Hybrid Wavelet-Fractal Image Compression Based on A Rate-Distortion Criterion

Jin Li

Sharp Labs. of America
5750 NW Pacific Rim Blvd., Camas, WA 98607

C.-C. Jay Kuo

Integrated Media Systems Center and Department of Electrical Engineering-Systems
University of Southern California, Los Angeles, California 90089-2564

ABSTRACT

A new hybrid wavelet-fractal coder (WFC) for image compression is proposed in this research. We show that the application of contractive mapping for interscale wavelet prediction in the wavelet domain offers bit rate savings in some regions. The prediction residue is then quantized and encoded by traditional wavelet coders. WFC allows the flexibility to choose either direct coding of wavelet coefficients or fractal prediction followed by residual coding to achieve a better rate-distortion (R-D) performance. A criterion of low complexity is derived to evaluate the R-D efficiency of fractal prediction. The superior performance of WFC is demonstrated with extensive experimental results.

Keywords: wavelet-fractal coder, image coding, fractal compression, wavelet compression.

1 INTRODUCTION

Research on theory and applications of fractals has been performed for more than thirty years. Pentland [12], among many others, showed that quite a few natural physical processes generated fractal-like surfaces and when the 3-D fractal surface projected to a plane, the resulting 2-D image had a self-similar property which characterized a fractal image. The application of the fractal (or self-similar) concept to practical image compression was first demonstrated by Jacquin [8], who proposed a block-based fractal coder. More improvements have been performed by many researchers at a later stage [6], [7].

Fractal compression is however distinctive from conventional transform-based coding methods in several aspects. First, it uses parameters of the contractive mapping [8] to encode an image rather than a direct coding of its content. Second, unlike the invertible transform used in transform-based coders, contractive mapping is an irreversible procedure. Third, quantization of contractive mapping parameters is not the main source of distortion. Compression artifact is primarily caused by the process of contractive mapping. As a result, the bit rate and image quality control for the fractal coder is difficult to perform. In spite of all achievements in fractal theory and fractal coder implementations, the rate-distortion (R-D) performance of fractal coders is disappointing. They can hardly compete with the state-of-art wavelet coders such as the embedded zerotree wavelet coder (EZW) proposed by Shapiro [14] and the layered zero coder (LZC) proposed by Taubman and Zakhor [15].

The relationship between the fractal and the transform-based coders has been recently unveiled by researches [2], [4]. It can be shown that contractive mapping actually explores redundancy between different image resolutions, and can be viewed as an interscale prediction in the wavelet domain. Thus, there should be a way to combine fractal and wavelet coding ideas for a better compression method. Rinaldo and Calvagno [13] proposed a predictive pyramid coder (PPC) by exploring the interscale redundancy. PPC performs a block-based prediction which predicts finer scale wavelet coefficients from those of coarser scales. The prediction of PPC is independent for each scale and directional subband, and the block size can be adjusted. It turns out that PPC bears little resemblance to the contractive mapping defined in traditional fractal coders. Also, its coding efficiency is still not as good as EZW and LZC (see Table 2 in Section 5).

In this paper, we attempt to shed more light on the relationship between contractive mapping and wavelet coding. Following the pioneering work in [2], [4], a more complete framework for fractal prediction in the wavelet domain will be developed. It includes three components: (1) the contractive operator that provides prediction across wavelet scales, (2) the scaling operator that shrinks prediction coefficients, and (3) the isometrical transformation which shuffles pixels within and across subbands. We will then integrate the wavelet domain contractive mapping with wavelet coders which leads to a new hybrid wavelet-fractal coder (WFC) with a superior R-D performance. Unlike conventional fractal coders where the whole image is encoded by fractal alone, we only apply fractal prediction to a selected part of the image whenever it is more efficient. To be more precise, an adaptive rule is developed to evaluate the perspective rate saving achieved by fractal prediction, which is then compared with the overhead bits required by prediction. Fractal prediction is only adopted when the rate saving is more than the overhead. Finally, both unpredicted wavelet coefficients and fractal prediction residue are quantized and entropy encoded by state-of-the-art wavelet coders. Experimental results show that fractal prediction is able to reduce the coding rate in a certain bit consuming area such as the texture region and improves the overall R-D performance of the coder.

The paper is organized as follows. We describe the implementation of contractive mapping in the wavelet domain in Section 2. Algorithmic details of WFC are presented in Section 3. The issue of estimating perspective rate saving based on a model-based R-D criterion is investigated in Section 4. Extensive experimental results are provided to compare the performance of WFC with various state-of-the-art fractal and wavelet coders in Section 5. Concluding remarks are given in Section 6.

2 INTERSCALE WAVELET PREDICTION VIA CONTRACTIVE MAPPING

2.1 Contractive Mapping in the Space Domain

The conventional fractal coder operating on spatial domain blocks is briefly reviewed below. Let us partition the image $\mathbf{f}$ into a set of non-overlapping square range blocks $\{\mathbf{r_j}, j = 1, \cdots, N_R\}$ of size $K_R \times K_R$. That is,

$$\bigcup_{j=1}^{N_R} \mathbf{r_j} = \mathbf{f}, \qquad \mathbf{r_i} \cap \mathbf{r_j} = \emptyset, \quad for \quad i \neq j, \tag{1}$$

$$\mathbf{r_j} = \{(x, y) \mid r_x - K_R/2 \leq x < r_x + K_R/2, r_y - K_R/2 \leq y < r_y + K_R/2\}. \tag{2}$$

The range block has to be matched with a domain block of size $K_D \times K_D$:

