LOW COMPLEXITY SEPARABLE MATCHING PURSUITS

Yuan Yuan, A.N. Evans and D.M. Monro

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

ABSTRACT

Methods of reducing the complexity of the matching pursuits algorithm with minimal loss of fidelity when coding displaced frame difference (DFD) images in video compression are investigated. A full search using 2D basis functions is used as a benchmark. The use of separable 1D bases greatly reduces the complexity, and significant further reductions are achieved by using only a 1D inner product search to locate the atom position, followed by a further 1D inner product search in the opposite direction to identify the second 1D basis function. To avoid ignoring significant structures orthogonal to the search direction, it is proposed to alternate the initial search direction between horizontal and vertical scanning. This produces a modest increase in distortion compared to the full 2D search, with a complexity reduction in excess of an order of magnitude.

1. INTRODUCTION

The key contribution of this work is in evaluating low complexity alternatives for implementing the Matching Pursuits (MP) algorithm for video coding without unacceptable loss of fidelity. MP iteratively decomposes a signal into a series of atoms selected from an over-complete set of non-orthogonal dictionary basis functions and was originally proposed for digital audio by Mallat and Zhang [1]. In image and video compression, MP has mainly found application in the coding of DFD images, first demonstrated by Neff and Zakhor [2]. Applied within a standard codec, such as H.263, it has been demonstrated to achieve significant improvements in fidelity [2].

MP is, however, very costly of computational resources. A full two-dimensional MP decomposition requires repeated inner product searches of all basis functions in the dictionary at each image position. At each step an "atom" specifying the position, sign, amplitude and basis function index is encoded, transmitted and subtracted from the image before the process repeats.

Although the code of an atom costs many bits, MP succeeds because one atom specifies many pixels in the decoded image compared to DCT or wavelet transform coding [3].

Although MP is highly efficient in terms of coding performance, its computational cost remains prohibitive and the development of reduced complexity schemes is an area of emerging research interest. In [2] Neff and Zakhor addressed this problem by the use of separable Gabor basis functions and performing the full inner product search only in the 16×16 window with greatest energy. Moschetti *et al* [7] go further and simply select the pixel position in the DFD image with the highest energy as the point at which to perform a full 2D inner product search, an approach that is very computationally efficient but introduces significant additional distortion.

Alternative work has addressed the problem of dictionary design, resulting in smaller dictionaries that can be efficiently constructed from elementary basis functions [4, 5]. More recently, the Schwarz inequality has been used to exclude a significant number of basis functions from the final inner product search, producing a 20%-35% reduction in the separable inner product calculations [6].

In this paper, we investigate reduced complexity MP algorithms based on separable basis functions that offer a favourable trade-off between distortion and complexity. We show that using a 1D inner product to select the pixel position at which to perform an additional 1D search in the opposite direction greatly reduces the complexity with only a modest increase in distortion. Further, a novel Alternating 1D Search algorithm is demonstrated to have significantly better distortion performance than the 1D search with a complexity that is between 13 and 16 times less than that of the Full 2D separable search.

2. REDUCED COMPLEXITY SEPARABLE MP ALGORITHMS

The basic theory of matching pursuits is well known and can be found in the literature [1-5]. We base our discussion of the complexity of our new algorithms on an initialisation cost followed by the three-step per atom model of [5]:

Initialise compute a full set of inner products
Repeat

1. *Find atom*. Full search or reduced complexity strategy.
2. *Atom Update*. Subtract quantized atom from image
3. *Repair*. Recompute required inner products only in atom footprint.

Until distortion or bit rate criterion met

The *Atom Update* step contributes negligible cost and so is not considered in our calculations. Assuming that the *Repair* step uses the same algorithm as step 1, we first estimate the complexity of a full 2D separable and Maximum Energy algorithms, as these represent the two extremes of complexity in terms of the information used to select the location of the atoms.

2.1. Full 2D separable inner product search

Here we take the initialisation cost as the number of multiply and accumulate operations required for a separable 2D inner product between the entire DFD image and all normalized 2D basis functions:

$$Init_{2DSep} = p(b\,w + b^2\,w\,) \quad (1)$$

where I = initialisation complexity, w = atom width and height (assuming all atoms have the same width and height), p = number of pixels in the region and b = number of 1D basis functions in the dictionary.

