Harvey A. Cohen, *Access and Retrieval from Image Databases Using Image Thumbnails*, Proc Int'l Conf Signal Processing Applications, ISSPA, Gold Coast, Australia, Aug 26-28, 1996, Vol 1, pp 427-8. Preprint

ACCESS AND RETRIEVAL FROM IMAGE DATABASES USING IMAGE THUMBNAILS

Harvey A. Cohen

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

ABSTRACT

The emerging role of thumbnail images in the selection of images for display from large image databases and via network access requires that the characteristics of thumbnails be seriously studied. We introduce a measure of effective compression $C_{eff}(p)$ that is a function of the probability p that thumbnail access will be followed by the display of a full-scale image. For credible values of p, we deduce the need for the development of thumbnail-based image compression schemes, which efficiently and effectively use thumbnail data as part of the code for the full-scale image. We briefly discuss design features of thumbnail-based variants of the block-oriented coding schemes of vector quantization, fractal coding, and JPEG.

1. THUMBNAIL IMAGES

Thumbnail images are now commonly used as WYSIWYG directory icons for the selection of images from directories and local databases. Notable systems for managing query of image databases, the QBIC system developed by Niblack and co-workers at IBM San Jose, [2] and the PhotoBook System developed at MIT Media Lab by Pentland et al, [3], feature arrays of small images for displaying the images most similar in content to the users query. A number of researchers, notably Picard[3], have stressed the important emerging role of very large image databases as in collections of compressed images stored on a CD ROM, or the far larger compilations available via the Internet. The needs for a user-friendly interface to these very large image albums and libraries requires the development of image compression schemes suitable and efficient for applications where the image thumbnails are more often required than full images, and full image display is invariably preceded by a data fetch of the thumbnail data. This entails that image compression schemes suitable for very large database applications should be so designed that thumbnail data is separately accessible, and that the information delivered in the thumbnail should be used with other image code for economical image synthesis or reconstruction.
In this paper these considerations lead us to introduce a measure of image compression that gives an effective compression for image data accessed via thumbnails.
We examine how traditional block-oriented vector quantization (VQ) and fractal coding based on Jacquin's scheme should be modified. Finally we discuss JPEG coding in this regard.

1.1 Gray-scale Thumbnails

Although image thumbnails are now commonplace, a literature search failed to locate any previous formal description of thumbnails. The following general definition of a thumbnail image is proposed:
An image thumbnail is produced after the partitioning of an image into rectangular (usually square) blocks of pixels, and constructing a thumbnail image comprising blocks, of single pixel-size or larger, of uniform pixel value. The blocks in the thumbnail are uniformly scaled with respect to those of the image.
In the simplest case of the uniform partitioning of an image into constant-sized blocks each of R rows and C columns the image block at (rR, cC) in the image may be denoted by its block row number r and block column c as B[r][c]: the corresponding block in the thumbnail contains pixels of gray-scale or colour index b[r][c]. For gray-scale images, b[r][c] is usually chosen as the block mean

$$b[r][c] = \Sigma B[r][c]_{ij} / RC$$

The smallest thumbnail image contains just one pixel of gray-scale b[r][c] to represent the block B[r][c], is here simply called the image thumb. The thumbnail image with blocks the same size as the corresponding blocks in the original image is here called the zoomed thumbnail.
Some display systems offer thumbnails based on sub-sampling the uncompressed image, eg using $b[r][c] = B[r][c]_{00}$. The MIT system PhotoBook [3] offers both block mean and sub-sampled thumbnails..

Where an image has been decomposed into regularly scaled blocks, as in quad-tree decomposition, the thumb is clearly the thumbnail scaled so that the smallest block is of single pixel size.

1.2 Colour Thumbs

For true-vision images, with true-vision display, thumbnails could in principle be produced with each component of the thumb the average of corresponding component (red, green, blue) of the pixels in each block
For blocked colour images, mean RGB component in each block may be likewise computed from the actual red, green, and blue components for each pixel in an RxC pixel block B as per:

$$m[r][c]_{colour} = \Sigma B[r][c]_{ij\ colour}/RC$$

For an indexed-colour environment, the most appropriate colour index to use in the thumb is index p in the palette available that is closest to m:
$\| palette[p] - m[r][c]\| \leq \| palette[q] - m[r][c]\|$ any colour index q. There is no confusion in referring to the thumb computed using this approximate mean as the colour-mean thumb. See http:\cs@latrobe.edu.au/~image/thumb.html

2. COMPRESSION EFFICIENCY WITH THUMBNAILS

The crudest measure of efficiency of image compression is the compression ratio, lets call it C, of the entire image, as it is transmitted. C measures the efficiency of retrieval of whole image from a data-base. If the user selects such an image for full-scale display via a thumbnail based interface, then the cost of accessing includes the transfer cost of all the thumbnails perused. We here define an effective compression as a function of the probability p that from a set of thumbnails a whole-image selection will be made. For simplicity suppose that one is dealing with a set of images all of the same size. Then, formally:
If p is the probability that a viewed thumbnail image will be selected for full-image display
T = data to be transferred to generate image thumbnail
F = data to be transferred for full-size image display
S = savings in using thumbnail data to generate display
The data required at a workstation to display a particular image is just T + F, but from the systems perspective the average amount of image data to be delivered for the display of each full-sized image is

$$\frac{1}{p} \quad T + F - S \quad \text{where } 0 \leq S \leq T$$

By definition of compression, the image data in the display is just C F. Hence the effective compression $C_{eff}(p)$ in terms of bytes in the display versus transferred bytes for probability p of full-image retrieval is

$$C_{eff}(p) = C\frac{\frac{pF}{T}}{1 - \frac{pS}{T} + \frac{pF}{T}}$$

The following are indicative values:

C	60	60	45	60	45
p	.0.1	0.1	0.1	0.1	0.1
$\frac{F}{T}$	.3	3	4	1	5
$\frac{S}{T}$	0	1.0	1.0	1	1.0
C Eff	13.86	15.0	13.86	6	16.5

The table shows the dominant significance of the ratio $\frac{F}{T}$ on the effective compression, and the significant impact of the savings S. One can conclude that it is most crucial to devise efficient thumbnail compression schemes, and that if the thumbnail information is utilised for the synthesis of the displayed image there will be significant advantages, which especially ought to be incorporated into network browsers.

