EMBEDDED DCT CODING WITH SIGNIFICANCE MASKING

N. K. Laurance and D. M. Monro

School of Electronic and Electrical Engineering, University of Bath
Claverton Down, BA2 7AY, England

e-mail: neil@ee.bath.ac.uk D.M.Monro@bath.ac.uk
Fax: +44 1225 826073
Internet: http://dmsun4.bath.ac.uk/

SUMMARY

We evaluate two schemes for significance switching of DCT coefficients for block-based embedded DCT image compression. Both schemes deliver their best compression at any PSNR or vice versa, within a data stream that can be terminated within a few bits of any point. An image is partitioned into equal sized blocks, and a fixed point DCT of each block is calculated. The coefficients are then passed through successive 'significance sweeps' of the whole image from the most significant down to the least significant coefficient bitplanes. The coded data stream includes bits to refine only the significant coefficients at each sweep. With each new sweep, newly significant coefficients may appear within a block, and the two switching schemes evaluated are efficient methods based on block addressing and block masking. Both methods give good compression when used losslessly to the fixed point DCT precision. The best coder outperforms the baseline JPEG method in PSNR at any compression, and is similar to state of the art wavelet coders.

1. BACKGROUND

The JPEG baseline method for still image coding uses the Discrete Cosine Transform (DCT) in a fixed 8x8 pixel partition [1, 2]. Through a linear quantization table and zigzag scanning of DCT coefficients, the redundancy and bandwidth characteristics of the DCT are exploited over a range of compressions. Recently, it has become clear that the JPEG coder is not effective at higher compression ratios compared to other methods such as wavelets, which have produced better results while having the advantage of being fully embedded [3, 4]. Recent success in wavelet coding relies on exploiting the structure of the Discrete Wavelet Transform (DWT) through significance switching in a tree structure. This success is a result of good structuring and quantization of the data stream, and not necessarily due to any inherent superiority of the DWT.

A new round of standardization now underway, JPEG 2000, calls for a single still image coder to produce lossless and lossy compression, while being fully embedded and giving improved compression performance. Monro and Sherlock [5] have shown how DCT coding can be extended to higher compressions by relaxing the JPEG restriction on block size. This paper shows how significance switching can be applied to achieve highly accurate embedded DCT coding. By applying significance switching in combination with the zigzag scanning normally used to structure DCT data streams, the efficiency of block based DCT compression is improved at any compression ratio, including lossless.

2. METHOD

2.1 Progressive DCT

We partition an image into blocks of edge size $N = 2^n$ pixels and calculate the fixed point 2D DCT of each block. This is done in the usual way, by applying the 1D DCT to rows and then columns

We use the same coefficient normalization as JPEG coders, so that the MSE significance of all coefficients is equal, and a compromise is struck between range and precision:

$$C_0 = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} F_k$$

$$C_n = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} F_k \cos\frac{(2k+1)\,n\pi}{2N} \quad , n = 1, \dots N-1$$

Because the DCT is orthogonal, the most significant bit planes of the coefficients are the most significant in terms of image distortion. Each bit in the same bit plane of each coefficent makes the same contribution to the total MSE. We can therefore describe a progressive DCT coder as:

```
Sweep DCT bit planes MSB to LSB
    Sweep all image blocks
        Switch on significant coefficients
        Send data bits for significant coefficients
    Next block
Next bit plane
```

At present the coder is implemented using a fixed block size, and the scanning of image blocks is in raster order. Normally, most of the energy of an image is compacted into the first few DCT coefficients. As in the JPEG method, zig-zag transmission of the coefficients will encounter the majority of the information early in the scanning process, particularly in the most significant bit planes.

To implement significance switching, we construct a binary mask for each image block, giving the status of its DCT coefficients, as in Figure 1. Initially all coefficients are switched off, and as each block is visited, newly significant coefficients are switched on before the data is sent. The decoder maintains a copy of the mask, and the mask must be updated as each block is visited. We have evaluated two ways of communicating this information to the decoder, namely, DPCM Address Switching and Mask Switching.

2.2 ZigZag DPCM Address Switching (ZZD)

In this method the coefficients are switched by a list of addresses sent in zigzag order. At first sight this would seem to imply a large overhead, for example 6 bpp to address coefficients 0 to 63 for an 8x8 block. However this is substantially reduced by sending the difference between the addresses and a stop symbol, all entropy coded. At high compression ratios the symbol streams are short, and since only new addresses are sent on each pass the scheme is efficient all the way to lossless coding.

