QUALITY ASSESSMENT OF VIDEO COMPRESSION USING WAVELET TRANSFORM

By

L. ESCALIN TRESA *   **M. SUNDARARAJAN ****

* Research Scholar, Sathyabama University, India.
** Principal, Alpha College of Engineering & Technology, India.

ABSTRACT

Video compression is the practice of reducing the size of video files while maintaining as much of the original quality as possible. To accomplish this, an application known as "codec" analyzes the video frame by frame, and breaks each frame down into square blocks known as "macroblocks". The codec then analyzes each frame, checking for changes in the macroblocks. Areas where the macroblocks do not change for several frames in a row are noted and further analyzed. If the video compression codec determines that these areas can be removed from some of the frames, it does so, thus reducing the overall file size. There are two types of video compression: lossy and lossless. Lossy compression results in a lower file size, but also yields a loss in quality. Lossless compression, on the other hand, produces a less compressed file, but maintains the original quality. Videos and images are one of the most important approaches to represent some data. Now-a-days all the communication processes are working on such media. The main problem with this kind of media is its large size. Also, this large data contains a lot of redundant information. Compressing the video helps to reduce the size and thus saves the transmission bandwidth and storage space to achieve a high compression ratio while preserving the video quality. DWT (Discrete Wavelet Transform) and DCT (Discrete Cosine Transform) are the most common video compression techniques. DCT has high energy compaction and requires less computational resources, DWT on the other hand is a multiresolution transformation. But the compression ratio that can be achieved is low. The proposed method uses a Hybrid DWT-DCT algorithm on motion compensated frame by taking the advantages of both methods. The performance of the proposed method can be evaluated using compression ratio, PSNR (Peak Signal-to-Noice-Ratio) and mean square error.

Keywords: Compression, Discrete Cosine Transform, Discrete Wavelet Transform, Motion Compensation, Motion Estimation.

INTRODUCTION

Video has been an essential part of entertainment and communication nowadays. But it requires a large amount of storage space and transmission bandwidth. For example, a movie of 90 minutes of 30 frames per second frame rate and resolution of 750*570 will require 2.78 GB. The storage and bandwidth requirement of this uncompressed movie is very high. If this can be compressed to a size of certain MBs, the storage requirements can be reduced to a great extent. Video compression helps in achieving this without affecting the quality of the video. Videos and images contain a large amount of redundant information. The main technique of video compression relies in discarding this redundant information to which the human eyes are insensitive. While compressing the video, there is a tradeoff between the video quality and the amount of compression achieved. If the compression achieved is high, the quality of the reconstructed video is low. Also, the latency in compressing the video should be considered. Thus a good compression scheme aims in obtaining a better compression while considering all the parameters.

Discrete Cosine Transforms (DCT) and Discrete Wavelet Transform (DWT) are the most commonly used methods for video compression. DCT has high energy compaction and requires less computational resources, DWT, on the other hand, is a multiresolution transformation. But the compression ratio that can be achieved is low. A hybrid

method by combining DWT and DCT can achieve better result than individually applying the two methods. Most of the existing DCT based and DWT based codecs achieve a low compression ratio and the computation time is more [1]. On the other hand, the hybrid approach has shown much better results in terms of compression ratio, latency, PSNR and the quality of the reconstructed frame.

1. Overview Of Used Compression Algorithms

1.1 Discrete Wavelet Transform

The wavelet transform of a signal gives its time frequency representation. On the wavelet transform which was developed to overcome the shortcomings' of the Short Time Fourier Transform (STFT) is a multiresolution transformation. This property of the Fourier transform helps to analyse different frequencies of signals at different resolutions. In contrast to a wave which is an oscillating function of time or space, the wavelets are localized waves and are non periodic. The energy of the wavelets is concentrated in time or space.

