IMPROVED MATCHING PURSUITS IMAGE CODING

Yuan Yuan and D. M. Monro

Department of Electronic and Electrical Engineering, University of Bath, Bath, BA2 7AY, United Kingdom
email { d.m.monro, eepyy } @bath.ac.uk

ABSTRACT

This paper reports improvements in compression of both inter- and intra-frame images by the Matching Pursuits (MP) algorithm. For both types of image, applying a 2D wavelet decomposition prior to MP coding is beneficial. The MP algorithm is then applied using various separable 1D codebooks. MERGE coding with Precision Limited Quantization (PLQ) is used to yield a highly embedded data stream. For inter-frames (residuals) a codebook of only 8 bases with compact footprint is found to give improved fidelity at lower complexity than previous MP methods. Great improvement is also achieved in MP coding of still images (intra-frames). Compared to JPEG 2000, lower distortion is achieved on the residual images tested, and also on intra-frames at low bit rates.

1. INTRODUCTION

The key contribution of this work is to achieve improved image coding by 2D Matching Pursuits (MP) for both still images (intra-frames) and motion compensated residual images (inter-frames). The improvements arise from carrying out a 2D wavelet decomposition transform prior to MP and using a picked codebook for good fidelity and low complexity [9].

MP iteratively decomposes a signal into a series of 'atoms' selected from an over-complete codebook of non-orthogonal basis functions, as first shown for audio by Mallat and Zhang [1]. In video coding MP has been used with motion compensated residual images, and when applied in a standard H.263 codec, Neff and Zakhor (N&Z) achieved significant PSNR improvement [2].

MP is computationally greedy; a full 2D search requires repeated inner product comparisons with all basis functions in the dictionary at each image position. At each step the position, sign, amplitude and codebook index of an atom is encoded and transmitted. The atom is then subtracted from the image before the process repeats. Many algorithmic improvements are known, and in previous work we have shown that the use of 1-D searching with 2D atoms can reduce the complexity [3].

Although the code of an atom costs many bits, MP succeeds because one atom covers many pixels compared to DCT or wavelet coding [4]. MP has only recently been found useful for still image coding [5, 6], in which the information is more evenly distributed than in residuals.

Alternative work has addressed the problem of dictionary design, resulting in smaller dictionaries that can be efficiently constructed from elementary basis functions [7, 8]. Monro [9] used a process of 'basis picking' to find efficient 1D dictionaries for images which are used separably in this present work.

2. HYBRID WAVELET/MATCHING PURSUITS

The coding efficiency of MP for residual images arises from the localised, sparse nature of the data produced by motion compensation. This present work performs a 2D wavelet decomposition (using the standard biorthogonal 9/7 wavelet) to localise the image data for both stills and residuals.

2.1. Matching Pursuits

The MP algorithm [1-6] can be described as a three-step per atom process:

Initialise Compute a full set of inner products between the image and all bases in a codebook.

Repeat

1. *Find an atom*. Full 1D or 2D search or reduced complexity strategies are possible [3].
2. *Image Update*. Subtract quantized atom from image.
3. *Repair Inner Products*. Recompute required inner products only in atom footprint.

Until distortion or bit rate criterion met.

The *Image Update* step contributes negligible cost. In a practical application hundreds or thousands of atoms are found per image, and the computation is dominated by the search and repair steps, particularly for 2D image coding.

In this work, we have used a full search to locate the optimum atom in each iteration. Although in previous work some investigators have restricted searching to local regions of high energy to reduce complexity, we have found that using a hierarchy of blocks whose true optimum inner products are recorded and updated is just as fast.

2.2. Precision Limited Quantization (PLQ)

Precision Limited Quantization (PLQ) was originally motivated by psychovisual considerations [10]. PLQ fits well with embedded coding schemes, and has been found useful in maximizing PSNR in MERGE coding both with audio and video data [9], as in this work. In Figure 1, if A is the unquantized amplitude of an atom, $S = sign(A)$, $F = \log_2|A|$, i.e. the First Significant Bit (FSB) of $|A|$, and $R =$ the remaining bits in the range 0 to $2^{PL-1} - 1$, for any precision limit $PL > 1$, including the FSB.

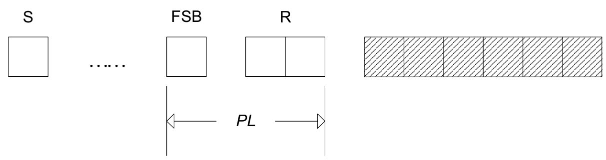

Figure 1. Precision Limited Quantization of an atom of amplitude A by the triple $\langle S, F, R \rangle$.

