Region of Interest Based Coding of 2-D and 3-D Magnetic Resonance Images

R. Srikanth
Department of Electrical Engineering
Indian Institute of Science, Bangalore
srikanth@ee.iisc.ernet.in

A.G. Ramakrishnan
Department of Electrical Engineering
Indian Institute of Science, Bangalore
ramkiag@ee.iisc.ernet.in

Abstract

We present a lossy to lossless wavelet based image compression scheme with progressive transmission and region of interest coding functionalities for 2D and 3D magnetic resonance images. Context based entropy coding is used to efficiently code significance and sign information of wavelet coefficients. Contexts for 2D and 3D schemes are built using a mutual information based context merging algorithm. The obtained compression performance for the 3D scheme of about 2.1 bits per pixel is comparable to the existing Multidimensional layered zero coding (MLZC) scheme [4].

1. Introduction

3D Magnetic Resonance Image (MRI) data, which contains multiple slices representing a part of a body, requires compression for efficient storage and transmission. Huge amount of such data is generated in hospitals which requires to be stored for future reference and study. Compression of medical data is also required in teleradiology applications where image data needs to be transmitted over the network. Lossless compression, progressive transmission and region of interest (ROI) are important functionalities for a compression scheme. Recent still image compression standard, JPEG 2000 which is a wavelet based scheme, provide such functionalities for 2D images but is out of scope for 3D images. 3D MR images have correlation both within and across the slices. Hence a compression scheme for 3D MR images should exploit this correlation. In this work, we present wavelet based compression schemes with the above functionalities for 2D and 3D MR images. We employ separable 2D and 3D integer wavelet transforms for decorrelating 2D and 3D images respectively. We exploit correlation within the wavelet subbands as against across the subbands as in Embedded Zerotree wavelet (EZW) and Set Partitioning in Hierarchical Trees (SPIHT) [2] coders. This is based on the recent studies of [1], [6] that exploitation of intraband correlation is better than interband correlation. JPEG 2000 which is based on Embedded Bit Coder with Optimal Truncation (EBCOT) of [7] also exploits intraband correlation. We also employ context based coding to efficiently code significance and sign information of wavelet coefficients. Contexts for 2D and 3D schemes are designed on a training data using the mutual information based context merging algorithm of [5]. We compare the results of our 2D and 3D schemes with that of Multidimensional layered zero coding (MLZC) [4].

2. Mask Generation for ROI Coding

A typical MR image consists of two parts:

1. Air part (background)
2. Flesh part (foreground)

The flesh part contains the useful clinical information which needs to be compressed without any loss. On the other hand, the air part does not contain any clinical information. This has been verified with the radiologist we are collaborating with. It is only noise and consumes unnecessary bit budget and impairs the performance of a compression scheme. We generate image masks in such a way that the flesh part is totally included and the pixel values in the air part are made zero. This mask is used as region of interest in our coding scheme. Morphological operations can be effectively used to generate image masks, which contain a value of '1' in the foreground and a value of '0' in the background. The original image is then multiplied with these masks to obtain "background noise free" images while keeping the information in the foreground intact. The algorithm for generating the mask is given below:

1. Binarize the image with a threshold decided by the histogram of the image.
2. Holes may be formed within the foreground. Close these holes using morphological 'closing' operation.
3. Background may contain spurious lines. Use morphological 'erode' operation to remove these lines.
4. The above erosion operation also erodes the boundary of the foreground region. To make sure that the mask spans the entire foreground region, use morphological

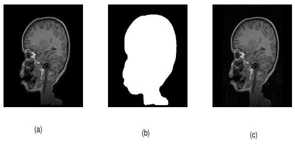

Figure 1. Suppression of background in an MR image using morphological operations. (a) original image (b) the generated mask (c) background suppressed image.

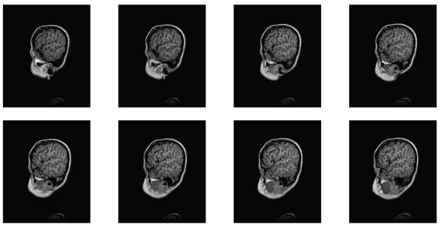

Figure 2. Sample 8 consecutive slices of Brain MR Images with 1mm thickness.

'thickening' operation to thicken the boundary of the foreground region.

