Spatial Video Compression using EZW, 3D-SPIHT, WDR & ASWDR Techniques

Abhishek Jain* 
Electronics & Communication Engineering 
SATI, Vidisha (M.P.), INDIA

Dr. Ashutosh Datar 
Biomedical Engineering 
SATI, Vidisha (M.P.), INDIA

Abstract— Over the past decade, the success of wavelets in image processing has contributed to its unprecedented popularity. This paper also realizes and evaluates the efficiency of wavelet based EZW (Embedded Zerotree Wavelet), 3D-SPIHT (3D-Set Partitioning in Hierarchical Trees), WDR (Wavelet Difference Reduction) and ASWDR (Adaptively Scanned Wavelet Difference Reduction) image compression techniques on a colored video. Some comparative results are obtained based on various performance parameters such as Peak Signal to Noise Ratio (PSNR), Mean Squared Error (MSE), Maxloop, Compression Ratio (CR) and Bit-Per-Pixel (BPP) ratio. These techniques sustain faithful compression and reproduction of video for all video formats.

Keywords— Video Compression, EZW, 3D-SPIHT, WDR, ASWDR, PSNR, MSE.

I. INTRODUCTION

Due to the increasing traffic caused by multimedia information and digitized form of representation of images and video; compression has become a necessity. Video compression is used to minimize the size of a video file without degrading the quality of the video. Over the past few years, a variety of powerful and sophisticated wavelet based schemes for image and video compression have been developed and implemented [1]-[7]. Some of the most promising are algorithms that minimize the amount of memory which the encoder or decoder must use [8], [9]. An algorithm, which is embedded and which minimizes PSNR is described in [10] (Rate-distortion Optimized Embedding). The discrete wavelet transform (DWT) [1], [2] has gained wide popularity due to its excellent decorrelation property. Many modern image and video compression systems embody the DWT as the intermediate transform stage. After DWT was introduced, several codec algorithms were proposed to compress the transform coefficients as much as possible but a compromise must be maintained between the higher compression and a good perceptual quality of image. Achieving much higher compression is simply not possible without discarding some perceptible information. Thus, the rate of compression is application dependent. Video coding for telecommunication applications has evolved through the development of the ISO/IEC MPEG-1, MPEG-2 and ITU-T H.261, H.262, H.263 video coding standards (and later enhancements of H.263 known as H.263+ and H.263++) and H.264 [11], [12] and has diversified from ISDN and T1/E1 service to embrace PSTN, mobile wireless networks, and LAN/Internet network delivery. Throughout this evolution, continued efforts have been made to maximize coding efficiency. The performance of these base coders generally degrades at low bit-rates mainly because of the underlying block-based Discrete Cosine Transform (DCT) scheme [13]. An extended analysis of motion compensated frame difference for block-based motion prediction error is described in [14]. This paper discusses four different approaches in compression having different algorithms using DWT. A performance evaluation is carried out between these techniques on the basis of some performance parameters. The purpose of this comparative study is to provide a basis for other innovative works in video compression for superior results.

The most powerful progressive method, Embedded Zerotree Wavelet (EZW) coding algorithm introduced by Shapiro [3] combines stepwise thresholding and progressive quantization, focusing on the more efficient way to encode the image coefficients in order to minimize the compression ratio. Among these, Spatial-Orientation Tree Wavelet (STW) [4] and Set Partitioning in Hierarchical Trees (SPIHT) [5] are found to be the more advantageous because of their different approach of encoding the wavelet transform. These wavelet based image compression algorithms (SPIHT and STW) are considered as refined versions of the seminal EZW algorithm. The 3D-Set Partitioning in hierarchical trees (3D-SPIHT) technique which is proposed by Kim and Pearlman [6] is the extended form of SPIHT coding algorithm, in which the relationship among coefficients lying in different frequency bands is based on octal tree structure rather than quad-tree structure. The most enhanced image compression algorithm is the Adaptively Scanned Wavelet Difference Reduction (ASWDR) algorithm proposed by Walker [15], [16]. ASWDR technique adjusts the scanning order used by Wavelet Difference Reduction (WDR) algorithm [17], so as to predict locations of new significant values. The WDR method employs a fixed ordering of the positions of wavelet coefficients.