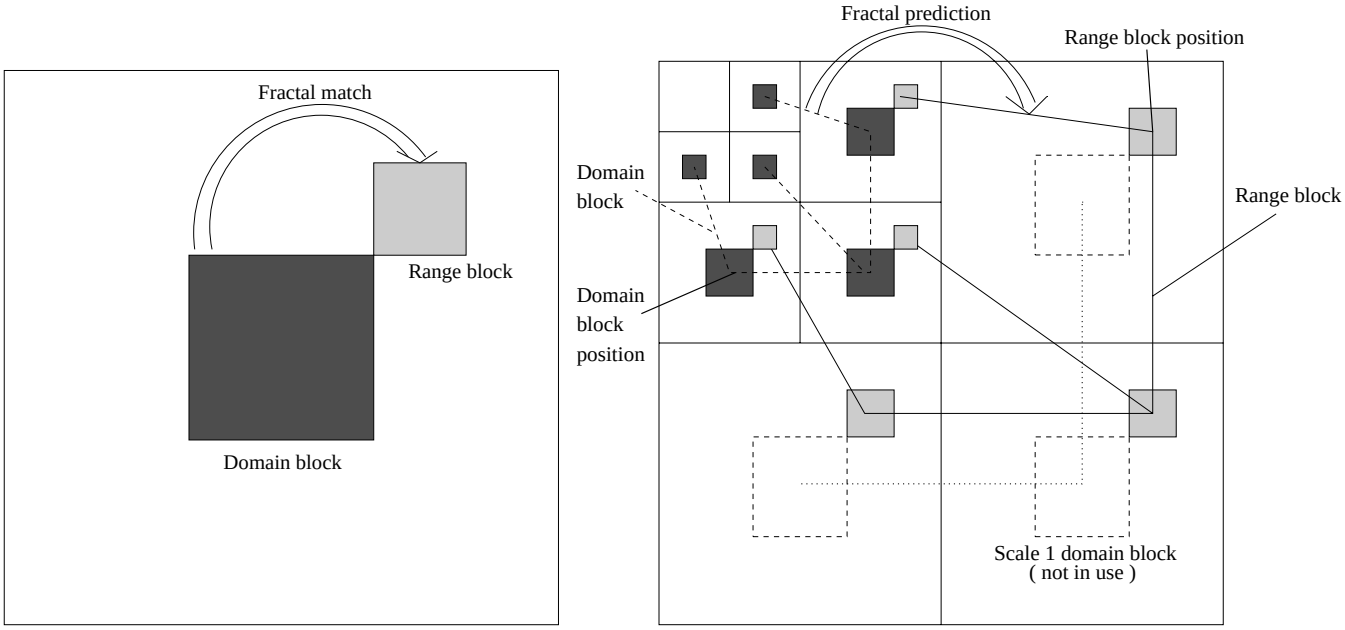
$$\mathbf{d_k} = \{(x, y) \mid d_x - K_D/2 \leq x < d_x + K_D/2, d_y - K_D/2 \leq y < d_y + K_D/2\}. \tag{3}$$

Centers of the range and domain blocks, i.e. (r_x, r_y) and (d_x, d_y), are used to denote their positions. Usually, the size of the domain block is twice of that of the range block $K_D = 2 \cdot K_R$. All domain blocks constitute a domain pool

$$\mathcal{D} = \{\mathbf{d_k}, k = 1, \cdots, N_D\}.$$

During the matching process, the domain block $\mathbf{d_k}$ is averaged and subsampled from size $K_D \times K_D$ down to size $K_R \times K_R$, and followed by isometrical transformation, contrast scaling and offset operations. Mathematically, the process can be written as

$$\hat{\mathbf{r_j}} = \tau_j(\mathbf{d_k}) = \alpha_j \Gamma_j(S_j(\mathbf{d_k})) + o_j, \tag{4}$$



(a) (b)

Figure 1: Contractive mapping in (a) the spatial domain, (b) the wavelet domain

where S_j, Γ_j, α_j and o_j represent the spatial contraction, isometrical transformation, contrast scaling and luminance offset, respectively. Isometrical transformation Γ_j includes identity, horizontal and vertical flip, diagonal reflection along the first axis (transpose) and the second axis (diagonal flip), and 90^o, 180^o, and 270^o rotations [8]. Fractal matching in the space domain is depicted in Fig. 1(a), and the corresponding operating chain of (4) is shown in Fig. 2(a). Due to the use of the spatial contraction operator S_j, this search process is also known as *contractive mapping*.

For each range block $\mathbf{r_j}$, it requires a high computational complexity to search the best matching domain block $\mathbf{d_k}$ for its position (d_x, d_y), isometrical transformation Γ_j, scaling parameter α_j and offset parameter o_j. To reduce the complexity, it is possible to search within a subset $\mathcal{D}_j$ of the entire domain pool for locality control and/or classified block matching. Once the range-domain match pair $\hat{\mathbf{r}}_\mathbf{j} = \tau_j(\mathbf{d_k})$ is determined, fractal decoding can be carried out iteratively.

2.2 Contractive Mapping in the Wavelet Domain

Fractal signals possess a self-similar property across scales while the wavelet transform provides an efficient multiresolution representation of signals. It is thus interesting to investigate the fractal structure in the wavelet domain. Let us perform the Haar wavelet transform on both range and domain blocks. Then, the chain of spatial contractive mapping operations shown in Fig. 2(a) can be converted to the chain of wavelet contractive mapping operations shown in Fig. 2(b). The averaging and subsampling operation S_j is equivalent to moving up the domain block by one scale in the wavelet domain, since the Haar transform is exactly the same as the combined averaging and subsampling operations. The isometrical transformation operator Γ_j corresponds to a pixel shuffling operation which may involve additional sign changes and inter-subband coefficient exchanges. The contrast scaling factor α_j is reduced by half because of the normalization process in the wavelet transform, and the offset o_j is related only to the difference between the dc components of the range and the domain blocks.