2.3 Zigzag Mask Switching (ZZM)

This scheme sends an explicit binary mask in zigzag order to switch on individual coefficients. At the early stages, this mask is very short, and in later passes only new coefficients have to be masked on at each pass.

Various methods of packaging the mask have been considered. The results reported here use a scheme in which a length symbol is sent before the mask. The length value

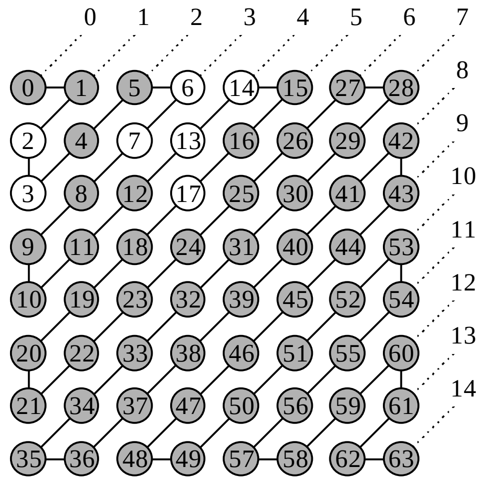

Figure 1. A significance switch labels individual DCT coefficients as on (clear) or off (shaded).

used is the manhattan depth of the highest order coefficient. For example in Figure 1, the highest order coefficient has zigzag address 17 and manhattan depth 5. Using manhattan depth will result in 3 extra mask bits being sent. However the manhattan depth requires 2 fewer bits at all block sizes than would a zigzag length. Recalling that the DCT packs most of the energy into a small subset of coefficients, leaving a large number of smaller coefficients, the manhattan depth is more efficient for small values. In either case the switching mask will contain mostly 'off' symbols, which can be efficiently

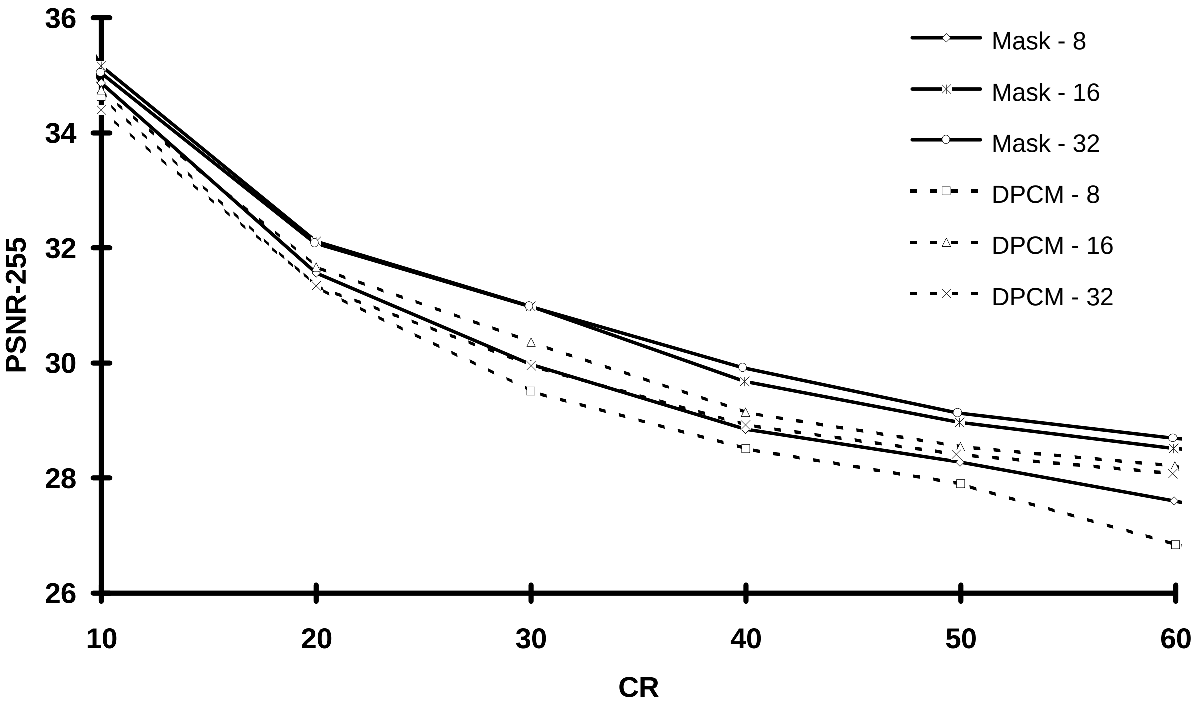

Figure 2. A comparison of the DCT significance switching schemes using different block sizes.