The *Find Atom* step requires a comparison between all basis functions at each pixel position giving a cost per atom of

$$FindAtom_{2DSep} = p \times b^2. \quad (2)$$

For the *Repair* step the inner product must be recomputed for an atom footprint of $(2w-2)^2$ pixels:

$$Repair_{2DSep} = (2w-2)^2(b\,w + b^2\,w). \quad (3)$$

2.2. Maximum energy search

Using the energy in the DFD image requires just 1 multiplication per pixel for initialisation:

$$Init_{Energy} = p. \quad (4)$$

Find Atom requires p comparisons to determine the pixel position with the greatest energy at which all inner products must be evaluated, giving

$$FindAtom_{Energy} = p + b^2 \times w^2 \quad (5)$$

and a repair cost of

$$Repair_{Energy} = (2w-2)^2. \quad (6)$$

2.3. Reduced complexity 1D search

The Full 2D separable MP algorithm requires all inner products to be evaluated in order to find the position to place each basis function and no additional calculations to determine which basis function to use. At the other extreme, the maximum energy method needs no inner products to find the position and an inner product with all basis functions to select which to use.

In this work we use a 1D inner product between a 1D scan of the DFD image and the set of normalized 1D basis functions to determine the atom position. The 1D initialisation cost is then

$$Init_{1D} = p \times b\,w. \quad (7)$$

The *Find Atom* step requires $p \times b$ comparisons to determine the pixel position and the first basis function. Rather than evaluating the inner products of all dictionary entries containing this first basis function, the second basis function is found by a further 1D inner product search in the orthogonal direction, giving a total (reduced) cost of

$$FindAtom_{1D} = p \times b + b \times w = b(p + w). \quad (8)$$

We have found that finding the second basis function using (8) produces a virtually identical performance to a Full 2D inner product search at the pixel position with all 2D basis functions containing the first basis but with reduced complexity.

The repair cost is also reduced by a factor of b in comparison to the full search

$$Repair_{1D} = (2w-2)^2\,bw. \quad (9)$$

2.4. Alternating 1D search

The final algorithm performs a 1D inner product search, as described in Section 2.3, in alternating horizontal and vertical directions to identify the atom position and first 1D basis function, followed by a 1D search in the opposite direction. This is intended to capture significant structures in horizontal and vertical directions, thus improving the convergence. Compared to the 1D search, the *Initialisation* and *Repair* complexity are doubled while the *Find Atom* is unchanged, giving:

$$Init_{Alt1D} = 2(\,p \times b\,w), \quad (10)$$

$$FindAtom_{Alt1D} = b(p + w), \quad (11)$$

	Initialise	Find Atom	Repair
Full 2D	8.52+e08	4.06+e07 α	1.21+e07 α
1D	4.06+e07	2.03+e06 α	5.78+e05 α
Alt. 1D	8.11+e07	2.03+e06 α	1.16+e06 α
Energy	1.01+e05	2.61+e05 α	1.44+e03 α

Table 1: Theoretical complexity of the algorithm coding α atoms using a codebook of size 20 on a CIF image.

$$Repair_{Alt1D} = 2 \times (2w - 2)^2 \, bw. \qquad (12)$$

2.5. Complexity Comparison

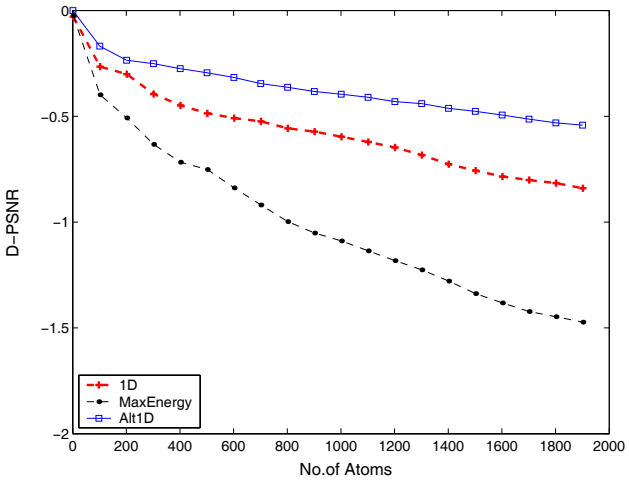
To give a feeling for the relative complexities of the searching options when dealing with a real image, the complexities for a full CIF image (352 x 288 pixels) coded with a separable dictionary of 20 1D basis functions of width 20 pixels are given in Table 1. In practice this is a realistic set of values, except that the widths of the basis functions will normally vary.