1.2 Discrete Cosine Transform

Discrete Cosine Transform is one of the most commonly used transformation techniques. It is the base for most of the available video and image compression standards like JPEG, MPEG1, MPEG2, and MPEG4. The main advantage of DCT is its simplicity in calculation and good energy compaction. But the compression ratio that can be achieved is very low.

The two dimensional DCT is relevant these days since the images and the video frames are two dimensional [2]. This can be obtained from 1dimensional DCT by applying it vertically and horizontally across a signal. The DCT is an orthogonal transform, which has an image independent basis functions, an efficient algorithm for computation, and good energy compaction and correlation reduction properties.

1.3 Hybrid DWT – DCT Transform

The Hybrid DWT DCT transform exploits the properties of both the DWT and DCT techniques and provides a better compression. The input frame obtained from the video is first converted into a 32*32 block. Each block is then transformed individually. The 32*32 block is converted into

16*16 after one level DWT and discarding all the coefficients except the LL (i.e. LH, HH, and HL). The second level of the 2 dimensional DWT is applied on the retained LL coefficients. And this yields a 8*8 block after discarding all the LH, HH, HL coefficients and preserving only LL. The DCT is applied on this block. After the transformation by DCT, lossy compression occurs and then quantization is applied on the DCT coefficients which rounds off the high frequency components to zero [3].

The reconstruction can be performed by the reverse process i.e. first the inverse quantization is done and then the IDCT is performed which yields an 8*8 block. The first level IDWT gives a 16*16 block and the second level of IDWT gives the 32*32 block. This process is applied for the entire image.

2. Motion Estimation

Motion estimation is the most computationally complex and resources-dependent step in the compression process. It plays a crucial role in the video compression since a fair amount of compression can be achieved at this stage itself. Motion estimation can be either by pixel based or block based methods. The pixel based method is more computationally complex [4]. In the block based methods, the algorithm finds minimum cost between the blocks of a frame. There are various parameters that can be used to find the cost. The mostly used and computationally less complex ones are Mean Absolute Difference (MAD) and Mean Square Error (MSE).

Mean absolute difference is given by

$$MAD = \frac{1}{n*n}\sum_{j=1}^{n-1}|cij - rij| \tag{1}$$

Where n is the size of the macro block

Cij is the current pixel and rij is the reference pixel

Mean square error

$$MSE = \frac{1}{n*n}\sum_{j=1}^{n-1}(cij - rij) \tag{2}$$

Where n is the size of the macro block

Cij is the current pixel and rij is the reference pixel

2.1 Adaptive Rood Pattern Search

The frames of a video are highly coherent. The probability of the neighboring frames to have more redundant information is very high and this property is being exploited

by the ARPS. The motion vector of the current macro block is found from the motion vector of the block immediate left to it. The algorithm proceeds by checking the rood pattern around the predicted motion vector in addition to it. This rood pattern search is always the first step and the search is always performed in the region where the probability of finding the motion vector is very high. The centre for the subsequent search is selected as the point with the least cost and the search algorithm is changed to Small Diamond Search Pattern (SDSP).

Here the search area is confirmed as the diamond with five points and the point with the least cost is found. The algorithm repeats until the least cost is the centre of the diamond pattern. If the least cost point is found to be the centre of the rood pattern, the algorithm is stopped at that point and thus the computational time is reduced.

The advantage of ADPS over Diamond Search (DS) is that if the least weighted point is at the centre of the rood, it directly jumps to SDSP whereas the DS wastes its time in doing LDSP. Here the only care is needed in avoiding multiple computations ie checking the points many times. Also, if the first predicted vector is far away from the centre point, it jumps to SDSP instead of doing LDSP.

Among the entire block matching algorithms, ARPS is the one with minimum search points and thus less computational complexity and also the PSNR values are much similar to that of exhaustive search [5].