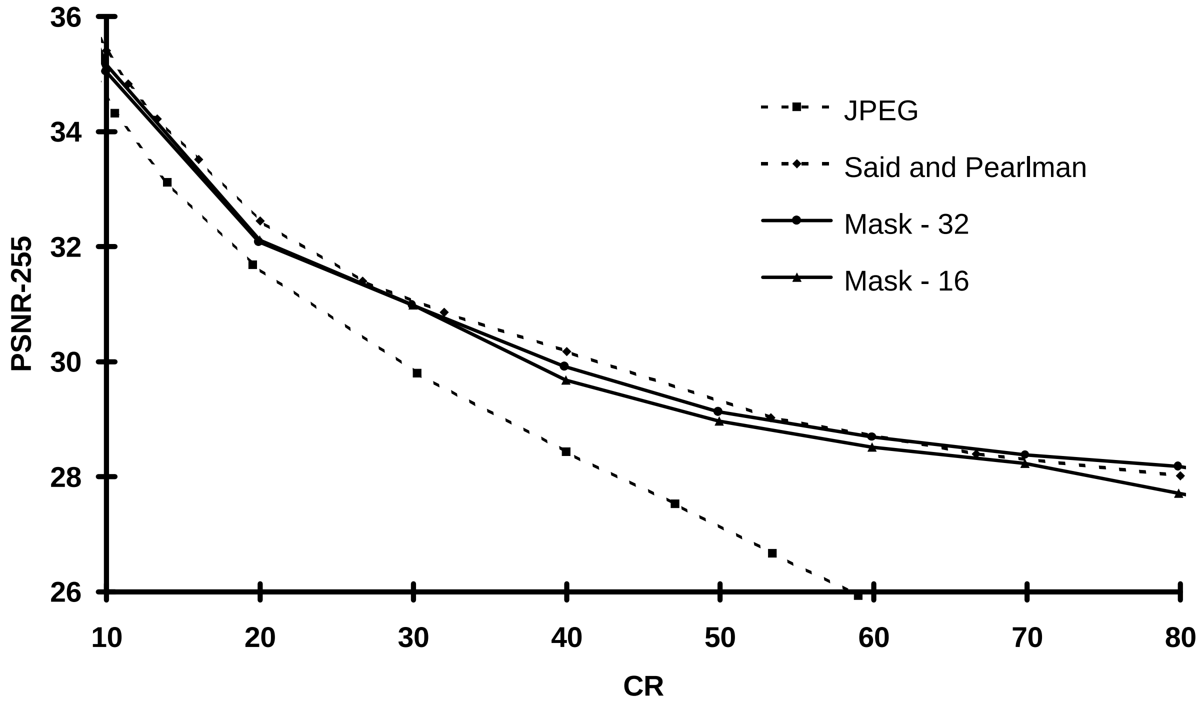

Figure 3. Mask switched DCT coder compared to JPEG and Said & Pearlman Wavelet coders.

compressed with an arithmetic coder.

2.4 Entropy coding

In our implementation, both ZZD and ZZM versions use first order predictive adaptive arithmetic coding. This is particularly useful for the mask data which has long runs of low entropy 'off' symbols. The Mask Switching coder has three synchronized streams, the manhattan depth, the mask data and the DCT data. The DPCM Address Switching coder uses two streams, one for the difference in zigzag addresses and the other for the DCT data bits. The synchronization is controlled so that each stream is within a few bytes of synchrony with the other, so that a data stream could be terminated at any time with optimal decoded PSNR.

3. RESULTS

For evaluation, we first of all applied the methods using 8x8, 16x16, and 32x32 blocks over a range of compression ratios to the Gold Hill test image. In Figure 2 we show the rate/distortion curves obtained for the two methods. As would be expected from the known performance of JPEG compared to the use of larger blocks [5], the larger blocks perform better at high CR, with the differences between block sizes becoming insignificant at lower CR. Ringing of the decoded images is visible at higher compression ratios. Because smaller blocks confine the spatial extent of ringing, as a general rule one would use the smallest block size that gives acceptable PSNR. We judged this to be 16x16 across the entire range.