Each domain or range block in the space domain is now decomposed into several subblocks across scales and directional subbands in the wavelet domain. For typical test images of size 512×512 in our experiments, we consider range and domain blocks of sizes 16×16 and 32×32, respectively. After performing a 5-level pyramid wavelet decomposition, A range block of size 16×16 is decomposed into 3 scale-1 wavelet subblocks $rw_1(i, j)$ of size 8×8, 3

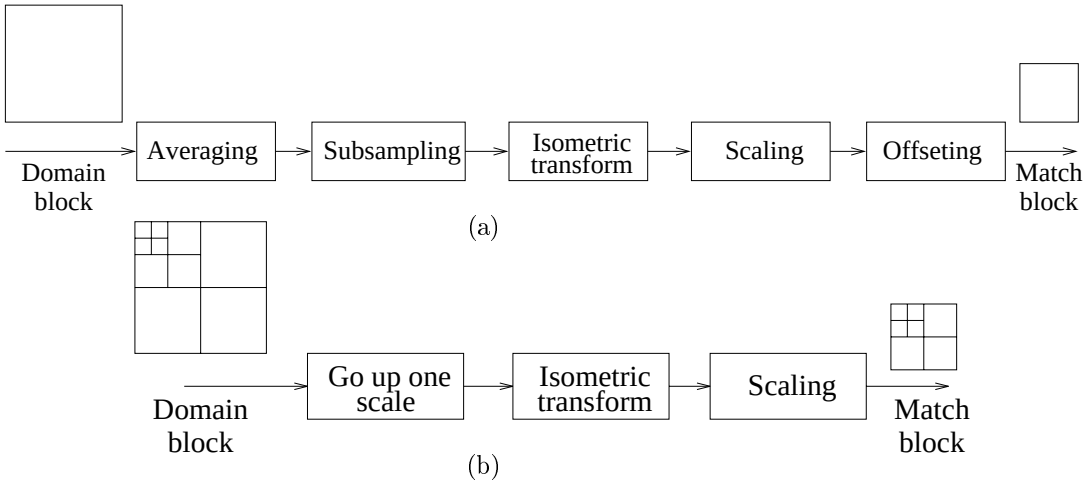


Figure 2: Flowchart of contractive mapping operations in (a) the spatial domain, (b) the wavelet domain

scale-2 wavelet subblocks $rw_2(i,j)$ of size 4×4, 3 scale-3 wavelet subblocks $rw_3(i,j)$ of size 2×2, 3 scale-4 wavelet subblocks $rw_4(i,j)$ of size 1×1, and 1 scale-5 dc subblock $rw_5(i,j)$ of size 1×1. Similarly, a domain block of size 32×32 is decomposed into scale-1, scale-2, scale-3, scale-4, scale-5 wavelet subblocks of size 16×16, 8×8, 4×4, 2×2 and 1×1 and denoted by $dw_d(i,j)$ with $d = 1, \cdots, 5$, respectively. By omitting dc components, contractive mapping in the wavelet domain establishes a predictive relationship of wavelet coefficients across scales. That is, for a given range block, we choose an optimum domain block which minimizes distances between $dw_d(i,j)$ and $rw_{d-1}(i,j)$ for scales $d = 5$, 4, 3, 2 simultaneously. The dc coefficients of $rw_5(i,j)$ are encoded and decoded directly. For coefficients of other scales, we can perform the following prediction. We use wavelet coefficients of $dw_d(i,j)$ of the best matched domain block to predict the wavelet coefficient in $rw_{d-1}(i,j)$ of the range block for $d = 5$, 4, 3, 2. This prediction procedure is called the *wavelet-fractal prediction*. The infinite decoding resolution property of a fractal coder is achieved by extrapolating wavelet coefficients from scales d (domain block coefficients) to $d-1$ (range block coefficients) repetitively for $d = 1, 0, -1, \cdots$.

Mathematically, the wavelet-fractal prediction can be written as

$$\hat{\mathbf{r_j}} = \tau_j(\mathbf{d_k}) = \alpha_j \Gamma_j(\mathbf{d_k}) \tag{5}$$

We do not need the contractive operator S_j since it is achieved by using the wavelet filter. In the wavelet domain, the size of the range block K_R of scale $l-1$ is equal to the size of the domain block K_D of scale l, i.e. $K_R = K_D = K$.

Although isometry is very important for efficient wavelet-fractal prediction, it was not thoroughly examined in previous wavelet-fractal work [4], [13], [2]. Here, we show that the space-domain isometrical transformation Γ_j can be performed in the wavelet domain as well. It involves pixel shuffling within the wavelet subblocks, pixel interchanging between HL and LH subbands and sign changes. The wavelet isometry operator Γ_j is derived from its corresponding operator in the space domain. Analysis shows that a wavelet isometry operator of 90^o counter clockwise rotation can be achieved through:

1. Isometrically transform all LL, LH, HL, HH subbands with the same 90^o counter clockwise rotation.
2. Interchange the LH and HL subbands.
3. Reverse the signs of the LH and HL subbands.

Similarly, we can derive the other 7 common types of wavelet isometric transformations and summarize the results in Table. 1. To conclude, when a wavelet block undergoes an isometric transform, all its subbands first undergo the same spatial isometric transform. Then if the applied isometry operators are transpose, diagonal flip, 90^o or 270^o rotation, the HL and the LH subbands are interchanged. Finally, we may need to reverse the signs of certain subbands.

Isometry	Spatial	Wavelet Domain		
	$r(i,j)$	$w_l^{lh}(i,j)$	$w_l^{hl}(i,j)$	$w_l^{hh}(i,j)$
Identity	$d(i,j)$	$w_{l+1}^{lh}(i,j)$	$w_{l+1}^{hl}(i,j)$	$w_{l+1}^{hh}(i,j)$
Flip up-down	$d(-i,j)$	$w_{l+1}^{lh}(-i,j)$	$-w_{l+1}^{hl}(-i,j)$	$-w_{l+1}^{hh}(-i,j)$
Flip left-right	$d(i,-j)$	$-w_{l+1}^{lh}(i,-j)$	$w_{l+1}^{hl}(i,-j)$	$-w_{l+1}^{hh}(i,-j)$
Transpose	$d(j,i)$	$w_{l+1}^{hl}(j,i)$	$w_{l+1}^{lh}(j,i)$	$w_{l+1}^{hh}(j,i)$
Diagonal flip	$d(-j,-i)$	$-w_{l+1}^{hl}(-j,-i)$	$-w_{l+1}^{lh}(-j,-i)$	$w_{l+1}^{hh}(-j,-i)$
90^o rotation	$d(j,-i)$	$-w_{l+1}^{hl}(j,-i)$	$w_{l+1}^{lh}(j,-i)$	$-w_{l+1}^{hh}(j,-i)$
180^o rotation	$d(-i,-j)$	$-w_{l+1}^{lh}(-i,-j)$	$-w_{l+1}^{hl}(-i,-j)$	$w_{l+1}^{hh}(-i,-j)$
270^o rotation	$d(-j,i)$	$w_{l+1}^{hl}(-j,i)$	$-w_{l+1}^{lh}(-j,i)$	$-w_{l+1}^{hh}(-j,i)$

Table 1: Correspondence of the isometry operator in the spatial and wavelet domains.