2.3 MERGE Coding

Each atom chosen by MP has a position in the data space with the *S*, *F* and *R* attributes and codebook index *K*. Lossless coding of the attributes is done by the MERGE algorithm, in which atoms are gathered into groups with all attributes in common, and the positions are signalled by run length coding. This works well with PLQ, which keeps the number of groups reasonably small. To reduce the number of groups further, the sign *S* of each atom is sent as one bit of side information which is efficient because positive and negative signs have near-equal probability.

If the Precision Limit is *PL*, the MERGE algorithm is:

For *F* (the FSB) from Maximum to Minimum
    For *R* (the amplitude Residual) from 2^{PL-1} to 0
        For each Basis Function *K* used
            Signal by Run Length Coding the position of each atom with attributes (*F, R, K*).
            Send the Sign *S* of the atom (1 bit)
        End of *K* (Basis Function) Group
    End of *R* (PLQ Residual) Group
End of *F* (FSB) Group

Maximum embedding is achieved by sending atoms in order of decreasing amplitude, with the codebook entry as the innermost loop. MERGE automatically compensates for variations in the frequency of occurrence of the attributes of the atoms and eliminates the need for entropy coding of them. The adaptive run length coding in MERGE adjusts to the statistics of the atom position.

2.4. Low to High Frequency Scanning of Atoms

In this work, a wavelet decomposition is made of both still (intra-) and residual (inter-) images. An effect of this, as with every decorrelating transform, is to skew the coefficient magnitudes towards low frequency sub-bands. (It is incorrect to speak of energy as the wavelet transform is not orthonormal and therefore not energy preserving.)

Virtually all image compression algorithms take advantage of this in scanning the coefficients from low to high frequencies. In this work, the MERGE coder scans the lowest frequency sub-bands first. In the case of 2D atoms, there is no advantage or disadvantage in particular scanning orders within bands, but the scanning order is inherited from the 1D work reported in [9], and is illustrated for 2 scales in Figure 2.

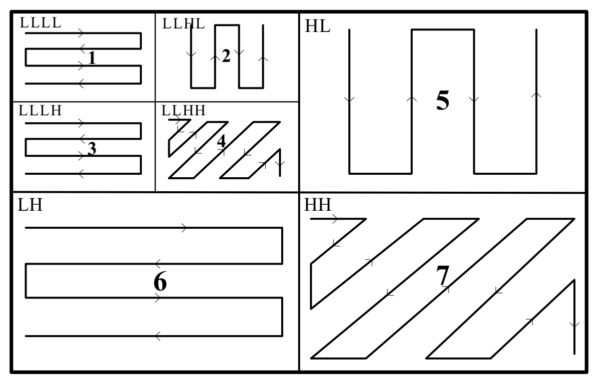

Figure 2. Low to high frequency Scan Order of Wavelet Coefficients.

3. CODING EXPERIMENTS

Several parameters affect the fidelity of images compressed by the hybrid Wavelet/MP method with MERGE/PLQ coding, including the choice and size of codebook, number of wavelet scales and the value of PL To determine the best settings, a reasonable set of values was assumed and each of the others was varied in turn.

The experiments were repeated with the best settings to confirm the best combination. Many experiments have been performed with consistent results.

3.1. Number of Wavelet Scales

With all the codebooks used and with all Precision Limits $1 \le PL \le 4$, improved rate/distortion is achieved on all the still (intra) images tested. Figure 3 illustrates the advantage of wavelet transformation on the Gold Hill image with the intra8 codebook and $PL = 2$.

The improvement on residuals is less dramatic, but in general either one or two wavelet scales is advantageous.

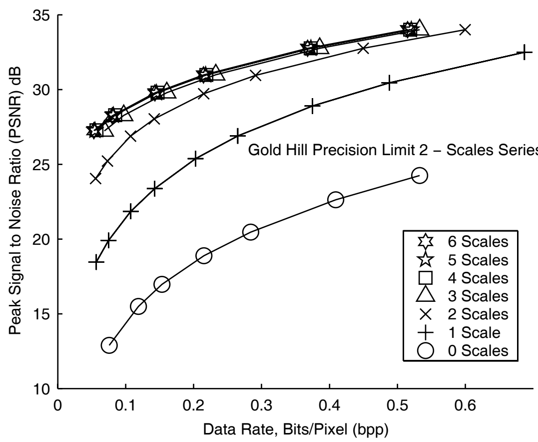

Figure 3. Varying the number of scales of wavelet decomposition in coding Gold Hill with the Intra8 codebook and *PL*=2.