5. Multiply the original image with the resulting binary mask.

Figure 1 shows an MR image, its mask and the image obtained after multiplication with the mask.

3. Integer Wavelet Transform

Lossless compression is not possible with conventional wavelet transforms because they map integer-valued image data to real-valued wavelet coefficients. For lossless image coding, we need a transformation which results in integer coefficients. This can be achieved by integer wavelet transforms. These transforms can be built using lifting schemes as explained in [3]. We need 2-D transforms for images which can be easily obtained by applying 1-D transform on rows and columns separately. Such transforms are called separable transforms. Likewise 3-D wavelet transform can be obtained by applying 1-D transform to the three dimensions sequentially. Figure 2 shows sample MR slices to be compressed and Fig. 3 shows 3-D separable integer wavelet decomposition of the above MR Images with 2 levels in spatial and 2 levels in temporal (across slices) domain

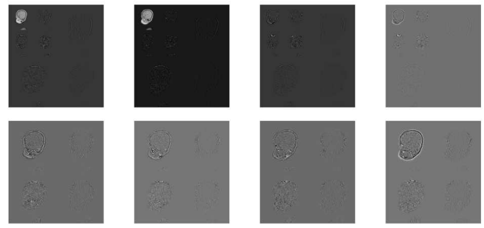

Figure 3. 3-D Separable Integer Wavelet Decomposition of the above MR Images with 2 levels in spatial and 2 levels in temporal (across slices) domain.

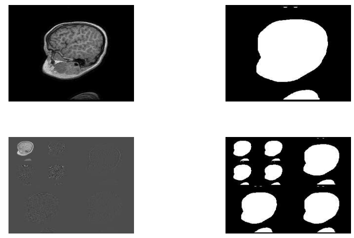

Figure 4. (a) MR image (b) its corresponding mask, bright regions show region of interest (c) 2-level wavelet decomposition of (a) and (d) 2-level decomposition of mask in (c) showing regions of interest in each subband.

4. Coding Scheme

4.1. 2D Scheme

We propose a new Region of interest (ROI) based coding scheme for 2D and 3D images. We exploit intraband correlation instead of interband correlation as in SPIHT or EZW algorithms. We first give the algorithm for 2D images which can be easily extended to 3D images. We use a fixed-block to exploit intraband correlation. The region of interest coding can be easily incorporated in intraband schemes by using the masks generated in section II. We use this mask to code only the regions of interest. In our work, we define ROI as the flesh part. To be able to identify the corresponding regions of interest in the wavelet domain, we also decompose the mask. Figure 4 (a) and (b) respectively show a MR image and its corresponding mask and (c) and (d) respectively show a 2 level wavelet decomposition of MR image and 2 level decomposition of mask. The ROI for each subband is clearly delineated by the decomposed mask.

1. Apply n-level 2-D separable integer wavelet transform to the given image and label all background pixels as "do not care" (DNC) symbols.
2. Tile the wavelet transformed image into $v \times v$ lattices.

3. For each lattice k having at least one non-DNC symbol, find the maximum absolute value w_{max} and ignore lattices with all DNC symbols. Let $T_k = \lfloor log_2 abs(w_{max}) \rfloor$ be the threshold of the kth lattice. Store these values in an array th.
4. Set the maximum of all the thresholds as the global threshold, T_g.
5. Scan the wavelet image starting from the lowest frequency band to the highest frequency band in zig zag manner. In each band, the lattices are scanned in a raster order. Lattices with all DNC elements are not scanned.
6. If $T_k < T_g$, the lattice is insignificant with respect to T_g and a '0' is recorded in the list $lstb$. If $T_k >= T_g$, the lattice is significant and the decoder needs to be informed of this. If this lattice is first time significant, a '1' is recorded in the list $lstb$. If the lattice is already significant, no information is sent to the decoder, since this lattice will be significant for the future lower global thresholds.
7. If the lattice is significant, check for the significance of each non-DNC coefficient in raster scan order. If the coefficient is significant, a '1' is appended to the significant list $lis1$ otherwise a '0' is appended. If the coefficient is positive significant, a '0' is appended to the sign list $lis2$ or a '1', if it is negative significant.
8. *Refinement Pass:* The current bit of the significant coefficients (at the previous threshold) is sent to the decoder through the list $lis3$.
9. Stop if the required bitrate is met or $T_g = 0$, otherwise, set $T_g = T_g - 1$ and go to step 5.