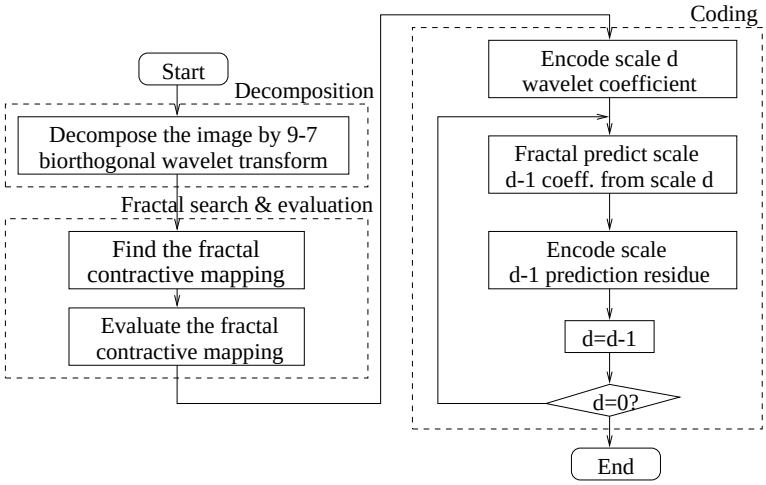


Figure 3: Framework of the fractal wavelet coder

The wavelet contractive mapping relation is derived through the Haar wavelet. Nevertheless, we extend it to other wavelet basis. In fact, the spatial fractal matching relation (4) is nothing more than an assumption that in the space domain, the range block resembles the contractive mapping of the domain block. Extending this assumption to its counterpart in the wavelet domain, we assume that the wavelet coefficients of coarse scale can be predicted by a wavelet contractive mapping of the fine scale. The wavelet contractive mapping involves the specification of the range-domain pair $\hat{\mathbf{r}}_\mathbf{j} = \tau_j(\mathbf{d_k})$, the wavelet isometry operator Γ_j, and the scaling operator α_j, as shown (5). This assumption, which does not constrain the wavelet basis to be haar, is the core of wavelet-fractal prediction.

There are several variations in implementing the prediction between wavelet subbands. One is to perform scale-dependent prediction. That is, we determine the best range-domain block pair for every two consecutive scales. Another possibility is to allow scale and directional-dependent prediction as done in PPC [13] to reduce the prediction error. However, they are not as efficient as the one described above in optimizing the rate-distortion performance.

3 WAVELET-FRACTAL CODER (WFC)

A new hybrid image coder called the wavelet-fractal coder is proposed in this section. Generally speaking, we first transform an image from the space domain to the wavelet domain. Then, wavelet coefficients are either encoded directly or predicted from coefficients from the previous scale via contractive mapping. For the latter case, both contractive mapping parameters and prediction residuals are encoded. The flowchart of the proposed wavelet-fractal coder (WFC) is depicted in Fig. 3. Detailed implementations are described below. Without loss of generality, we consider the case that the image to be compressed is of size 512×512 in the following discussion.

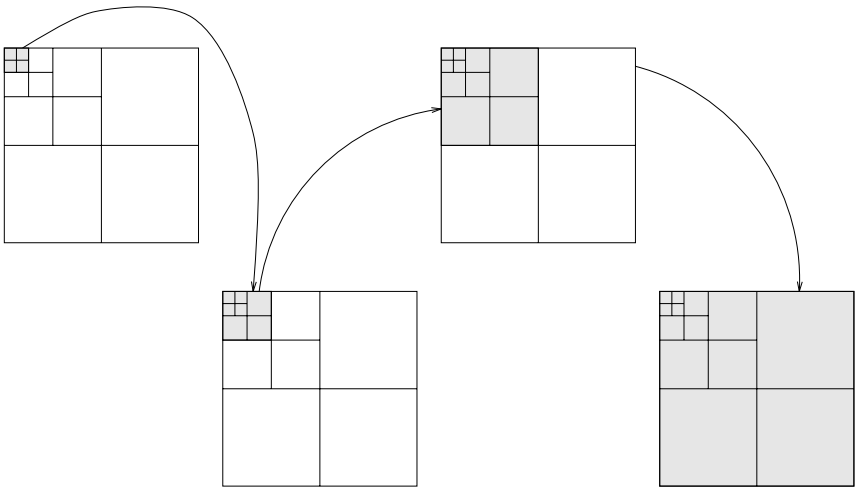


Figure 4: Fractal wavelet coding process

1. Wavelet decomposition
   We decompose the image with a 5-level 2-D pyramidal wavelet transform. The biorthogonal 9-7 tap spline filter-II [1], with the symmetric boundary extension [10] is used in this work. With the symmetric boundary extension, we reduce the signal discontinuity effect across boundaries to enhance the coding efficiency. For an image of size 512×512, the LL band of the coarsest level is of size 16×16. We partition the image into a union of range blocks of size 16×16. Coefficients in each wavelet subband are split into range blocks of different sizes as explained in Fig. 1(b). This process is equivalent to the decomposition of the original range block $\mathbf{r_j}$ of size 16×16 in space domain into multiple subblocks of different scales and directional subbands.

