A NON-SEPARABLE LIFTING APPROACH FOR 3D IMAGE COMPRESSION

David Montgomery*, Abbes Amira and Fionn Murtagh
School of Computer Science
Queen's University Belfast
*d.montgomery@qub.ac.uk

ABSTRACT

The three-dimensional wavelet transform is extremely important for image and video processing. This paper presents a number of three dimensional non-separable wavelet transforms each of which are obtained using a lifting scheme. The performances of the new wavelets in terms of psycho-visual reconstruction quality and peak signal-to-noise ratio are compared to tensor product wavelets in a lossy image compression application.

1. INTRODUCTION

During the last decade wavelets have taken there place at the forefront of research for the development of image and video processing applications. These wavelet based approaches have outperformed existing strategies in many areas including image denoising, segmentation and most notably compression. Results have shown that wavelets offer improved compression ratios, and produce less visually detrimental artefacts when compressing at low bit rates than previous approaches such as the Discrete Cosine Transform (DCT), and for this reason wavelet based coding is being introduced into emerging standards such as JPEG 2000 [1].

Leading researchers have exposed a wide variety of wavelet bases with their own individual mathematical properties suitable for different applications. Authors such as Shapiro [2], Pearlman [3,4], and Taubman [5] have lead to the development of novel, efficient algorithms for wavelet based encoding which exploit the hierarchical structure of the transform, and also most recently, Sweldens, Daubechies and Calderbank have developed an alternate method for the construction of wavelets using an approach known as lifting [6,7]. This method does not require auxiliary memory unlike the computation of the discrete wavelet transform through convolution. It is the aim of this paper to use a combination of these ideas to develop a number of novel non-separable wavelet transforms for use in the field of three-dimensional image compression.

The remainder of this paper is structured as follows. Section 2 begins by introducing the one-dimensional lifting based LeGall 5/3 wavelet transform. Section 3 outlines the proposed framework for non-separable lifting. In section 4 lossy coding results for both separable and non-separable wavelet transforms are analysed. Finally we conclude with a brief summary of our results and some closing remarks in section 5.

2. THE LIFTING SCHEME

2.1. Introduction

The lifting scheme is an efficient approach which is more flexible than the convolution methodology and can be used to define a wavelet basis on an interval or on an irregular grid without using the concept of the Fourier transform. All classical wavelets can be generated using the lifting strategy [6,7].

Initially for the one dimensional case the *lazy wavelet transform* is applied to the signal. This process splits the samples into even and odd components. The resulting two sets are closely correlated and so it is only natural that given one set, e.g. the odd, one can be a good predictor for the other set. Although this prediction will normally be satisfactory it will not be necessary to keep this information in both sets, we store only the part of the set which is not predictable or the prediction error. This is known as *dual lifting*. This is useful; however, we loose some vital properties such as the mean value of the signal. This mathematical property can be restored using what is known commonly as a *primal lifting* step, where the second set is updated with data from the new subset. Another interesting and useful property of the lifting scheme is the ease of inversion. The inverse transform can be calculated by simply reversing the order of the operations and inverting the signs in the lifting steps.

2.1. Mathematical Background

To date two wavelet bases have been specified for inclusion into the JPEG 2000 ISO standard. These are the LeGall 5/3 filter for lossless compression and the Daubechies 9/7 filter for lossy compression. JPEG 2000 will use lifting for wavelet construction.

Mathematical formulations for the construction of wavelets using the 5/3 LeGall filter [6] are shown below where x is the initial input signal and $s_0[n] = x[2n]$ and $d_0[n] = x[2n+1]$.

$$\begin{cases} d[n] = d_0[n] - \left\lfloor \frac{1}{2}(s_0[n+1] + s_0[n]) \right\rfloor \\ s[n] = s_0[n] + \left\lfloor \frac{1}{4}(d[n] + d[n-1]) + \frac{1}{2} \right\rfloor \end{cases}$$

The flooring function is required for the mapping of integers to integers for lossless compression. Using this lifting scheme special care must be used at the signal boundaries for the wavelets to explicitly live on the discrete set where the data is defined.

