BASIS PICKING FOR MATCHING PURSUITS IMAGE CODING

D. M. Monro
Department of Electronic and Electrical Engineering, University of Bath,
Bath, BA2 7AY, United Kingdom
Fax: +44 1225 386073 e-mail d.m.monro@bath.ac.uk

ABSTRACT

The novel 'Basis Picking' algorithm is applied to select dictionaries of 1D basis functions for coding of image data by Matching Pursuits. For both motion compensated residual images and normal still images, bases are picked with a hybrid Wavelet/Matching Pursuits image codec using 1D scanning. By successively adding bases one at a time from a set of 1289 candidates according to their ranked Signal to Residual Ratio (SRR) performance, effective codebooks are constructed. These outperform traditional selection by ranked frequency of use from the same candidate bases over a range of compressions of both residuals and still images. This holds for both the training sets used for picking and other test images. Picked bases are also capable of good quality compression of still images, which was not previously thought to be feasible by Matching Pursuits.

1. INTRODUCTION

The key contribution of this paper is to apply an effective method for designing basis function dictionaries (the codebooks) for use with the Matching Pursuits (MP) coding algorithm. This is shown for a hybrid Wavelet/MP image codec using 1D atoms, but the technique would apply to all data types for which MP might be used. As well as being effective for motion compensated residual images, good results are also achieved with still images.

MP was first introduced by Mallat and Zhang for digital audio [1]. Neff and Zakhor achieved improved low bit rate video coding by MP for motion compensated residual images within an H.263 video codec [2].

In MP coding a dictionary of basis functions is repeatedly searched for the inner product of largest magnitude within the data set. The position, sign, quantized amplitude and dictionary index of the selected basis form the code of an 'atom'. The atom is subtracted from the data before the process is repeated on the remaining residual. Often the bases are an over-complete, non-orthogonal set of Gabor functions, and the importance of the codebook in the performance of the algorithm is well established [3, 4, 5].

This work is part of the Bath University MP (BUMP) project with the following overall objectives:

1. Reduced complexity of matching [6].
2. Reduced bit rate for coding atom parameters [7].
3. Improved codebooks (this paper.)
4. Extend to normal images (this paper) and audio.

2. HYBRID 1D WAVELET/MP CODEC

The codec used here employs a multiscale wavelet decomposition using the well known biorthogonal 9/7 filterbank, followed by 1D scanning of the wavelet coefficients. This has the advantage that MP coding of a 1D signal is considerably less computationally intensive than 2D MP image coding. In addition recent work with images has shown that 1D atoms can be used advantageously in image searching and coding [6].

For coding, a multiscale wavelet decomposition is applied before MP approximation. Experiments show that 2 scales for residuals and 5 scales for still images are a good choice. For coding, the sub-bands are scanned to give a 1D signal, to which 1D MP is applied. Experiments show that the best scanning direction is 'with the grain', following the low frequency detail, rather than 'across the grain', encountering high frequency detail such as edges. For certain bands there is no preferred direction. However mixing the scan order prevents streakiness which is visible in still frames with a dominant direction. A good scheme for both PSNR and visual quality mixes Horizontal (H), Vertical (V) and Diagonal (D) directions with scanning back and forth (Z for Zig-zag), in the mode HVHD-Z, as illustrated by Figure 1 for a 2 scale decomposition.

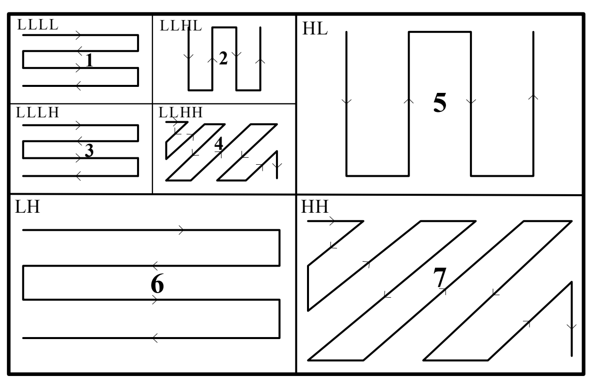

Figure 1. HVHD-Z Scan Order of Wavelet Coefficients.

3. BASIS PICKING

Two codebooks are designed, one for still images (intra frames) in which the Gold Hill luminance (Y) 704 x 576 image is the training set, and the other for motion compensated residual frames (intra frames) in which a composite of three residuals from the sequences Foreman, Container, and Bus are used, each of size 352 x 288. It was not feasible to use larger training sets because picking runs later in the process took up to 12 hours.