II. EMBEDDED ZEROTREE WAVELET (EZW)

The EZW algorithm was one of the first algorithms to show the full power of wavelet-based image compression. The other three algorithms were built upon the fundamental concepts that were first introduced with EZW.
An embedded coding is a process of encoding the transform magnitudes that allows for progressive transmission of the compressed image. In EZW, the root location is marked by encoding only one symbol for the output as described in [3], [4]. Consequently in EZW, the zerotrees provide narrow descriptions of the locations of insignificant values. Zerotrees allow for a concise encoding of the positions of significant values that result during the embedded coding process. The embedding process used by EZW is called bit-plane encoding [3]. The main advantage of this encoding is that the encoder can terminate the encoding at any point, thereby allowing a target bit rate to be met exactly. To arrive at a perfect reconstruction, the process is repeated after lowering the threshold, until the threshold has become smaller than the smallest coefficient to be transmitted. Similarly, the decoder can also stop decoding at any point resulting in the image that would have been produced at the rate of the truncated bit stream. The information of transmission of the coefficient positions is very much necessary. Indeed, without this information the decoder will not be able to reconstruct the encoded signal (although it can perfectly reconstruct the transmitted bit stream). It is in the encoding of the positions where the efficient encoders are separated from the inefficient ones. EZW encoding uses a predefined scan order to encode the position of the wavelet coefficients. Several scan orders are possible, as long as the lower sub bands are completely scanned before going on to the higher sub bands. Thus, EZW provides an efficient way to encode the coefficients in order to achieve higher compression. Consequently, this algorithm yields excellent results without any pre-stored tables or codebooks, training, or prior knowledge of the image/frame source.

III. 3D-SET PARTITIONING IN HIERARCHICAL TREES (3D-SPIHT)

Set partitioning in hierarchical trees (SPIHT) is an image compression algorithm that exploits the inherent similarities across the subbands in wavelet decomposition [18] of an image. The SPIHT algorithm is used to the multi-resolution pyramid after the sub-band/wavelet transformation is performed. The embedded coding property of SPIHT allows exact bit rate control without any penalty in performance. The same property also allows exact MSE distortion control. SPIHT codes the individual bits of the image wavelet transform coefficients following a bit-plane sequence. Thus, it is capable of recovering the image perfectly by coding all bits of the transform. The SPIHT video coding system is shown in Fig. 1.

![SPIHT Video Coding System](image)

The 3D-Set Partitioning in hierarchical trees (3D-SPIHT) technique which is proposed by Kim et al. [6], [7] is extended from the above known SPIHT coding algorithm. It is a simple and efficient wavelet zero tree image coding algorithm which has been proved its efficiency with high performance, precise rate control and its real-time capability in compression of video. The video coder is fully embedded, so that a variety of monochrome or color video quality can thus be obtained with a single compressed bit stream. Consequently, the compression process can be terminated at any desired rate [19].

The wavelet coefficients are considered as a collection of spatial orientation trees where each tree is formed of coefficients from all sub bands belonging to the same spatial location in an image. The wavelet coefficients are scanned column wise then line wise, from low subbands to high subbands. After that an iterative 3D-SPIHT algorithm selects an initial threshold based on the largest wavelet coefficient [20]. When the largest coefficient magnitude in the set is greater than or equal to the selected threshold, a tree wavelet coefficient set is significant. In the 3D-SPIHT algorithm, there are two important passes: sorting pass and refinement pass [7]. A recursive partitioning is realized on the tree. So the position of significant coefficient in the descendants of the considered coefficient is identified [21]. In SPIHT, the relationship among coefficients lying in different frequency bands is based on quad-tree structure, while the one is based on octree structure [7] in 3D-SPIHT.

IV. WAVELET DIFFERENCE REDUCTION (WDR)

One of the defects of SPIHT is that it only implicitly locates the position of significant coefficients. This makes it difficult to perform operations, such as region selection on compressed data, which depend on the exact position of significant transform values. Region selection, also known as region of interest (ROI), means selecting a portion of a compressed image, requires increased resolution. Such compressed data operations are possible with the Wavelet Difference Reduction (WDR) algorithm of Tian and Wells [10], [15]. The term difference reduction refers to the way in which WDR encodes the locations of significant wavelet transform values. In WDR, the output from the significance pass consists of the signs of significant values along with sequences of bits which concisely describe the precise locations of significant values.

The WDR algorithm is a very simple procedure. A wavelet transform is first applied to the image, and then the bit-plane based WDR encoding algorithm [16], [17] for the wavelet coefficients is carried out. WDR mainly consists of five steps which include initialization, updating threshold, significance pass, refinement pass [15] and the repetition of steps 2 to 4 until the bit budget [16], [17] is reached.
V. ADAPTIVELY SCANNED WAVELET DIFFERENCE REDUCTION (ASWDR)

It is one of the most enhanced image compression algorithms proposed by Walker [15], [16]. The ASWDR algorithm aims to improve the subjective perceptual qualities of compressed images and improve the results of objective distortion measures. The ASWDR algorithm is a simple modification of the Wavelet Difference Reduction (WDR) algorithm [17].

