EZW AND SPIHT IMAGE COMPRESSION TECHNIQUES FOR HIGH RESOLUTION SATELLITE IMAGERIES
K Nagamani 1 and AG Ananth 2

Address for Correspondence
1 Lecturer, R V College of Engineering, Bangalore-560059.
2 Professor, R V College of Engineering, Bangalore-560059.
E Mail knmsm_03@yahoo.com

Abstract
Image compression methods employing wavelet transforms have been successfully implemented to provide high compression rates while maintaining good image quality. The significance map is a binary decision indicates if a coefficient of a 2-D discrete wavelet transform (DWT) has a zero or nonzero quantized value. The coding of the significance map, or in other words, the positions of those coefficients that will be transmitted as nonzero values is one of the important aspects of low bit rate image coding. This results in a considerable improvement in encoding the significance map, and hence, a higher efficiency in compression. Applying the DWT coefficients and using EZW (Embedded Zero tree wavelet) and SPIHT (Set Partitioning in Hierarchical Trees) coding techniques the compression ratios and PSNR are determined for a standard LENA Image and high resolution Satellite urban image (SatImg). The results obtained for EZW coding are compared with that of SPIHT coding for the same set of images. The results show that it is possible to achieve higher compression ratios ~8 and PSNR ~29.20 for SPIHT coding compared to EZW coding where the compression ratio ~1.07, PSNR ~13.07 can be achievable for applications relating to satellite urban imageries. Both the techniques indicate that maximum compression ratios are achievable for standard Lena Image for an acceptable quality of the image. The results are presented and discussed in the paper.

I INTRODUCTION
The Discrete Wavelet Transform (DWT) techniques have been devised by Creusere (1997) to decompose discrete time signals. The analysis scheme is named as sub-band coding. Burt (1983) defined a technique very similar to sub-band coding and named it pyramid coding which is also known as multi-resolution analysis. Later Vetterli et al (1990) made some improvements to the sub-band coding scheme, removing the existing redundancy in the pyramid coding scheme. A detailed coverage of the discrete wavelet transform and multi-resolution analysis can be found in a number of articles available in the literature. (J M Shapiro, 1992, 1993: R. A. DeVore et al 1992, A. S. Lewis and G. Knowles 1992, Bryan Usevitch, 2001). A time-scale representation of a digital signal is obtained using digital filtering techniques. In the discrete case, filters of different cutoff frequencies are used to analyze the signal at different scales. The signal is passed through a series of high pass filters to analyze the high frequencies, and it is passed through a series of low pass filters to analyze the low frequencies. The resolution of the signal, which is a measure of the amount of detail information in the signal, is changed by the filtering operations, and the scale is changed by up-sampling and down-sampling (sub-sampling) operations. Sub-sampling a signal corresponds to reducing the sampling rate, or removing some of the samples of the signal.

II SOFTWARE DEVELOPMENT AND IMAGE ANALYSIS
Wavelet transform provides a hierarchical signal representation, each coefficient corresponds to a spatial area and a frequency range. It is identical to a hierarchical sub-band system, where the sub-bands are logarithmically spaced in frequency and represent octave-band decomposition. At each level, the signal can be further decomposed into a coarser approximation and a corresponding added detail. At the first stage of wavelet transform, the original image is divided into four sub-bands. Each sub-band is further decomposed by separable vertical and horizontal filters. This process is continued till the final level is reached.

A threshold T is defined and used to compare with all coefficients. A wavelet coefficient x is insignificant if \(|x|<T\), and is significant if \(|x|>T\). If a wavelet coefficient in a coarser scale is insignificant with respect to the given threshold value T, then all the coefficients of the same orientation in the same spatial location at finer scales are likely to be insignificant with respect to T.