The ZZM (Mask Switching) version is superior at high CR, and also slightly better at low CR. In Figure 3 we compare the 16x16 block size embedded ZZM switched DCT with JPEG and a wavelet coder by Said and Pearlman [4], generally regarded as one of the best available. As can be seen the DCT coder is competitive in PSNR terms with this state of the art wavelet coder, and superior to JPEG at all compression ratios. Figure 4 shows detail from the CCITT test image Gold Hill after the whole luminance (Y) component was compressed to 0.2bpp. The dramatic improvement in visual quality over JPEG of both significance switched schemes can be seen.

Finally we tested the compression achieved when the coders were used to reconstruct the image from all the DCT data, so that the only loss is that caused by fixed point truncation in the DCT. Table 1 gives the results for Gold Hill and also for Lenna, for which comparisons of lossless compression are available in the literature.

4. DISCUSSION

The philosophy of significance switching is that the overheads introduced will be compensated for by the savings in not transmitting bits for small coefficients until they are switched on. Good performance was expected at high compression, perhaps surprising was excellent low compression ratio and lossless performance.

The ZZM (masked) switch using a manhattan depth symbol is the better method. It is always better than JPEG, and very close to the Said and Pearlman wavelet.

	Gold Hill	Lenna
Method	bpp	bpp
JPEG	5.35	4.95
RRN 95 [6]	-	4.42
ZZM 8	4.84	4.51
ZZM 16	4.76	4.41
ZZM 32	4.85	4.44
ZZD 8	5.41	4.97
ZZD 16	5.31	4.85
ZZD 32	5.49	5.06

Table 1. Lossless performance of Zigzag Masked (ZZM) and Zigzag DPCM (ZZD) significance switching for block sizes 8, 16, and 32. The DCT coefficients are truncated to 12 bits. Otherwise recovery is exact.

The lossless performance on Lenna of the 16x16 ZZM DCT coder also exceeded that of any coder tried by Ranganathan et al [6], leading us to conclude that, unexpectedly, ZZM 16 is also a state of the art lossless coder. It is in addition a state of the art coder at any compression ratio, giving better performance than JPEG throughout, and close to the Said and Pearlman wavelet coder. Indeed Mask 32 gives higher PSNR over part of the range.

A decoder in a multimedia system which uses a progressive, embedded coder can begin to reveal an image as soon as transmission commences. This is an advantage often claimed for wavelets, but significance switched DCTs also have this capability.

As well as being significant in their own right, these results are important in the context of the renewed standardization efforts for still image coding. While emerging technologies such as fractals [7] and wavelets [3, 4] should be carefully considered on their merits, the DCT has the advantage of being widely used and understood as a result of its adoption as the core technology of JPEG and all current versions of MPEG. While still capable of further tuning, ZZM 16 is a progressive, embedded compression method using a single technology which gives leading edge results over the entire range of compression from lossless to very low bit rates.

Our results show that a significance switched DCT coder is capable of meeting the requirements of future image compression standards in an evolutionary manner.

5. REFERENCES

[1] WALLACE, G.K.: 'The JPEG Still Picture Compression Standard', 1991, *Comm. ACM*, **34**, (4), pp. 30-44.

[2] PENNEBAKER, W.B., and MITCHELL, J.L.: 'JPEG Still Image Compression Standard', (Van Nostrand Reinhold, New York, 1993).

[3] SHAPIRO, J.: 'Embedded Image Coding Using Zerotrees of Wavelet Coefficients', *IEEE Transactions on Signal Processing*, 1993, **41**, (12), pp. 3445-3462.

[4] SAID, A. and PEARLMAN, W.A.: 'Image Compression Using the Spatial-Orientation Tree', *IEEE International Symposium on Circuits and Systems*, 1993, (694), pp. 279-282.

[5] MONRO, D.M., and SHERLOCK, B.G.: 'Optimal Quantization Strategy for DCT Image Compression', *IEE Proceedings on Signal Vision and Image Processing*, (in publication).

[6] RANGANTHAN, N., ROMANIUK, S. G., and NAMUDURI, K. R.: 'A Lossless Image Compression Algorithm Using Variable Block Size Segmentation', *IEEE Transactions on Image Processing*, 1995, **4**, (10), pp. 1396-1407.

[7] MONRO, D. M., 'A Hybrid Fractal Transform', Proc ICASSP 1993, pp. V, 169-172.

(a) (b) (c)

Figure 4. Detail from the CCITT image Gold Hill Y component at 0.2 bpp. (a) JPEG PSNR 28.4
(b) ZZD 16 - DPCM Address Switched 16x16 DCT PSNR 29.1 (c) ZZM 16 - Mask Switched 16x16 DCT PSNR 29.7.