Adaptive Encoding & Decoding of Compressed Video Using SPIHT Algorithm

K M Amritha, Nithin S S

Abstract— **Video compression is the heart of digital television set top boxes, DSS, HDTV decoders, DVD players, video conferencing, internet video and other applications. These benefit in the fact that they require less storage space for archived video information and less bandwidth for transmission. An efficient approach to distributed video coding is the Transform domain Wyner-Ziv (TDWZ) video coding. Existing system use optical flow based motion re-estimation technique and a generalized reconstruction algorithm for video coding. Based on this findings a new method is proposed where wavelet transform instead of the discrete cosine transform followed by compression using Set Partitioning In Hierarchical Tree (SPIHT) have been proposed. SPIHT algorithm is a fast and efficient technique for compression also the main advantage of wavelet transform over discrete cosine transform is that it has both time and frequency localization ability, which can give a better performance in compression.**

Index Terms— **Distributed Video Coding (DVC), Optical flow, SPIHT algorithm, Transform domain Wyner-Ziv (TDWZ), wavelet transform.**

I. Introduction

In computer science and information theory, data compression, source coding, or bit-rate reduction involves encoding information using fewer bits than the original representation. The process of reducing the size of a data file is referred to as data compression. Compression can be either lossy or lossless. Lossless compression reduces bits by identifying and eliminating statistical redundancy. No information is lost in lossless compression. Lossy compression reduces bits by identifying unnecessary information and removing it. Compression is useful because it helps reduce resource usage, such as data storage space or transmission capacity. Data compression is subject to a space- time complexity trade-off. For instance, a compression scheme for video may require expensive hardware for the video to be decompressed fast enough to be viewed as it is being decompressed, and the option to decompress the video in full before watching it may be inconvenient or require additional storage. The design of data compression schemes involves trade-offs among various factors, including the degree of compression, the amount of distortion introduced (e.g., when using lossy data compression), and the computational resources required to compress and decompress the data. However, the most important reason for compressing data is that more and more we can share data. The web and its underlying networks have limitations on bandwidth that define the maximum number of bits or bytes that can be transmitted from one place to another in a fixed amount of time. In lossy data compression schemes, some loss of information is acceptable. Dropping nonessential detail from the data source can save storage space. Lossy data compression schemes are informed by research on how people perceive the data in question. For example, the human eye is more sensitive to subtle variations inluminance than it is to variations in color. Distributed video coding(DVC), which is a new coding paradigm for video compression based on Slepian wolf and Weiner ziv theorem. Slepian wolf theorem states that for lossless coding of two or more correlated sources, the optimal rate achieved when performing joint encoding and decoding can theoretically be reached by doing separate encoding and joint decoding. Weiner ziv theorem shows that this result still holds for lossy coding under the assumption that the sources are jointly Gaussian.

In recent years, DVC has attracted the interest of many researchers. Transform Domain Weiner Ziv (TDWZ) video coding is one popular approach to DVC. Then came the DISCOVER (DIStributed COding for Video sERvices) codec which introduced improvements on TDWZ coding efficiency. The coding efficiency of DVC critically depends on the quality of side information generation and noise residue. In typical system the encoder is made responsible for exploiting the redundancies by predicting the current frame to be coded from previously coded information. Then the residual between the frame to be coded and its prediction is transformed, quantized and coded. But here, the decoder is responsible for generating the prediction, thereby relieving the encoder from this complex task. In DVC, the prediction generated at the decoder is called the side information (SI). Earlier the side information were generated using block based method which have the drawbacks like – blocking artifacts (discontinuities at the block border) and less prediction accuracy. But in SING (Side Information and Noise Learning) codec, which came after DISCOVER, the SI is generated using optical flow based method which compensates the weakness of an Overlapping Block Motion Compensation (OBMC).

Therefore to overcome these drawbacks a TDWZ codec using Motion Reconstruction Re estimation (MORE) was proposed where

- an optical flow based motion re estimation to update side information and residue frames

Manuscript received June, 2015.

***K M Amritha**, Applied Electronics & Communication Systems, Nehru College of Engineering & Research Centre, Pampady, Thrissur, India, 9947378988*

***Nithin S S**, Applied Electronics & Communication Systems,Nehru College of Engineering & Research Centre, Thrissur, India,*

- a motion compensation to generate a more accurate estimate of correlation noise.
- a generalized reconstruction algorithm

are all integrated into SING codec.

The rest of this paper is organized as follows. In section II, which presents the proposed system with SPIHT algorithm and section III presents the simulation results

II. PROPOSED SYSTEM

The main objective of the paper is to create a codec with high coding efficiency, better quality of compression when compared to the existing methods.