3. Huffman Coding

The entropy coding used here is the Huffman coding. This is a lossless compression technique which assigns a prefix code known as the Huffman code to the input signals. The basic idea of the Huffman coding is to assign to each of the source symbol alphabet a fixed number of bits which does not exceed the memory capability [6]. The length of the applied bits depends on the amount of information contained in the source symbol. Thus the main idea of the Huffman coding is to replace each of the source symbols with a simpler one and is carried out step by step. This process is repeated until only two symbols for which (0, 1) is the optimal code.

4. The Proposed Video Codec

The proposed algorithm aims to achieve a high compression ratio with less latency while preserving the original video quality by optimizing the available methods for video compression. A tradeoff between the different performance parameters is to be considered. The block diagrams of the proposed encoder and the decoder are given in Figures 1 and 2 respectively.

4.1 Encoder

The input video of any format is converted into frames. The frames are processed one after the other. For the intra frame (first frame), motion estimation and compensation are not performed. They are transformed by hybrid DWT-DCT transform and are reconstructed back by the inbuilt decoder at the encoder and are stored in the memory. From the second frame onwards (inter frames), the motion estimation and compensation is performed. The motion estimation is done by the adaptive rood pattern search algorithm by finding the difference between the current and the reference frames and then the second frame is predicted from the first frame and then the error frame between the first and the second frame is generated. The transform for the inter frames is applied to these kind of error frames that are generated. After quantization, these frames are entropy coded with Huffman coding and the generated bit streams are transmitted to the decoder side. The size of the compressed frame is reduced and thus there is a saving in the storage space [7]. Figure 1 shows the Encoder block diagram.

4.2 Decoder

The encoded bit stream is first inverse variable length coded and the decoded bit streams are first inverse cosine transformed and further inverse wavelet transformation is applied to the obtained coefficients. To this decompressed image, intensity motion compensation is applied. The motion compensation performed at the encoder side removes the high intensity pixels in the image and are reconstructed at this stage. The intensity motion
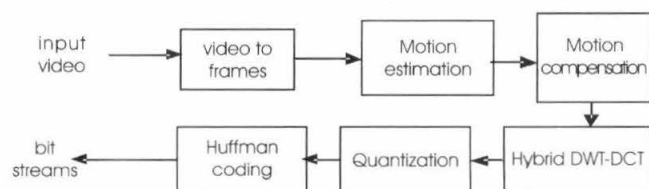

Figure 1. Encoder block diagram

method by combining DWT and DCT can achieve better result than individually applying the two methods. Most of the existing DCT based and DWT based codecs achieve a low compression ratio and the computation time is more [1]. On the other hand, the hybrid approach has shown much better results in terms of compression ratio, latency, PSNR and the quality of the reconstructed frame.

1. Overview Of Used Compression Algorithms

1.1 Discrete Wavelet Transform

The wavelet transform of a signal gives its time frequency representation. On the wavelet transform which was developed to overcome the shortcomings' of the Short Time Fourier Transform (STFT) is a multiresolution transformation. This property of the Fourier transform helps to analyse different frequencies of signals at different resolutions. In contrast to a wave which is an oscillating function of time or space, the wavelets are localized waves and are non periodic. The energy of the wavelets is concentrated in time or space.

1.2 Discrete Cosine Transform

Discrete Cosine Transform is one of the most commonly used transformation techniques. It is the base for most of the available video and image compression standards like JPEG, MPEG1, MPEG2, and MPEG4. The main advantage of DCT is its simplicity in calculation and good energy compaction. But the compression ratio that can be achieved is very low.

The two dimensional DCT is relevant these days since the images and the video frames are two dimensional [2]. This can be obtained from 1dimensional DCT by applying it vertically and horizontally across a signal. The DCT is an orthogonal transform, which has an image independent basis functions, an efficient algorithm for computation, and good energy compaction and correlation reduction properties.

