To appear in *Proc. International Conference on Image Processing*, ICIP-99, Oct. 25–28, 1999, Kobe, Japan

REGION-BASED IMAGE COMPRESSION USING FRACTALS AND SHAPE-ADAPTIVE DCT

Kamel Belloulata[†], *Ryszard Stasiński*[‡] *and Janusz Konrad*[†]

[†] INRS-Télécommunications, 16 Place du Commerce, Verdun, QC, H3E 1H6, Canada
[‡] Høgskolen i Narvik, Lodve Langes gt. 2, N-8501 Narvik, Norway

ABSTRACT

The significant effort to provide region-based compression and functionality within the MPEG-4 standard is not paralleled in the still-image compression domain. In this paper, we propose an approach to fractal coding of still images that is truly region-based. Unlike previous fractal compression methods the proposed approach compresses an image region-by-region based on a prior segmentation, very much like in MPEG-4; individual regions can be decoded without full image decoding. The method performs the domain/range block matching in frequency domain using a shape-adaptive discrete cosine transform. Experimental results evaluating the performance of the approach are shown.

1. INTRODUCTION

Region-based image coding has two advantages over standard block-based coding: it allows manipulation of image objects without complete decoding of the bit stream and it often improves the quality of coded images. In a region-based coding scheme, a segmentation of the image into regions (*alpha planes* in MPEG-4 [1]) is known before encoding and is transmitted along with the object texture. This segmentation is exploited during encoding to insure that an independent decoding/manipulation of individual objects can be performed at the receiver. It is this type of functionality that we seek to achieve with a fractal-based still image coder.

It has been known for quite some time that a fractal image coder performs better, in terms of compression ratio and/or image quality, for variable-size than for fixed-size blocks [2]. One possibility is to use a quadtree image partitioning [3] where initially large blocks are progressively sub-divided by thresholding the coding error; range and domain blocks remain rectangular. This approach is outperformed by methods allowing non-rectangular range and domain blocks [4, 5]. For example, block merging can be applied to the initial rectangular blocks to arrive at an arbitrarily-shaped (AS) partitioning of the image into AS range blocks [6]. This method outperforms quadtree-based fractal coding [3] at the cost of significantly increased computational complexity. Note that both methods create an image partitioning during the encoding and thus cannot handle a prior image partitioning (alpha planes) as is done in MPEG-4. In this sense, the methods are texture-adaptive rather than region-based.

In order to further improve the coding performance, hybrid schemes using fractal coding of discrete cosine transform (DCT) coefficients [7, 8] or wavelet coefficients [9] have been proposed. In the former case, first all range and codebook (domain) blocks are transformed using DCT. Then, for each range block the corresponding codebook block is found in the DCT coefficient domain by scaling and adding an offset. To reduce the total number of transformations and to adapt block size to texture variance, hierarchical coding schemes using variable size of range blocks in the DCT domain have been proposed [8].

In this paper, we propose a *true* region-based fractal coding scheme; objects can be defined by a prior segmentation (alpha planes) and coded independently of each other by fractals applied to DCT coefficients. In the proposed method, the range blocks remain rectangular. Since some of them are located at object boundaries, thus encompassing pixels from two or more objects, in order not to mix such pixels we propose to use the shape-adaptive DCT (SA-DCT) algorithm [10] instead of the standard DCT. For example, for a two-object boundary block two separate SA-DCT transforms are performed, while for a non-boundary block (fully enclosed in an object), standard DCT is used. The same procedure is applied to the domain blocks. Certain modifications to reduce computational complexity are proposed as well.

2. CURRENT FRACTAL/DCT CODING

Let I be the image to be coded, $\mathcal{R} = \{r_1, \ldots, r_N\}$ be the set of N non-overlapping range blocks partitioning I, and $\mathcal{D} = \{d_1, \ldots, d_M\}$ be the collection of M over-

lapping domain blocks forming another partition of I. For each r_i we have to find a domain block d_j and a transformation w_i that jointly minimize the distance $\varepsilon(r_i, w_i(d_j))$, where $\varepsilon(.,.)$ is a measure of distortion or *dissimilarity* and w_i is a local affine transformation.