3.2 Best Precision Limit (PL)

The MERGE coder functions best when the number of merged groups is not excessively large. This means relatively small codebooks work well with small values of *PL*. For example if *PL*=2 with a codebook of 8 bases, there are 8x8x2 = 128 merged groups per FSB. *PL*=2 is usually the optimum choice except at low bit rates on residual images, when *PL*=1 is comparable or occasionally slightly better. These differences are not as dramatic as those achieved by the wavelet decomposition. Figure 4 shows the rate/distortion performance when various values of *PL* are used in coding Gold Hill with 5 scales of wavelet decomposition.

3.3. Codebooks

In Figures 5 and 6 can be seen the effect of improved codebooks in compressing both still and residual images. The Intra8 and Resid8 codebooks were generated by Basis Picking as described by Monro [9] on 1D scans of images. In this work they are applied separably to form 2D codebooks, i.e. a 1D codebook of 8 bases produces a 2D codebook of 64 combinations. The Intra8 codebook consists of the first 8 picked bases on the Gold Hill image, and is best for still images when tested on Barbara2 and Boats. Resid8 is the first 8 picked bases from a composite of three residuals [9], and is not significantly better than Intra8 or D1 for residuals.

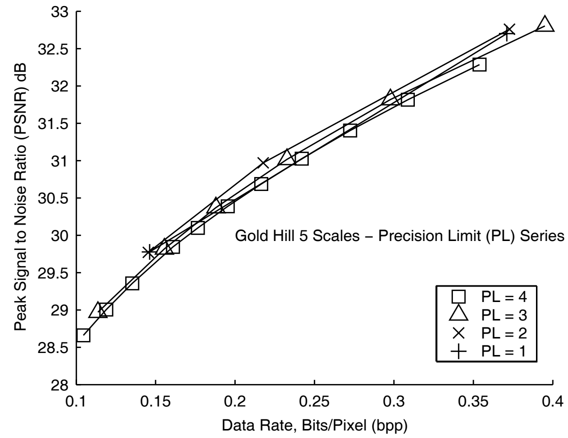

Figure 4. Varying the Precision Limit (PL) in coding Gold Hill with 5 Wavelet Scales using codebook Intra8.

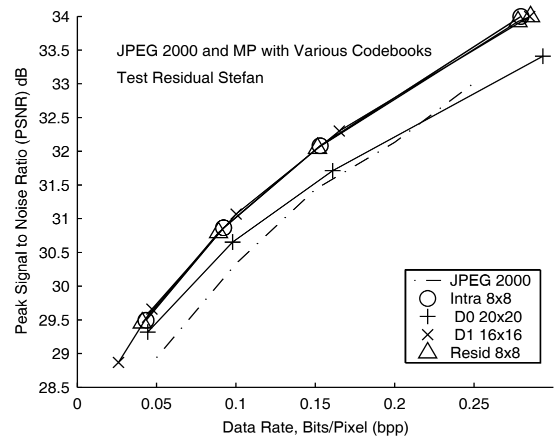

Figure 5. PSNR results for the Stefan residual.

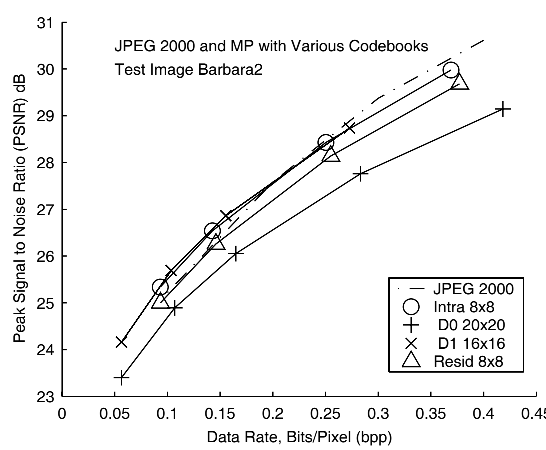

Figure 6. PSNR results for the Barbara2 Image.

Figure 5 shows the rate-distortion performance of the improved MP method on an inter-frame (residual) from the Stefan sequence. The PSNR figures are derived from the fully decoded frame, in which the decoded residual is added to the prediction. Also shown are the results for a full 2D search with the N&Z 400 atom codebook D0 [2] and the 256 atom Czerepinski et al codebook D1 [7]. For comparison JPEG 2000 (Kakadu V 3.2) and the Intra8 codebook are also included. The Stefan residual was not part of the training set for either Intra8 or Resid8.

Figure 6 shows the results for the Barbara2 intra-frame (still image). The same comparisons with other codebooks are included as in Figure 5. None of the codebooks were trained on Barbara2.

