Cite as: Harvey A. Cohen, *Fractal Image Coding for Thumbnail-based image Access* Proc. Int'l Conf Signal Processing Applications, ISSPA, Gold Coast, Aug 26-28, 1996, Vol 1 pp 158-161.

Fractal Image Coding for Thumbnail Based Image Access

Harvey A. Cohen
Computer Science and Computer Engineering
La Trobe University
Melbourne Victoria Australia 3083
Email: H.Cohen@latrobe.edu.au
Fax: +61 3 9479 3060

EXTENDED DRAFT VERSION FOR ISSPA -96

ABSTRACT

A variant of block-oriented fractal coding termed thumb fractal is introduced that essentially is the Jacquin-Barnsley block-oriented fractal code for the (signed) difference image between the target image and the image thumbnail exploded to full-size. Thumb fractal code involves fewer block parameters than other fractal coding, as data from the image thumbnail is utilised in image generation. Using domain block to range block size ratio of 4 for thumb fractal code yields a super-fast converging codec, with initial image the exploded thumb. This codec is especially suitable for thumbnail-based image access, where image retrieval is preceded by thumbnail access and diaply.

In thumbnail-based retrieval from image databases the display of any fully featured (FF-) image at a workstation is commonly preceded by its selection from a display of reduced-size, thumbnail image (T- image). For Clip-Art browsing, thumbnail-based retrieval is now the norm. However, thumbnail-based retrieval has not been built into commercial net-work browsers because available image codecs were not designed for this purpose. Elsewhere[1] we have analysed the efficiency of thumbnail-based image retrieval, and pointed out that effective implementation of thumbnail-based retrieval requires that the data supplied for the T-image be efficiently utilised in the production of the FF-image.

In this paper we describe two algorithms for image coding that are thumbnail based, with the image thumbnail serving to provide correction tiles through the use of the fractal transform, as introduced by Jacquin,[2] [3], Barnsley[4], and Fisher[5]. The first of these coding algorithms, thumbail fractal coding, is related to proposals of Øien and Lepsøy [6] *while the other, called thumb VQFT (vector quantization via Fractal Transform) is an entirely novel variant of vector quantization. The data presented over a range of domain block interblocks, demonstrate the capability of these two thumb codecs, applied to single-leval image decomposition into 4x4 blocks. Using a fixed encoding interblock, for a single iteration, thumb VQFT is generally superior to thumb fractal, but on the second fractal iteration effective convergence is reached with PSNR superior to thumb VQFT coding. The fractal codec, which is very fast to decode, encodes essentially as efficiently as related fractal codecs, but will prove far more efficient for image based retrieval. However, these conclusions are clearly applicable to multi-level image blocking, such as quadtree.*

2. Thumbnail based VQ - Mathematical Analysis

The basic idea of thumbnail VQ is to use the zoomed-out thumbnail as a first approximation to the decoded image, and to use this image itself as a source of tiling blocks that are to supply corrections. The correction tiles can be determined either by sub-sampling or by averaging within the zoomed-out thumbnail, and subtracting from each pixel the block mean to yield a zero-sum tile, which may be rotated/reflected and/or contrast scaled. Note that this specification is compatible with multi-level partitioning schemes, as the familiar quadtree.

In this presentation, for clarity, we shall only detail thumbnail VQ where the image is partitioned only into blocks of the one size, 4x4. For gray-scale images, the thumbnail used is based on the mean-value of pixels within each such 4x4 block of the original image. The codebook IS precisely the zoomed-out thumbnail, or FT image, for which the 'natural' thumbnail has been expanded to be the same size as the original image.

Using the notation above, where the original image block is denoted by B[r][c], the first approximation to each 4x4 image block is the zoomed-out thumbnail,

$$b[r][s] \begin{pmatrix} 1111 \\ 1111 \\ 1111 \\ 1111 \end{pmatrix} \quad \text{where} \quad b[r][c] = \frac{1}{16} \sum_{i,j} B[r][c]_{i,j}$$

The second approximation aims to determine the 4x4 correction

$$C[r][c] = B[r][c] - b[r][s] \begin{pmatrix} 1111 \\ 1111 \\ 1111 \\ 1111 \end{pmatrix}$$

which is a block with zero sum:

$$\Sigma_{ij} C[r][c]_{ij} \ 0 \text{ all } r,C.$$

Cite as: Harvey A. Cohen, *Fractal Image Coding for Thumbnail-based image Access* Proc. Int'l Conf Signal Processing Applications, ISSPA, Gold Coast, Aug 26-28, 1996, Vol 1 pp 158-161.

