Video Compression Using H 264 Standard

T. GOPI NADH REDDY[1], S. MAHESH BABU[2], P SAI MAHESH[3], SK SUBHANI[4], P HEMANTH YADAV[5]

[1, 2, 3, 4, 5] *Student, Vasireddy Venkatadri Institute of Technology, A. P, India*

Abstract- There is a huge demand for video data transmission in telecommunication series, entertainment industry. Here the problem is digital video data rate are very large which consumes more band width, storage. For this reasons video must be compressed for easy transmission and compact storage. So video compression standards have been developed to reduce picture redundancies and allowing information to be transmitted and stored in efficient manner. Even though there is rapid progress in mass storage capacity, processor speed and communication system performance demand for static storage capacity and data transmission bandwidth continue to outstrip the capabilities of available technologies. The growth of various multimedia-based web applications increasing rapidly so they need efficient way for transmitting data with less transmission bandwidth. For this video data must be compressed. Here we are using H 264 standard for compressing the video data. This standard provides high compression ratio and better quality compare to other techniques.

Indexed Terms- H 264, DCT, IDCT, Compression ratio

I. INTRODUCTION

There are different compression algorithms which were proposed and established as standards, such as joint photographic experts group (JPEG) standard and discrete cosine transform based coding, discrete wavelet-based coding. [1] However, every standard and algorithms have their complexities .Implementing such algorithms in hardware is usually expensive and consumes lot of power Discrete wavelet transform (DWT) is the fastest computation of wavelet transform. [2] It is a simple and lessen computational time and resources. In DWT, a time scale representation of signal is obtained. This signal is to be analyzed and passed through filters having different cut off frequencies at different scales. DWT decompose the signals into different sub-bands with both frequency and time information to obtain a high compression ratio. It supports progressive image transmission.

Another method that has gained a great deal of popularity in recent years is Discrete cosine transform(DCT) which is well known and commonly used for image compression and video compression. DCT converts the pixels in a picture into sets of spatial frequencies. During quantization less significant frequencies are discarded.

Since we know that video compression are often achieved by using image compression on each frame of the video and applying techniques such as motion estimation and compensation, it is always a better idea to have some knowledge about how to apply motion estimation to achieve compression. Several algorithms have been developed for motion estimation. Generally, two types are mostly discussed they are:

1. pixel-based motion estimation
2. block-based motion estimation

In pixel-based motion estimation, the motion vectors of each pixel in the image is determined. This estimation is based on the fundamental principle that is intensity of a pixel remains constant even displaced, hence it is also called as the optimal flow method. Block-based motion estimation has become popular for its fast approach. Also, computational complexity of this model is low compared to the pixel-based motion estimation. This model takes an assumption that the image is made up of significant objects in translational model

II. PROPOSED METHOD

[3] Here we are using H 264 standard for video compression. H264 is advanced standard for video compression, the process of converting digital video into a format that takes less capacity when it is stored

or transmitted. It provides better quality video with same bit rate and better compression ratio compare to other standards. This standard uses discrete cosine transform (DCT) approach which convert the pixel intensity values into spatial frequencies. This technique can be implemented without increasing the complexity of the system and it also provide flexibility to perform vast operations. Mostly this standard is used in applications such as television, DVD-video, videoconferencing and internet video streaming. [4] Following figure shows the encoding and decoding processes and highlights the parts that are covered by the H.264 standard.

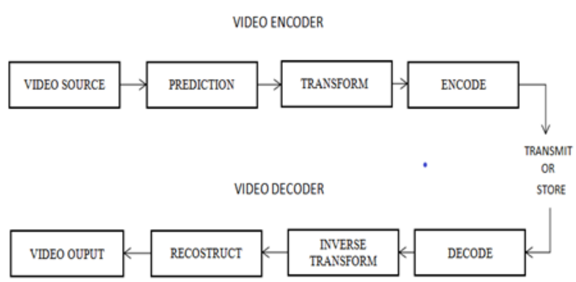


Figure: H264 standard

III. ENCODER PROCESS

A. Prediction

The encoder processes a frame of video in units of Macro blocks i.e. 16x16 displayed pixels. And prediction of these Macro blocks is based on the previous coded data, either from the current frame or from other frames that are already transmitted. Encoder subtract the predicted information from the current Macro block to form residual.

Prediction methods used by the H 264 are more flexible than other standards, getting accurate predictions and efficient compression. Intra prediction uses 16x16, 4x4 size blocks for predicting the Macro block from surrounding, previously coded pixels within the same frame. Inter prediction also uses same block sizes to predict pixels in current frame from similar regions in already coded frame.

B. Discrete cosine transform

Discrete cosine transform is an intra-Frame compression type because it applies DCT to each frame and then compresses each frame. After all the frames were compressed, the compressed video output is obtained. Whose size is smaller compare to original video. In DCT method, the compression occurs in three steps.

Step1: [5] The image is converted into frames and then DCT is applied on each frame, which converts entire pixel values into frequencies. As human eye is sensible only to low frequencies, high frequencies were discarded.

Step2: During quantization, which is the primary source of data loss, the DWT terms are divided by a quantization matrix, which considers human visual perception. [6] The human eyes are more reactive to low frequencies than to high ones. Higher frequencies end up with a zero entry after quantization and the domain was reduced significantly.

Step3: Finally, coding technique is applied for compression.

DCT is simple and symmetrical, it is easy to be implemented compared to the other compression techniques.