Although the initialisation costs are generally higher than the Find Atom and Repair steps, in practice thousands of atoms are coded and the per atom cost of the two latter stages will dominate the overall complexity.

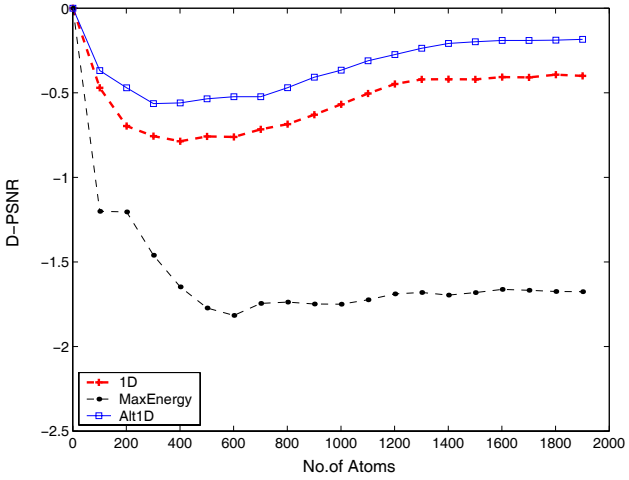
3. CODING EXPERIMENTS

To estimate the distortion introduced by finding sub-optimal atoms at sub-optimal locations, CIF-sized DFD frames from the Stefan, Foreman and Mobile sequences were encoded using 1 to 2000 atoms by each method and the PSNR was recorded. The basis dictionary and quantisation method used were those of [2] and the bit cost per atom is assumed to be constant. Figure 1 shows the additional distortion produced by the 1D Search, Alternating 1D Search and Maximum Energy search, when compared to the optimal Full 2D search.

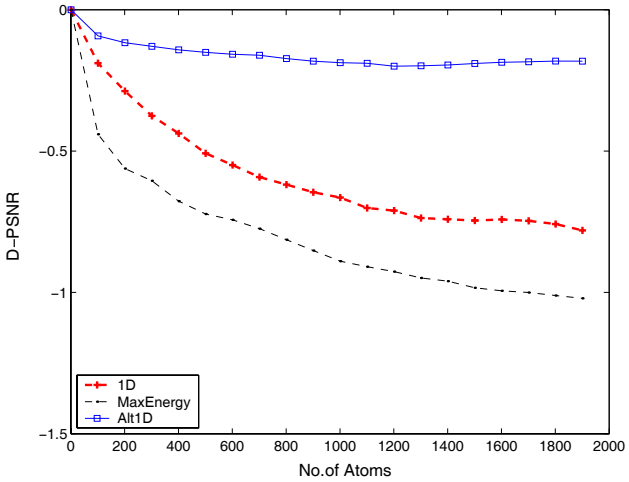
The Maximum Energy Search introduces the greatest distortion for all sequences and for realistic numbers of atoms (>500) has a degradation well in excess of 1 dB. Although the 1D search performs better, it still has a loss of 0.5 dB from 500 atoms upwards, increasing with the number of atoms for 2 of the 3 sequences. The distortion performance of the Alternating 1D search is significantly better than the other search methods. On the Mobile sequence its distortion is within 0.2 dB of the optimal Full 2D search. For the Foreman sequence the distortion peaks at –0.56 dB at 400 atoms but then improves to within 0.2 dB for 1400 atoms upwards. The distortion is worst on the Stefan sequence but still considerably better than the two other sub-optimal search strategies.