This correction block is coded by searching through 16x16 blocks within the code-book, and compressing each to a 4x4 size which is called Z. If the mean pixel within Z is called $\overline{z}$, then the 4x4 block

$$Y(0) = Z - \overline{z}\begin{pmatrix}1111\\1111\\1111\\1111\end{pmatrix}$$

has zero mean and is a suitable tile, as are all tiles derived from Y(0) by any of the 8 eight transformations $T(\aleph)$ which are the product of a rotation by a multiple of 90 degrees, and a reflection about a diagonal.

$$Y(\aleph) = T(\aleph)\, Y(0)\, T^{-1}(\lrcorner) \qquad (\lrcorner=0, \dots, 7)$$

Any linear multiple of the eight $Y(\aleph)$ is a potential tile. Hence the coding problem becomes the determination of the $Y(\aleph)$ so as to minimise

$$\| C[r][c] - aY(\aleph) \|^2$$

For a fixed C[r][c] and $Y(\aleph)$, the minimisation of this error is precisely the determination of the slope of the line of best fit (least-squares) of elementary regression theory, so that

$$a = \frac{16\Sigma\, C_{ij} Y(\lrcorner)_{ij}}{\Sigma Y_{ij}^{2}}$$

where the summations are over the 16 pixels in the 4x4 blocks. (The index $\aleph$ is superfluous in the denominator). Hence choosing a in accord with this formula the block error for $Y(\aleph)$ is after elementary algebra reduced to

$$\Sigma\, C_{ij}^{2} - \frac{16\left(\Sigma\, C_{ij} Y(\lrcorner)_{ij}\right)^2}{\Sigma Y_{ij}^{2}}$$

Encoding of the block B[r][c] thus is a two stage process

(a) Computing the block mean b[r][c] to be used in the thumbnail

(b) Determing for the position (I,J) within the exploded thumbnail, and the transformations $\aleph$ of the minimum of

$$1 - \frac{16\left(\Sigma\, C_{ij} Y(\lrcorner)_{ij}\right)^2}{\Sigma\, C_{ij}^{2}\, \Sigma Y_{ij}^{2}}$$

(the index $\aleph$ is deliberately omitted from the denominator). I,J together with $\aleph$ and the contrast a given by

$$a = \frac{16\Sigma\, C_{ij} Y(\lrcorner)_{ij}}{\Sigma Y_{ij}^{2}}$$

give the VQ code for the image block rc,

3. Thumbnail based Fractal coding

The basic idea of thumbnail fractal coding is to apply the methods of fractal coding to the signed image that is the difference between the exploded thumbnail, the FT image, and the target image. This implies that both the domain blocks and range blocks in the signed difference image have zero sum. Hence in fractal coding according to the Fisher variant of Jacquin coding, the gray-scale in the range block must be linear (with no offset) in the gray-scale of the compressed domain block.

The overall distinction between thumb-nail fractal coding and thumb VQFT is clearly shown in Fig 1,2.

4. Experimental Results

Using 4x4 blocks (range in fractal terminolgy), with 16x16 domain blocks, thumbnail-based code was generated for a range of inter-block distances, which in fact corresponds to a range of compressions:

For 256x256 Lena decoding data is as follows:

Thumb VQFT

Exploded thumbnail has PSNR = 23.7932

After tiling with correction tiles

Inter-block	16,16	9,9	3,3	2,2	1,1
PSNR	30.291	30.697	32.636	33.363	33.621

Thumb Fractal Coding

Exploded thumbnail has PSNR = 23.7932

PSNR on subsequent fractal iterates

Inter-block	16,16	9,9	3,3	2,2	1,1
1st Iterate	30.291	30.689	31.845	32.156	32.195
2nd Iterate	30.277	31.537	33.210	33.724	34.254

5. Discussion

It was found that the PSNR of images generated via thumb VQFT coding was generally better than for the first iterate of thumb fractal coding, but that on the next iterate the thumb fractal was superior. There was no further significant improvement to the thumb fractal decoded image on further iterates.

It is noted that for (16,16) interblock, the thumb fractal and the thumb VQFT code is identical, so that the VQ decode and the first iterate of the fractal decode are identical. This situation will arise for interblocks that are a multiple of the range block size.

6. Conclusion.

The implementations described demonstrate that the concept of thumb-nail based coding [1] can be directly applied to fractal encoding and a new variety of vector. As far as overall image compression is concerned, we have shown that there is no loss of bits; however, in application to image based retrieval these thumb codes are as discussed [1] far more efficient than their non-thumb counterparts.

The experimental data presented here has been limited to single level coding, but clearly the same conclusions will

Cite as: Harvey A. Cohen, *Fractal Image Coding for Thumbnail-based image Access* Proc. Int'l Conf Signal Processing Applications, ISSPA, Gold Coast, Aug 26-28, 1996, Vol 1 pp 158-161.

hold for multi-level coding where much greater compression applies.