3. THUMBNAIL-BASED IMAGE CODECS

In this section we show practical means of designing image codecs which are thumbnail based, making effective use of the data contained within an image thumbnail. Attention is confined to block oriented codecs, where the image thumb provides a representative value for each block.

.

3.1 Vector quantization

Image coding via vector quantization requires the use of a code book whose contents are best called tiles, though often, with mathematical emphasis, called vectors. The coding algorithm involves the partition of an image into blocks, and the determination for each block in the image of the block-code, namely the number of the tile in the code-book that is closest, usually in the least squares sense. Classic references are [4][5].

What is introduced here is a thumb VQ code comprising four parts instead of the traditional three:

(a) Header information
(b) Code - set of tile numbers for each block
(c) Thumb code: a table of mean gray-scale or colour-mean for each tile in the code book
(d) The code book - an album of tiles/vectors

To generate the image thumb requires the use of parts (a) (b) and (c); to generate the full scale image requires in addition the codebook (d). Using the formula for effective compression above, one can readily estimate the considerable savings for typical situations.

The thumb code (c) is relatively small, and adds little to overall coded image size. The extra storage costs for the thumb code can be made negligible, at computational cost: for gray-scale images, knowledge of the block-mean, enables the more significant bits of one pixel to be determined from the pixel values in the block.

In sum, for regular blocked, or regularly graded blocks as in quad-tree, VQ coding is readily augmented to facilitate thumb-based retrieval.

3.2 Thumbnail-based Fractal Coding

In block-oriented fractal coding, as introduced by Jacquin[6] [7] the image is partitioned into non-overlapping blocks, called range blocks, precisely as in vector quantization The fractal code for each image block, is simply the location of a larger block within the image, termed a domain block, together with the parameters of the transformation by which the range block may be produced from the corresponding domain block.

The basic idea of thumbnail based fractal coding is to use the zoomed-out thumbnail as a first approximation to the decoded image, and to iteratively improve this blocked image using (signed) tiles that have zero pixel sum. Because of this zero mean, the domain block code is smaller than in Jacquin-type coding, as the pixel gray-scale transformation is simply scaled (homogenous linear transform). Thus this variety of coding is truly thumb-nail based, and also has the advantage of very rapid convergence.[8].

3.3 JPEG Coding and Thumbs

In base-line JPEG[1] coding of gray-scale images, the mean of each 8x8 block, the DC component for the block, is stored separately, but is not readily accessible. It is clear that with this thumb material separately accessible, would be advantageous for thumb-based retrieval. For colour images, the different scale for colour and chromaticity leads to some modest complications which will be discussed elsewhere. The basic point is that the path ahead is here.

It is important to note that although there is provision for a thumb within the JFIF variant of JPEG,[9] this refers to a RGB thumb, which is stored in addition to the usual JPEG data.. The overall conclusion is that with some relatively modest alteration to the standard, thumbnail-based JPEG coding could be achieved, with the introduction of coders and decoders that would be backwards compatible.

4. CONCLUSIONS

Maximising the speed of image retrieval is essential to user satisfaction. in network access, where image selection is (almost) invariably preceded by the viewing and selection of its thumbnail. The traditional measure of image data compression is not adequate to describe the efficiency of such thumbnail-based retrieval, so a new measure, C_{eff} has been devised. Our analysis leads to proposals for a range of thumbnail-based codecs vector quantization, and JPEG. Images illustrating these codecs and issues of thumb quality are on the WWW page:
hppp://cs@latrobe.edu.au/image/thumbs

References

[1] Gregory K. Wallace, *The JPEG Still Picture Compression Standard*, 1991, C.A.C.M., 34(4): 30:44.

[2] Wayne Niblack and Myron Flickner, *Find Me the Pictures That Look Like This: IBM's Image Query Project,* Advanced Imaging, April 1993, pp 32-35.

[3] A. Pentland, R.W. Picard and S. Sclaroff, *Photobook: Tools for content-base manipulation of image databases,* in Proc Society for Optical Engineers, SPIE Storage and Retrieval of Image and Video Databases II, San Jose, Calif, Feb 1993.

[4] Lloyd, S.P. *Least squares quantization in PCM*, Bell Lab Memo, July 1957; also in IEEE Trans Info Theory, 'IT-28, 129-137 (1982).

[5] Max,, J. *Quantizing for minimum distortion*, IRE Trans Info Theory, IT-6(1), 7-12 (1960)

[6] 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

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

[8] Harvey A. Cohen, *Fractal Image Coding for Thumbnail Based Access*, (in Press) Proceedings ISSPA 96.

[9] Eric Hamilton, *JPEG File Interchange Format Version 1.01*, Aug 20, 1991, C-Cube Microsystems, Inc, 399A West Trimble Road, San Jose, CA 95131