The lists $lis1$, $lis2$, $lis3$ and $lstb$ can be further losslessly compressed by employing arithmetic coding. The mask information is sent as a side information to the decoder by using differential coding of the coordinates of the contour of the mask. At the decoder, the mask is simply generated by filling the regions with in the contour. The side information required is about 0.01 bits per pixel (*bpp*). Since the most important coefficients (with higher thresholds T_k) are coded before the least important ones (with lower thresholds), there will be an ordering of wavelet coefficients resulting in progressively transmittable bit stream. The decoder can stop at any step and reconstruct the image that is best at that level. The image reconstructed at $T_g = 0$ will be identical to the original image and hence results in lossless compression. At low thresholds, more blocks will be insignificant and a high value of v results in good compression. But at low thresholds, more coefficients will be significant and if v is high, the cost for sending individual significance of each non-DNC coefficient is high. This impairs compression performance at low thresholds. Hence there is a trade off in the choice of lattice size v. We choose $v = 4$ which is based on our experimentation with different values of v.

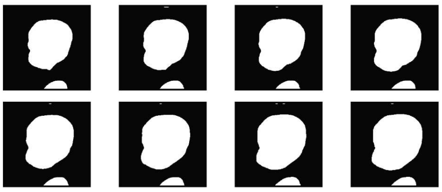

Figure 5. Masks showing regions of interest of MR images in 2

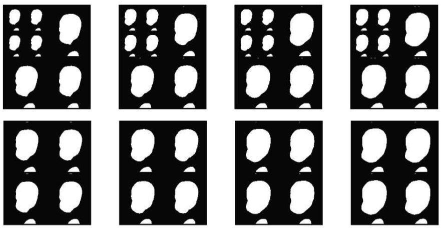

Figure 6. 2-level decomposition of masks of figure 5 showing regions of interest in each subband

4.2. 3D Scheme

The above scheme can easily be extended to volumetric MR images. The idea is to exploit both intraframe and interframe correlation in these images and hence to achieve higher compression ratios. The motivation behind applying 3-D transform is to decorrelate the images across the slice direction in addition to spatial decorrelation so that decorrelated images can be efficiently coded. Region based coding is achieved by decomposing the masks similar to the 2D case. Figure 5 shows masks of MR images of the Fig. 2. Figure 6 shows 2 level decomposition of masks along both slice and spatial directions. The decomposed masks in various subbands delineate regions of interest. Tile the wavelet images by cuboids of size $v \times v \times vt$ (typically $v = 4$ and $vt = 2$). Apply the algorithm described in the previous section to obtain a progressive bit stream. The coding performance can be further improved by using context-based entropy coding of significance and sign lists.

4.3. Context-based Entropy Coding

Arithmetic coders are not only more efficient than Huffman coders but also have the advantage that source modeling is independent of coding. If the symbols to be coded are assumed to be independent of each other, then the lower bound on the number of bits per symbol is given by the first order entropy of the source to be coded. Arithmetic coders asymptotically achieve this bound. But in practice, the symbols to be coded are not independent of each other. For example, if a wavelet coefficient w is insignificant, then we can expect the neighbouring coefficients are also insignificant. Also, significance of a coefficient can be predicted using the significance information of its neighbours. Similar observations can be drawn for signs of the coefficients. Using this correlation, one can reduce the lower bound on the achievable bit rate using context-based coding. Let S be the source symbol to be coded and $H(S)$ denote the first order entropy. Let C be the context or conditioning event and $H(S|C)$ be the conditional entropy of S given the context C. Then one can show that $H(S) \leq H(S|C)$ with equality if and only if S and C are independent. The quantity $I(S;C) = H(S) - H(S|C)$ is called mutual information. The greater the mutual information, the lower is the conditional entropy and better is the coding performance. One can maximize the mutual information by carefully choosing the contexts. Generally contexts are defined based on the already coded discrete symbols . Therefore the number of possible contexts are finite. Theoretically, the mutual information can be maximized by choosing as large number of contexts as possible. But in practice, the number of symbols to be coded are finite because of which some contexts may have few symbols. Since the probabilities of symbols required for entropy coding are estimated on fly, lesser number of symbols result in poor estimates of probabilities. This increases the coding cost. This problem is called as context dilution. To avoid this, we need to limit the number of contexts without much decrease in mutual information. The process of choosing a fixed number of contexts from the possible contexts is called as context quantization.

