A Survey on Efficient Memory in Video Codec using Hybrid Algorithm

T.Bernatin, Priya Rani, Alka Kumari
Department of ETCE, Sathyabama University. Chennai, India

Abstract— *Now-a-days video is one of the important components in multimedia communication. In various applications such as broadcasting news over satellite and other channel, storage of data etc, video communication can be found. The unnecessary video data is removed so that the data can be sent efficiently and with less memory storage. For compressing the video, hybrid algorithm is used which is a combination of DWT and modified SPIHT algorithm. Discrete Wavelet Transform (DWT) is an efficient tool for compression because it offers higher compression ratio and minimizes the computation energy. DWT is widely used because of its capability of displaying images at both high and low frequency resolutions. SPIHT (Set Partitioning in Hierarchical Trees) provides good image quality with high PSNR, effectively optimized for progressive videos and is a fully embedded codec, so it is considered as the most effective and simpler algorithm. This paper analyzes a survey on various performance of SPIHT such as Peak Signal to Noise Ratio (PSNR), Picture Quality Scale (PQS), Least Mean Square Error (LMSE), and Compression Ratio (CR).*

Keywords— DWT, SPIHT, PSNR, LMSE, PQS, CR.

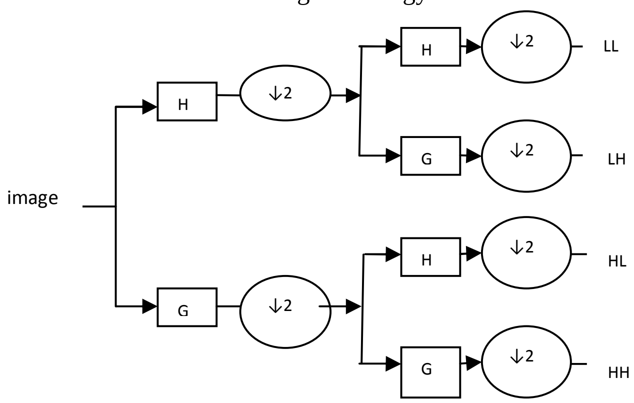
I. INTRODUCTION

These days, people use a variety of multi-media devices including cell phones, camcorders, digital cameras and even latest professional-level digital SLRs which contributes to inadequate transmission bandwidth of network and memory storage for device. So, the video compression concentrates the problem of reducing the amount of data required to represent a digital image. The main objective is to reduce the storage size as much as possible so that it can be transferred easily without losing the quality and information and to obtain significant computational time saving.

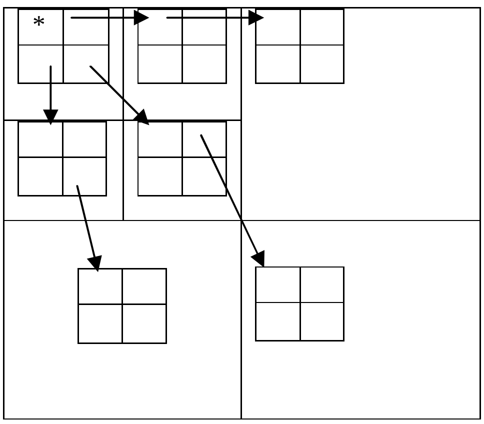
Codec is used for compressing and decompressing the video information. The work of codec is that they take digital media data and either compress them (for transport and storage) or decompress them (for viewing and transcoding). The first step in video codec is that an input video is converted into still images for processing. This image is separated into small blocks. Then discrete wavelet transform (DWT) is applied on each block which decomposes the still image into different sub-bands for getting pixel coefficient of the still image. DWT is a high resolution transformations and it spans the low frequency parts (approximation parts) of the signal as the LL sub-band can be seen by human vision. Now, SPIHT (Set Partitioning in Hierarchical Trees) algorithm is introduced which is a significant improvement of EZW (embedded zero-tree wavelet) algorithm. It provides highest image quality, fully embedded coded file and fast encoding/decoding. By the combination of DWT and modified SPIHT algorithm, we get the performance of high PSNR (peak signal to noise ratio) which is the most widely used video quality metric and high SSIM (Structural similarity index) which is a new metric that shows better results than PSNR. It also gives high compression ratio which is a term to quantify the reduction in data representation size and high picture quality scale which is a system for rating image quality based upon features of images that affect their perception by human eye. The quality of resultant video is relevant concern in video compression and thus to improve the quality Picture Quality Scale is used.

