THUMBNAIL-BASED IMAGE CODING UTILISING THE FRACTAL TRANSFORM

Harvey A. Cohen

Computer Science and Computer Engineering
La Trobe University
Melbourne Victoria Australia 3083
Email: H.Cohen@latrobe.edu.au

ABSTRACT

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). The 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 image codecs based on the fractal transform that are thumbnail-based; one, called thumb fractal is a variant of block-oriented fractal coding; the other, vector quantization via fractal transform, VQFT. For both codecs signed correction tiles are derived via the fractal transform. Both codecs have as initial image the exploded thumbnail, to which VQFT applies correction tiles, while thumb fractal iteratively applies similar correction tiles to a rapid convergence. Experimental data is presented for fixed range block size of 4x4, but is indicative of results for other image blocking schemes, such as quadtree.

1. INTRODUCTION

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). The user's demand for rapid display of the FF-image after selection from thumbnail display is undoubtedly the primary consideration in image codec design. The latency involved is composed of data-retrieval time, together with image generation time. Especially for image database access via a network (LAN or InterNet), data retrieval time itself is vitally significant, so that codecs which should be called thumbnail-based, need to be developed. The current situation is that thumbnail-based retrieval is inefficiently performed by commercial network browsers because the available image codecs were not designed for this purpose; yet such browsers cache recent data,

A little analysis clarifies the need for thumbnail-based codecs: If the data required for production of the thumbnail is T, that for the FF-image is F, whilst the savings in the use of the T data to generate the FF image is S, with $0 \leq S \leq T$, then the user will receive (via database acess) the full data required for FF-display in a time reduced by the factor (1 - S/F). From the systems perspective, where there is a certain probability p, say, that a given thumbnail access will be followed by the access and display of the FF image, the savings in data transmission when S is (almost) equal to T are reduced [1].

ThIn this paper we introduce two thumbnail-based codecs based on the fractal transform, both of which take as an initial display image the image thumbnail exploded to full image size. The basic idea is to use the zoomed-out thumbnail as a first approximation to the decoded image, and to use the difference between the target image and the zoomed-out thumbnail as a source of tiling blocks that are to supply corrections. These tiling blocks are zero sum signed blocks computed using the fractal transform. However one of the codecs, called VQFT, is very properly an instance of vector quantization, as a single tiling of correction blocks serves to complete the FF-image. The other codec, involves an iteration process, and is essentially a variant of block-oriented fractal coding as introduced by Jacquin. 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 coding algorithm, termed thumb fractal coding, is related to proposals of Øien and Lepsøy [6]. The data presented over a range of domain block interblocks, demonstrate the capability of this thumb codec, applied to single-level image decomposition into 4x4 blocks. The fractal codec, which is very fast to decode, encodes essentially as efficiently as related fractal codecs, but is inherently far more efficient for thumbnail based retrieval.

2. THUMB FRACTAL CODEC

The basic idea of the thumb fractal codec is to use the zoomed-out thumbnail as a first approximation to the decoded image, and to use the difference between the target image and the zoomed-out thumbnail 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,

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

The fractal 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:

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

This correction block is coded by searching through 16x16 blocks within the target image, 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.

$$Y(\aleph) = T(\aleph)\,Y(0)\,T^{-1}(\aleph) \quad (\aleph=0, \ldots, 7)$$

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

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

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

$$\Sigma C_{ij}^{2} - \frac{16(\Sigma C_{ij} Y(\aleph)_{ij})^{2}}{\Sigma Y_{ij}^{2}}$$

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 the following (the index $\aleph$ is deliberately omitted from the denominator).:

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

I,J together with $\aleph$ and the contrast a given by

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

give the 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. However, where the inter-block is chosen so that domain block blocks are composed of an exact number of range blocks, the two codecs are essentially indistinguishable.

4. EXPERIMENTAL RESULTS

In the experiments reported here, the range blocks were uniformally 4x4, and the domain block size was 16x16. We found the best matching domain block via exhaustive search for a range of interblocks from (16, 16) to (1,1). The best match involved selecting transform (one of 8 available) requiring 8-bits, together with a contrast, for which 6 bits was used, although less bits might be satisfactory.

Thumb VQFT - 256x256 Lena

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	33.36	33.621

Thumb fractal - 256x256 Lena

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

No significant improvement to PSNR in subsequent iterations.

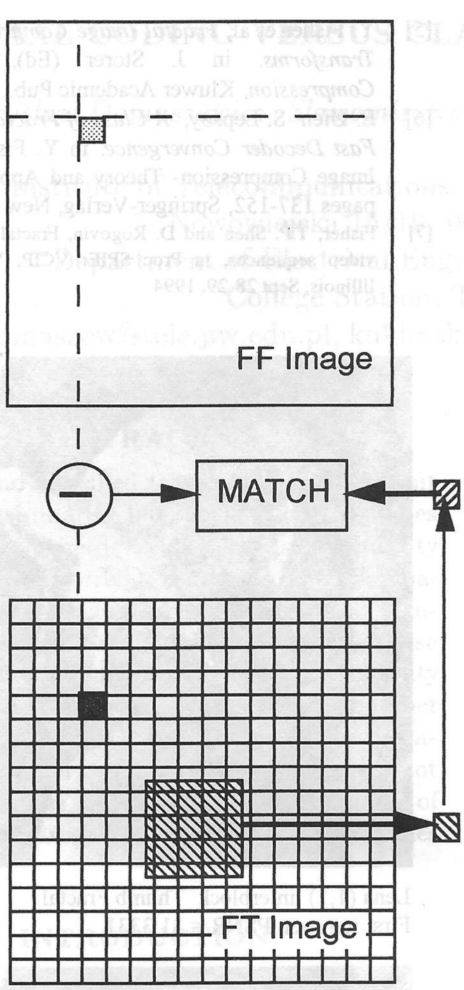

Fig 1. Encoding of Thumb VQFT.
The code for a block in the FF image (top) serves to compute the difference between the indicated block and the corresponding uniform block in the exploded thumb, FT image (lower). The optimum code is found by searching through the FT image for a larger block, which after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