C. Bitstream Encoding

[7] During video coding process number of values must be encoded to form a compressed bit stream. These values include:

1. Quantized coefficients
2. Information that enables decoder to re-construct the prediction
3. Information about structure of compressed data and compression tools used in encoding process.
4. Information about video sequence.

These values and syntax elements are converted into binary form using run-length coding or variable length coding i.e. Huffman coding. These encoding methods produce compact and efficient binary representation of the information. The encoded Bitstream can be stored and transmitted.

IV. DECODER PROCESS

A. Bitstream decoding

Decoder receives the compressed encoded bitstream, the encoded bit stream contains all the information that a decoder needs to decode the bitstream for reconstruction of video. [8] It decodes each of the syntax and extract the information such as quantized coefficients, prediction information etc. This information is used in reversing the coding process and reconstruct a sequence of video images.

B. Re-scaling and Inverse transform
After decoding the bitstream the quantized transform coefficient is re-scaled. Re-scaling is the process in which each coefficient is multiplied by an integer value to restore its original scale. An inverse transform combine the standard patterns, weighted by re-scaled coefficients to reconstruct each block of residual data. The inverse discrete cosine transform (IDCT) decodes all the spatial frequencies into their respective pixel intensity values. IDCT based decoding is the basis for image and video decompression. In H 264, after re-scaling and zig-zag positioning the input is given to the IDCT. An 8x8 block of input have values in rage from -2048 to 2047 and output values range from -256 to 255. This information helps in the reconstruction.

C. Reconstruction
For each macro block, the decoder frames an indistinguishable prediction to the one made by the encoder. The prediction is added to the residual. Which is decoded to reproduce a decoded macro block which would then be able to be shown as some portion of a video frame.

V. EXPERIMENTAL RESULTS

Following parameters will describe the degradation of video after the compression.

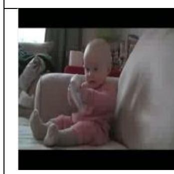
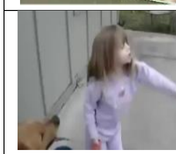
1. Peak Signal-to-Noise Ratio (PSNR): PSNR measures the amount of degradation after compression. PSNR is defined as ratio between the maximum possible power of a signal and the power of corrupting noise.

$$\text{PSNR} = 10 * \log \frac{255^2}{\text{MSE}}$$

2. Mean Square Error (MSE): The MSE describes the cumulative squared error between the compressed image to the original image. The mean square error is the average of the squared errors between original and estimated readings in a data sample.

3. Compression Ratio (CR): The compression ratio is defined as the ratio of compressed date rate to uncompressed data rate:

CR= Original Video Size/Compressed Video Size
The below table perfectly demonstrates the performance of the proposed system in terms of PSNR and Compression Ratio.

Input Video Frame	Video File Name	PSNR (dB.)	Compression Ratio
	Baby.avi	60.58	8.576
	TennisBall.avi	64.28	9.036
	Girl.avi	58.62	8.75

The Visual Representation of GUI that we used in our proposed System is shown in figure below:

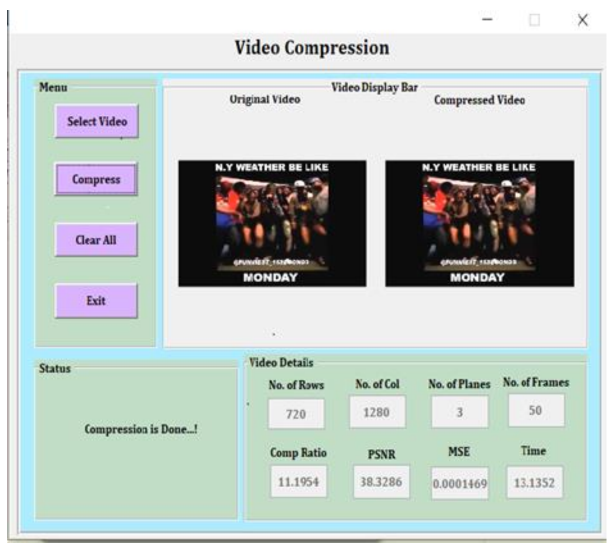


Fig : Original video and compressed video

CONCLUSION

The proposed adaptive standard is mainly meant for mobile communication and internet video streaming. This algorithm further can be extended to work with 3D video images. Since the standard eliminates particular insignificant band of frequencies that saves the computations it can be implemented in areas where battery constraints and resources are economical. If the parameter values applied are increased and enables to lower the number of bits transmitted, then the compression ratio is high

REFERENCES

[1] Peter Symes, Digital Video Compression, McGraw Hill Professional, 2004
[2] Vasudev Bhaskaran, Image and Video Compression Standards,: Algorithms and Architectures.
[3] White Paper: An Overview of H.264 Advanced Video Coding Iain Richards Vcodex
[4] Detail Block Diagram: http://inst.eecs.berkeley.edu/~ee290t/sp04/lectures/video_coding.pdf
[5] DCT: http://www.onlinejournal.in/IJIRV3I3/121.pdf
[6] Quantization: https://shodhganga.inflibnet.ac.in/bitstream/10603/123449/9/09_chapter%202.pdf
[7] Bit Stream Encoding: http://www.videosurveillance.co.in/H.264.pdf
[8] Bit Stream Decoding:https://www.vcodex.com/h264avc-4x4-transform-and-quantization/
[9] Quality measures base on PSNR: http://en.wikipedia.org/ wiki/Peak_signal-to-noise_ratio