III EZW ALGORITHM
The EZW output stream will have to start with some information to synchronize the decoder. The minimum information required by the decoder is the number of
wavelet transform levels used and the initial threshold, assuming that always the same wavelet transform will be used. Additionally the image dimensions and the image mean. Sending the image mean is useful by removing it from the image before coding. After imperfect reconstruction the decoder can then replace the imperfect mean by the original mean. This can increase the PSNR significantly. The first step in the EZW coding algorithm is to determine the initial threshold. If we adopt bit-plane coding then our initial threshold $t_0$ will be:

$$t_0 = 2^{\left\lfloor \log_2 \left( \max_{x,y} \left| f(x,y) \right| \right) \right\rfloor}$$

Here $\max(.)$ means the maximum coefficient value in the image and $(x,y)$ denotes the coefficient. With this threshold we enter the main coding loop.

Normally two passes are used to code the image: a dominant pass and a subordinate pass. The dominant pass finds pixel values above a certain threshold, and the subordinate pass quantizes all significant pixel values found in this and all previous dominant passes previous.

After each dominant pass, a subordinate pass is then performed on the subordinate list which contains all pixel values previously found to be significant. The subordinate pass performs pixel value quantization which achieves compression by telling the decoder with a symbol roughly what the pixel value is instead of exactly what the pixel value is. Since the initial threshold is one-half the maximum magnitude of all pixel values for the first dominant pass, then in the first subordinate pass only two ranges are specified in which a significant pixel value could lie: the upper half of the range between the maximum pixel value and the initial threshold, or the lower half of the same range. A pixel value in the upper half of the range gets coded with the symbol upper (for upper part of the range), while a pixel value in the lower half gets coded with the symbol lower. A pixel value found to be in a particular range is quantized, from the decoders viewpoint, to the midpoint of that range. Upon subsequent subordinate passes the threshold has been cut in half and so there are twice as many ranges as the last subordinate pass plus two new ranges corresponding to the new lower threshold. By reading the subordinate symbol corresponding to a significant pixel and knowing the threshold, the decoder is able to determine the range in which the pixel lies and reconstructs the pixel value to the midpoint of that range. Thus from the decoders viewpoint the rough estimate of a significant pixel's value is getting more refined and accurate as more subordinate passes are made. So, the subordinate passes quantize pixel values to a two symbol alphabet which then get encoded by using an adaptive arithmetic coder as described by Witten et al ( ) thus achieving compression.

For every pass a threshold is chosen against which all the coefficients are measured. If a wavelet coefficient is larger than the threshold it is encoded and removed from the image. If it is smaller it is left for the next pass. When all the wavelet coefficients have been visited the threshold is lowered and the image is scanned again to add more detail to the already encoded image. This process is repeated until all the coefficients have been encoded completely.

**IV. SPIHT Algorithm**

The SPIHT (Set Partitioning in Hierarchical Trees) algorithm is a refined version of EZW algorithm. It can perform better at higher compression ratios for a wide variety of images than EZW. The algorithm uses a partitioning of the trees in a manner that tends to keep insignificant coefficients together in larger subsets.

Before the algorithm is explained, certain notations are to be made familiar. The data structure used by the SPIHT algorithm is similar to that used by the EZW algorithm. The wavelet coefficients are again divided into trees originating from the lowest resolution band (band I). The coefficients are grouped into 2 x 2 arrays that, except for the coefficients in band I, are offsprings of a coefficient of a lower resolution band. The coefficients in the lowest resolution band are also divided into 2 x 2 arrays. However, unlike the EZW case, all but one of them are root nodes. The coefficient in the top-left corner of the array does not have any offspring.

The trees are further partitioned into four types of sets, which are sets of coordinates of the coefficients:
- $O(i,j)$ This is the set of coordinates of the offsprings of the wavelet coefficient at location $(i,j)$.
- $D(i,j)$ This is the set of all descendants of the coefficient at location $(i,j)$.
- $H$ is set of all root nodes.
- $L(i,j)$ This is the set of coordinates of all the descendants...
of the coefficient at location except for the immediate offsprings of the coefficient at location \((i,j)\).