A new method is proposed where an adaptive lifting based wavelet technique for video compression is incorporated. The original video is transformed using adaptive lifting based CDF 9/7 wavelet transform (Cohen-Daubechies-Feauveau wavelet are the historically first family of biorthogonal wavelet) and CDF 9/7 followed by compression using Set Partitioning In Hierarchical Tree algorithm (SPIHT) and the performance was compared with the popular traditional (TDWZ). SPIHT is an compression algorithm that exploits the inherent similarities across the sub bands in a wavelet decomposition. The algorithm codes the most important wavelet transform coefficient first and then transmits the bits so that an increasingly refined copy of the original copy of the original input can be obtained progressively. The performance metric Peak Signal to Noise Ratio (PSNR) for the reconstructed video will be computed. The proposed adaptive lifting algorithm gives better performance than traditional wavelet transforms. Lifting allows incorporating adaptivity and nonlinear operators into the transform. This proposed method will efficiently represent the edges and will be a promising for video compression. The proposed adaptive methods will reduce edge artifacts and ringing and will give an improved PSNR than the traditional (TDWZ). The flow chart of the proposed system is shown in Fig.1. SPIHT algorithm is a fast and efficient technique for compression. The main advantage of wavelet transform over discrete cosine transform is that it has both time and frequency localization ability, which can give a better performance in compression.

A. SPIHT Algorithm

This SPIHT method deserves special attention because it provides the highest image quality, progressive image transmission, fully embedded coded file, simple quantization algorithm, fast coding/decoding, completely adaptive, lossless compression, exact bit rate coding & error protection. SPIHT belongs to the next generation of wavelet encoders, employing more sophisticated coding. In fact, SPIHT exploits the properties of the wavelet-transformed images to increase its efficiency. SPIHT wins in the test of finding the minimum rate required to obtain a reproduction indistinguishable from the original. The SPIHT advantage is even more pronounced in encoding color images, because the bits are allocated automatically for local optimality among the color components, unlike other algorithms that encode the color components separately based on global statistics of the individual components. SPIHT represents a small "revolution" in image compression because it broke the trend to more complex (in both the theoretical and the computational senses) compression schemes. While researchers had been trying to improve previous schemes for image coding using very sophisticated vector quantization, SPIHT achieved superior results using the simplest method: uniform scalar quantization. Thus, it is much easier to design fast SPIHT codec. The SPIHT process represents a very effective form of entropy-coding. A straightforward consequence of the compression simplicity is the greater coding/decoding speed. The SPIHT algorithm is nearly symmetric, i.e., the time to encode is nearly equal to the time to decode. (Complex compression algorithms tend to have encoding times much larger than the decoding times). 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 (every single bit of it) by coding all bits of the transform. However, the wavelet transform yields perfect reconstruction only if its numbers are stored as infinite-precision numbers. SPIHT allows exact bit rate control, without any penalty in performance (no bits wasted with padding).

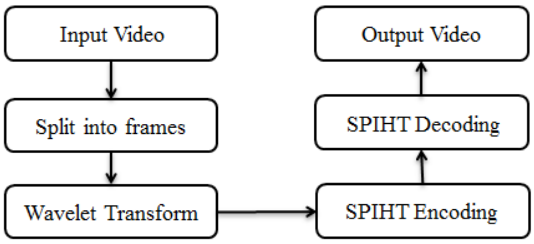

Fig.1: Flow chart of proposed system

The SPIHT technique is based on three concepts:
(i) Partial ordering of the transformed image elements by magnitude, with transmission of order by a subset partitioning algorithm that is duplicated at the decoder,
(ii) Ordered bit plane transmission of refinement bits,
(iii) Exploitation of the self-similarity of the image wavelet transform across different scales.

The partial ordering is a result of comparison of transform element (coefficient) magnitudes to a threshold. An element is significant or insignificant with respect to a given threshold, depending on whether of not it exceeds that threshold. The crucial part of coding process is that the way subsets of coefficients are partitioned and how the significance information is conveyed. The coding is actually done to the array

$$\mathbf{c} = \Omega(\mathbf{p}),$$

where $\Omega(.)$ represents a unitary hierarchical subband transformation and $\mathbf{p}$ is the original image. The decoder initially sets the reconstruction vector to zero and updates its components according to the coded message. After receiving the value (approximate or exact) of some coefficients, the decoder can obtain a reconstructed image.

$$\mathbf{p} = \Omega^{-1}(\mathbf{c})$$

B. Set Partitioning Sorting Algorithm

Main task of sorting pass in each iteration is to select those coefficient that satisfy

$$2^n \leqq |c_{i,j}| \leqq 2^{n+1}$$

with n decremented in each pass. Given n, if $|c_{i,j}| \geqq 2^n$ then the coefficient is significant; otherwise it is called insignificant. The sorting algorithm divides the set of pixels into partitioning subsets τ_m and performs the significance test.