A common approach is to project the higher dimensional context onto a linear space (real line) and find appropriate quantization levels on the real line by optimizing certain cost function [8]. Qunatization levels are generally obtained either by using dynamic programming or by employing Lloyd-Max quantizer. These schemes are suboptimal because projection of higher dimensional contextual information onto a linear space results in information loss. In this work, we use the method proposed by [5]. This method operates directly on contexts C for finding the required number of contexts. This method is based on iteratively merging pairs of contexts which result in minimum mutual information reduction. This scheme is based on the result that merging of contexts result in decrease of mutual information. This scheme is used to form 2D and 3D contexts for coding the significance list $lis1$. We employ the contexts for sign list $lis2$ used in EBCOT. We first explain the context merging algorithm.

Let Y be the random variable representing the data symbols to be coded and let X be the context. The mutual information is defined as

$$I(X;Y) = \sum_x \sum_y p(x,y) log_2 \frac{p(x,y)}{p(x)p(y)} \qquad (1)$$

where, $p(x,y)$ is the joint probability distribution and $p(x)$ and $p(y)$ are respective marginal distributions of the random variables X and Y. Let P be the initial number of contexts and F be the required number of contexts. The context merging algorithm described in [5] is as follows:

1. Compute the joint probability distribution $p(x,y)$ from the training data.
2. Find the contexts x_i and x_j such that after combining them, the reduction in $I(X;Y)$ is minimum. This is accomplished by pair wise search.
3. Merge x_i and x_j into one context.
4. Set $P = P - 1$. Repeat steps 2 and 3 until $P < F$.

In [5], this algorithm was used for finding contexts for zero coding (ZC) primitive in JPEG2000. For each wavelet coefficient, eight first order neighbours are used to form a context. Each element assumes 1 if it is significant wrt a threshold or 0 if it is insignificant. Therefore, there will be 256 possible contexts. The above context merging algorithm is used on a training data to reduce the number of contexts. It was found that the coding performance with 3 or 4 contexts was as good as 9 JPEG2000 coding contexts. We use similar strategy to define contexts for significance map coding for 2D and 3D schemes. In the 2D scheme, we use eight first order neighbours to define a context. Let $w(i,j)$ be a wavelet coefficient whose significance wrt to a threshold T_k needs to be coded. Let i,j be the coordinates of the coefficient. The first order neighbourhood of $w(i,j)$ denoted by $N(i,j)$, is given by:

N(i,j) = [w(i-1,j-1), w(i-1,j), w(i-1,j+1), w(i,j-1), w(i,j+1), w(i+1,j-1), w(i+1,j), w(i+1,j+1)]

Each element of $N(i,j)$ is set to 0 or 1 depending on whether its magnitude is less than or equal to T_k. The first four elements of $N(i,j)$ are scanned before $w(i,j)$ and hence they form a causal neighbourhood. The last four elements are scanned after $w(i,j)$ and they form noncausal neighbourhood. To be able to identify the contexts at the decoder, we use threshold the T_k for the causal elements and T_{k+1} for the noncausal neighbourhood. Since the significance wrt to previous threshold T_{k+1} is already known to the decoder in the case of noncausal neighbourhood. We use any eight images from the database as training images and form 256 contexts and estimate joint and marginal probabilities of symbols 0 and 1. We then apply the above context merging algorithm to reduce the number of contexts to 8. We also use the same contexts for the other MRI dataset called MR-MRI data set.

For the 3D scheme, we define 3D context as combination of 8 first order neighbours and two elements in the previous and future frames at the same spatial location. Let $w(i,j,l)$ be the wavelet coefficient whose significance needs to be coded. We define 3D neighbourhood $N(i,j,l)$ of the location i,j,l as follows:

N(i,j,l) = [w(i-1,j-1,l), w(i-1,j,l), w(i-1,j+1,l), w(i,j-1,l), w(i,j+1,l), w(i+1,j-1,l), w(i+1,j,l), w(i+1,j+1,l), w(i,j,l-1), w(i,j,l+1)]

Again $w(i,j,l-1)$ belongs to the causal neighbourhood and $w(i,j,l+1)$ belongs to the noncausal neighbourhood and the thresholds are used accordingly so that decoder will also be able to recover the contexts. Now there will be 1024 possible contexts from which we select 8 contexts using the context merging algorithm. We use any 8 consecutive MR images from the data set as training images to built the required contexts. These contexts are used for the two data sets. We do not use any contexts for refinement list $list3$ as the statistical redundancy is small for refinement bits.

