this alternative affine transform parametrization, we have the same scaling parameter as in the conventional case, while the other transform parameter, g, is now shown to be the range block mean $\bar{r}$. This parameter is only a property of the range block and therefore can be encoded separately. We shall hereafter refer to this alternative affine parameter as $\bar{r}$. Note that this alternative parametrization is the same as that derived by Øien and Lepsøy [19], but our formulation simplified the derivation considerably.

By comparison, the unconstrained least squares result for the conventional affine transform method gives

$$s = \frac{\langle R - \bar{r}\mathbf{I}, D - \bar{d}\mathbf{I}\rangle}{\|D - \bar{d}\mathbf{I}\|^2}, \quad g = \bar{r} - s\bar{d}. \quad (5)$$

Note how the luminance offset mixes the range block mean with the scaling coefficient. This mixing is inefficient and is the main reason for its inferiority to the alternative affine parameter (see justification in Section 3).

The minimal distortion for the alternative affine transform method can be simplified to

$$E(R,D) = \|R - \bar{r}\mathbf{I}\|^2 - \frac{\langle R - \bar{r}\mathbf{I}, D - \bar{d}\mathbf{I}\rangle^2}{\|D - \bar{d}\mathbf{I}\|^2} = \|R - \bar{r}\mathbf{I}\|^2 - s^2\|D - \bar{d}\mathbf{I}\|^2, \quad (6)$$

which is, of course, the same as in the conventional transform case.

So far, this unconstrained formulation gives a simple basis for fractal compression; namely, we seek minimizing the distortion over $D \in \Omega$ by maximizing the normalized inner product

$$\frac{|\langle R - \bar{r}\mathbf{I}, D - \bar{d}\mathbf{I}\rangle}{\|D - \bar{d}\mathbf{I}\|}. \quad (7)$$

The corresponding optimal scaling parameter can then be quantized.

However, in practice, this scheme is unsatisfactory since we need to impose constraints on the scaling in order to ensure contractivity. It is better to work with a set of quantized fractal parameters $\{s_i\}$ and $\{\bar{r}_k\}$ and then to consider the minimization of the best matching error [8].

2.2. Full search fractal compression

Let $\mathscr{S} = \{s_i\}$ and $\mathscr{R} = \{\bar{r}_k\}$ be sets of quantized fractal parameters, then the full search scheme considers the minimization of

$$\hat{E}(R,D) = \|R - s_i(D - \bar{d}\mathbf{I}) - \bar{r}_k\mathbf{I}\|^2 = \|R - \bar{r}_k\mathbf{I}\|^2 + s_i^2\|D - \bar{d}\mathbf{I}\|^2 - 2s_i\langle R - \bar{r}_k\mathbf{I}, D - \bar{d}\mathbf{I}\rangle. \quad (8)$$

But

$$\langle R - \bar{r}_k\mathbf{I}, D - \bar{d}\mathbf{I}\rangle = \langle R - \bar{r}\mathbf{I}, D - \bar{d}\mathbf{I}\rangle = -s\|D - \bar{d}\mathbf{I}\|^2. \quad (9)$$

Let us define E_{quant} to be the additional matching error due to the quantization of the affine parameters. Using Eqs. (5), (7) and (8), we have

$$E_{\text{quant}} = \hat{E}(R) - E(R). = (s_i - s)^2\|D - \bar{d}\mathbf{I}\|^2 + n^2(\bar{r} - \bar{r}_k)^2. \quad (10)$$

Eq. (10) tells us how to choose the discretized values of $\mathscr{S} = \{s_i\}$ and $\mathscr{R} = \{\bar{r}_k\}$, namely, by the separate Lloyd–Max quantization of the two affine parameters. Moreover, by minimizing Eq. (10) with the constraint of using a specified number of bits for the quantization of both parameters, we have derived the optimal bit allocation scheme in [27]:

$$b_1 = \left\lceil \frac{b_{\text{total}} + \log_2\left(d_{\bar{r}}/\sqrt{\overline{\text{var}(D)}}\right)}{2} \right\rceil,$$

$$b_2 = b_{\text{total}} - b_1 \tag{11}$$

where b_1 is the number of quantization bits allocated for the range block mean, and b_2 that for the scaling; b_{total} is the total number of quantization bits used. $\overline{\text{var}(D)}$ is the average variance for the domain blocks in the pool. For example, for the Lena image, when eight bits are used, then the optimal allocation is two for scaling and six for range block mean. By contrast, there is as yet no such optimal bit allocation scheme for the conventional affine transform parameters. The best result in this direction for the conventional transform case is that due to Hartenstein et al. [9], which assumes an elliptic shape for the error functions with a constant ratio between the lengths of the major and minor axes. But this result is sub-optimal as different error functions have different shapes [9].

3. Hybrid fractal-LPC compression

There has been tremendous progress in fractal compression since the pioneer work of Barnsley and Jacquin in the late 1980s. As the encoding time complexity issues are gradually being solved (see, e.g. [28]), there is a steady growth of applications of fractals, especially in hybrid systems, e.g. [14,15,31]. However, such fractal hybrid systems tend to be rather difficult to analyze. Here, we shall consider a relatively simple hybrid system and show how the analysis of such a hybrid can be facilitated by the adoption of the alternative parametrization as discussed in Section 2.