So,

\[
L(i,j) = D(i,j) - O(i,j)
\]

A set \(D(i,j)\) or \(L(i,j)\) is said to be significant if any coefficient in the set has a magnitude greater than the threshold.

The algorithm makes use of three lists: the list of insignificant pixels (LIP), the list of significant pixels (LSP), and the list of insignificant sets (LIS). The LSP and LIS lists contain the coordinates of coefficients, while the LIS contains the coordinates of the roots of sets of type \(D\) or \(L\) .

The initial value of the threshold is given as 

\[
n = \log_2 C_{\text{max}}
\]

Where, \(C_{\text{max}}\) is the maximum magnitude of the coefficients to be encoded. The LIP list is initialized with the set \(H\). Those elements of \(H\) that have descendants are also placed in LIS as type \(D\) entries. The LSP list is initially empty.

In each pass, the members of the LIP are first processed, then the members of LIS. This is essentially the significance map encoding step. In the refinement step the elements of LSP are processed. Each coordinate contained in LIP is examined first. If the coefficient at that coordinate is significant (i.e., it is greater than \(2^n\)), a 1 is transmitted, followed by a bit representing the sign of the coefficient (1 for positive, 0 for negative). Then that coefficient is moved to the LSP list. If the coefficient at that coordinate is not significant, a 0 is transmitted. After examining each coordinate in LIP, the sets in LIS are examined. If the set at coordinate \((i,j)\) is not significant, a 0 is transmitted. If the set is significant, a 1 is transmitted. If the set is of type \(D\), each of the offsprings of the coefficient at that coordinate is checked. For each coefficient that is significant, a 1 is transmitted, the sign of the coefficient, and then the coefficient is moved to the LSP. For the rest a 0 is transmitted and their coordinates are added to the LIP. If this set is not empty, it is moved to the end of the LIS and marked as type \(D\). This new entry into the LIS has to be examined during this pass. If the set is empty, the coordinate \((i,j)\) is removed from the list. If the set is of type \(D\), each coordinate in \(D(I,j)\) is added to the end of the LIS as the root of a set of type \(D\). Again, note that these new entries in the LIS have to be examined during this pass. Then \((i,j)\) is removed from the LIS.

Once each of the sets in the LIS (including the newly formed ones) is processed, a refinement step is started. In the refinement step each coefficient that was in the LSP prior to the current pass is examined and output the \(n\)th most significant bit of \(|c_{ij}|\). The coefficients that have been added to the list in this pass are ignored because, by declaring them significant at this particular level, the decoder has already been informed of the value of the \(n\)th most significant bit. This completes one pass. Depending on the availability of more bits or external factors, \(n\) is decremented by one and the process continues.

**Peak Signal to Noise Ratio (PSNR):**

The PSNR is calculated with the following formula

\[
\text{PSNR} = 10 \log_{10} \left( \frac{\text{Max grey level} \times MN}{\sum_{xy} \left| g(x,y) - f(x,y) \right|^2} \right)
\]

Where  \(g(x,y)\) is the compressed image, \(f(x,y)\) is the raw image , \(M\) is the image width, \(N\) is the image height and max. gray level is the max. value of \(f(x,y)\). A Max grey level = 255 has been used (as there are 0 to 255 grey levels represented with 8 bits in the BMP format images).

**V Results and Discussion**

For the present analysis two sets of images Lena (Standard) and SatImg (High resolution satellite image of a Urban Scene) are selected. Both the input image files are used in BMP format. The image sizes were considered to be 256*256 pixels. For both the images discrete wavelet transform (DWT) coefficients have been determined. Considering a threshold value of 8 for EZW and SPIHT algorithm and using Huffman coding, the images were compressed and reconstructed. The compression ratios and the calculated PSNR values determined for the two sets of images Lena and SatImg are shown in Table 1 and Table 2.