First, all range and codebook blocks are transformed *via* DCT to arrive at $R_i = DCT(r_i)$ and $D_j = DCT(d_j)$. Let (u, v) be 2-D frequency of a DCT coefficient and let ϕ be a contractivity operator that maps a 16×16 domain block $D_j(u, v)$ onto an 8×8 range block $R_i(u, v)$, while retaining only the low-frequency part of the spectrum of D_j [7]. The local affine transformation w_i is usually expressed using the difference in range/domain block positions (i and j), the contractivity operator ϕ, a scaling factor s_i, one of isometries $\{l_i^n\}_{n=1}^8$ and an offset o_i, all applied in the frequency domain. Then, $w_i(D_j(u, v))$ can be also written as $s_i(l_i^n(\phi(D_j(u, v))))+ o_i$. The eight isometries $\{l_i^n\}_{n=1}^8$ are similar to the ones used typically in the spatial case (mirror reflections, 90^o rotations, etc.) but adapted to the DCT coefficient domain [11].

The domain block index j, the scaling factor s_i and the isometry l_i^n are found by minimizing the mean-squared error between R_i and $w_i(D_j)$. However, in order to precisely represent the mean local intensity, we do not apply the offset o_i to $w_i(D_j)$ but instead we directly transmit the DC value of the range block r_i:

$$\widehat{R}_i(u, v) = \begin{cases} R_i(0, 0) & \text{if } u = v = 0, \\ s_i(l_i^n(\phi(D_j(u, v))) & \text{otherwise,} \end{cases} \quad (1)$$

where $\widehat{R}$ denotes the range block information to be transmitted.

To encode the block r_i, an exhaustive search for index j (codebook block d_j) and for an isometry l_i^n is performed in the frequency domain that, jointly with s_i, minimize $\varepsilon(r_i, w_i(d_j))$. For each combination of j and l_i^n, the scaling factor is computed as follows:

$$s_i = \frac{\sum_{u,v} l_i^n(\phi(D_j(u, v)))R_i(u, v)}{\sum_{u,v}[l_i^n(\phi(D_j(u, v)))]^2}. \quad (2)$$

3. NEW FRACTAL/SA-DCT CODING

In the fractal/DCT scheme described above the boundary range blocks contain pixels from an object *and* a background (or another object). Thus, independent decoding of objects is not possible and the coding quality may suffer since pixels in a boundary block may have very different characteristics if located on different sides of the boundary. By applying the standard DCT to such a block, spectral properties of these pixels are mixed up making the search for a good range-domain correspondence unreliable.

To alleviate both difficulties, we propose to apply the shape-adaptive DCT [10] to each segment $\mathcal{S}$ of the boundary range and domain blocks. As shown in Fig. 1, the basic concept of the SA-DCT is to perform vertical DCTs of the active pixels[1] followed by horizontal DCTs of the vertical DCT coefficients with the same frequency index. The final coefficients are located in the upper-left corner of each block. The number of the SA-DCT coefficients is identical to the number of active pixels. Since the segment shape is transmitted, the decoder can perform the inverse SA-DCT. In simulations, we used a slightly modified SA-DCT [12].

The non-boundary blocks are treated as described in Section 2. However, to assure object-based functionalities, we limit the search for domain blocks to the object from which the processed range blocks come. The approach is illustrated schematically in Fig. 2. Since the boundary blocks, both range and domain, consist of two or more segments, we process each segment independently. First, we apply the SA-DCT to each segment (Fig. 3) and then we find all parameters of the transformation w_i for each segment of the block. In principle, each segment of a boundary range block should be matched against all domain blocks (boundary and non-boundary) of an object. However, note that the basis functions and the corresponding coefficients of a segment $\mathcal{S}$ (Fig. 1) undergoing SA-DCT are strongly dependent on segment shape. Thus, they are, in general, quite different from basis functions (coefficients) of a full domain block. Consequently, even if same-index coefficients are similar ($R_i(u, v) \approx D_j(u, v)$), their corresponding basis need not. Since little performance gain can be expected from a full-object search (confirmed experimentally to be around 0.1dB), we propose to match each boundary (non-boundary) range block only against boundary (non-boundary) domain blocks with the additional benefit of reduced computational complexity.

[1] Active pixels are those that belong to the segment of interest ($\mathcal{S}$) and passive (undefined) are those that do not (are in $\mathcal{U}$).