For natural images, the range block mean intensities are highly correlated spatially, and this observation forms the basis for the LPC of the DC component in JPEG's DCT transform coding. Indeed, our range block means are precisely the DC component in DCT compression. Thus we can expect to achieve further compression by entropy encoding the prediction error of the range block mean intensities, yielding a hybrid fractal-LPC scheme. Note that this scheme differs from the hybrid LPC-IFS scheme proposed by Nappi and Vitulano [18]. In their case, the coding scheme is primarily an LPC followed by further fractal compression of the prediction error, whereas our scheme is primarily a fractal compression, followed by LPC compression of the DC component.

One of the advantages of our hybrid scheme is that the LPC part can be implemented using the JPEG DC compression protocol, which has been well developed for general applications. Moreover, this hybrid scheme may be used for thumbnail-based retrieval systems, with the DC component used for thumbnail display, and which can be improved to yield the fully featured image by iterating the fractal codes on the thumbnail as the initial image.

To test the feasibility of this hybrid and to facilitate theoretical comparison to the conventional affine transform parametrization case, we consider the simple previous-pixel LPC. That is, the prediction of each range block mean is given by the previous range block, viz. $\hat{\bar{r}}(i,j) = \bar{r}(i,j-1)$, where $\hat{\bar{r}}(i,j)$ is the predicted mean intensities for the (i,j)th range block. The histogram for the prediction error (i.e. the difference between successive range block means) is plotted in Fig. 1. Similarly, we apply the same predictive coding to the luminance offset in the conventional affine transform case. The corresponding histogram is given in Fig. 2.

Clearly, the predictive error for the range block mean intensities are much more concentrated at zero than that for the luminance offset, which is closer to a triangular shape than to the Laplacian distribution that the range block means appear to exhibit. Similar differences in distributions are observed for the baboon image, as given in Figs. 3a and b. We can understand the difference

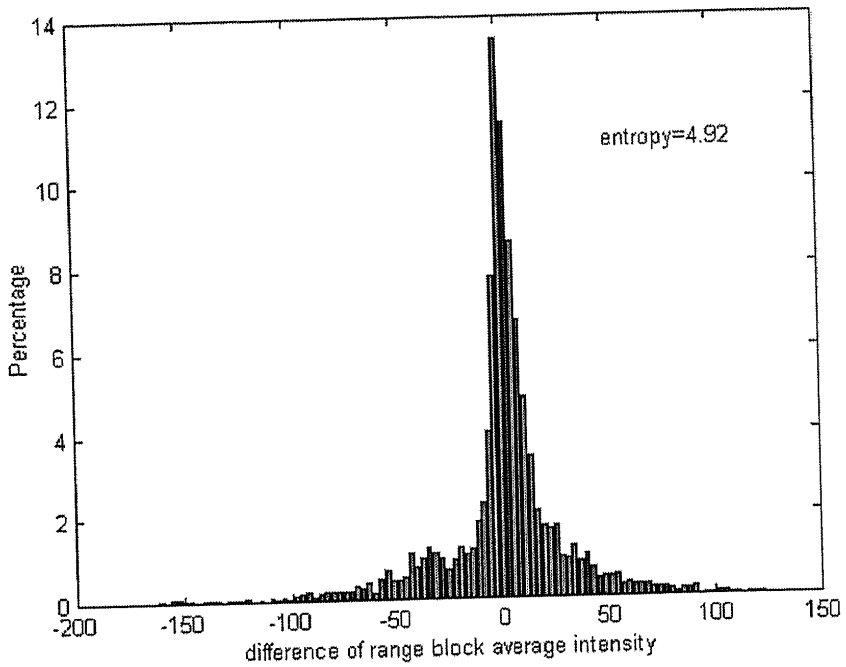

Fig. 1. Histogram of difference of range block average intensity, s-bits = 2, r-bits = 8.

by referring to the relationship between the affine parameters in Eq. (5), which gives the relationship in prediction error as follows:

$$g - \hat{g} = \bar{r} - s\bar{d} - (\hat{\bar{r}} - \hat{s}\hat{\bar{d}}) = \bar{r} - \hat{\bar{r}} - (s\bar{d} - \hat{s}\hat{\bar{d}})$$
$$\Rightarrow \mathrm{var}(g - \hat{g}) = \mathrm{var}((\bar{r} - \hat{\bar{r}}) - (s\bar{d} - \hat{s}\hat{\bar{d}})). \quad (12)$$

But from Eqs. (5) and (7), we see that the search for the optimal matching domain block and its associated scaling depends only on the respective blocks normalized to be zero-mean, namely, $R - \bar{r}\mathbf{I}$ and $D - \bar{d}\mathbf{I}$. Thus we may reasonably assume that the range block mean (and its prediction error) is statistically independent of the scaling and the domain block mean (and their prediction errors). Thus Eq. (12) leads to the inequality:

$$\mathrm{var}(g - \hat{g}) = \mathrm{var}((\bar{r} - \hat{\bar{r}}) - (s\bar{d} - \hat{s}\hat{\bar{d}}))$$
$$\Rightarrow \mathrm{var}(g - \hat{g}) = \mathrm{var}((\bar{r} - \hat{\bar{r}}) + \mathrm{var}(s\bar{d} - \hat{s}\hat{\bar{d}}))$$
$$\Rightarrow \mathrm{var}(g - \hat{g}) > \mathrm{var}(\bar{r} - \hat{\bar{r}}). \quad (13)$$

Thus the variance of the prediction error for the luminance offset is always greater than that for the range block mean. With reference to the comment after Eq. (5), the mixing effectively broadens the distribution of the error considerably.