The WDR algorithm employs a fixed ordering of the positions of wavelet coefficients but the ASWDR method employs a varying order which aims to adapt itself to specific image features. ASWDR adjusts the scanning order so as to predict locations of new significant values. The scanning order of ASWDR dynamically adapts to the locations of edge details in an image, and this enhances the resolution of these edges in ASWDR compressed images. Thus, ASWDR exhibits better perceptual qualities, especially at low bit rates, than WDR and SPIHT compressed images preserving all the features of WDR.

The ASWDR on an image/frame is executed by a step by step procedure described below [22]:

\[
\text{Step 1: A wavelet transform is performed on the discrete image/frame } f_{[j,k]}, \text{ producing the transformed image/frame } [j,k].
\]

\[
\text{Step 2: A scanning order for the transformed image is chosen, } [j,k] = a(m). \text{ The transform values are scanned via a linear ordering, } m = 1,2,3,...,X \text{ where } X \text{ is the number of pixels. In [4], row-based scanning is used in the horizontal subbands and column-based scanning is used in the vertical subbands with the zigzag scanning order through subbands from higher scale to lower scale [5].}
\]

\[
\text{Step 3: In this step an initial threshold } T \text{ is chosen. The } T \text{ is chosen in such a way that at least one transform value has magnitude less than or equal to } T \text{ and all transform values have magnitudes less than } 2T.
\]

\[
\text{Step 4: (Significance pass). The positions for new significant values are recorded as depicted in [16]. These new significant indices are then decoded using difference reduction [17], [23].}
\]

\[
\text{Step 5: (Refinement pass). Record the refinement bits, the next significant bits, for the old significant transform values. This generation of refinement bits is also known as standard bitplane encoding which is utilized by all embedded codecs [5], [15].}
\]

\[
\text{Step 6: (New scanning order). For the level containing the all-lowpass subband, the indices of the remaining insignificant values are used as the scan order at that level. The scan order at level } k \text{ is used to create the new scan order at level } k-1 \text{ as follows: Run through the significant values (i.e. the parent values) at level } k \text{ in the wavelet transform. Each parent value induces a set of four child values for all the levels except the last. The last level induces three child values as described in the spatial-orientation tree definition in [11]. At level } k-1, \text{ the insignificant values are enclosed in the first part of the scan order lying among these child values. Now again run through the insignificant values at level } k \text{ in the wavelet transform. This provides the insignificant values enclosed in the second part of the scan order lying among the child values induced by these insignificant parent values. This new scanning order for level } k-1 \text{ is further used to create the new scanning order for level } k-2, \text{ until all levels are exhausted [22].}
\]

\[
\text{Step 7: Divide the present threshold by 2. Repeat Steps 4-6 until either all the levels are exhausted or a given distortion metric [16] is fulfilled.}
\]

VI. PERFORMANCE PARAMETERS

A. Mean Squared Error (MSE) and Peak Signal to Noise Ratio (PSNR)

The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error. The phrase peak signal-to-noise ratio, often abbreviated as PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale.

The PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codecs. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs it is used as an approximation to human perception of reconstruction quality, therefore a higher PSNR would normally indicate that the reconstruction is of higher quality [24]. The mean squared error (MSE) for the two m×n images I (i, j) and K (i, j) where one of the images is considered a noisy approximation of the other is defined as:

\[
\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left[ I(i,j) - K(i,j) \right]^2
\]

The PSNR is defined as:

\[
\text{PSNR} = 10 \cdot \log_{10} \left( \frac{\text{MAX}_i^2}{\text{MSE}} \right)
\]

Here, \( \text{MAX}_i \) is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with B bits per sample, \( \text{MAX}_i \) is \( 2B-1 \). For color images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three.

B. Maxloop

The Maxloop indicates the number of steps for a particular compression algorithm [24].
C. Compression Ratio (CR) and Bit-Per-Pixel (BPP)

A measure of achieved compression is given by the Compression Ratio (CR) and the Bit-Per-Pixel (BPP) ratio. CR indicates that the compressed image is stored using CR % of the initial storage size while BPP is the number of bits used to store one pixel of the image [22]. Thus the formulation for CR is given by:

\[
CR = \frac{\text{Compressed Frame Size}}{\text{Uncompressed Frame Size}} \times 100
\]

For a grayscale image the initial BPP is 8. For a truecolor image the initial BPP is 24, because 8 bits are used to encode each of the three colors (RGB color space).