A theoretical aspect of this work has been to bring into sharp relief the relation between fractal coding and vector quantization. The situation is rather more subtle than implied by Fisher et al who has described fractal coding as self-VQ [7]. It was found that the PSNR of images generated via thumb VQFT coding was generally better than for the first iterate of thumb fractal coding, but that on the next iterate the thumb fractal was superior.There was no further significant improvement to the thumb fractal decoded image on further iterates

.**References**

[1] Harvey A. Cohen, *Access and Retieval from Image Databases using Image Thumbnails,* Preprint Jan, 1996

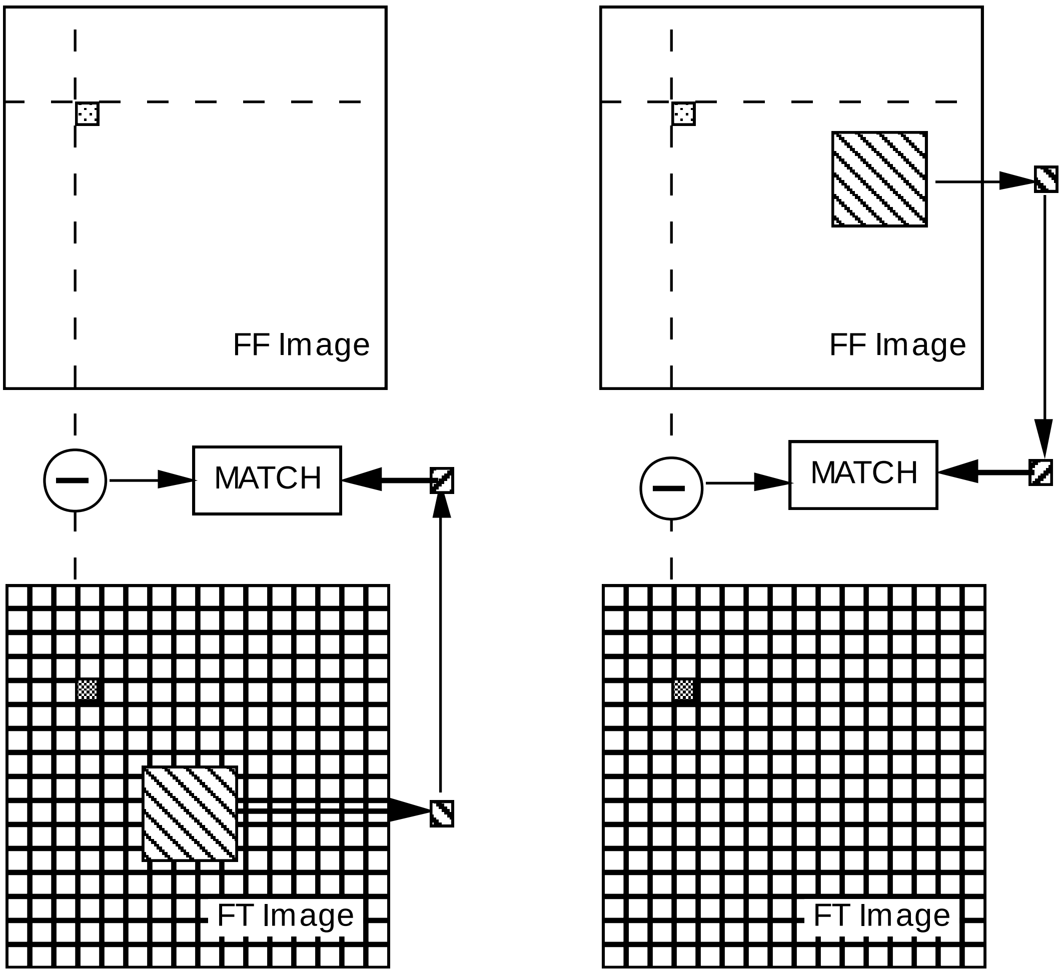
[2] A.E. Jacquin, *A novel fractal block-coding technique for digital images,* Proc Int'l Conf Acoustics, Speech and Signal Processing, ICASSP'90, pp 2225-2228

[3] A.E.Jacquin, *A novel fractal block-coding technique for digital images,* IEEE Trans. Image Processing, 1992, Vol 1, No 1, pp 18-30.

[4] M.F. Barnsley, *Methods and apparatus for image compression via Iterated Function Systems,* US Patent No. 4,941,199, 10 July 1990. + later patent

[6] E. Øien S. Lepsøy, *A Class of Fractal Image Coders with Fast Decoder Convergence,* in "Fractal Image Compression- Theory and Application", Chapter 8. pages 137-152, Springer-Verlag, New York.

[7] Fisher, T.P. Shen and D. Rogovin, Fractal (self VQ) encoding of video sequences, in Proc SPIE: VCIP, Vol 2304-16, Chicago, Illlinois, Sept 28-29, 1994.



Cite as: Harvey A. Cohen, *Fractal Image Coding for Thumbnail-based image Access* Proc. Int'l Conf Signal Processing Applications, ISSPA, Gold Coast, Aug 26-28, 1996, Vol 1 pp 158-161.

Fig 1. Encoding of Thumb VQFT.
The code for a block in the FF image (top) serves to compute the difference between the indicated block and the corresponding uniform block in the exploded thumb, FT image (lower). The optimum code is found by searching through the FT image for a larger block, which after normalising to have zero sum, is compressed, and transformed, is the best match.

Fig 2. Encoding of Thumb Fractal.
The code for a block in the FF image (top) serves to compute the difference between the indicated block and the corresponding uniform block in the exploded thumb, FT image (lower). The optimum code is found by searching through the FF image for a larger block, which after normalising to have zero sum, is compressed, and transformed, is the best match.