1.3 Hybrid DWT – DCT Transform

The Hybrid DWT DCT transform exploits the properties of both the DWT and DCT techniques and provides a better compression. The input frame obtained from the video is first converted into a 32*32 block. Each block is then transformed individually. The 32*32 block is converted into

16*16 after one level DWT and discarding all the coefficients except the LL (i.e. LH, HH, and HL). The second level of the 2 dimensional DWT is applied on the retained LL coefficients. And this yields a 8*8 block after discarding all the LH, HH, HL coefficients and preserving only LL. The DCT is applied on this block. After the transformation by DCT, lossy compression occurs and then quantization is applied on the DCT coefficients which rounds off the high frequency components to zero [3].

The reconstruction can be performed by the reverse process i.e. first the inverse quantization is done and then the IDCT is performed which yields an 8*8 block. The first level IDWT gives a16*16 block and the second level of IDWT gives the 32*32 block. This process is applied for the entire image.

2. Motion Estimation

Motion estimation is the most computationally complex and resources-dependent step in the compression process. It plays a crucial role in the video compression since a fair amount of compression can be achieved at this stage itself. Motion estimation can be either by pixel based or block based methods. The pixel based method is more computationally complex [4]. In the block based methods, the algorithm finds minimum cost between the blocks of a frame. There are various parameters that can be used to find the cost. The mostly used and computationally less complex ones are Mean Absolute Difference (MAD) and Mean Square Error (MSE).

Mean absolute difference is given by

$$MAD = \frac{1}{n*n}\sum_{j=1}^{n-1} |c_{ij} - r_{ij}| \qquad (1)$$

Where n is the size of the macro block

c_{ij} is the current pixel and r_{ij} is the reference pixel

Mean square error

$$MSE = \frac{1}{n*n}\sum_{j=1}^{n-1} (c_{ij} - r_{ij}) \qquad (2)$$

Where n is the size of the macro block

c_{ij} is the current pixel and r_{ij} is the reference pixel

2.1 Adaptive Rood Pattern Search

The frames of a video are highly coherent. The probability of the neighboring frames to have more redundant information is very high and this property is being exploited

by the ARPS. The motion vector of the current macro block is found from the motion vector of the block immediate left to it. The algorithm proceeds by checking the rood pattern around the predicted motion vector in addition to it. This rood pattern search is always the first step and the search is always performed in the region where the probability of finding the motion vector is very high. The centre for the subsequent search is selected as the point with the least cost and the search algorithm is changed to Small Diamond Search Pattern (SDSP).

Here the search area is confirmed as the diamond with five points and the point with the least cost is found. The algorithm repeats until the least cost is the centre of the diamond pattern. If the least cost point is found to be the centre of the rood pattern, the algorithm is stopped at that point and thus the computational time is reduced.

The advantage of ADPS over Diamond Search (DS) is that if the least weighted point is at the centre of the rood, it directly jumps to SDSP whereas the DS wastes its time in doing LDSP. Here the only care is needed in avoiding multiple computations ie checking the points many times. Also, if the first predicted vector is far away from the centre point, it jumps to SDSP instead of doing LDSP.

Among the entire block matching algorithms, ARPS is the one with minimum search points and thus less computational complexity and also the PSNR values are much similar to that of exhaustive search [5].

3. Huffman Coding

The entropy coding used here is the Huffman coding. This is a lossless compression technique which assigns a prefix code known as the Huffman code to the input signals. The basic idea of the Huffman coding is to assign to each of the source symbol alphabet a fixed number of bits which does not exceed the memory capability [6]. The length of the applied bits depends on the amount of information contained in the source symbol. Thus the main idea of the Huffman coding is to replace each of the source symbols with a simpler one and is carried out step by step. This process is repeated until only two symbols for which (0, 1) is the optimal code.

4. The Proposed Video Codec

The proposed algorithm aims to achieve a high compression ratio with less latency while preserving the original video quality by optimizing the available methods for video compression. A tradeoff between the different performance parameters is to be considered. The block diagrams of the proposed encoder and the decoder are given in Figures 1 and 2 respectively.