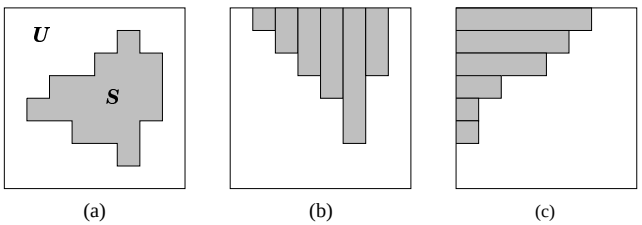


Figure 1: Illustration of SA-DCT: (a) arbitrarily-shaped region; (b) vertical alignment followed by vertical 1-D DCTs; (c) horizontal alignment followed by horizontal 1-D DCTs.

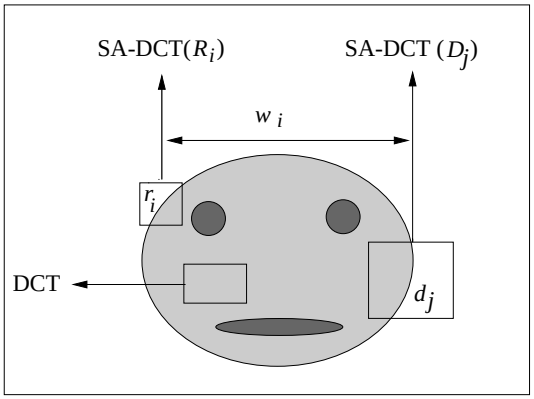


Figure 2: Schematic illustration of the proposed fractal/SA-DCT coding scheme.

To find the best domain block d_j for a given range block r_i, we used the following distortion measure:

$$\varepsilon = \sum_{(u,v)\in\mathcal{P}_k} (R_i^k(u,v) - s_i(l_i^n(\phi(D_j(u,v)))))^2,$$

$\mathcal{P}_k$ (Fig. 3) is an SA-DCT-transformed segment $\mathcal{S}_k$ (k-th segment of boundary range block r_i) and $R_i^k(u,v)$ is an SA-DCT coefficient in $\mathcal{P}_k$. The domain coefficients outside each domain-block segment are set zero. Clearly, contributions to ε are only made at indices within range-block segment $\mathcal{P}_k$. If at (u,v) a coefficient in $\mathcal{P}_k$ exists while it does not exist in the segment of the domain block under consideration, then a significant contribution to ε is made. In consequence, smaller distortions will result if either range and domain block segments have similar shape or if coefficients are small. Thus, good matches are found between similarly-shaped segments or segments with little spectral content, either an acceptable solution. In case a domain-block segment fully englobes $\mathcal{P}_k$, it is likely to be rejected if it significantly differs in shape/size from $\mathcal{P}_k$ since the same-index coefficients should be quite different.

Although initially we had implemented a search for j and l_i^n independently for each segment of a block, extensive testing has demonstrated only minimal performance gain at high computational complexity compared to forcing the same j and l_i^n for all segments of a range block. We opted for the latter solution in the final experiments.

4. EXPERIMENTAL RESULTS

We compared the performance of the proposed hybrid fractal/SA-DCT scheme with standard fractal coding in the spatial domain and with the hybrid fractal/DCT scheme (Section 2). We used identical coding parameters in all experiments, including the same quantizer and Huffman coding of the transformation parameters.

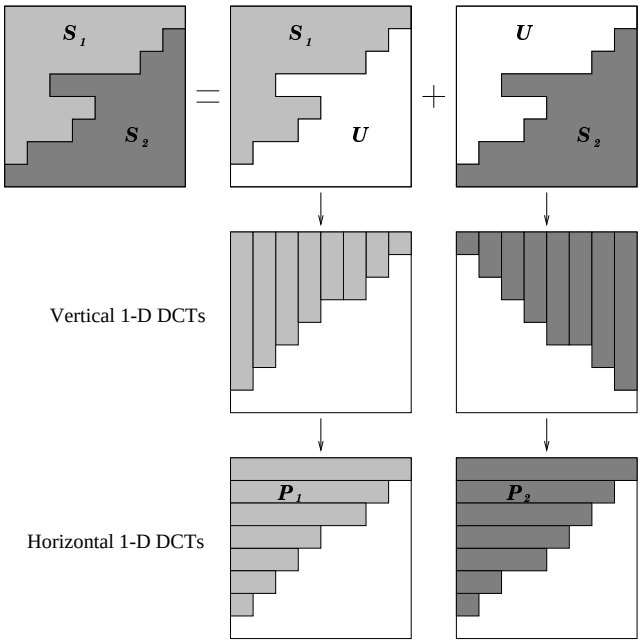
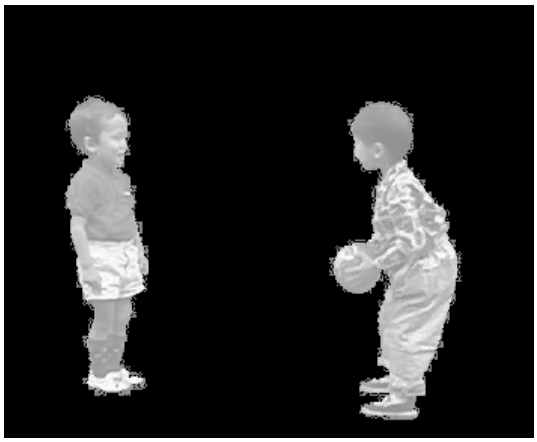