To construct a codebook, a dictionary of 1289 candidate bases is created spanning a range of parameters. The basis picking algorithm described in pseudocode is:

Initialize
Fixed Dictionary Size = 0
Candidate Dictionary Size = 1289
Repeat
For AtomNo = 1 to MaxAtoms
For CandidateNo = 1 to CandidateDictionarySize
Code the training set with the fixed dictionary and the selected CandidateNo
End (Of considering Candidates)
End(Of coding Atoms)
Add Best Candidate to Fixed Dictionary
Remove Best Candidate from Candidate Dictionary
Until finished building the codebook

This process is called 'Basis Picking', and is similar to a method introduced for trimming motion vector fields in video compression [8]. It has also been applied by the author to construct codebooks for a wavelet based MP audio codec. A basis is picked by ranking the Signal to Residual Ratio (SRR) of the candidates after coding a certain number of atoms (*MaxAtoms*), and considering their rankings at selected stages corresponding to useful compressions encountered in practical image and residual compression using the hybrid Wavelet/MP codec.

For these experiments, the Candidate bases were

$$g_k = \left(\exp\left(\frac{-\pi t^2}{\sigma_k}\right) \right)^{0.25} \cos\left(\frac{\pi f_k t}{w_k} + \phi_k \right)$$

where the codebook index is k and $t \in [-w_k, ..., w_k]$.

Numbers of samples $(2w_k + 1) \in [3,5,7,9,11,13,15]$

Basis frequencies $f_k \in [0,1..., w_k]$

Phase shifts $\phi_k \in [(0,0.5,1.0,1.5,2)\pi/4]$

Attenuation factors $\sigma_k \in [1,2,4,8,12,16,20,24]$

The unit impulse (Dirac function) is also included in the candidates. A number of candidates are zero at the sampling points, and are eliminated. The remaining 1289 bases are normalized to unit sum of squares. For the test composite residual image, the first 9 atoms selected are shown normalised in Figure 2, with their parameters given in Table 1. For still (intra) images, the first 12 atoms selected are shown normalised in Figure 3, with their parameters given in Table 2.

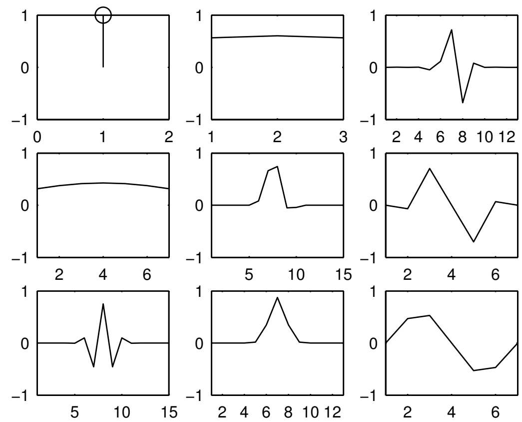

Figure 2. The first 9 bases selected by Basis Picking for the residual test set. The parameters are given in Table 1.

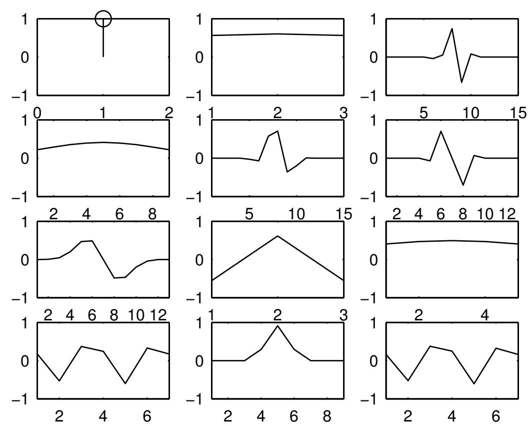

Figure 3. The first 12 bases selected by Basis Picking for the Gold Hill image. The parameters are given in Table 2.

3. POPULARITY ALGORITHM

For comparison, the data is coded with various numbers of atoms using the full 1289 Candidate basis functions. The most frequently occurring basis functions after coding a range of numbers of atoms were ranked and used as a codebook. This 'Popularity Contest' approach has often been used in codebook design [2, 3]. It was observed that the ranked frequency of choice of the most popular atoms is quite stable from 5,000 atoms onwards. The residual codebook of Table 3 was selected from these results and is shown normalized in Figure 4. A codebook based on popularity for the still image was also found but is not shown for reasons of space. It was included in the application to test data in Figure 5.