(a) Stefan frame 24 predicted from frame21



(b) Foreman frame 18 predicted from frame 15



(c) Mobile frame 18 predicted from frame 15

Figure 1. PSNR loss relative to Full 2D searching for DFD frames from 3 CIF sequences. The Full 2D search is 0dB and the graphs show the loss in fidelity using other methods.

	Stefan	Foreman	Mobile
Full2D	1730	427	5970
Alt1D	1990	537	6182
1D	2187	597	7310
Max Energy	2616	851	8120

Table 2: Number of atoms required to achieve a PSNR of 31dB for Stefan, Foreman and Mobile sequences.

Figure 2 shows the complexity of the 4 search methods, calculated using the values in Table 1 and the number of atoms required by each search method to reach a distortion of 31 dB (tabulated in Table 2) for the same 3 sequences as before. For clarity, the computational complexity is plotted on a log scale.

As expected the Full 2D search is the most complex and requires the fewest atoms. The Maximum Energy search is the least complex but requires between 36% and 99% more atoms than the Full 2D search, reducing its complexity advantage and giving lower compression.

The complexity of the 1D Search is only slightly less than that of the Alternating 1D Search. Despite having lower initialisation and repair costs it requires more atoms to reach the requisite distortion at lower compression.

4. DISCUSSION AND CONCLUSIONS

We have assumed throughout that the cost of sending 2D atoms by our various methods is similar. This has enabled us to carry out complexity and PSNR comparisons by coding by each method to the same number of atoms.

Several alternatives for reduced complexity have been investigated involving sub-optimal searching strategies. The novel Alternating 1D Search has been shown to offer a significant reduction in complexity while still producing distortion performance that is tolerably close to that of a full 2D search. In implementing a practical MP coder, this would be our method of choice.

We have also investigated a double 1D search in which the maximum of a 1D horizontal and vertical search is used to locate the atom position, followed by a 1D search in the other direction. However, the PSNR performance of this Double 1D search was only marginally better than Alternating 1D Search and it would require the search direction to be communicated to the decoder. This would violate the assumption of equivalent bit cost per atom, so in practice its distortion at a given bit rate is expected to be higher than the Alternating 1D.

The sub optimal searching and atom selection approach presented is conceptually different from that of Neff and Zakhor [2] and therefore the two techniques could be combined to produce further savings in complexity, albeit with additional distortion.

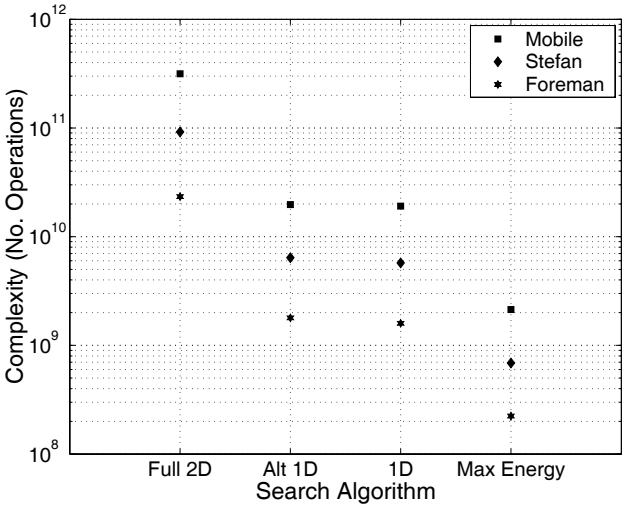


Figure 2. Search Algorithm Complexity to Reach a Distortion of 31 dB. The Alt 1D and 1D methods have similar complexity at this PSNR, but Table 2 shows that the 1D would require 10% to 18% more atoms..

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, Feb. 1997.

[3] 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.

[4] 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, Oct. 2000.

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

[6] B. Jeon, and S. Oh, "Fast Matching Pursuit With Vector Norm Comparison", IEEE Trans. Circuits Syst. Video Technol., 13(4) pp. 338-342, April 2003.

[7] F. Moschetti, L. Granari, P. Vandergheynst and P. Frossard, "New dictionary and fast atom searching method for matching pursuit representation of displaced frame difference", IEEE Int. Conf. on Image Processing, Rochester, USA, vol. III, pp. 685-688, Sept. 2002.