Figure 3: Example of application of the SA-DCT to a boundary block with two segments.

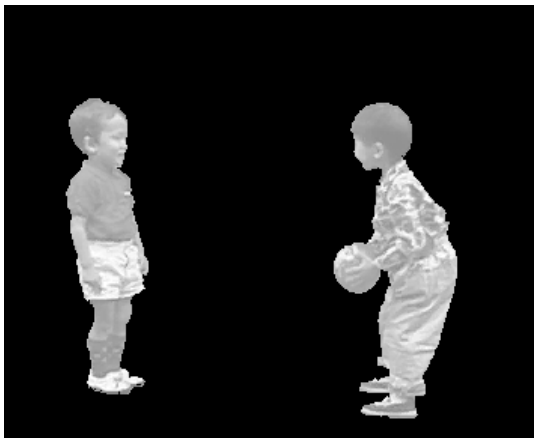
As the test images we used frames from MPEG-4 test sequences *Children* (CIF) and *Cyclamen* (SIF) with the associated segmentations. To perform the algorithm comparison in a likely situation of region manipulation by a user, in addition to encoding the original images we also encoded images with the background replaced by zero intensity (black). This scenario emulates a "cut-and-paste" operation permitted by this coding approach.

Table 1 shows the performance and the CPU time (166MHz Ultra-1 Sparc station) for the three schemes in the case of encoding the main object against black background. The reconstructed images are shown in Fig. 4. All boundary and non-boundary blocks were used to calculate the bit rate and PSNR of the luminance component. The coding cost for shape information was not included in the rate calculation but it has marginal impact on the performance [10]. For both images, the hybrid fractal/SA-DCT scheme outperforms the other schemes and that with the lowest computational effort. The performance gain for *Cyclamen* is larger due to the complex (long) boundary of the coded object where the SA-DCT excels. The improvement is also subjective as can be seen in Fig. 4. The zero-background case is handled well by the fractal/SA-DCT scheme since all segments of a block are processed independently, unlike in the other two approaches.

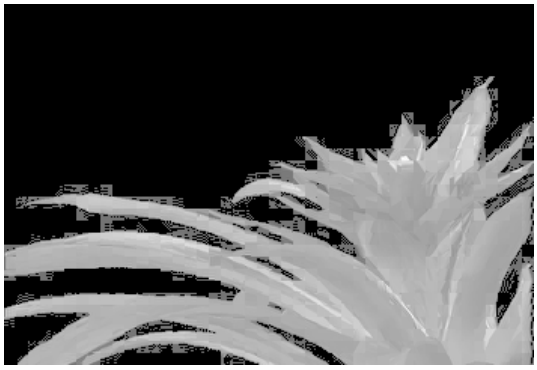
Table 2 and Fig. 5 show the results when encoding



(a) Fractal/DCT (34.08 dB at 0.47 bpp)



(b) Fractal/SA-DCT (35.25 dB at 0.49 bpp)



(c) Fractal/DCT (28.00 dB at 0.31 bpp)

(d) Fractal/SA-DCT (32.75 dB at 0.29 bpp)

Figure 4: Iteratively reconstructed foreground objects with background set to zero (black).