4.1 Encoder

The input video of any format is converted into frames. The frames are processed one after the other. For the intra frame (first frame), motion estimation and compensation are not performed. They are transformed by hybrid DWT-DCT transform and are reconstructed back by the inbuilt decoder at the encoder and are stored in the memory. From the second frame onwards (inter frames), the motion estimation and compensation is performed. The motion estimation is done by the adaptive rood pattern search algorithm by finding the difference between the current and the reference frames and then the second frame is predicted from the first frame and then the error frame between the first and the second frame is generated. The transform for the inter frames is applied to these kind of error frames that are generated. After quantization, these frames are entropy coded with Huffman coding and the generated bit streams are transmitted to the decoder side. The size of the compressed frame is reduced and thus there is a saving in the storage space [7]. Figure 1 shows the Encoder block diagram.

4.2 Decoder

The encoded bit stream is first inverse variable length coded and the decoded bit streams are first inverse cosine transformed and further inverse wavelet transformation is applied to the obtained coefficients. To this decompressed image, intensity motion compensation is applied. The motion compensation performed at the encoder side removes the high intensity pixels in the image and are reconstructed at this stage. The intensity motion
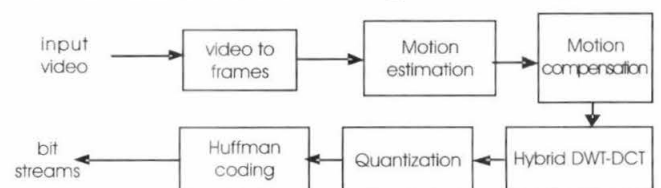

Figure 1. Encoder block diagram

compensated image has a better visual quality than the decompressed image [8]. Figure 2 shows the Decoder block diagram.

5. Performance Criterion

5.1 Compression Ratio

The reduction in the size of data due to compression is given by compression ratio. It gives the amount of compression performed. Space savings obtained by compression can be also shown by compression ratio

CR = Discarded Data / Original Data

5.2 Peak Signal to Noise Ratio (PSNR)

PSNR is the ratio between the powers of the reconstructed image to that of the noise in the image [9]. It is used to measure the quality of the reconstructed image. It is given by

$$PSNR = 10 * \log_{10}(255^2 / MSE) \qquad (3)$$

Where MSE is mean square error

5.3 Mean Square Error (MSE)

MSE measures the average of the squared error in the reconstructed image. It is the difference between the square of the original and the reconstructed image.

$$MSE = \frac{1}{(m*n)} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2 \qquad (4)$$

6. Simulation Results

The input frame is shown in Figure 3. Figures 4 and 5 shows the compressed and decompressed frames of DWT. The input frame i.e. the current frame is compared with the reference frame and motion estimation is performed on these frames. This process yields the motion vector. The motion vectors are used for the compensation to yield the predicted frames. The rest of the process functions on these predicted frames. After motion compensation, wavelet transform is applied on the predicted frame. Here, one
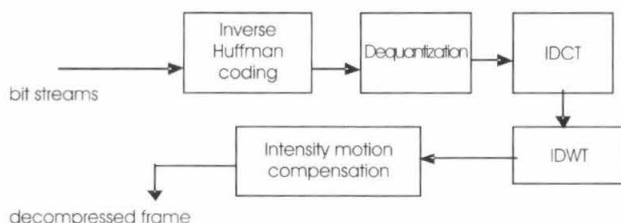