3. PROPOSED APPROACH

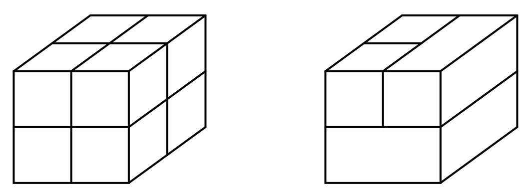
3.1. Introduction

Typically for wavelet based multi-dimensional compression a tensor product of one-dimensional wavelet transforms are used. However, the same lifting description can be applied directly for multi-dimensional data. We propose three separate three-dimensional non-separable transforms based on the Haar, LeGall 5/3 and Daubechies 9/7 biorthogonal filters.

The motivation for using non-separable wavelet transforms for image compression is to attempt to minimise or eliminate any blocky artefacts commonly associated with tensor product based wavelet transforms. These visually detrimental artefacts are particularly evident in the case of Haar wavelets as they are characteristic functions of squares. Non-separability allows for more degrees of freedom in design and has filters which are better adapted to the human visual system.

3.2. Non-Separable Framework

Our approach is inspired by the work of Uytterhoeven and Bultheel who applied a similar construction based on the Red-Black Gauss-Seidel technique to two-dimensional data for use in the field of image denoising [8]. However, for the case of three-dimensional data a number of problems arise. We must have at least a three step decomposition for the Mallat pyramidal representation to hold. To obtain this the molecular structure at each lifting stage must change which is certainly not the case for two-dimensional data. The technique initially splits the data into two subsets. Unlike the one-dimensional approach where even-odd splitting occurs, in the three-dimensional case checkerboard splitting occurs with red and black cubes. The next step is to predict values of the red subset from its nearest horizontal and vertical neighbours in the black set. The elements in the black subset are replaced by the prediction errors. Next we update the red subset from the new black subset values to preserve the average value of the data. To achieve a lower resolution approximation of the original data it is necessary to consider the diagonal neighbours also. For this we must partition our new red subset into two new sets, say blue and yellow. Similarly to the previous resolution level the yellow subset elements are predicted using linear interpolation based on nearest neighbours. The blue elements are then updated using the yellow coefficients to preserve the average value. To obtain the correct decomposition structure the final step involves splitting the blue set into green and white subsets. The result of this process is a green set representing a low resolution version of the original three-dimensional image, and a white set representing the detail information (cfr. 2). The next step of the three-dimensional decomposition will be performed only on the green set. The decomposition will use periodic symmetric extension at the signal boundaries as the wavelet decompositions are derived from 1d filters which are symmetric [1]. The following section outlines the procedure necessary for decomposition using the three-dimensional non-separable LeGall 5/3 filter. The Haar and Daubechies cases mentioned later can also be derived from this framework.



(a) (b)

Figure 1: Subband structures after the 3D;
(a) Separable decomposition.
(b) Non-separable decomposition

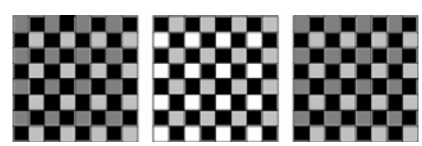


Figure 2: Splitting procedure for 3 slices.
(a) black squares – black subset
(b) light grey squares – yellow subset
(c) dark grey squares – green subset
(d) white squares – white subset

3.3. The LeGall 5/3 Filter

Horizontal/Vertical Lifting

1. The 3D image is split into red cubes and black cubes.

2. The black cubes are predicted by linear interpolation of the six neighbouring red cubes.

$$x_{i,j,k} = x_{i,j,k} - \left[\frac{1}{6}(x_{i-1,j,k} + x_{i,j-1,k} + x_{i,j+1,k} + x_{i+1,j,k} + x_{i,j,k+1} + x_{i,j,k-1}) \right]$$

3. The red cubes are then updated using the previously calculated black cubes to preserve the mean.

$$x_{i,j,k} = x_{i,j,k} + \left[\frac{1}{12}(x_{i-1,j,k} + x_{i,j-1,k} + x_{i,j+1,k} + x_{i+1,j,k} + x_{i,j,k+1} + x_{i,j,k-1}) + \frac{1}{2} \right]$$