5. Results and Discussion

We apply the 2D and 3D algorithms on 256×256, 8 bit saggittal MR images provided by NIMHANS (NIMHANS data set) and standard MR-MRI data set of [4] . We first generate the masks required to define the regions of interest using the mask generation algorithm defined in section 2. In this work, we assume the "flesh" part as the region of interest. We do not code the background image which does not contain any useful clinical information. We send the coordinates of the contour of the mask to the decoder as a side information. The coordinates are coded using differential coding and requires about 0.01 *bpp* to send the mask information. At the decoder, the mask is reproduced by filling the region within the contour by $1's$. We compare the performances of coding schemes with and without region of interest. We apply different integer wavelet transforms for 2D and 3D schemes and compare their performances. In the 3D scheme, we use the same filter for both spatial and axial (across the slices) decomposition. We built contexts for 2D and 3D schemes using 8 successive images from one data set. We then use these contexts for other images. In the 3D case, we can apply the compression scheme on the whole volume. But this does not allow access of particular images within the volume. Also, large memory size is required to buffer all the images in the volume. A simple solution to both of these problems is to apply 3D scheme on a group of images (GOI). We define group sizes of 8 and 16 and compare their performances. Finally, we compare the performances of these schemes with the MLZC and 3D EZW schemes of [4] on the data set MR-MRI.

5.1. Performance evaluation for different filters

Table 1 gives average lossless compression results in bits per pixel (bpp) for 2D and 3D schemes with and without context based coding on the NIMHANS dataset. We apply 2-level spatial decomposition for 2D scheme and 2-level spatial and axial decomposition for 3D schemes. The group size in the 3D scheme is 8. We fix block sizes as 4×4 and $4 \times 4 \times 2$ for 2D and 3D schemes respectively. All the filters give comparable performances. 3D schemes perform better than the 2D schemes and context-based modeling improves the compression performance by about 0.1 *bpp*. The improvement in compression performance of 3D schemes is about 0.1*bpp* over the 2D schemes.

Table 1. Lossless compression in (bpp) of 2D and 3D wavelet schemes for different wavelet filters on NIMHANS dataset. 2D-WOC:2D scheme without context coding,2D-WC: 2D scheme with context, 3D-WOC: 3D scheme without context, 3D-WC: 3D scheme with context.

Filter	2D-WOC	2D-WC	3D-WOC	3D-WC
(2,2)	2.35	2.2	2.18	2.12
(4,4)	2.31	2.18	2.17	2.1
(4,2)	2.33	2.18	2.18	2.1
(2+2,2)	2.33	2.18	2.18	2.1
(6,2)	2.33	2.19	2.18	2.11
(9,7)	2.36	2.21	2.21	2.13

Table 2. Lossless compression (in bpp) of 2D wavelet based schemes with and without region of interest on NIMHANS dataset. 2D-WOC-REC: 2D scheme without ROI and Context coding, 2D-WC-REC: 2D scheme without ROI and with Context coding.

Filter	2D-WOC-REC	2D-WC-REC	2D-WOC	2D-WC
(4,4)	3.12	2.35	2.31	2.18

5.2. Rectangular vs Region of Interest Coding

Tables 2 and 3 respectively give the performance measures in *bpp* on the NIMHANS dataset for 2D and 3D schemes with and without region of interest coding. In the later case, we define entire volume as region of interest and we call the coding scheme as "Rectangular Coding Scheme". The filter used is $(4,4)$, with 2 levels of spatial and axial decomposition and with the same block sizes as mentioned above. Using the region of interest coding, performance of 2D improves by about 0.2 *bpp* over the rectangular coding. In case of rectangular coding, the incorporation of context based coding improves the coding performance by 0.7 *bpp*. This can be attributed to the efficient coding of the significance map which contains large number of zeros due to the background.

In the 3D case, using the region of interest coding improves the coding performance by 0.2 *bpp* over the rectangular coding. In the rectangular coding, the use of context based coding improves the performance by about 0.3 *bpp*.