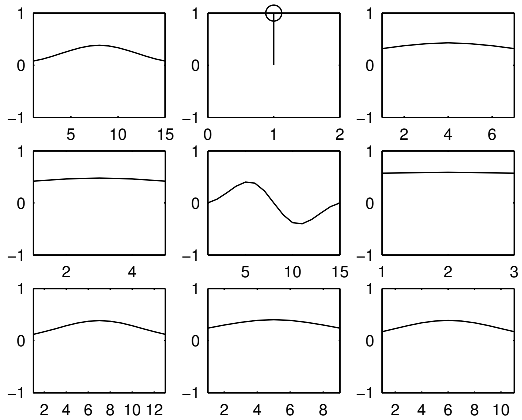

Figure 4. The top 9 bases for residuals ranked by Popularity. Parameters are given in Table 3.

4. APPLICATION TO TEST SETS

In Figures 5 and 6, Picked and Popular Bases are applied to test images with the hybrid Wavelet/MP codec. Atoms are quantized by Precision Limited Quantization (PLQ) [7, 9], and coded by an embedded bit plane method called MERGE, in which the atoms are grouped by combined attributes of dictionary index and amplitude for efficient transmission by Run Length Coding.

In Figure 5, Bases are added one at a time in repeatedly compressing a residual image from the Stefan sequence (not part ot the training set) at 0.2 bpp with 2 wavelet decomposition scales, and the still image Barbara2 at 0.4 bpp with 5 scales. Picked bases are added in the order picked, and Popular bases in order of popularity.

From Figure 5, it was decided that 8 picked bases made a good codebook for residuals, and 10 picked bases for images. In Figure 6 these codebooks are applied across a range of bit rates to the same test data.

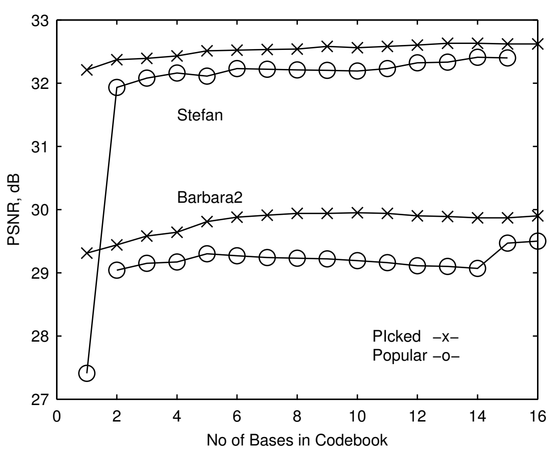

Figure 5. The PSNR of a test residual at 0.2 bpp and a still image at 0.4 bpp with successive Picked and Popular bases added to the codebook.

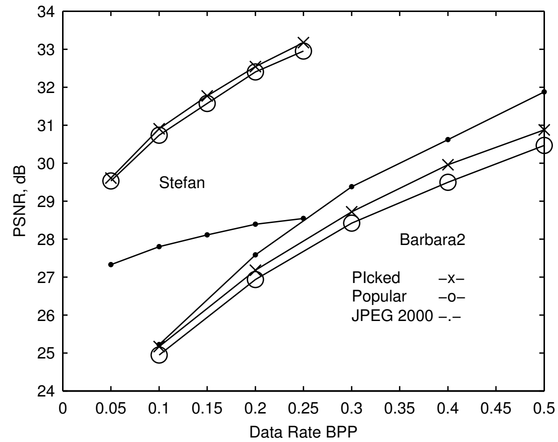

Figure 6. Rate/distortion graphs for the test residual from the Stefan sequence with 8 picked bases for residuals, and the still image Barbara2 with 10 picked bases for images using MERGE/PLQ coding of the atoms.

Picked No k	Samples $2w_k+1$	Frequency f_k	Phase $(*\pi/4)$	Attenuation σ_k
1	1	0	0	1
2	3	0	0	12
3	13	5	1.5	1
4	7	0	0	24
5	15	6	0	2
6	7	2	2	1
7	15	6	0	2
8	13	1	0	1
9	7	1	2	20

Table 1. The first 9 Picked Bases for residuals.