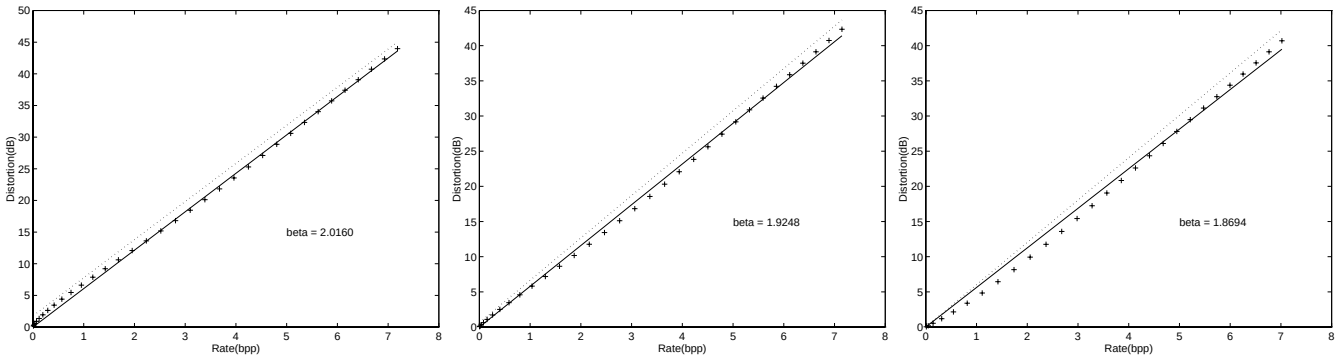
2. Search for optimal wavelet-fractal prediction
   For each range block $\mathbf{r_j}$, its best matching domain blocks $\mathbf{d_k}$ is searched. As indicated in Section 2.2, each domain block is also constituted of multiple subblocks across scales and directional subbands. During the search, the domain block $\mathbf{d_k}$ first undergoes the wavelet isometry Γ_j and scaling α_j, the resultant domain subblock is then used to generate prediction for its corresponding range subblock one scale down. We compare the perspective coding rate without fractal prediction and with fractal prediction. The domain block $\mathbf{d_k}$ which gives the maximum coding rate saving is recorded with its position (d_x, d_y), isometry operator Γ_j, and scaling factor α_j.

   We adopt exhaustive search in this research. However, various fast search algorithms [3] can also be applied with a suboptimal rate-distortion performance but a much lower computational complexity.

3. Prediction efficiency evaluation
   As a result of Step 2, we know the position (d_x, d_y), isometry operator Γ_j, and scaling factor α_j of the best matched domain block $\mathbf{d_k}$ for each range block $\mathbf{r_j}$. Then, we evaluate whether fractal prediction for this range block is efficient or not in comparison with direct coding. This involves again the estimation of the perspective coding rate with and without fractal prediction at a certain quality level. Details of this step will be given in this and next sections.

4. Wavelet coding assisted by fractal prediction
   After the prediction efficiency evaluation, we proceed the wavelet coding assisted by fractal prediction in a top-down manner as depicted in Fig. 4. Our coding process starts at scale d. There is no fractal prediction for this scale. We first encode the scale d wavelet coefficients with LZC down to the terminal significant threshold T. We then move on to scale d-1, and use fractal to predict scale d-1 wavelet coefficients from scale d.

   For the non fractal predicted blocks, the prediction value is set to be 0 and the prediction residue is just the same as the original wavelet coefficients. We use again LZC to encode both the prediction residue and the non fractal predicted wavelet coefficients in scale d-1 down to the same terminal significant threshold T. The process repeats until scale 1 wavelet residue has been encoded.



(a) (b) (c)

Figure 5: Rate-distortion performance of coding of the generalized Gaussian pdf with shape parameter (a) $\gamma = 0.7$, (b) $\gamma = 1$ (Laplacian), (c) $\gamma = 2$ (Gaussian), where the dotted line is the Shannon Lower Bound (SLB) of the source, the '+' symbols denote experimental points, and the solid line indicates the approximating rate-distortion function with $\beta = 2.0160, 1.9248, 1.8694$, respectively.

For both fractal search and fractal efficiency evaluation, we use a rate-distortion (R-D) criterion. Suppose without fractal prediction, the range block $\mathbf{r_j}$ needs $R_b(j)$ bits in wavelet coding, and with fractal prediction of overhead bits $R_o(j)$, it only needs $R_a(j)$ bits to wavelet encode the prediction residue, the rate saving of fractal can be calculated as:

$$R_s(j) = R_b(j) - R_a(j) - R_o(j), \tag{6}$$

Apparently, the optimum range-domain match is the one that offers the maximum rate saving, and only for those range blocks with rate savings greater than 0 should the fractal prediction be adopted. As a consequence, we use an adaptive wavelet-fractal prediction, and record the fractal status $s(j)$ of range block $\mathbf{r_j}$ as

$$s(j) = \begin{cases} 1, & R_s(j) > 0, \\ 0, & R_s(j) \leq 0 \end{cases} \tag{7}$$

All fractal statuses $s(j)$ are laid according to the position of the range block, and form a small binary image. This *fractal status image* is encoded using an arithmetic coder with a 2 bit context consists of the fractal statuses of its left and up block. For the range block which uses wavelet-fractal prediction, the descriptions of its contractive mapping are further encoded which include the position of the domain block (d_x, d_y), the type of the isometry operator Γ_j, and the quantized scaling factor α_j.

We also use a prescreening process which directly assigns range blocks with $R_b(j) < R_o(j)$ as fractal inefficient or $s(j) = 0$, for its rate saving $R_s(j)$ will be negative no matter what $R_a(j)$ is. This prescreening process speeds up the wavelet-fractal search by about 2-4 times at the middle coding rate.

In this research, we choose the layered zero coder(LZC) [15] developed by Taubman and Zakhor to encode the prediction residue. We have made small modifications of LZC to fit it for WFC. The essence of LZC is that the wavelet coefficients are quantized successively into layers, so that the quantization outputs for every layer are simply binary '0's or '1's which can be compressed by the context adaptive arithmetic coder [5] very efficiently. A *terminal significant threshold* T, which is the quantization step size of the final layer, controls the coding quality of LZC. Details of LZC can be found in [15] and [9]. We select LZC for its outstanding R-D performance and its simplicity in implementation. However, we by no means restrict the residue coder to be LZC, other wavelet coding approaches can be used as well.