Table 1: Coding results for zero background (black)

	Children		
Fractal/Spatial	33.90 dB	0.45 bpp	213 s
Fractal/DCT	34.08 dB	0.47 bpp	225 s
Fractal/SA-DCT	35.25 dB	0.49 bpp	114 s
	Cyclamen		
Fractal/Spatial	28.12 dB	0.27 bpp	272 s
Fractal/DCT	28.00 dB	0.31 bpp	298 s
Fractal/SA-DCT	32.75 dB	0.29 bpp	140 s

Table 2: Coding results for full image

	Children		
Fractal/Spatial	27.08 dB	0.52 bpp	1185 s
Fractal/DCT	27.16 dB	0.55 bpp	1214 s
Fractal/SA-DCT	26.58 dB	0.56 bpp	957 s
	Cyclamen		
Fractal/Spatial	31.40 dB	0.53 bpp	821 s
Fractal/DCT	31.51 dB	0.53 bpp	920 s
Fractal/SA-DCT	30.95 dB	0.62 bpp	411 s

the full images. Note the slightly worse performance of the fractal/SA-DCT scheme (about 0.5dB PSNR penalty with respect to the other approaches), however again a significant gain in the computational efficiency. In the case of encoding full images, the proposed scheme suffers as the search for the best domain block is limited to an object unlike in the standard schemes that perform a full-image search. This, however, is counterbalanced by the computational efficiency and the ability to independently transmit and decode each object of an image.

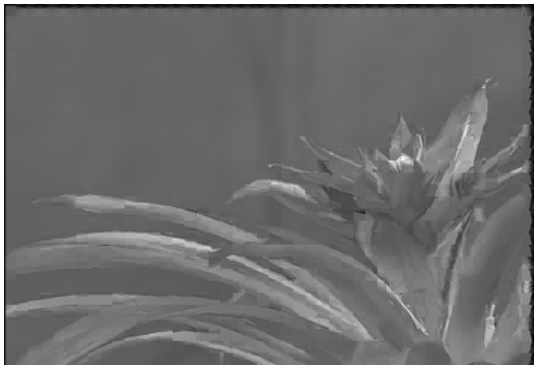
5. REFERENCES

[1] Special issue on MPEG-4, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, Feb. 1997.

[2] A. Jacquin, "Image coding based on a fractal theory of iterated contractive image transforma-

(a) Fractal/DCT (27.16 dB at 0.55 bpp)

(b) Fractal/SA-DCT (26.58 dB at 0.56 bpp)



(c) Fractal/DCT (31.51 dB at 0.53 bpp)

(d) Fractal/SA-DCT (30.95 dB at 0.62 bpp)

Figure 5: Iteratively reconstructed full images.

tions," *IEEE Trans. Image Process.*, vol. 1, pp. 18–30, Jan. 1992.

[3] Y. Fisher, "Fractal encoding with quadtrees," in *Fractal Image Compression: Theory and Applications to Digital Images* (Y. Fisher, ed.), pp. 55–77, Springer-Verlag, 1995.

[4] L. Thomas and F. Deravi, "Region-based fractal image compression using heuristic search," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 823–838, 1995.

[5] D. Saupe, M. Ruhl, R. Hamzaoui, L. Grandi, and D. Martini, "Optimal hierarchical partitions for fractal image compression," in *Proc. IEEE Int. Conf. Image Processing*, vol. III, pp. 737–741, Oct. 1998.

[6] M. Breazu and G. Toderean, "Region-based fractal image compression using deterministic search," in *Proc. IEEE Int. Conf. Image Processing*, vol. III, pp. 742–746, Oct. 1998.

[7] Y. Zhao and B. Yuan, "A hybrid image compression scheme combining block-based fractal coding and dct," *Signal Process., Image Commun.*, vol. 8, pp. 73–78, Mar. 1996.

[8] K. Barthel, J. Schuttemeyer, T. Voye, and P. Noll, "A new image coding technique unifying fractal and transform coding," in *Proc. IEEE Int. Conf. Image Processing*, vol. III, pp. 112–116, Oct. 1994.

[9] K. Belloulata, A. Baskurt, H. Benoit-Cattin, and R. Prost, "Fractal coding of subbands with an oriented partition," *Signal Process., Image Commun.*, vol. 12, pp. 243–252, June 1998.

[10] T. Sikora and B. Makai, "Shape-adaptive DCT for generic coding of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 59–62, Feb. 1995.

[11] R. Bracewell, K. Chang, A. Wang, and Y. Wang, "Affine theory for two-dimensional Fourier transform," *Electron. Lett.*, vol. 29, no. 3, p. 304, 1993.

[12] P. Kauff and K. Schüür, "An extension of shape-adaptive DCT (SA-DCT) towards DC separation and ΔDC correction," in *1997 Picture Coding Symposium*, pp. 647–652, Sept. 1997.