We now attempt to characterize the broadening by making the following reasonable assumptions about the distributions of the above parameters:

1. the scaled domain block mean $s\bar{d}$ is uniformly distributed (from 0 to D, since we only consider positive scaling)
2. $\bar{r}$ and $s\bar{d}$ are statistically independent (as remarked above)
3. the prediction error $y = (\bar{r} - \hat{\bar{r}})$ has a Laplacian distribution $(1/2\sigma)\mathrm{e}^{-|y|/\sigma}$

Assumptions 1 and 3 are supported by empirical results, while assumption 2 is plausible since in our formulation, the search for the best matching domain blocks is made independent of the block means. Moreover, we have calculated the correlation for $\bar{r}$ and $s\bar{d}$ for various images, and their low values (e.g. for the Lena image, the correlation is 0.0951) are consistent with our assumption of independence.

Using the above assumptions, we now proceed to analyze the distribution of the predictive error $x = g - \hat{g}$ via the following two lemmas.

Lemma 1. We first note that since $\hat{s}\hat{\bar{d}}$ is just the translated copy of $s\bar{d}$, it has the same uniform

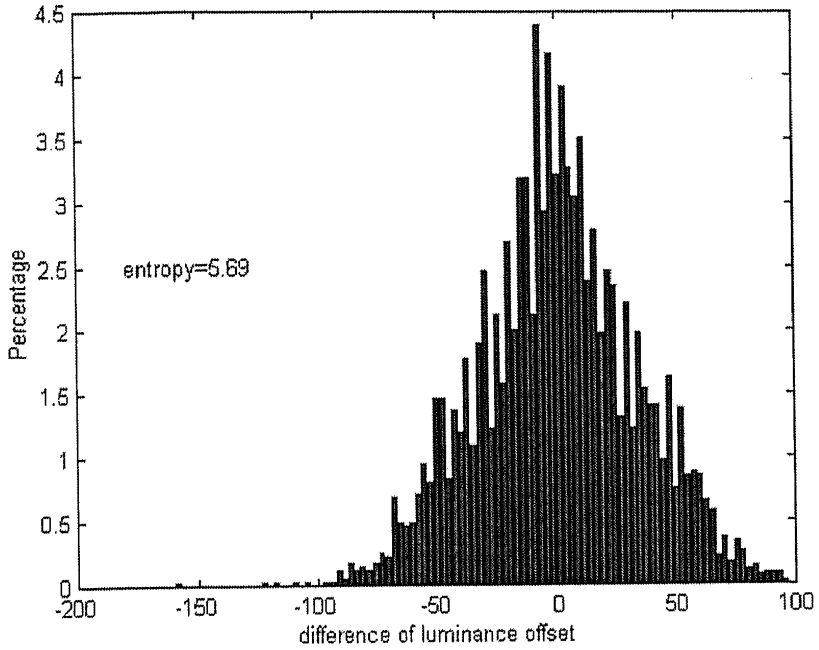

Fig. 2. Histogram of difference of the luminance offset, s-bits = 2, g-bits = 8.

distribution as $s\vec{d}$, namely,

$$f_1(x) = \frac{1}{D} \quad \text{for } x \in [0, D].$$

Thus $-\hat{s}\hat{\vec{d}}$ has the distribution

$$f_2(x) = \frac{1}{D} \quad \text{for } x \in [-D, 0].$$

If we further assume that $s\vec{d}$ is white (i.e. no spatial correlation), then the distribution for $z = (s\vec{d} - \hat{s}\hat{\vec{d}})$ is given by the convolution of the two uniform distributions:

$$p_z(z) = \int_{-\infty}^{\infty} f_1(z - x) f_2(x)\, dx.$$

Consider the case $z > 0$, then we have

$$p_z(z) = \int_{-D}^{0} f_1(z - x) \frac{1}{D}\, dx.$$

Now

$$f_1(z - x) = \frac{1}{D} \Leftrightarrow 0 < z - x < D$$
$$\Leftrightarrow z - D < x < z \text{ and } z < D.$$

Hence, combining the upper limit of x at 0, we have

$$p_z(z) = \int_{z-D}^{0} \frac{1}{D}\frac{1}{D}\, dx,$$
$$= \frac{1}{D^2}[0 - (z - D)] = \frac{D - z}{D^2}, \quad z < D.$$

The case for $z < 0$ is similar, yielding the final result

$$p_z(z) = \frac{D - |z|}{D^2} \quad \text{for } |z| < D. \tag{14}$$

Lemma 2. Now consider the prediction error for the luminance offset

$$x = g - \hat{g} = (\bar{r} - \hat{\bar{r}}) - (s\vec{d} - \hat{s}\hat{\vec{d}}).$$