4. DISCUSSION AND CONCLUSIONS

Improvements to image coding by MP have been achieved through the use of a wavelet decomposition prior to MP coding. MP has previously been known to work well with residual images, which are noisy with isolated islands of significant data. The wavelet transform applied to still images achieves a similar effect. The effectiveness of smaller codebooks of reduced basis width may arise from the multiscale nature of the wavelet transform.

For intra-frames this improvement is very significant, e.g. as much as 15 dB at low bit rates in Figure 3. For inter-frames the advantage of the wavelet transform is less dramatic. However the smaller codebooks used in this work allow compression to the same fidelity at lower complexity than previous methods.

The results of our work, not all of which are shown, lead us to recommend that a wavelet decomposition should precede MP coding, 1 or 2 scales for inter-frames and 4 or 5 scales for intra-frames. The MERGE coder with PL=2 has given best results with codebooks Intra8 for intra-frames and Resid8 for inter-frames.

This research suggests the feasibility of using the embedded MERGE coder to code both intra- and inter-frames by MP in a motion compensated image coder. With much reduced complexity, the rate/distortion performance would be improved at considerably lower computational cost than in earlier work [2, 7].

All our results outperform JPEG 2000 at low bit rates. At present the run length coder in MERGE is based on Golomb coding. It could be expected that application of adaptive context based entropy coding to the run length coding within the MERGE coder could improve the results still further, as could other forms of prediction.

Our codebooks are highly effective. Interestingly, Intra8 is as good as Resid8 on residuals, although Resid8 is not as effective on stills. The 20 basis codebook (D0) used in the original work of Neff and Zakhor [2] is the least effective, while the 16 Basis codebook (D1) of Czerepinski et al [7] is very little different from either Intra8 or Resid8 on residuals. Based on these results, for simplicity we would use Intra8 for all 2D applications.

The rate/distortion gain with small codebooks comes because the cost of sending atoms is reduced, so that more atoms can be are used at a given bit rate even though on average each individual atom improves the PSNR by less.

Use of these codebooks also offers reduced computational cost. Intra8 and Resid8 applied separably in 2D contain only 8x8 bases of maximum size 15 pixels, compared to 16x16 bases of maximum size 25 in D1. The repair stage, done separably in 2D, dominates the MP algorithm, and its complexity scales as the square of the number of 1D bases and the cube of the basis size [3]. Therefore our results show that for equivalent fidelity an MP coder using Intra8 or Resid8 will involve 1/18.5 of the computations in the repair stage compared to D1 (1/80 compared to D0 whose largest basis is 35.)

5. REFERENCES

[1] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries", IEEE Trans. Signal Processing, vol. 41, pp. 3397–3415, Dec. 1993.

[2] R. Neff and A. Zakhor, "Very low bit rate video coding based on matching pursuits", IEEE Trans. Circuits and Systems for Video Tech., vol. 7, pp. 158–171, 1997.

[3] Y. Yuan, A. N. Evans and D. M. Monro, "Low complexity Separable Matching Pursuits", IEEE Int. Conf. Acoustics Speech Signal Process. (ICASSP 2004), Montreal, May 2004.

[4] W. Poh and D. M. Monro, "Comparison of residual compression methods in motion compensated video", IEEE Int. Workshop on Multimedia Signal Processing, Virgin Islands, Dec. 2002.

[5] P. Frossard, P. Vandergheynst and R. M. Figueras I Ventura, "High flexibility scalable image coding", Proceedings of VCIP 2003, July 2003.

[6] R. M. Figueras I Venrura, P. Vandergheynst, P. Frossard and A. Cavallaro, "Color image scalable coding with matching pursuit", IEEE Int, Conf Acoustics, Speech, Signal Process (ICASSP 2004), Montreal, May 2004.

[7] P. Czerepinski, C.J. Davies, N. Canagarajah, and D. Bull, "Matching pursuits video coding: dictionaries and fast implementation", IEEE Trans. Circuits Syst. Video Technol., 10 (7) pp. 1103-1115, 2000.

[8] R. Neff and A. Zakhor, "Matching Pursuit Video Coding-Part I: Dictionary Approximation", IEEE Trans. Circuits Syst. Video Technol., 12 (1), pp. 13-21, 2002.

[9] D. M. Monro, "Basis Picking for Matching Pursuits Image Coding", IEEE Int. Conf. Image Process. (ICIP2004), Singapore, October 2004.

[10] D. M. Monro, J-L Aufranc, M. A.Bowers and W Poh, "Visual embedding of wavelet transform coefficients", IEEE Int. Conf. Image Process. (ICIP 2000), Sept. 2000.