II. DISRETE WAVELET TRANSFORM

Discrete Wavelet Transform has become very versatile image processing tool and is widely used for compression purpose as it helps in manipulating video compression and progressive video transmission. In video compression, the digital videos must be accumulated and recovered in an effective and proficient manner. DWT displays data at different resolutions which means it breaks the data into different coefficient. They are approximate coefficient and detailed coefficient; out of which detailed coefficient is discarded due to high frequency rate and approximation coefficient is used for compression method at different resolution and has highest energy than other.



The still images are decomposed into different sub bands images using wavelet transform, namely, low-low (LL), high-low (HL), low-high (LH), and high-high (HH) in order to get pixel coefficient of an image. The LH sub band represents the vertical data of the image; HL sub band represent the horizontal part of the image; HH sub band represent the diagonal information about the image. The significant part of spatial domain is contained in LL sub bands. The image presents in LL sub band corresponds information visible at low resolution level and image present in that of LH, HL, and HH correspond to information available at high resolution level and which means more image data. The LL sub band is further decomposed into sub bands to get detail about that part of image.



The moving images is framed into still images and the images are decomposed into low pass filter and high pass filter along rows and columns and results are down sampled by a factor 2 of each image. For one image, we will get two sub images corresponding low frequency and high frequency along both rows and columns.

III. SET PARTITIONING HIERARCHICAL TREES

Set Partitioning in Hierarchical Trees (SPIHT) is an efficient and simple algorithm with many inherent characteristics such as it is self adaptive, completely embedded, accurate rate control, idempotent. SPIHT is an extended version of EZW (Embedded Zero tree Wavelet) coding and is one of best method used for coding. SPIHT supports

- Unrestricted dimensions images
- Multi resolution encoding or decoding
- Modularity
- 8,16, or larger bit images.

The video is first converted using wavelet transform and then information about wavelet coefficient is transmitted. Decoder uses received signal for reconstructing the wavelet and recovering the video and thus an inverse wavelet transform is performed. SPIHT is a method for encoding and decoding the wavelet transform of a video. In the course of encoding and transmitting data regarding wavelet transform, it will be possible to perform an inverse transformation on wavelet for decoder and reconstruct the original video. The coding process consist of two passes: sorting pass and refinement pass. Each and every node of the tree is recognized by the pixel coordinate which corresponds to a pixel. Its direct successors are called off spring, corresponds to the pixels of the identical spatial orientation in the next superior level of the pyramid.

Each node, has either no off spring or four off springs, which always form a group of closest pixel.

The pixels in the peak level of the pyramid are the tree roots and are also grouped in 2*2 adjacent pixels and in each group of four, one of them does not have any off springs.

Example of SPIHT algorithm is explained below-

21	**7**	**10**	**13**
8	**6**	**5**	**4**
5	**-6**	**3**	**2**
3	**4**	**2**	**0**

Using three passes at the encoder, transmitted bit stream is generated, and then bit stream is decoded.

1) During first pass

The value of n is 4. The lists at the encoder are:

LIP- {(0,0)→21; (0,1)→7; (0,2)→10; (0,3)→13}
LIS- {(0,1)→7; (0,2)→10; (0,3)→13}
LSP- {}

The element of LIP is examined. At (0,0), the coefficient is greater than 16, therefore it is a significant bit and a 1 is transmitted. If the coefficient is positive then it is indicated by 0 and the coordinate (0,0) is moved as the first entry in LSP.