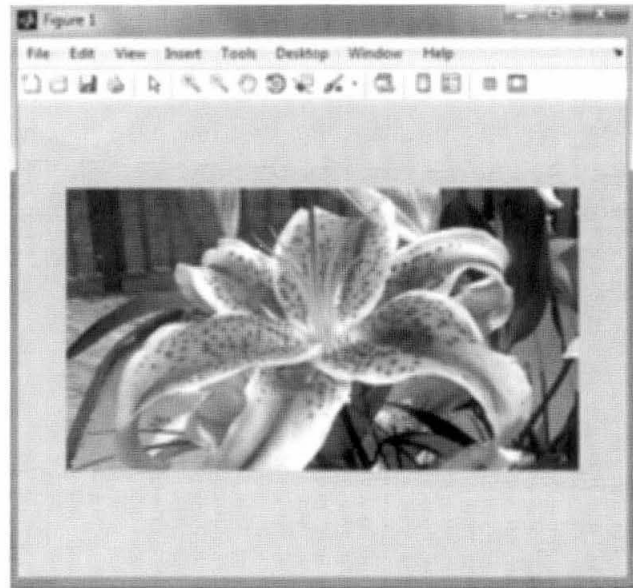Figure 2. Decoder block diagram

Figure 3. Input frame

Figure 4. Compressed frame(DWT)

Figure 5. Decompressed frame(DWT)

level two dimensional wavelet transform is used and the filter used is biorthogonal. After the wavelet transform, cosine transformation is applied on the wavelet transformed frame. After the hybrid transform, the

quantization is performed and then entropy coding is done. The compressed frame is shown in Figure 6. After the compression, the process is reversed for decoding. First, the inverse VLC and then dequantization is done. Then the IDCT is performed and then the IDWT is done. The decompressed frame obtained is shown in Figure 7. The motion compensation at the encoder results in loss of some intensity which is regained at the decoder through intensity based motion compensation. Figure 8 shows the comparison of the compression ratio, PSNR and MSE of DWT and the hybrid method. It is clear that the proposed hybrid scheme system can achieve a high compression ratio of 96% and the PSNR values are also good. From the decompressed frame, it is clear that the reconstructed frame is having better visual quality.

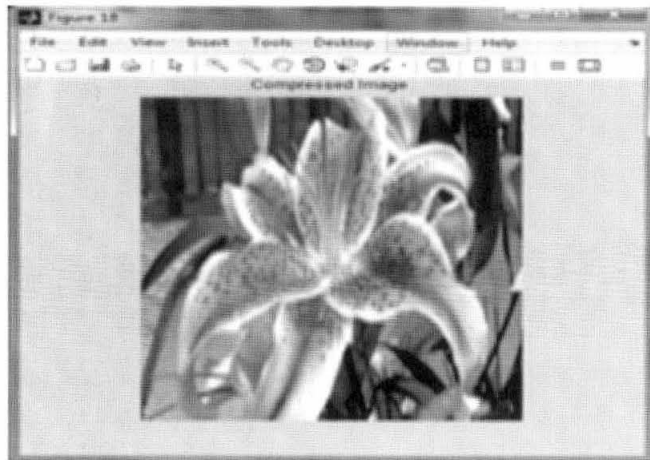Table 1 shows the parameter value comparison of DWT and hybrid DWT-DCT in which both has same compression

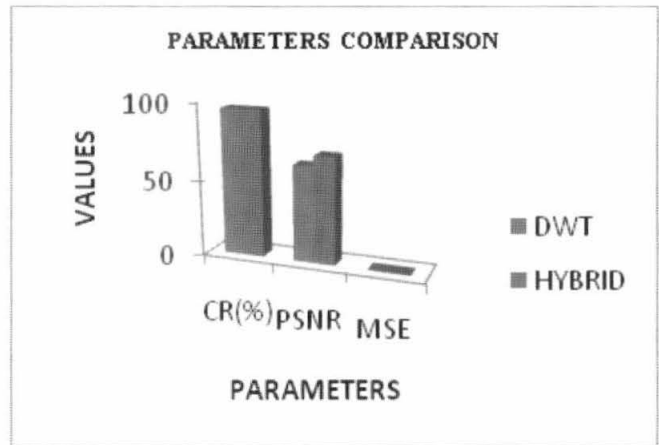Figure 6. Compressed frame(HYBRID)

Figure 7.Decompressed frame (HYBRID)