4 MODEL BASED RATE-DISTORTION CRITERION

A key component of WFC is the estimation of the perspective coding rate for a particular region in the wavelet domain with and without fractal prediction. A model based rate-distortion (R-D) criterion is presented in this section to achieve this task.

Image Coder	Exp 1		Exp 2		Exp 3		Exp 4		Exp 5	
	Rate(bpp)	PSNR(dB)	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR
JPEG	-	-	-	-	0.1855	28.61	0.3785	33.34	0.7643	36.68
FRAC	-	-	-	-	0.2175	30.71	0.4477	33.40	0.7626	35.92
PPC	-	-	0.18	31.2	0.26	32.78	0.37	34.0	-	-
EZW	0.0359	25.79	0.0825	28.66	0.1816	31.76	0.3694	34.93	0.7574	38.11
LZC	0.0359	26.33	0.0825	29.27	0.1816	32.51	0.3694	35.60	0.7574	38.63
WFC	0.0359	26.42	0.0825	29.41	0.1816	32.68	0.3694	35.84	0.7574	39.02

Table 2: Performance comparison for the Lena image.

The criterion is based on the fact that the R-D performance of LZC on generalized Gaussian source can be closely approximated by

$$D = \sigma^2 2^{-\beta R}, \tag{8}$$

where β is called the coding efficiency parameter. In Fig. 5, we verify (8) by plotting the experimental R-D performance of the layered zero coder (LZC) versus the Shannon lower bound (SLB) for the generalized Gaussian source with the following pdf:

$$P(x) = ae^{-[b|x|]^\gamma}, \qquad a = \frac{b \cdot \gamma}{2\Gamma(1/\gamma)} \qquad b = \sigma^{-1}\left[\frac{\Gamma(3/\gamma)}{\Gamma(1/\gamma)}\right]^{1/2}. \tag{9}$$

where γ is a shape parameter and σ^2 is the variance of the source. A generalized Gaussian pdf is completely determined by its shape parameter γ and variance σ^2. Coefficients in a wavelet subband are often modeled as Laplacian pdf, which is in fact a specific case of the generalized Gaussian pdf with the shape parameter $\gamma = 1$. It is worthwhile to point out that the coding efficiency parameter β is independent of the variance of the source. Suppose that the variance of the source is increased by a factor of k^2, i.e. $\sigma_i^{'2} = k^2\sigma_i^2$. If we increase the terminal significant threshold of LZC accordingly by a factor of k, i.e., $T^{'} = k \cdot T$, LZC will encode the new source with exactly the same coding rate $R_i^{'} = R_i$ but k^2 the distortion $D_i^{'} = k^2 D_i$. Consequently, we have

$$D_i^{'} = \sigma_i^{'2} 2^{-\beta R_i^{'}}. \tag{10}$$

Clearly, the coding efficiency parameter is unaffected by the change of the source variance. Although the shape of the source pdf does affect the coding efficiency, the influence is small. Empirically, as we change the shape parameter of the generalized Gaussian source from 0.7 to 2.0, β only changes from 1.8694 to 2.0160. What is more, although the variances of different subbands vary greatly, the pdf shapes are usually stable and close to Laplacian, therefore the R-D efficiency parameter β can be approximated by a constant for all wavelet subbands.

In LZC, coding is controlled by selecting the terminal significant threshold T. LZC spends very few bits for coefficients with variances smaller than $T^2/12$, and encodes other coefficients with quantization error $T^2/12$. Note that such strategy achieves approximately the optimal bit allocation [11]. The *threshold distortion* $D_t = T^2/12$ controls the LZC coding, and is very close to the coding mean squared error D_{MSE}. Suppose that the target image quality is $PSNR$. We can set the codec control parameter of LZC, i.e. the threshold distortion D_t, as

$$D_t \approx D_{MSE} = 255^2 \cdot 10^{-PSNR/10}, \qquad \text{and set} \qquad T = \sqrt{12 D_t} = \sqrt{12} \cdot 255 \cdot 10^{-PSNR/20}.$$

We can then estimate the perspective coding rate of a range block. As a range block is composed of several subblocks of different scales and directional subbands, it can be viewed as a compound of N i.i.d. sources with sizes S_1, S_2, $\cdots$, S_N and variances σ_1^2, σ_2^2, $\cdots$, σ_N^2. Suppose the variances of subblocks of a range block $\mathbf{r_j}$ without fractal prediction are denoted by $\sigma_{b,1}^2$, $\sigma_{b,2}^2$, $\cdots$, $\sigma_{b,N}^2$, and they decrease to $\sigma_{a,1}^2$, $\sigma_{a,2}^2$, $\cdots$, $\sigma_{a,N}^2$ with fractal prediction. We can calculate the perspective coding rate with and without wavelet-fractal prediction, respectively, as

$$R_b(j) = \frac{1}{\beta}\sum_{i=1}^{N} S_i \log_2 \frac{\max\{\sigma_{b,i}^2, D_t\}}{D_t}, \qquad R_a(j) = \frac{1}{\beta}\sum_{i=1}^{N} S_i \log_2 \frac{\max\{\sigma_{a,i}^2, D_t\}}{D_t}. \tag{11}$$

Image Coder	Exp 1		Exp 2		Exp 3		Exp 4		Exp 5	
	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR	Rate	PSNR
	Image Barbara									
LZC	0.0353	22.84	0.1335	25.24	0.3351	29.46	0.6679	33.98	1.1761	38.55
WFC	0.0353	22.88	0.1335	25.82	0.3351	29.90	0.6679	34.19	1.1761	38.77
	Image Lighthouse									
LZC	0.0508	22.74	0.1457	25.79	0.3322	29.59	0.6743	34.02	1.2163	38.53
WFC	0.0508	23.19	0.1457	26.40	0.3322	30.18	0.6743	34.26	1.2163	38.76
	Image Town									
LZC	0.0702	19.44	0.2040	22.36	0.5136	26.15	1.0211	30.91	1.7566	36.11
WFC	0.0702	19.56	0.2040	22.53	0.5136	26.56	1.0211	31.17	1.7566	36.25
	Image Wood									
LZC	0.0747	14.86	0.2173	17.99	0.5775	21.55	1.1882	26.79	2.1581	32.15
WFC	0.0747	15.19	0.2173	18.09	0.5775	22.30	1.1882	26.92	2.1581	32.20
	Image Baboon									
LZC	0.0168	19.57	0.0865	21.05	0.3504	23.96	0.8935	27.86	1.7706	32.94
WFC	0.0168	19.55	0.0865	21.15	0.3504	24.24	0.8935	28.39	1.7706	33.46

Table 3: Performance comparison for test images: Barbara, Baboon, Wood, Town and Lighthouse.