Picked No k	Samples $2w_k+1$	Frequency f_k	Phase $(*\pi/4)$	Attenuation σ_k
1	1	0	0	1
2	3	0	0	12
3	15	6	1.5	1
4	9	0	0	20
5	15	3	1	2
6	13	4	2	1
7	13	1	2	4
8	3	1	0	8
9	5	0	0	16
10	7	2	1.5	20
11	9	1	0	1
12	7	2	1.5	16

Table 2. The first 12 Picked Bases for Gold Hill.

Picked No k	Samples $2w_k+1$	Frequency f_k	Phase $(*\pi/4)$	Attenuation σ_k
1	15	0	24	1
2	1	0	0	1
3	7	0	0	24
4	5	0	0	24
5	15	1	2	24
6	3	0	0	24
7	13	0	0	24
8	9	0	0	24
9	11	0	0.5	24

Table 3. The 9 most popular bases for residuals.

5. DISCUSSION AND CONCLUSIONS

The Basis Picking algorithm is an effective method of codebook design. It has been demonstrated with a particular 1D hybrid Wavelet/MP image codec, but would apply to all types of data with which MP might be used, including 2D. Because picking was done with unquantized atoms, further work would be required to determine if quantization affects the picked results.

The poor performance of the Popularity method is surprising. Possibly this is because good but similar bases compete with one another and are not ranked as highly as they should be. Examining the lack of diversity revealed in the bases of Figure 2 and Table 2 reinforces this view.

The unit impulse, or Dirac function was the first basis picked. It also appears high on the popularity list and its occurrence as the second most popular atom for residuals causes a large jump in PSNR as is evident in Figure 5. Apart from that, the top of the popularity rankings are dominated by zero and low frequency bases. The picked lists have greater diversity and are more effective.

Very few bases appear on more than one list. Two notable exceptions are the Dirac function and the second basis picked, which is the same for still images and residuals. There is a scattering of zero frequency bases in the picked list, and some of those have high popularity rankings also.

To find the optimum codebook from a range of candidates is a combinatorial problem. Picked codebooks are not strictly optimum, either for the training sets or images in general, but they are superior to the Popular ones. The MATLAB program to find atoms with the Gold Hill image finds 1000 atoms per second using 10 bases on a 2.8 GHz Pentium 4. To pick the 10th basis by coding 20,000 atoms requires coding the image 1280 times, or just over 7 hours, which is tolerable given that a codebook need only be designed once. However to find the best 16 from all combinations would require the codec to be used ${}_{16}C_{1289}$ times, which is unrealistic. Even to find the best 2 bases combinatorially would require 1289*1288 codings of about 4 seconds each, i.e. about 40 days.

Although it is unrealistic to search for the optimum set of bases, it is expected that only modest improvement is possible from this set of candidates on this data.

The training set was small, to keep computer time to a tolerable level. The rankings with both algorithms were stable (but different) over a wide range of numbers of atoms, so perhaps only a few thousand atoms are required to select a short list of a few hundred basis functions by the Popularity algorithm. There is a risk that good bases could be missed, for example the Popularity algorithm never selected the second picked basis, either for residuals or stills. Despite this risk, popular bases could be used as candidates for Basis Picking on a relatively small number of atoms. With this approach larger and more representative training sets could be used. It would also be feasible to design different Basis Picked codebooks for different categories of data.

6. 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 Syst. Video Techol., vol. 7, pp. 158–171, Feb. 1997.

[3] P. Czerepinski, C. Davies, N. Canagarajah and D. Bull, "Matching pursuits video coding: dictionaries and fast implementation", IEEE Trans. Circuits Syst. Video Technol., vol. 10, No. 7, pp. 1103-1115, Oct. 2000.

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

[5] F. Moschetti, L. Granari, P. Vandergheynst and P. Frossard, "New dictionary and fast atom searching method for matching pursuit representation of displaced frame difference", Proc. IEEE Int. Conf. Image Process. (ICIP 2002), vol. III, pp. 685-688, Sept. 2002.

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

[7] D. M. Monro and W. Poh, "Improved coding of atoms in matching pursuits", Proc. IEEE Int. Conf. Image Process. (ICIP 2003), Sept. 2003.

[8] S. Tredwell and A. N. Evans, "A sequential vector selection algorithm for controllable bandwidth motion description encoding", IEEE Sympos. Intell. Multimedia, Video & Speech Process., Hong Kong, May 2001.

[9] 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.