$$\max_{(i,j)\in\tau_m}\left\{\left|c_{i,j}\right|\right\} \geq 2^n \ ?$$

If the result is NO, then all coefficients in τ_m are insignificant & τ_m is said to be insignificant. And if the result is YES, then some coefficients in τ_m are significant and τ_m is said to be significant. And this τ_m is partitioned by encoder & decoder. This partitioning is repeated until all significant sets are reduced to 1. Since the result of each significance test is a single bit written on a compressed stream, the number of test must be minimized. Therefore to achieve this goal a Spatial Orientation Tree (SOT) is used.

C. Spatial Orientation Tree

Spatial Orientation Tree (SOT) exploits the relationship between wavelet coefficients in different subband pyramid & it is used to create & partition the set τ_m. Fig.2 shows the parent –offspring dependencies in SOT. Each node of the tree corresponds to a pixel and it is identified by the pixel coordinate. Its direct descendants (offspring) correspond to the pixels of the same spatial orientation in the next finer level of the pyramid. The tree is defined in such a way that each node has either no offspring (the leaves) or four offspring, which always form a group of 2x2 adjacent pixels.

In Fig.2, the arrows are oriented from the parent node to its four offspring. The pixels in the highest level of the pyramid are the tree roots and are also grouped in 2 x 2 adjacent pixels. In each group, all of 4 coefficients (except the marked one) becomes the root of SOT. In general, a coefficient at (i, j) is the parent of 4 coefficient at location (2i,2j), (2i+1,2j), (2i,2j+1), (2i+1,2j+1).

The following sets of coordinates are used to present the new coding method:

- $O(i, j)$: set of coordinates of all offspring of node (i, j)
- $D(i, j)$: set of coordinates of all descendants of the node (i, j)
- $H(i, j)$: set of coordinates of all spatial orientation tree roots.

$$L(i, j) = D(i, j) - O(i, j)$$

The set partitioning rules are simply the following:

- The initial partition is formed with the sets $\{(i, j)\}$ and $D(i, j)$, for all $(i, j) \in H$.
- If $D(i, j)$ is significant, then it is partitioned into $L(i, j)$ plus the four single-element sets with $(k, l) \in O(i, j)$.
- If $L(i, j)$ is significant, then it is partitioned into the four sets $D(k, l)$, with $(k, l) \in O(i, j)$.

D. SPIHT Coding

Significance information are stored in 3 ordered lists- list of insignificant sets (LIS), list of insignificant pixels (LIP), and list of significant pixels (LSP). In all lists each entry is identified by a coordinate (i, j), which in the LIP and LSP represents individual pixels, and in the LIS represents either the set $D(i, j)$ or $L(i, j)$. During the sorting pass, the pixels in the LIP are tested, and those that become significant are moved to the LSP. Similarly, LIS is tested in sequential order, and when a set is found to be significant it is removed from the list and partitioned. In each iteration, when the coordinates C(i,j) are moved to LSP it is known to both encoder & decoder that

$$2^n \leqq |c_{i,j}| \leqq 2^{n+1}$$

Therefore, the best value that the decoder can give ĉ(i,j) is between 2^n & 2^n+1. During refinement pass, for each entry (i, j) in the LSP, except those included in the last sorting pass (i.e., with same n), output the nth most significant bit of $|c_{i,j}|$.

We can summarize the coding into four steps as followed:
Step 1: Initialization
Output n= $[\log_2(\max_{i,j}\{|Ci,j|\})]$
Set LSP={Ø}
Set LIP={(i,j)ϵH} & LIS={D(i,j), (i,j)ϵH}
Step 2: Sorting Pass
For each entry in the LIP, output the significance (“1” if significant, “0” if not significant). If found significant, remove it from the LIP and add to the LSP.
For each entry in the LIS, output the significance. If found significant, output its sign. Perform the set partitioning using any one of the rules. According to the significance, update the LIS, LIP and LSP.
Step 3: Refinement Pass
For each entry in the LSP, except those which are added during the sorting pass with the same n, output the nth most significant bit.
Step 4: Quantization and update pass
n is decremented by 1 & the steps-2, 3 and 4 are repeated until n = 0.
The decoder steps are exactly identical. Only the output from the encoder will be replaced by the input to the decoder.

E. Wavelet Transform

Wavelet compression is a form of data compression well suited for image compression (also video compression and audio compression). The goal is to store image data in as little space as possible in a file. Wavelet compression can be either lossless or lossy. This means that the transient elements of a data signal can be represented by a smaller amount of information than would be the case if some other transform, such as the more widespread discrete cosine transform. The main advantage of wavelet transform over discrete cosine transform (DCT) is that it has both time and frequency localization ability, which result in better performance in compression. Thus, researchers have paid

ISSN: 2278 – 1323

much attention to wavelet construction and proposed well-known wavelet bases.