In the list of LIP the next three coefficients are insignificant and thus a 0 is transmitted. The contents of LIS is examined and at the location (0,1) (10,13,5,and 4), none of the coefficient are significant at this point of threshold, a 0 is transmitted. All the coefficients at location (1,0) and (1,1) are insignificant at this point of threshold, so a 0 is transmitted.

8 bits are transmitted at the end of this pass

10000000

the three lists are

LIP- {(0,1)→7; (1,0)→8; (1,1)→6}
LIS- {(0,1)D; (1,0)D; (1,1)D}
LSP- {(0,0)→21}

D is the coordinates of descendants of (i,j).

2) During second pass

The value of n is decremented to 3 and threshold becomes $2^3 = 8$.

Now, content of LIP is examined and at location (0,1) and (1,1) two elements are insignificant so, a 0 is transmitted and at location (1,0) one element is significant so, a 1 is transmitted.

The content of LIS is examined and the coefficient at location (0,1) are 10;13;5;4, out of which two are significant and thus set (0,1) is significant and thus a 1 is transmitted. At location (1,0) and (1,1) all the elements are insignificant and thus a 0 is transmitted.

The content of LSP is examined during refinement pass from the previous pass and there exists only one element with value 21. The third most significant bit of 21 is 1, so we transmit a 1
(21_{10} = 10101 has the bits: $b4 = 1$, $b3 = 0$, $b2 = 1$, $b1= 0$, $b0 = 1$).

13 bits are transmitted in this pass:

0001010100001

And the lists are:

LIS-{(0,1)→7;(1,1)→6;(1,2)→5;(1,3)→4}
LIP-{(1,0)D;(1,1)D}
LSP-{(0,0)→21;(0,2)→10;(0,3)→13}

3) Now n is decremented to 2 and threshold becomes 2^2=4.

Bits sent during this pass is given as-

10111010101101100110000010

At the end of this pass is listed as-

LIP- {(2,2)→3;(2,3)→2;(3,0)→3;(3,2)→2;(3,3)→0}
LIS- {}
LSP-{(0,0)→21; (0,1)→7; (0,2)→10; (0,3)→13; (1,0)→8; (1,1)→6; (1,2)→5; (1,3)→4; (2,0)→5; (2,1)→6; (3,1)→4}

DECODING

Each list is initialized by decoder as an encoder:

LIP-{(0,0); (0,1); (0,2); (0,3)}
LIS-{(0,1)D; (1,0)D; (1,1)D}
LSP-{}

- Decoding, first pass

  After the bit streams 10000000 are received, the decoder can change the list to

  LIP-{(0,1); (1,0); (1,1)}
  LIS-{(0,1)D; (1,0)D; (1,1)D}
  LSP-{(0,0)}

ISSN: 2349-6495

The reconstruction of the image at this point is

19	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

- Decoding, second pass

After receiving the bit string 0001010100001 the lists are decoded by the decoder-

LIP-{(0,1); (1,0); (1,1); (1,2); (1,3)}

LIS-{(1,0)D; (1,1)D;}

LSP-{(0,0); (0,2); (0,3)}

The reconstruction of the image at this point

23	0	12	12
0	0	0	0
0	0	0	0
0	0	0	0

- Decoding, third pass

After receiving the bit string 10111010101101100110000010 the decoder can change the lists to

LIP-{(2,2);(2,3);(3,0);(3,2);(3,3)}

LIS-{}

LSP-{(0,0);(0,1);(0,2);(0,3);(1,0);(1,1); (1,2);(1,3);(2,0);(2,1);(3,1)}

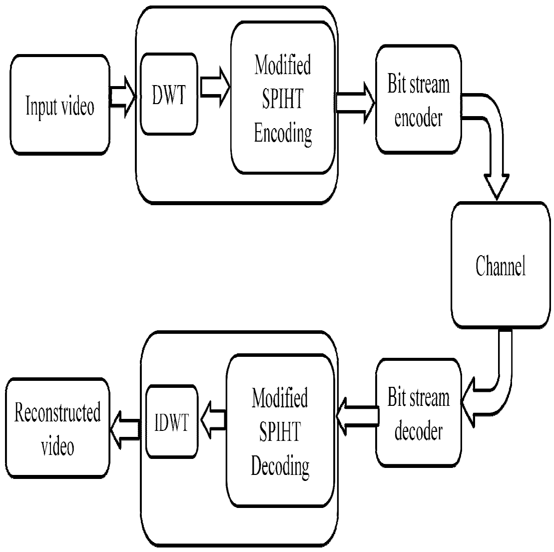
The reconstruction of the image at this point is