Diagonal Lifting I

4. Partition the red cubes into blue and yellow subsets.

5. The yellow cubes are predicted by linear interpolation of the eight neighbouring blue cubes.

$$x_{i,j,k} = x_{i,j,k} - \left[\frac{1}{8}(x_{i+1,j,k+1} + x_{i-1,j,k+1} + x_{i+1,j+1,k} + x_{i+1,j-1,k} + x_{i-1,j+1,k} + x_{i-1,j-1,k} + x_{i+1,j,k-1} + x_{i-1,j,k-1}) \right]$$

6. The blue cubes are then updated using the previously calculated yellow cubes to preserve the mean.

$$x_{i,j,k} = x_{i,j,k} + \left[\frac{1}{16}(x_{i+1,j,k+1} + x_{i-1,j,k+1} + x_{i+1,j+1,k} + x_{i+1,j-1,k} + x_{i-1,j+1,k} + x_{i-1,j-1,k} + x_{i+1,j,k-1} + x_{i-1,j,k-1}) + \frac{1}{2} \right]$$

Diagonal Lifting II

7. Partition the blue cubes into green and white subsets.

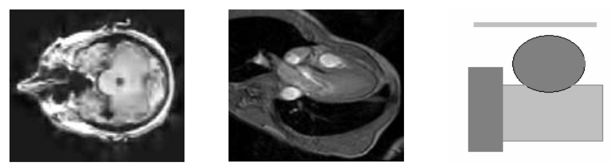
8. The white cubes are predicted by linear interpolation of the four neighbouring green cubes and the green cubes are then updated using these white values to preserve the mean.

4. RESULTS

In this case we use SPIHT encoding to discuss the suitability of both separable and non-separable wavelet transforms for three-dimensional image compression.

SPIHT is a wavelet based image compression coder which is based on three underlying principles: (1) the exploitation of the hierarchical structure of the wavelet transform, by using a tree based organization of the coefficients; (2) partial ordering of the transformed coefficients by magnitude, with the ordering data not explicitly transmitted but recalculated by the decoder; and (3) ordered bit plane transmission of refinement bits for the coefficient values. This leads to a compressed bitstream in which the most important coefficients (regardless of location) are transmitted first, the values of all coefficients are progressively refined, and the relationship between coefficients representing the same location at different scales is fully exploited for compression efficiency [3,4].

The non-separable Haar, LeGall 5/3 and Daubechies 9/7 biorthogonal filters have been tested using a number of 8bpp monochromatic images where the first two are magnetic resonance image data and the third image is a simulated three-dimensional data cube. A three level wavelet decomposition in employed in this study with no partitioning.



(a) (b) (c)

Figure 3: Slice of each data set. (a) Human brain, (b) Human heart and (c) Simulated imagery.

The tables that follow show a comparison between both approaches, in terms of the rate distortion level versus the peak signal to noise ratio for lossy compression performance. The definition of the peak signal to noise ratio as employed in this study is,

$$PSNR = 20\log_{10}\left(\frac{2^P - 1}{\sqrt{MSE}}\right)$$

where,

$$MSE = \frac{1}{LMN}\sum_{i=0}^{L-1}\sum_{j=0}^{M-1}\sum_{k=0}^{N-1}(\hat{x}[i,j,k] - x[i,j,k])^2$$

$x[i,j,k]$ is the original three-dimensional image with dimensions L x M x N and P bpp and $\hat{x}[i,j,k]$ is the reconstructed image.

Results show that in terms of PSNR and reconstruction quality the inclusion of a diagonal lifting step into the decomposition is beneficial. The psycho-visual quality of the reconstructed image is substantially improved with less visually detrimental effects being apparent. In the case of the two MR images banding is reduced along with the number of visual artifacts as shown in figure 4.

Bit Rate (bpp)	IMAGE-BRAIN PSNR(dB)					
	Separable			Non-Separable		
	Haar	LeGall 5/3	Daub 9/7	Haar	LeGall 5/3	Daub 9/7
0.125	22.50	25.10	25.62	22.49	25.13	25.87
0.250	25.01	27.75	28.39	25.03	28.87	29.01
0.500	28.18	33.16	32.06	28.27	33.27	33.57
1.000	31.67	36.11	36.14	32.10	36.99	37.03
2.000	35.23	42.24	40.97	36.85	42.77	41.15

Table 1: Results for 3D Brain scan.