VII. RESULTS AND DESCRIPTION

The original video is split in the form of frames which is then compressed by these algorithms for various Maxloop. The wavelet used in our experiment is biorthogonal spline wavelet 4.4 (bior 4.4) [25], [26]. The challenge of compression methods is to find the best compromise between a low compression ratio and a good perceptual result or it mainly application dependent. Therefore, Maxloop for these algorithms are selected on the basis of the application and the required Compression Ratio (CR) or the Peak Signal to Noise Ratio (PSNR). Simulation results for these techniques are carried out for different number of Maxloops. The video sequence tested is the standard colored video of Miss America which is in avi format. The general and video configurations of the video obtained in MATLAB are shown in Table 1.

<table>
<thead>
<tr>
<th>TABLE I. CONFIGURATIONS OBTAINED IN MATLAB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General Configuration</strong></td>
</tr>
<tr>
<td>Duration</td>
</tr>
<tr>
<td>Name</td>
</tr>
<tr>
<td>Tag</td>
</tr>
<tr>
<td>Type</td>
</tr>
<tr>
<td><strong>Video Configuration</strong></td>
</tr>
<tr>
<td>Bits Per Pixel</td>
</tr>
<tr>
<td>Frame Rate</td>
</tr>
<tr>
<td>Height</td>
</tr>
<tr>
<td>Number of Frames</td>
</tr>
<tr>
<td>Video Format</td>
</tr>
<tr>
<td>Width</td>
</tr>
</tbody>
</table>

The simulation results of video compression by executing the Embedded Zerotree Wavelet (EZW), 3D-Set Partitioning in hierarchical trees (3D-SPIHT), Wavelet Difference Reduction (WDR) and Adaptively Scanned Wavelet Difference Reduction (ASWDR) algorithms are realized. Based on these results, various comparisons are obtained on the basis of PSNR, MSE, CR and BPP for different Maxloops.

The original frame is shown in the Fig. 2 and compressed frames are shown in Fig. 3, Fig. 4, Fig. 5 and Fig. 6 with the number of Maxloops and with different attained values of performance parameters:

The simulation results of video compression by executing the Embedded Zerotree Wavelet (EZW), 3D-Set Partitioning in hierarchical trees (3D-SPIHT), Wavelet Difference Reduction (WDR) and Adaptively Scanned Wavelet Difference Reduction (ASWDR) algorithms are realized. Based on these results, various comparisons are obtained on the basis of PSNR, MSE, CR and BPP for different Maxloops.

The original frame is shown in the Fig. 2 and compressed frames are shown in Fig. 3, Fig. 4, Fig. 5 and Fig. 6 with the number of Maxloops and with different attained values of performance parameters:

![Fig. 2. Original Frame no.1](image)

![Fig. 3. Frame no.1 compressed by using EZW algorithm for Maxloop 11, CR=1.0077, BPP =0.2419, MSE=7.4255 and PSNR=39.4235 dB](image)
Fig. 4. Frame no.1 compressed by using 3D-SPIHT algorithm for Maxloop 11, CR=0.6552, BPP =0.1573, MSE=9.3952 and PSNR=38.4018 dB

Fig. 5. Frame no.1 compressed by using WDR algorithm for Maxloop 11, CR= 1.0679, BPP = 0.2563, MSE=7.4255 and PSNR=39.4235 dB

Fig. 6. Frame no.1 compressed by using ASWDR algorithm for Maxloop 11, CR= 1.0380, BPP = 0.2491, MSE=7.4255 and PSNR=39.4235 dB

Fig. 7 shows the variation in the CR for EZW, 3D-SPIHT WDR and ASWDR algorithms with the corresponding Maxloops.

Fig. 7. Average value of CR for EZW, 3D-SPIHT WDR and ASWDR techniques with different Maxloops
Fig. 8 shows the variation in PSNR for EZW, 3D-SPIHT WDR and ASWDR algorithms with the corresponding Maxloops.

Fig. 8. Average value of PSNR for EZW, 3D-SPIHT WDR and ASWDR techniques with different Maxloops

From Fig. 7 and Fig. 8, it is very much clear that the techniques with higher PSNR and lesser MSE values for a given Maxloop are in the following order:

1) EZW
2) WDR/ASWDR
3) 3D-SPIHT

On the contrary, order of these techniques with lesser CR (higher compression) and BPP values for a given Maxloop is as follows:

1) 3D-SPIHT
2) EZW
3) ASWDR
4) WDR

The comparative tabulation for all these algorithms are shown in Table 2. The attained values of PSNR, MSE, CR and BPP are shown in this table along with the Maxloops for EZW, 3D-SPIHT WDR and ASWDR techniques.