21	7	10	14
7	5	5	5
7	-7	5	0
0	5	0	0

IV. HYBRID ALGORITHM

The hybrid algorithm is a combination of Discrete wavelet transform (DWT) and modified Set partitioning in hierarchical tress (SPIHT) algorithm. These techniques are most popular for video compression and they have its own advantages and disadvantages. In DWT, wavelets are discretely sampled and the key advantage is that both frequency and location information are captured. It gives good compression ratio without losing more data of image but if we will do more than one level then we will get more compression ratio but the regenerated image is not similar to actual image. MSE is greater if DWT applies more than one level. It's not always true that we get better output in DWT. And it needs more processing power.

SPIHT (Set partitioning in hierarchical trees) algorithm is much powerful and efficient image compression technique as these have fully embedded coding property enabling easy progressive transmission of information. It provides high PSNR (peak signal to noise ratio) and better image quality. It is optimized for progressive image transmission and provides powerful error correction. SPIHT is completely adaptive and thus helps in fast encoding and decoding. SPIHT is the most efficient technique as it yields all the above objectives simultaneously. Though SPIHT is having many advantages over other compression techniques still it have many disadvantages like it is very unprotected to bit corruption, as a single bit error introduces remarkable image distortion depending on its location. SPIHT has a property of bit synchronization because leaking in single bit transmission leads to total misinterpretation from the side of the decoder. So, now when we will combine the advantages of both DWT and SPIHT algorithm, then we can overcome the disadvantages of both the techniques to get a better outcome.



The input video is first separated into several parts called images then each image is divided into different frames called sub-bands using the DWT technique. Each image is divided into approximation and detailed (vertical, horizontal, diagonal) sub-signal and then apply quantization on each frame, followed by modified SPIHT algorithm to encode the bit stream. Now it is transmitted through the channel and then decoded back through bit stream decoder. The SPIHT decoding is used to decode the values and inverse DWT is applied on different blocks to form a single image. At last, each images recombine to form a video.

PERFORMANCE ANALYSIS

The performance analysis of efficient memory in the video codec is based on certain parameters. These parameters can be calculated simply and quickly and provides a good way to evaluate the video quality.

I. Peak Signal To Noise Ratio (PSNR)

This is a ratio between the maximum possible power of a signal and power of the noise that affects the fidelity of its representation. PSNR can be expressed in the terms of logarithmic decibel scale. This is commonly used for measuring the quality of reconstruction of lossy compression codecs. When comparing the compression codecs, PSNR is an estimate to human observation of reconstruction quality. It examines the differences for every single pixel of the image. PSNR is defined via MSE (Mean squared error).

II. Mean Squared Error (MSE)

The mean squared error of an estimator measures the average of the squares of the errors, that is, the difference between what is being estimated and the estimator. The difference occurs because the estimator doesn't describe for information that could produce more precise estimate. Given a noise-free m×n monochrome image I and its noisy approximation K, MSE can be defined as:

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

The PSNR (in dB) can is calculated as:

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

III. Structural Similarity Index (SSIM)

The measurement of video quality is based on original uncompressed or undistorted video using a technique called Structural Similarity Index. In simple words, to find the similarity between two videos, SSIM is used. The methods like PSNR and MSE can be improved by using this method which is unpredictable with human eye perception. SSIM consider video degradation as apparent transform in structural information. . Structural information is the scheme that the pixels have strong inter-dependencies especially when they are spatially close and they carry significant data about the structure of the objects in the visual scene..