The first bracketed term on the right is assumed to have the Laplacian distribution $g(y) = (1/2\sigma)e^{-|y|/\sigma}$, while the second term has the triangular distribution found in Lemma 1. Assuming the two terms are independent, then the distribution for the prediction error for the luminance offset is given by the convolution

$$p_x(x) = \int_{-\infty}^{\infty} g(x - z) p_z(z)\, dz.$$

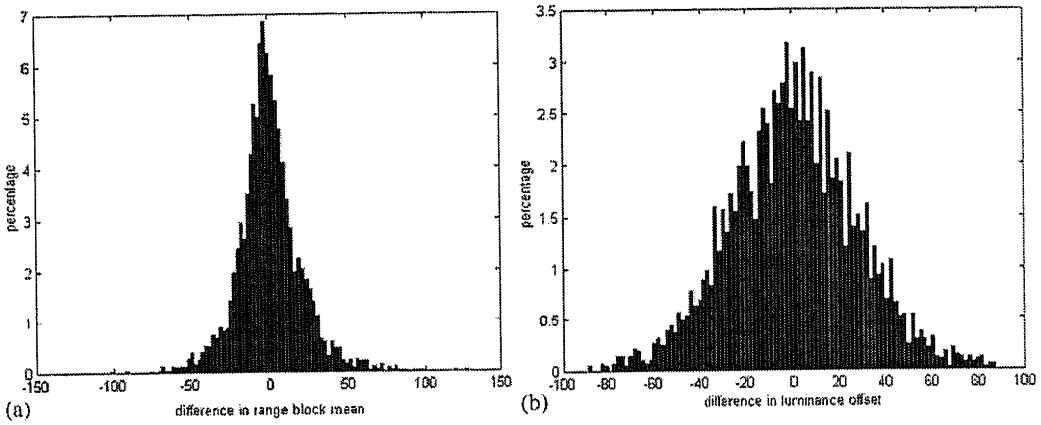

Fig. 3. (a,b): Histograms of the difference of the second affine parameter (s-bits = 2, g-bits = 8) for the Baboon image. The first corresponds to the results for the range block mean, while the second corresponds to the conventional luminance offset.

Consider the case for $x < -D$, we have

$$p_x(x) = \int_{-\infty}^{\infty} g(x-z)p_z(z)\,dz = \int_{-D}^{0} \frac{1}{2\sigma} e^{-|x-z|/\sigma}\frac{D+z}{D^2}\,dz + \int_{0}^{D} \frac{1}{2\sigma} e^{-|x-z|/\sigma}\frac{D-z}{D^2}\,dz$$

$$= \frac{1}{2\sigma D^2}\left(\int_{-D}^{0} e^{(x-z)/\sigma}(D+z)\,dz + \int_{0}^{D} e^{(x-z)/\sigma}(D-z)\,dz\right) = \frac{1}{2\sigma D^2}\begin{pmatrix}[-\sigma(D+z)e^{(x-z)/\sigma}]_{-D}^{0} - [\sigma^2 e^{(x-z)/\sigma}]_{-D}^{0} \\ +[-\sigma(D-z)e^{(x-z)/\sigma}]_{0}^{D} + [\sigma^2 e^{(x-z)/\sigma}]_{0}^{D}\end{pmatrix}$$

$$= \frac{\sigma}{2D^2}(e^{(x-D)/\sigma} + e^{(x+D)/\sigma} - 2e^{x/\sigma}) = \frac{\sigma e^{x/\sigma}}{D^2}\left(\cosh\left(\frac{D}{\sigma}\right) - 1\right).$$

And for $-D < x < 0$, we have

$$p_x(x) = \int_{-\infty}^{\infty} g(x-z)p_z(z)\,dz = \int_{-D}^{0} \frac{1}{2\sigma} e^{-|x-z|/\sigma}\frac{D+z}{D^2}\,dz + \int_{0}^{D} \frac{1}{2\sigma} e^{-|x-z|/\sigma}\frac{D-z}{D^2}\,dz$$

$$= \frac{1}{2\sigma D^2}\left(\int_{-D}^{x} e^{-(x-z)/\sigma}(D+z)\,dz + \int_{x}^{0} e^{(x-z)/\sigma}(D+z)\,dz + \int_{0}^{D} e^{(x-z)/\sigma}(D-z)\,dz\right)$$

$$= \frac{1}{2\sigma D^2}\begin{pmatrix}[\sigma(D+z)e^{-(x-z)/\sigma}]_{-D}^{x} - [\sigma^2 e^{-(x-z)/\sigma}]_{-D}^{x} \\ +[-\sigma(D+z)e^{(x-z)/\sigma}]_{x}^{0} - [\sigma^2 e^{(x-z)/\sigma}]_{x}^{0} \\ +[-\sigma(D-z)e^{(x-z)/\sigma}]_{0}^{D} + [\sigma^2 e^{(x-z)/\sigma}]_{0}^{D}\end{pmatrix} = \frac{1}{2\sigma D^2}\begin{pmatrix}\sigma(D+x) - \sigma^2(1-e^{-(x-D)/\sigma}) \\ -\sigma D e^{x/\sigma} + \sigma(D+x) - \sigma^2(e^{x/\sigma}-1) \\ +\sigma D e^{x/\sigma} + \sigma^2(e^{(x-D)/\sigma} - e^{x/\sigma})\end{pmatrix}$$

$$= \frac{1}{2\sigma D^2}(2\sigma(D+x) + \sigma^2(e^{-D/\sigma}2\cosh(x/\sigma) - 2e^{x/\sigma}) = \frac{D+x}{D^2} + \frac{\sigma}{D^2}(e^{-D/\sigma}\cosh(x/\sigma) - e^{x/\sigma}).$$