The rate saving of fractal can be calculated via

$$R_s(j) = R_b(j) - R_a(j) - R_o(j) = \frac{1}{\beta}\sum_{i=1}^{N} S_i \log_2 \frac{\max\{\sigma_{b,i}^2, D_t\}}{\max\{\sigma_{a,i}^2, D_t\}} - R_o(j). \tag{12}$$

Note that the rate saving $R_s(j)$ is insensitive to the change of the threshold distortion D_t. In particular, if D_t is smaller than variances of all subblocks, $R_s(j)$ is independent of the value of D_t. The computation of rate saving with (12) is simple and straightforward. The primary computational cost lies in the estimation of variances $\sigma_{b,i}^2$ and $\sigma_{a,i}^2$. We may implement operator $\frac{1}{\beta}\log_2 x$ as a lookup table for fast hardware implementation,

5 EXPERIMENTAL RESULTS

In this section, we compare the performance of the proposed wavelet-fractal coder (WFC) with several typical fractal and wavelet coders. We adopt the fractal parameter set which maximizes the coding rate saving. The search range covers all the image with step size 1. The number of isometry operators is 8, and the quantization precision for scaling factor Q_α is 2^{-4}. In comparison with traditional fractal coders in [8], [7] which adopt a search step size of 8 or 4, we choose a more refined step size in WFC. This is due to the fact that fractal prediction is only used for fractal efficient regions in WFC so that there is no need to compromise the fractal inefficient regions. The comparison coders include: the fractal coder (FRAC) of Jacobs [7], the predictive pyramid coder (PPC) of Rinaldo [13], the embedded zerotree wavelet coder (EZW) of Shapiro [14] and the layered zero coder (LZC) of Taubman and Zakhor [15]. We also show results of JPEG as an another reference. Test images used in the experiment are Lena, Barbara, Baboon, Wood, Town and Lighthouse of size 512×512. Experimental results for FRAC and PPC coders are directly taken from Table. 2 in [7] and Fig. 8 in [13], respectively.

Experimental results are shown in Table. 2 and the comparing R-D curves are plotted in Fig. 6 for test image Lena. The performance of FRAC [7] is not good even compared with JPEG. Although it outperforms JPEG at low bit rates, it is inferior to JPEG at the middle to high bit rate range. PPC is substantially better than JPEG, but it can hardly compete with the state-of-the-art wavelet coders such as EZW or LZC. WFC demonstrates a superior performance in comparison with all other coders. It outperforms LZC by 0.1-0.4dB, EZW by 0.6-0.9dB, PPC by 1.5-2.0dB, JPEG by 2.3-4.0dB and FRAC by 2.0-3.1dB. Several Lena images are given in Fig. 7, where (a) is the original image, (b) and (c) are results of JPEG at 0.3785bpp and WFC at 0.3694bpp, respectively. WFC-coded

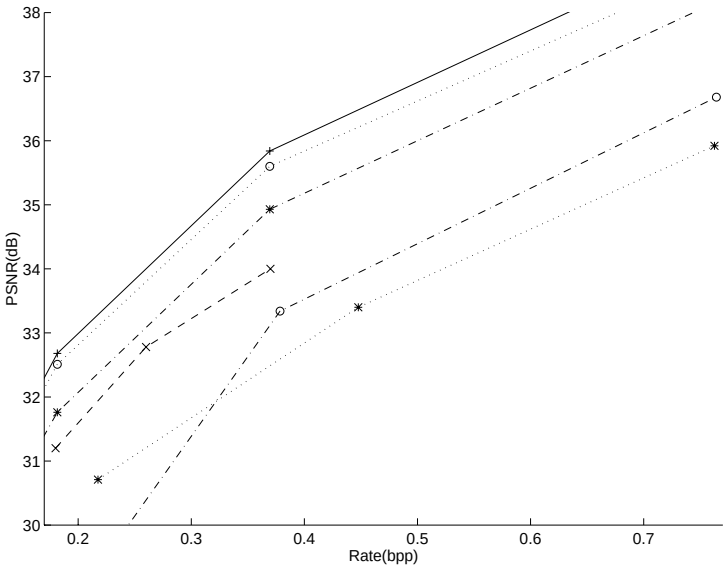


Figure 6: Rate-distortion performances for Lena using the (a) Fractal wavelet coder (FWC, the solid line with '+'), (b) Layered zero coder (LZC, the dotted line with 'o'), (c) Embedded zerotree wavelet coder (EZW, the dash dotted line with '*'), (d) Predictive pyramid coder (PPC, the dashed line with 'X'), (e) Conventional fractal coder (FRAC, the dotted line with '*'), (f) JPEG (the dash dotted line with 'o').

Lena looks much better than the JPEG-coded one subjectively. The difference in subjective appearance of WFC- and LZC-coded Lena is more subtle. Generally speaking, WFC-coded Lena appears better in the texture areas. Subjective qualities of the two coders are about the same in the edge and smooth regions. This can be partially explained by the fact that most fractal-efficient range blocks are in texture regions or around sharp edges. Range blocks using fractal prediction are marked with $\oplus$ in Fig. 7(d). At that particular bit rate, about 8% of the entire image is applied with fractal prediction.