First a wavelet transform is applied. This produces as many coefficients as there are pixels in the input (i.e., there is no compression yet since it is only a transform). These coefficients can then be compressed more easily because the information is statistically concentrated in just a few coefficients. This principle is called transform coding. After that, the coefficients are quantized and the quantized values are entropy encoded and/or run length encoded. The wavelet transform can provide us with the frequency of the signals and the time associated to those frequencies, making it very convenient for its application in numerous fields.

Wavelet Transform in general produces floating point coefficients. The biorthogonal wavelet transform implemented in a lifting based scheme reduces the computational complexity. Lifting based wavelet transform consists of Splitting, Lifting and Scaling modules.

- Splitting: X divided into 2; even & odd indexed samples of X.
- Lifting: Prediction operation P, estimate Xo(n) & Xe(n) results in an error signal d(n); d(n) updated by applying to the update operation U & resulting signal combined with Xe(n) to estimate s(n).
- Scaling: Normalization factor applied to s(n) & d(n) to produce XL1 &XH1 respectively.

Cohen-Daubechies-Feauveau (CDF) wavelet are the historically first family of biorthogonal wavelets, these are not the same as the orthogonal Daubechies wavelets, and also not very similar in shape and properties. Lifting scheme of biorthogonal transform 9/7 goes through 4 lifting steps & 2 scaling steps. Fig.3 shows forward CDF 9/7 discrete wavelet transform using adaptive lifting scheme. First lifting step (predict step 1) is applied to original row of samples and the results safely overwrite the odd samples in the original signals for use in the next lifting step. Similarly, this is done for four lifting steps.

First scaling step is applied to the results from the third lifting step. The results then safely overwrite the results from the third lifting step. The results from this step are the wavelet coefficients.

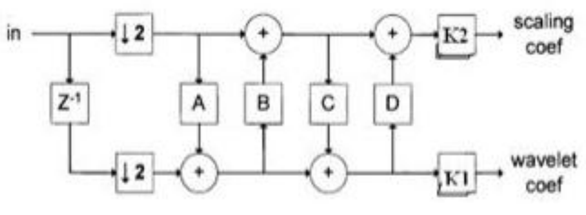

Fig. 3 Forward CDF 9/7 discrete wavelet transforms using adaptive lifting scheme

III. Simulation Results

The performance of the proposed system is evaluated by the average PSNR value. The existing method when evaluated for a test sequence *Crowd* had a PSNR of 29 dB. The proposed system has a PSNR of 40 dB. Thus this proposed system is more efficient than the existing method and will have better quality of compression.

IV. Conclusion

In this paper, we have presented a codec using SPIHT compression algorithm which will have more coding efficiency and better quality of compression when compared to the existing methods.

Acknowledgment

We wish to acknowledge Murughananthan and other contributors for developing this project and making it successful.

REFERENCES

[1] B. Girod, A.M. Aron, S. Rane and D.Rebollo Monedero, "Distributed video coding," Proc.IEEE, vol.93, no.1, pp.71-83, Jan 2005.

[2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," IEEE Trans. Inf. Theory, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[3] H. V. Luong, L. L. Raket, X. Huang, and S. Forchhammer, "Side information and noise learning for distributed video coding using optical flow and clustering," IEEE Trans. Image Process., vol. 21, no. 12, pp. 4782–4796, Dec. 2012.

[4] A Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," IEEE Trans. Circuits Syst. Video Technol., vol. 2, no. 1, pp. 91-95, Mar. 1992.

[5] Mounir Kaanichel, Beatrice Pesquet- Popescul, Amel Benazza, Adaptive lifting scheme with sparse criteria for image coding.

[6] D. ˇCern´a, V. Finˇek, discrete cdf 9/7 wavelet transform for finite-length signals

[7] Baochen Jiang1, Aiping Yang1, Chengyou Wang and Zhengxin Hou1, "Implementation of Biorthogonal Wavelet Transform Using Discrete Cosine Sequency Filter", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 6, No. 4, August, 2013.

[8] Ali Kadhim Al-Janabi," Low Memory Set-Partitioning in Hierarchical Trees Image Compression Algorithm", International Journal of Video&Image Processing and Network Security IJVIPNS-IJENS Vol:13 No:02

K M Amritha, M-Tech, Nehru College of Engineering & Research Centre, pampady, Thrissur.

Nithin S S, Assistant Professor, Nehru College of Engineering & Research Centre, Pampady, Thrissur.