The two corresponding cases for $x > 0$ are similar, yielding the final result

$$p_x(x) = \begin{cases} \dfrac{D-|x|}{D^2} + \dfrac{\sigma}{D^2}\left(e^{-D/\sigma}\cos h\left(\dfrac{x}{\sigma}\right) - e^{-|x|/\sigma}\right) & \text{for } |x| \leqslant D, \\ \dfrac{\sigma e^{-|x|/\sigma}}{D^2}(\cos h(D/\sigma) - 1) & \text{for } |x| > D. \end{cases} \tag{15}$$

In practice, the standard deviation for the prediction error $y = (\bar{r} - \hat{\bar{r}})$ is usually much smaller than the dynamic range for the scaled domain block mean $\Rightarrow \sigma \ll D$. Thus the distribution in (15) reduces to

$$p_x(x) \approx \begin{cases} \dfrac{D - |x|}{D^2} & \text{for } |x| \leqslant D, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

which is a triangular distribution as remarked earlier in this section.

The Laplacian distribution for the prediction error of the range block mean has variance $2\sigma^2$, while the triangular distribution (Eq. (16)) for the prediction error of the luminance offset has variance $D^2/6$. Hence in terms of variances, the prediction error for the luminance offset is about $(1/12)(D/\sigma)^2$ times that for the range block mean. Thus for images for which the assumptions hold, we now have theoretical justification for the superiority of the alternative transform over the conventional transform.

4. Experimental results

All the programmes were implemented in Visual C++ 5.0 on a Dell Pentium II 233 MHz PC, and most of the experiments were carried out on the ubiquitous Lena image (256×256). Some additional results were generated from the baboon image. The range block size is 4×4, and the domain block skip is 4. The allocation of bits to code the two affine parameters follows the formula given in Eq. (11).

In Section 4.1, we compare the decoding performance of the two affine transform parametrizations. Then in Section 4.2, we evaluate the performance of our hybrid Fractal-LPC compression scheme.

4.1. Fractal decoding

In Section 2, we have provided theoretical justification for preferring the range block mean to the luminance offset for encoding. However, it may be argued that the alternative affine transform in Eq. (2) would lead to a slower decoder since the domain block mean needs to be calculated at each iteration: the alternative transform takes about 30% more computation per iteration for the decoding. However, in this sub-section, we shall see that the alternative transform still leads to better overall decoding performance because it has better convergence properties.

In this experiment, we used two bits for the scaling parameter, eight bits for the range block mean intensity, and nine bits for the luminance offset for the Lena image. In Figs. 4 and 5, we see the effect of the alternative affine transform parametrization on the convergence rate of the fractal decoding. The curves showed the PSNR of the decoded images after various iterations. In Fig. 4, we used a blank image as the initial image, and the conventional affine transform required about 10 iterations before convergence whereas our alternative affine transform converged after just three iterations. Since each iteration is about 30% faster in the conventional case, the decoding time for the alternative transform works out to be about 40% that for the conventional case. This result is consistent with the faster convergence property of Øien and Lepsøy's decoder based on orthogonal projection [19].

In Fig. 5, the initial image is now a gray image with intensity of 128. This difference in the initial image lead to an improvement in the convergence for the conventional affine transform, achieving convergence by about five iterations. For our alternative affine transform, there was virtually no change in the fractal decoding. These results showed that our alternative affine transform parametrization leads to superior decoding that is robust to the initial image used, even when more bits were being used in the conventional affine transform case!

The first three decoded images using the conventional and the alternative affine transforms are shown in Figs. 6a–f. The superiority of the alternative affine transform is obvious, and its independence to the initial image can be explained as follows. In the first iteration, the range block mean intensities already provide a coarse level approximation for the original image and this is true for any initial image: the coarse level

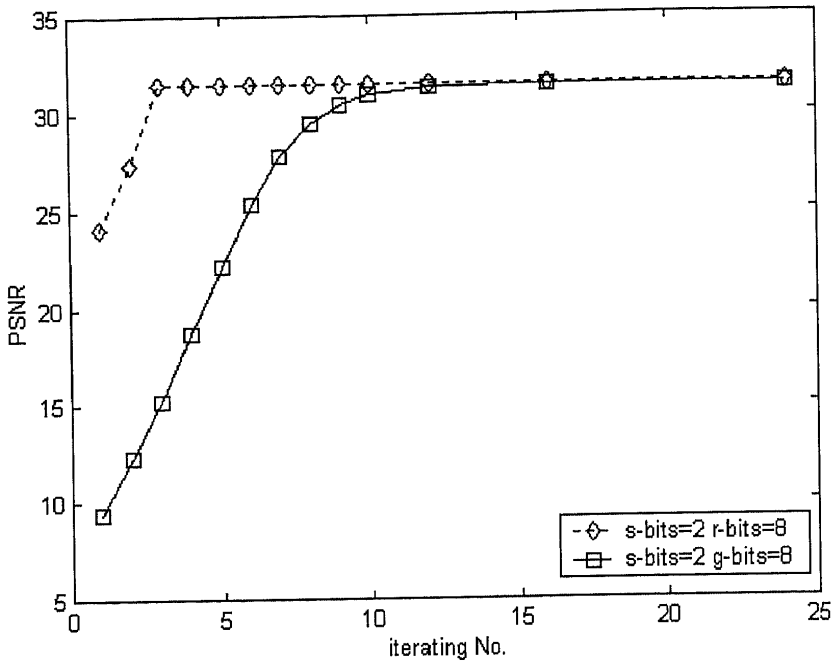

Fig. 4. Convergence of fractal decoding using a blank image as the initial image.