<table>
<thead>
<tr>
<th>Maxloop</th>
<th>Algorithm</th>
<th>PSNR (dB)</th>
<th>MSE (%)</th>
<th>CR</th>
<th>BPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maxloop 9</td>
<td>EZW</td>
<td>32.2676</td>
<td>38.5763</td>
<td>0.2714</td>
<td>0.0674</td>
</tr>
<tr>
<td></td>
<td>3D-SPIHT</td>
<td>31.2819</td>
<td>48.4302</td>
<td>0.1743</td>
<td>0.0418</td>
</tr>
<tr>
<td></td>
<td>WDR</td>
<td>32.0290</td>
<td>40.7797</td>
<td>0.2792</td>
<td>0.0670</td>
</tr>
<tr>
<td></td>
<td>ASWDR</td>
<td>32.0290</td>
<td>40.7797</td>
<td>0.2719</td>
<td>0.0642</td>
</tr>
<tr>
<td>Maxloop 10</td>
<td>EZW</td>
<td>35.9744</td>
<td>16.43</td>
<td>0.5234</td>
<td>0.1256</td>
</tr>
<tr>
<td></td>
<td>3D-SPIHT</td>
<td>34.8124</td>
<td>21.4825</td>
<td>0.3460</td>
<td>0.0830</td>
</tr>
<tr>
<td></td>
<td>WDR</td>
<td>35.5649</td>
<td>18.0545</td>
<td>0.5693</td>
<td>0.1366</td>
</tr>
</tbody>
</table>
In this paper the simulation results of video compression are obtained by applying Embedded Zerotree Wavelet (EZW), 3D-Set Partitioning in hierarchical trees (3D-SPIHT), Wavelet Difference Reduction (WDR) and Adaptively Scanned Wavelet Difference Reduction (ASWDR) algorithms. The simulation results show that all these algorithms sustain faithful compression and reproduction of the video, preserving the picture quality. ASWDR is having lesser value of CR than WDR technique. Thus, it achieves high compression than WDR while retaining all of the important features of WDR such as low complexity, region of interest (ROI) capability and progressive SNR capability.

The results also show that the PSNR and MSE values are better in EZW as compared to 3D-SPIHT, WDR & ASWDR method. Thus, it can be used in the applications where higher perceptual quality is fruitful than the CR since it is providing higher PSNR values for less Maxloops. In the contrast, 3D-SPIHT gives less CR (higher compression) and BPP values than EZW, WDR & ASWDR techniques which is the main requisite in compression. So, 3D-SPIHT is advantageous in the applications where lower BPP is necessitated. The result also illustrates that EZW is providing a better compromise between lesser CR and higher PSNR among all the four different approaches.

In future, the choice of wavelet or the use of multiwavelets with some other compression techniques can also lead for an enhanced outcome. The recently introduced curvelet transform can also be used in conjunction with wavelets for better results.

REFERENCES
Abhishek Jain received his B. E. degree in Electronics & Communication Engineering from Rajiv Gandhi Technical University, Bhopal and the M.Tech. (Hons.) degree in Digital Communication from Barkatullah University, Bhopal (M.P.), India. He also became a lifetime member of IETE and qualified GATE in 2011. He has more than 5 years of teaching experience and presently he is working as an Assistant Professor in the Department of Electronics & Communication Engineering, SATI (Degree), Vidisha (M.P), India.

Ashutosh Datar was born in Vidisha, India on 12th October 1967. He has received his B.E. (Electronics & Instrumentation) degree from Samrat Ashok Technological Institute, Vidisha (M.P.) India, and M.Tech. (Biomedical Engineering) degree from Institute of Technology, Banaras Hindu University (BHU) Varamasi, in 1989 and 1998, respectively. He obtained his Doctoral degree from Rajiv Gandhi Proudyogiki Vishwavidyalaya, (State Technological University of Madhya Pradesh), Bhopal India in 2012.

He is presently working as an associate professor in Department of Biomedical Engineering, Samrat Ashok Technological Institute, Vidisha (M.P.) India.

Mr. Datar is a life member of IE(1), ISTE, BMESI, Instrumentation Society of India. His current research interests include digital signal processing, multirate signal processing, filterbanks, Bio-medical signal processing and Medical image processing.