In the next experiment, we perform a more thorough comparison of WFC and LZC, which is the wavelet residue coder of WFC. Our objective is to investigate the performance gain of fractal prediction. The experiments are performed on images Barbara, Baboon, Wood, Town and Lighthouse. The experimental results are shown in Table. 3. In general, the gain of WFC over LZC is in the range of 0.0-0.8dB. The actual gain depends on the characteristics of image and the operating bit rate. The gain is more for images with a lot of textured patterns and encoded at a higher bit rate. Note that fractal prediction is used adaptively in WFC. For the extreme case that all range blocks are not efficient for fractal prediction, we only lose bits for the coding of a blank prediction status image. It is also worthwhile to point out that there is only one case among 30 comparisons (i.e. Baboon image coded at 0.0168bpp) that WFC is inferior to LZC by 0.02dB. For all other 29 cases, WFC performs better with the PSNR gain varying from 0.0 to 0.8dB. The better R-D performance is however achieved by a higher computational cost.

There is one additional interesting observation from the above experiments. As the contractive mapping provides a tool for interscale wavelet prediction, it is not efficient for all image regions. When integrated with the state-of-the-art wavelet coder (LZC), approximately 8% of the entire image region in Lena benefits from the use of fractal prediction. Fractal efficient regions are usually the ones that consume more coding bits in typical wavelet coders. For example, LZC devotes around 17% of the total coding bits to the 8% fractal efficient region in Lena. Clearly, the incorporation of an additional fractal prediction unit in the wavelet coder does provide a better R-D performance.

6 CONCLUSIONS AND EXTENSIONS

In this research, the fractal coder is integrated with the wavelet coder and constructs a wavelet-fractal coder (WFC) with a superior R-D performance. Fractal prediction used in WFC includes: the contractive operator that provides prediction across wavelet scales, the scaling operator that shrinks prediction coefficients, and the isometry

operator which shuffles pixels within and across the subbands. As in DPCM or motion compensated video coding, the residue of fractal prediction is further encoded by the state-of-the-art wavelet coder. A rule of low complexity was developed to estimate the overhead information required and the bit rate saving achieved by using fractal prediction so that it is only applied to fractal efficient regions. This rule optimizes the R-D performance and depends only on the variance of the source and the codec control parameter, i.e. the threshold distortion. Wavelet coefficients in other regions are encoded directly with successive quantization and context-based adaptive arithmetic coding. The R-D criterion can also be used to determine the optimal fractal parameter set which maximizes the bit rate saving. WFC outperforms JPEG, typical fractal and wavelet coders by 2.3-4.0dB, 1.5-3.1dB and 0.1-0.9dB, respectively. The result clearly demonstrates that the incorporation of fractal prediction provides a substantial performance improvement for the wavelet coder.

Currently, WFC uses a constant range-domain block size K. Further performance improvement is expected by varying the block size in fractal prediction. Another shortcoming of WFC is that it is computationally intensive in the encoding part. For practical applications, it is important to reduce the complexity of fractal search. This can be achieved by either reducing the size of the domain pool, or adopting fast fractal search schemes [3] which greatly reduce the search complexity with little additional degradation. It is worthwhile to perform a thorough study on the tradeoff of complexity reduction and the degradation of the R-D performance.

7 REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, Vol. 1, pp. 205–230, Apr. 1992.

[2] G. Caso and C.-C. J. Kuo, "New results for fractal/ wavelet image compression," in *SPIE Visual Communications and Image Processing '96*, vol. 2727, (Orlando, FL), pp. 536–547, Mar. 1996.

[3] G. Caso, P. Obrador, and C.-C. J. Kuo, "Fast methods for fractal image encoding," in *SPIE Visual Communications and Image Processing '95*, vol. 2501, (Taiwan), pp. 583–594, 1995.

[4] G. Davis, "Adaptive self-quantization of wavelet subtrees: a wavelet-based theory of fractal image compression," in *Wavelet Applications in Signal and Image Processing III, SPIE Vol. 2569* (A. F. Laine, M. A. Unser, and M. V. Wickerhauser, eds.), (San Diego, CA), pp. 294–306, July 1995.

[5] D. L. Duttweiler and C. Chamzas, "Probability estimation in arithmetic and adaptive huffman entropy coders," *IEEE Trans. on Image Processing*, Vol. 4, pp. 237–246, Mar. 1995.

[6] Y. Fisher (ed.), *Fractal Image Compression: Theory and Applications*, Springer-Verlag, New York, 1995.

[7] E. W. Jacobs, Y. Fisher, and R. D. Boss, "Image compression: a study of the iterated transform method," *Signal Processing*, Vol. 29, pp. 251–263, 1992.

[8] A. E. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Trans. on Image Processing*, Vol. 1, No. 1, pp. 18–30, Jan. 1992.

[9] J. Li, P.-Y. Cheng, and C.-C. J. Kuo, "On the Improvements of Embedded Zerotree Wavelet (EZW) Coding," in *SPIE: Visual Communication and Image Processing'95*, vol. 2501, (Taipei, Taiwan), pp. 1490–1501, May 1995.

[10] J. Li, P.-Y. Cheng, and C.-C. J. Kuo, "A Wavelet Transform Approach to Video Compression," in *SPIE: Wavelet Applications II*, vol. 2491, (Orlando, FL), pp. 1107–1118, Apr. 17-21 1995.

[11] J. Li, P.-Y. Cheng, and C.-C. J. Kuo, "Embedded wavelet packet image coder With fast rate-distortion optimized decomposition," in *SPIE: Visual Communication and Image Processing'97*, vol. 3204, (San Jose, CA), Feb. 1997.

[12] A. P. Pentland, "Fractal-Based Description of Natural Scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 661–674, Nov. 1984.

[13] R. Rinaldo and G. Calvagno, "Image coding by block prediction of multiresolution subimages," *IEEE Trans. on Image Processing*, Vol. 4, pp. 909–920, July 1995.

[14] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing*, Vol. 41, No. 12, pp. 3445–3462, Dec. 1993.

[15] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. on Image Processing*, Vol. 3, No. 5, pp. 572–588, Sep. 1994.

(a) (b)

(c) (d)

Figure 7: The Lena experimental results. We show (a) the original Lena, (b) the JPEG coded Lena at 0.3785bpp, 33.34dB, (c) the wavelet fractal (WFC) coded Lena at 0.3694bpp, 35.84dB, (d) the fractal efficient regions (marked with $\oplus$).