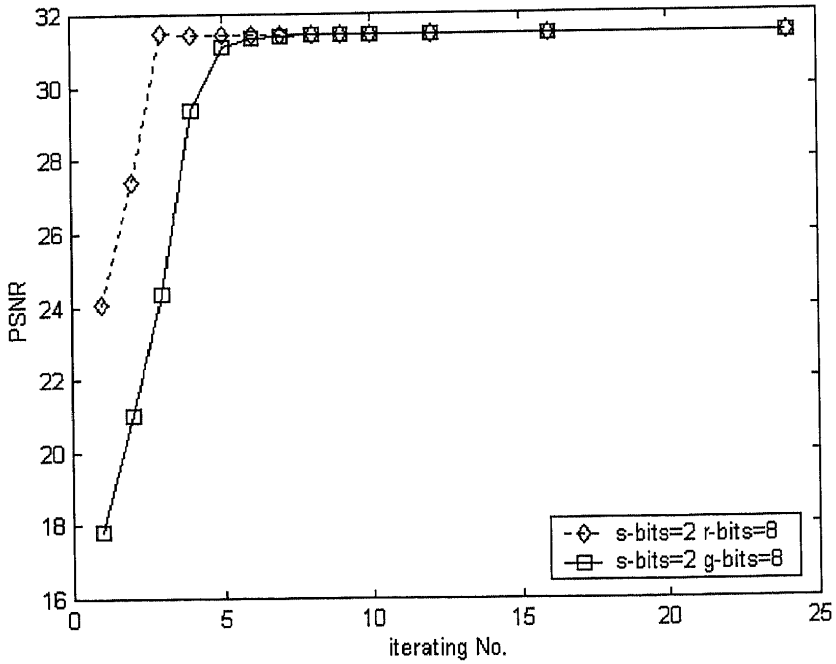

Fig. 5. Convergence of fractal decoding using a uniformly gray image (at intensity of 128) as the initial image.

description is already encoded by one of the affine parameter in the alternative affine transform. On the other hand, for the conventional affine transform, no such coarse level description is contained in the luminance offset, which mixes the range block mean intensities with the scaling.

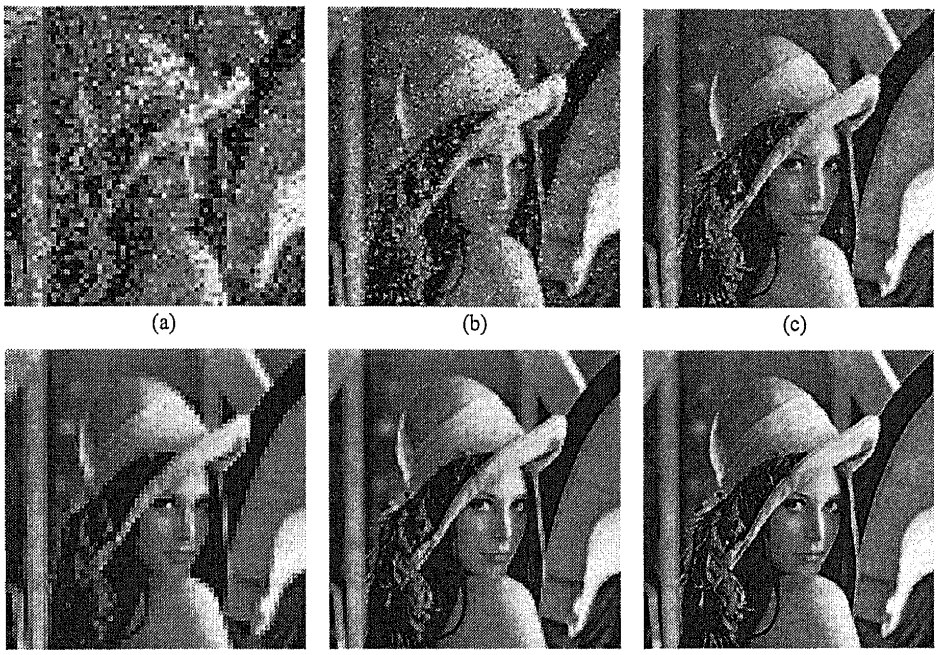

Fig. 6. (a–c): First three decoded images using the conventional affine transform; and (d–f): first three decoded images using the alternative affine transform.

In all subsequent experiments, the PSNR results will be derived from fractal decoding using the uniformly gray (intensity = 128) image as initial image and iterated for eight times.

4.2. Entropy-encoding

The sharp peak at zero prediction error for the range block mean in Figs. 1–3 suggests that we may obtain further compression by entropy encoding the prediction error. We thus implemented the basic fractal compression scheme as outlined in Section 3, with additional entropy encoding of the range block mean intensity. That is, we apply arithmetic coding (AC) to the prediction error of this alternative affine parameter. We shall refer to this scheme as LP + AC (Linear Prediction followed by AC). Similarly, we implemented the same LP + AC scheme for the luminance offset in the conventional affine transform case. The results are summarized in Tables 1 and 2. In terms of encoding time, there is practically no difference to the basic fractal scheme. The additional entropy encoding is very fast and is negligible in comparison to the fractal encoding part. The compression quality of course remains the same since the LP + AC coding is a lossless compression scheme.