The SSIM of a video is calculated on different windows. The evaluation of two windows x and y of general size N×N is:

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2-\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)}$$

With

- μ_x the average of x;
- μ_y the average of y;
- σ_x^2 the variance of x;
- σ_y^2 the of variance y;
- σ_{xy} the covariance of x and y;
- $c_1=(k_1L)^2$, $c_2=(k_2L)^2$ two variables to stabilize the division with weak denominator;
- $L$the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel}-1$);
- $k_1=0.01$ and $k_2=0.03$ by default.

The resulting decimal value of SSIM lies between -1 and 1.

III. Compression Ratio (CR)

The method used to compute the decrease in data representation size is termed as compression ratio. Mathematically, it is defined as the ratio between the uncompressed size and compressed size.

IV. Picture Quality (PQS)

PQS is a system for rating image quality based upon features of images that affect their perception by human eye.

PERFORMANCE ANALYSIS OF HYBRID ALGORITHM

Techniques used	Compression Ratio	PSNR	MSE
DWT	3.23	50.31	4.16
SPIHT	3.19	34	40
HYBRID	5.32	30.72	55.07

V. CONCLUSION

This paper presents the combination of DWT and modified SPIHT algorithm for video compression. DWT is a high resolution transformation. SPIHT gives good signal to noise ratio, on demand information sorting and is good quality matrix. The advantages of both DWT and SPIHT combinedly gives the proposed hybrid algorithm, a better performance such as PSNR value is high, compression ratio is high, better picture quality and lossless information.

REFRENCES

[1] T.Bernatin, G.Sundari "A Survey of Video Compression Using Video Codec" IETE 45th Mid Term Symposium on Broadband-Technologies and Services for Rural India MTS 14, 4th and 5th April, 2014.

[2] T.Bernatin , G.Sundari "Video compression based on Hybrid transform and quantization with Huffman coding for video codec" International conference on control, instrumentation, communication and computational technologies (ICCICCT), pp 476-480,July -2014,IEEE.

[3] Mr. M.B. Bhammar, Prof. K.A. Mehta "Performance Improvement of SPIHT Algorithm using Hybrid Image Compression technique" ISSN: 0975 – 6779| Nov 12 to Oct 13 | Volume – 02, Issue – 02.

[4] Shriram P Hegde, S. Ramachandran "Implementation of DWT-SPIHT Algorithm for Image Compression" PISER 14, Vol.02, Issue: 04/06 July-August; Bimonthly International Journal ISSN 2347-6680 (E).

[5] M.Atheeshwari, K.Mahesh "Video Compression Techniques – A Comprehensive Survey" International Journal of Advanced Research in Computer Science and Software Engg, Volume 4, Issue 1, January – 2014 ISSN: 2277 128X.

[6] T.Swetha, L.Anju, "Image Compression Using Improved Seam Carving With Hybrid Compression Technique" International Journal of Advanced Research Vol. 3, Special Issue 2, April 2014.

[7] 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.

[8] Amir Said, William 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, Volume 6 June 1996.

[9] S.Vijayadheeswar Reddy, I.V.G.Manohar, Dr.Rajeyyagari Sivaram "High Throughput Image Compression Using SPIHT with Huffman Encoding" International Journal of Advanced Trends in Computer Science and Engineering Volume 1, No.3, July – August 2012 ISSN No. 2278-3091.

[10] Chandandeep Kaur, Sumit Budhiraja "Improvements of SPIHT in Image Compression- Survey" International Journal of Emerging Technology and Advanced Engineering, ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 3, Issue 1, January 2013.

[11] Nivedita, Pardeep Singh, Sonika Jindal "A Comparative Study of DCT and DWT-SPIHT" IJCEM International Journal of Computational Engineering & Management, Vol. 15 Issue 2, March 2012.

[12] Arvind Kourav, Dr. Ashutosh Sharma "Comparative analysis of wavelet transform algorithms for image compression" International Conference on Communication and Signal Processing, April 3-5, 2014, India.