Bit Rate (bpp)	IMAGE-HEART PSNR(dB)					
	Separable			Non-Separable		
	Haar	LeGall 5/3	Daub 9/7	Haar	LeGall 5/3	Daub 9/7
0.125	23.22	26.22	26.28	23.16	26.34	26.36
0.250	26.17	28.62	29.14	26.14	28.81	29.33
0.500	29.32	34.11	34.16	29.33	34.56	34.63
1.000	31.96	37.81	37.99	32.40	37.98	38.51
2.000	36.10	43.49	42.18	36.75	43.60	43.01

Table 2: Results for 3D Heart scan.

Bit Rate (bpp)	IMAGE-SIMULATION PSNR(dB)					
	Separable			Non-Separable		
	Haar	LeGall 5/3	Daub 9/7	Haar	LeGall 5/3	Daub 9/7
0.125	25.08	26.15	26.22	24.41	26.11	25.87
0.250	26.65	28.03	29.67	25.55	27.65	29.01
0.500	31.18	31.99	33.42	30.10	30.45	33.57
1.000	34.89	36.12	37.89	33.34	35.67	37.03
2.000	40.27	41.04	41.03	38.90	40.78	40.04

Table 3: Results for 3D simulated image.

(a) Original image -

(b) Separable at 0.5bpp

(c) Non-separable at 0.5bpp

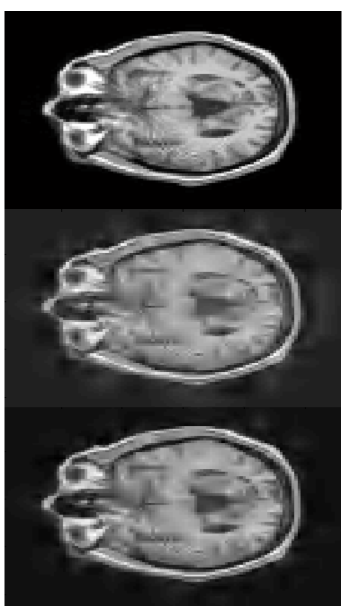


Figure 4: Reconstruction quality of brain image (slice 35) after separable and non-separable decompositions with LeGall 5/3 filter.

Reconstruction quality is particularly good at high bit rates, however at lower distortion levels the LeGall and Haar non-separable decompositions introduce some visual artifacts. This may be due to the smoothness of their equivalent 1d filters. This was not evident when using the Daubechies 9/7 non-separable wavelet transform. The results we acquired for the simulated image were not as successful. This is due to the nature of the simulated imagery. Its structure is blocky and therefore the Haar transform performed reasonably, particularly for the separable case.

5. CONCLUSIONS

Several reversible integer-to-integer three-dimensional non-separable wavelet transforms have been introduced. They have been compared with their separable counterparts on the basis of their lossy compression performance using the set partitioning in hierarchical trees encoding algorithm. Results indicate that the three stage non-separable wavelet transforms perform more efficiently for compression purposes in terms of the quality of the reconstructed image.

6. REFERENCES

[1] JPEG 2000 final committee draft ISO/IEC FCD15444-1 : 2000 (V1.0, 16 March 2000)

[2] J. M. Shapiro, “Embedded image coding using zerotrees of wavelets coefficients,” *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445-3462, Dec. 1993.

[3] A. Said and W. A. Pearlman, “A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp.243-250, June 1996.

[4] Y. S. Kim and W. A. Pearlman, "Stripe-Based SPIHT Lossy Compression of Volumetric Medical Images for Low Memory Usage and Uniform Reconstruction Quality," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Istanbul, Turkey, June 2000.

[5] David Taubman, “High Performance Scalable Image Compression with EBCOT,” *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158 – 1170, July 2000.

[6] A. R. Calderbank, Ingrid Daubechies, Wim Sweldens and Boon-Lock Yeo, “Lossless Image Compression Using Wavelet Transforms That Map Integers To Integers.” *Proceedings of the IEEE International Conference on Image Processing*, Vol. I, pp. 596599, October 1997.

[7] Wim Sweldens, “The Lifting Scheme: A New Philosophy in Biorthogonal Wavelet Constructions.” *Wavelet Applications in Signal and Image Processing III*, pp. 68--79, Proc. SPIE 2569, 1995.

[8] G. Uytterhoeven and A. Bultheel, “The red-black wavelet transform”, *Proceedings of IEEE Benelux Signal Processing Symposium*, March 1998 (pp. 191--194).