Table 1
Results of LP + AC using the alternative affine transform

Number of bits used for quantization of affine parameters				PSNR
Total	Scaling	Range block mean	Range block mean (using LP + AC)	
7	2	5	3.83	30.79
8	2	6	4.75	31.29
9	2	7	5.68	31.44
10	2	8	6.58	31.46

As expected, the alternative affine transform outperformed the conventional affine transform. The third column in Tables 1 and 2 denote the number of bits allocated for the quantization of the range block mean/luminance offset in the basic fractal compression scheme, while the fourth

column denotes the average bits required in the LP + AC scheme. The average savings in quantization bit for the range block mean ranges from 1.17 to 1.42 bits per range block, increasing linearly with the number of bits used. For the luminance offset, the savings ranged from 0.67 to 1.01 bits per range block. The difference is graphically illustrated in Fig. 7, which shows that the alternative affine transform parameter consistently yields higher savings of about 0.5 bit per range block, while achieving better PSNR decompressions!

Table 2
Results of LP + AC using the conventional affine transform

Number of bits used for quantization of affine parameters				PSNR
Total	Scaling	Luminance offset	Luminance offset (using LP + AC)	
7	2	5	4.29	30.86
8	2	6	5.33	31.16
9	2	7	6.31	31.34
10	2	8	7.25	31.42
11	2	9	7.99	31.33

5. Conclusions

We have conducted theoretical and empirical comparative studies between the conventional affine transform used in fractal compression with an alternative affine transform parametrization. We showed that the range block mean intensity is a superior affine transform than the conventional luminance offset. Specifically, we showed that using the alternative affine transform parametrization leads to faster convergence of fractal decoding that is robust to the initial image used in the decoding. This is particularly important in fractal compression applications such as the multimedia encyclopaedia Encarta [3] which would include images of all types and thus would benefit from the improved decompression speed without worrying about which initial image to use.

The alternative affine transform parametrization facilitated a new bit allocation theory, which yielded better compression quality in terms of PSNR at the same compression ratio. Moreover, the alternative affine transform can be easily adopted in all existing fractal compression schemes, and hence can provide additional

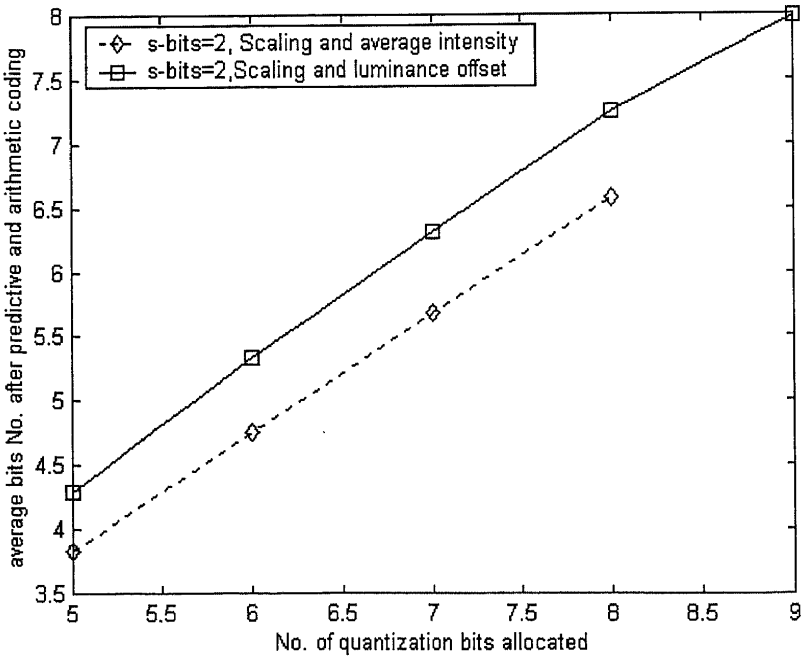

Fig. 7. Quantization bits for variable length coding versus fixed length coding.

enhancement at the negligible cost associated with the simple task of modifying the affine transform codes.

Finally, the alternative affine parameter facilitated analysis of a simple hybrid system made up of a fractal compression and a simple previous-pixel predictive coding followed by AC on the range block mean intensity. This hybrid fractal compression scheme achieved the same PSNR with lower bit-rate at essentially the same encoding time. Clearly, we can expect to do better still by carrying out more sophisticated predictive-coding schemes. However, since the optimal predictor for the range block mean is not necessarily the optimal predictor for the luminance offset, thus the theoretical comparison for such schemes for the two different parametrizations become much more complicated, and will be the focus of our future research.

References

[1] M. Ancis, W. Buchwald, P. Giusto, D. Daniele, D. Schmidt, Fractal zooming of thumbnails for progressive image coding, in: C.-C.J. Kuo, S.-F. Chang, S. Panchanathan (Eds.), Multimedia Storage and Archiving Systems III, Proceedings of the SPIE, Rhode Island, USA, Vol. 3527, 1998, pp. 541–549.

[2] B. Bani-Eqbal, Enhancing the speed of fractal image compression, Opt. Eng. 34 (6) (June 1995) 1705–1710.