Table 3. Lossless compression (in *bpp*) of 3D wavelet based schemes with and without region of interest on NIMHANS dataset. 3D-WOC-REC: 3D scheme without ROI and Context coding, 3D-WC-REC: 3D scheme without ROI and with Context coding.

Filter	3D-WOC-REC	3D-WC-REC	3D-WOC	3D-WC
(4,4)	2.61	2.3	2.17	2.1

Table 4. Comparative performance of 3D schemes for group sizes of 8 and 16 on NIMHANS dataset. $(4,4)$ filter is used for both the group sizes.

GOI size	3D-WOC	3D-WC
8	2.17	2.1
16	2.16	2.09

5.3. Effect of Group Size on the Coding performance

We evaluate the effect of group size on the coding performance of the 3D scheme. We use two group sizes 8 and 16 with 2-levels of spatial and axial decomposition and with the same block sizes as mentioned earlier. Table 4 shows the performance comparison in *bpp* for the two group sizes. The increase in performance with a group size of 16 over the group size of 8 is very minimal. Increase in group requires buffering of more slices and hence more memory requirements. Hence we fix the group size to 8.

5.4. Overall Comparison

We compare the performances of 2D and 3D schemes MLZC and 3D EZW of [4]. For the 2D and 3D wavelet schemes, the performances are given for $(4,4)$ biorthogonal filter, with 2 levels of spatial and axial decomposition and a GOI size of 8. Table 5 gives the performance measures in *bpp* and the best performance is highlighted. This scheme also performs better than MLZC and 3D EZW on the MR-MRI data set. Although the MLZC algorithm provides region of interest coding functionality, the performance is given for lossless mode where the entire volume is considered as region of interest. Therefore, comparison should be made between the MLZC and the 3D scheme without region of interest coding. As shown in Table 5, the performances are similar.

6. Conclusions

We presented a new wavelet based coding schemes for 2D and 3D MR image compression. A 2D and 3D separable integer wavelet transforms are respectively used to

Table 5. Lossless compression Results (in *bpp*) of wavelet based schemes. $(4,4)$ filter is used for wavelet based schemes.

Data Set	3D-WOC	3D-WC-REC	3D-WC	MLZC	3D EZW
MR-MRI	1.94	2.1	**1.83**	2.14	2.27

decorrelate 2D and 3D images. Correlation within the subbands is exploited using fixed size lattices. These schemes provide important functionalities required for teleradiology, like progressive transmission, region of interest coding and lossless compression. The performance of coding schemes is further improved by incorporating context based coding for the significance and sign maps. 2D and 3D contexts are constructed using a mutual information based context quantization algorithm. The achieved lossless compression performance of the 3D scheme is comparable to those of MLZC and 3D EZW algorithms [4].

Acknowledgements: We sincerely thank Prof. P.N.Jayakumar, NIMHANS, Bangalore and Dr. Gloria Menegaz for providing data for this work.

References

[1] Adrian Munteanu, Jan Cornelis, "Wavelet-Based Lossless Compression of Coronary Angiographic Images", *IEEE Trans. Med. Imag.*, vol.18, no.3, pp.272-281, 1999.

[2] A.Said and W.pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circuits Syst. Video Technol.*, vol.6, pp 243-250, 1996.

[3] Wim Sweldens " The Lifting Scheme: A custom-design construction of Biorthogonal Wavelets", *Journal of Applied and computational Harmonic Analysis*, vol.3, pp 186-200, 1996.

[4] Gloria Menegaz and Jean-Philippe Thiran "Lossy to Lossless Object-Based Coding of 3-D MRI data", *IEEE Trans. Image Processing*, vol.11, pp 1053-1061, sept-2002.

[5] Zhen Liu and Lina J. Karam, "Context Formation by Mutual Information Maximization", *Proc. ICIP 2002*, vol.3 pp 89-92.

[6] J.Liu and P.Moulin, "Analysis of interscale and intrascale dependencies between image wavelet coefficients", *Proc. ICIP 2000*, vol.1 pp 669-671.

[7] D.Taubman, "High Performance Scalable Image Compression with EBCOT", *IEEE Trans. Image Processing*, vol.9, pp 1158-1170, July 2000.

[8] X.Wu, "Context quantization with fisher discriminant for adaptive embedded wavelet image coding", *Data Compression Conference, 1999.*