Figure 8. Comparison of Parameters for DWT and Hybrid DCT-DWT

WAVELET	CR(%)	PSNR	MSE
DWT	96	63	0.0057
HYBRID	96	70	0.0278

Table 1. Parameter values of DWT and Hybrid DWT-DCT

TRANSFORM	FORMAT	5 Sec	10 Sec	15 Sec
DWT	AVI	15.8370	13.7676	28.6915
	MPEG	19.5329	16.5800	15.4361
	WMV	17.1009	17.5217	20.9782
HYBRID	AVI	229.6213	457.0950	245.5272
	MPEG	228.7956	404.6031	270.2712
	WMV	262.9929	314.9858	337.7835

Table 2. Execution time of DWT and Hybrid DCT-DWT

ratio, better PSNR and minimum error for hybrid method.

Also in this work, video of different formats like AVI, MPEG and WMV is executed and the time taken to execute videos of different seconds is given in Table 2.

Conclusion

The input video is converted into frames and the size of the frames is chosen as per the requirement. Using adaptive rood pattern search, the motion vectors are found and it is compensated globally. Wavelet decomposition is applied to the compensated image and DCT is applied. The resultant hybrid compressed frame is quantized and entropy coded with Huffman coding. The encoded bit stream is first inverse Huffman coded and IDCT transformed. It is further IDWT transformed and intensity compensated. Also the hybid DWT-DCT is compared with DWT. Hybrid method gives a high cmpression ratio and a better reconstruction with minimum error.

References

1]. Thazni Aziz, D. Raveena Judie Dolly, (2012). Motion Estimation and Motion Compensated Video Compression Using DCT And DWT, *International Journal of Emerging Technology and Advanced Engineering*, (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Vol. 2, Issue 12, .

2]. Kohtaro Asai, Tokumichi Murakami, Shuichi Yamagishi, Akira Minezawa, Yusuke Itani, (2011). "New Video Coding Scheme Optimized for High-Resolution Video Sources", *IEEE Journal of selected topics in signal processing*, Vol. 5, No. 7.

3]. Mayank Nema, Lalita Gupta, N.R. Trivedi, (2012). Video Compression using SPIHT and SWT Wavelet", *International Journal of Electronics and Communication Engineering*, Vol. 5, No 1.

4]. Dishant Khosla, Amandeep Kaur, (2012). "Design of Hybrid Compression Model using DWT-DCT-HUFFMAN Algorithms for Compression of Bit Stream", *International Journal of Engineering Research & Technology (IJERT)*, Vol. Issue 5.

5]. MuzhirShaban, Al-Ani1 andTalal Ali Hammour, (2011).

"Video Compression Algorithm Based on Frame Difference Approaches", *International Journal on Soft Computing (IJSC)*, Vol.2, No.4.

[6]. Limin Liu, Member, Zhen Li, and Edward J. Delp, (2009). "Efficient and Low-Complexity Surveillance Video Compression Using Backward-Channel Aware Wyner–Ziv Video Coding", *IEEE transactions on Circuits and Systems for Video technology*, Vol. 19, No. 4.

[7]. Ram Lautan Verma, Deepti Ojha, Parul Gupta, Manisha Gupta, "Excellence of dct- dwt a hybrid approach for image compression", AIET, Lucknow, U.P., India.

[8]. Sandhya Sharma and Sarabjeet Kaur, (2013). Image Compression using hybrid of DWT, DCT and Huffman Coding, *International Journal for Science and Emerging*, ISSN No. (Online):2250-3641 Technologies with Latest Trends 5(1): 19-23.

[9]. Sandhya Sharma, Urvashi Bhat, (2013). Image Compression using an efficient hybrid algorithm, *Journal on Today's Ideas - Tomorrow's Technologies*, Vol. 1, No. 1 pp. 45–50.

ABOUT THE AUTHORS

Research Scholar, Sathyabama University, India.

Principal, Alpha College of Engineering & Technology, India.