[3] M. Barnsley, L. Barnsley, Fractal image compression, in: Image Processing: Mathematical Methods and Applications, Clarendon Press, Oxford, 1997, pp. 183–210.

[4] M. Barnsley, L. Hurd, Fractal Image Compression, AK Peters, Wellesley, MA, 1993.

[5] M. Barnsley, A. Jacquin, Applications of recurrent iterated function systems to images, SPIE Visual Commun. Image Process. 1001 (1988) 122–131.

[6] H.A. Cohen, Thumbnail-based image coding utilizing the fractal transform, Proceedings of the ICIP-96 IEEE International Conference on Image Processing, Lausanne, September 1996.

[7] G.M. Davis, Awavelet-based analysis of fractal image compression, IEEE Trans. Image Process. 7 (2) (1998) 141–154.

[8] Y. Fisher (Ed.), Fractal Image Compression: Theory and Applications, Springer, Berlin, 1995.

[9] H. Hartenstein, D. Saupe, Kai-Uwe Barthel, VQ-encoding of luminance parameters in fractal coding schemes, Proceedings of the ICASSP-97, Munich, Germany, April 21–24, Vol. 4, 1997, pp. 2701–2704.

[10] J.E. Hutchison, Fractals and self-similarity, Indiana Univ. Math. J. 3 (5) (1981) 713–747.

[11] A.E. Jacquin, Image coding based on a fractal theory of iterated contractive image transform, IEEE Trans. Image Process. 1 (1) (1992) 18–30.

[12] A.E. Jacquin, Fractal image coding: a review, Proc. IEEE 81 (10) (1993) 1451–1465.

[13] L.M. Kaplan, C.C. Jay Kuo, Texture segmentation via haar fractal feature estimation, J. Visual Commun. Image Representation 6 (4) (1995) 387–400.

[14] T. Kim, R.E. Van Dyck, D.J. Miller, Hybrid fractal zerotree wavelet image coding, Signal Processs.: Image Commun. 17 (2002) 347–360.

[15] J. Li, C.C. Kuo, Image compression with a hybrid wavelet-fractal coder, IEEE Trans. Image Process. 8 (6) (June 1999) 868–874.

[16] J. Marie-Julie, H. Essafi, Using the fractal transform for image and indexing and retrieval by content, Proceedings of the Conference Fractals in Engineering, Arcachon, June 1997.

[17] D.M. Monro, F. Dudbridge, Fractal block coding of images, Electron. Lett. 28 (11) (1992) 1053–1055.

[18] M. Nappi, D. Vitulano, Linear prediction image coding using iterated function systems, Image Vision Comput. 17 (1999) 771–776.

[19] G.E. Øien, S. Lepsøy, A class of fractal image coders with fast decoder convergence, in: Y. Fisher (Ed.), Fractal Image Compression: Theory and Applications, Springer, Berlin, 1995, pp. 153–175.

[20] G.E. Øien, S. Lepsøy, T.A. Ramstad, An inner product space approach to image coding by contractive transformations, Proceedings of the ICASSP-91, Toronto, Canada, 1991, pp. 2773–2776.

[21] G.E. Øien, G. Narstad, Fractal compression of ECG signals, in: Y. Fisher (Ed.), Conference Proceedings of the NATO ASI Fractal Image Encoding and Analysis, Trondheim, July 1995, Springer, Berlin, Heidelberg, 1998.

[22] D.C. Popescu, H. Yan, MR image compression using iterated function systems, Magn. Resonance Imag. 11 (1993) 727–732.

[23] J. Puate, F. Jordan, Using fractal compression scheme to embed a digital signature into an image, in: T. Chiueh, A.G. Tescher (Eds.), Video Techniques and Software for Full-Service Networks, Proc. SPIE, Saratoga, USA, 2915 (January 1997) 108–118.

[24] D. Saupe, The futility of square isometries in fractal image compression, Proceedings of the ICIP-96 IEEE International Conference on Image Processing, Lausanne, September 1996.

[25] B. Schouten, P. de Zeeuw, Feature extraction using fractal codes, in Visual 99; visual information and information systems, Third International Conference, Springer, Berlin, 1999, pp. 483–492.

[26] T. Tan, H. Yan, Object recognition using fractal neighbor distance: eventual convergence and recognition rates, Proceedings of the ICPR2000, Barcelona, Spain, Vol. 2, 2000, pp. 785–788.

[27] C.S. Tong, M. Pi, Fast fractal image encoding based on adaptive search, IEEE Trans. Image Process. 10 (9) (September 2001) 1269–1277.

[28] C.S. Tong, M. Wong, Adaptive approximate nearest neighbour search for fractal image compression, IEEE Trans. Image Process. 11 (6) (June 2002) 605–615.

[29] B.E. Wohlberg, G. de Jager, A review of the fractal image coding literature, IEEE Trans. Image Process. 8 (12) (1999) 1716–1729.

[30] C.M. Wood, General data compression algorithm for space images using fractal techniques, in D.L. Crawford, E.R. Craine (Eds.), Instrumentation in Astronomy, Vol. VIII, Proc. SPIE 2198 (1994) 1336–1341.

[31] N. Zhang, H. Yan, Hybrid image compression method based on fractal geometry, Electron. Lett. 27 (5) (1991) 406–408.