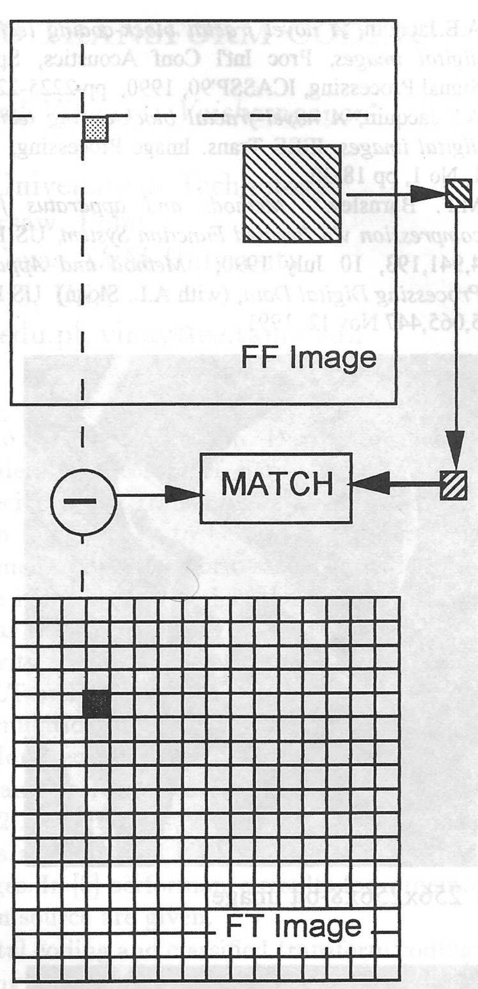

Fig 2. Encoding of Thumb Fractal.
The code for a block in the FF image (top) serves to compute the difference between the indicated block and the corresponding uniform block in the exploded thumb, FT image (lower). The optimum code is found by searching through the FF image for a larger block, which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

4. Conclusion.

The implementations described demonstrate that the concept of thumbnail based coding [1] can be directly applied to both block-oriented fractal encoding., and to a novel form of VQ using the fractal transform pixel of the FF image. As was required, T data of the thumbnail replaced data otherwise required for Ff image code. There were incidental advantages of very rapid convergence of the fractal decoding process. The experimental data presented here has been limited to single level coding, but clearly the same conclusions will hold for multi-level coding where much greater compression applies.

Note that our choice of 16x16 domain blocks in conjunction with 4x4 range blocks is in agreement with proposals of Øien and Lepsøy [6] for faster decoding. In fact. the basis for this design choice was to ensure that 4x4 range blocks, for any location of the boundaries of the domain block within an image, all elements of the compressed domain block could have different values - absolutely vital for vector quantization. The fractal coding scheme of Øien and Lepsøy [6], which involves the use of projection operators is related in other aspects, but lacks the explicit concern with thumbnail-based retrieval.

Finally we remark that although fractal coding has previously been described as self-VQ [7]; this study highlights the theoretical distinction between fractal and VQ coding.

.References

[1] Harvey A. Cohen, *Access and Retieval from Image Databases using Image Thumbnails,* to appear, Proceedings ISSPA 96, 26-28 August, 1996, Gold Coast, Queensland, Australia

[2] A.E.Jacquin, *A novel fractal block-coding technique for digital images,* Proc Int'l Conf Acoustics, Speech and Signal Processing, ICASSP'90, 1990, 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 System,* US Patent No. 4,941,193, 10 July 1990; *Method and Apparatus for Processing Digital Data,* (with A.L. Sloan) US Patent No. 5,065,447 Nov 12, 1991

[5] Y. Fisher et al, *Fractal Image Compression using Iterated Transforms,* in J. Storer (Ed), *Image and Text Compression,* Kluwer Academic Pub., pp 35-61, 1989

[6] E. Øien S. Lepsøy, *A Class of Fractal Image Coders with Fast Decoder Convergence,* in Y. Fisher (editor) "Fractal Image Compression- Theory and Application", Chapter 8. pages 137-152, Springer-Verlag, New York. (1994).

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

Lena 256x256x8-bit image

Lena exploded 4x4 Thumb
PSNR = 23.7975

Fig 2 Examples of first and second iterate for thumb fractal code, together with the target image and exploded thumbnail for 256x256x256 gray-scale lena. The image is regularly 4x4 blocked. The minimum interblock distance of domain blocks available for coding is specified

Lena (1,1) interblock Thumb Fractal
First iteration PSNR = 31.3331

Lena (1,1) Interblock Thumb Fractal Decode
Second iteration PSNR = 32.6760
[No significant improvement on later iterates].

These and other images demonstrating the coding of images via the thumb fractal and VQFT codecs may be viewed on the author's home page on the WWW:
http://www.cs.latrobe.edu.au/~image/harvey.html