**Table 1: Comparison of EZW and SPIHT algorithm for Lena Image**

<table>
<thead>
<tr>
<th>TECHNIQUE</th>
<th>RATE</th>
<th>COMP.RATIO</th>
<th>PSNR(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EZW</td>
<td>2.40</td>
<td>3.34</td>
<td>27.79</td>
</tr>
<tr>
<td>SPIHT</td>
<td>1</td>
<td>8</td>
<td>61.53</td>
</tr>
</tbody>
</table>

**Table 2: Comparison of EZW and SPIHT algorithm for Satellite Urban Image**

<table>
<thead>
<tr>
<th>TECHNIQUE</th>
<th>RATE</th>
<th>COMP.RATIO</th>
<th>PSNR(dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EZW</td>
<td>7.47</td>
<td>1.07</td>
<td>13.07</td>
</tr>
<tr>
<td>SPIHT</td>
<td>1</td>
<td>8</td>
<td>29.20</td>
</tr>
</tbody>
</table>
Table 1 and 2 clearly indicates that the Compression ratios and PSNR values achievable for Lena image are higher than that of Satellite Urban image. Where as the Compression ratio and PSNR values for SatImg (Table. 2) shows lower values indicating the presence of higher background noise for Satellite Images.

Further the Table 2 demonstrates that for satellite urban images the SPIHT coding techniques exhibits better Compression ratios ~ 8.0 and PSNR values ~29.20. Whereas EZW coding techniques indicates Compression ratio of ~1.07 and PSNR ~ 13.07 for the SatImg. For threshold value of 8 the original image and reconstructed Lena Image from EZW and SPIHT techniques are shown in Figures 1a, 1b and 1c and for SatImg are shown in Figure 2a 2b and 2c. It is clearly evident from Figure 1c that the Lena image reconstructed from SPIHT coding (1c) is better than the image reconstructed from EZW coding (1b). Similarly it is clearly seen from Figure 2c that the SatImg reconstructed form SPIHT coding (2c) has a better representation of the original image (2a) when compared to the reconstructed image from EZW coding (2b). These results suggest that for Compression of satellite urban images the SPIHT coding techniques are more effective than the EZW techniques. In case of SPIHT coding One could achieve maximum compression ratios ~8 and PSNR ~ 29.20 for satellite urban imageries and achieve good image quality.

VII CONCLUSION

• The DWT techniques applied for EZW and SPIHT coding for the compression of images indicate that the PSNR values determined for standard Lena Images are normally higher than satellite images.
• Using DWT and SPIHT coding techniques it is possible to achieve high compression ratio ~ 8 and PSNR values ~ 29.20 for Satellite Urban Imageries when compared to EZW coding.

ACKNOWLEDGEMENT

The Authors are thankful to Dr Ganesh Raj Dy Director, ISRO for providing the satellite images, Sri Ananth Upadhaya for the support in the software development and Head, Dept of Telecommunication, RV Engineering college for encouragement in carrying out the present research work.

REFERENCE

5. A. S. Lewis and G. Knowles, “Image Compression
Using the 2-D Wavelet Transform”, IEEE Trans. on Image Processing, Vol. 1, No. 2, pp. 244-250, April (1992)


ABOUT THE AUTHORS

Nagamani. K was born on 30/05/1973 in Bangalore, India. She received B.E degree in 1996 from Bangalore University and M.Tech degree in 2006 from Visveswaraih Technical University, Belgaum, India.

She is pursuing PhD. Presently working as Lecturer in Telecommunication at R.V. College of Engineering Bangalore. Her research interests include Digital compression, Embedded System design.

Dr.A.G. Ananth was born on 3/11/1947, at Bangalore, India. He received M.Sc degree in Nuclear Physics in 1969 from Bangalore University, India.

In 1975, the Physical Research Laboratory Ahmedabad awarded him PhD degree in Space physics. He served as Deputy Director at Indian space research organization, Bangalore. Presently working as Professor in Telecommunication Department of R.V College of Engineering, Bangalore. His research interests include Space physics, signal processing, Image compression, MIMO systems and satellite technology.