ISSN 2349-6193
Impact Factor (PIF): 2.243

IJESMR

International Journal OF Engineering Sciences & Management Research

ANALYSIS OF VIDEO COMPRESSION IN MPEG-2 STANDARD

*** Mr.V.R.Prakash, Mr.Preetam**
Hindustan University, Chennai

KEYWORDS: MCVC'S, DVD, HDTV.

ABSTRACT
All current video coding standards are motion-compensated video coders (MCVC's), where the current frame is predicted using a previously reconstructed frame and motion information, which needs to be estimated. The most common approach to exploit the temporal redundancy is motion-compensated prediction. MPEG-2 bit stream is basically just a series of coded frames one after the other. There are headers and time stamps to help decoders align audio and scrub through the bit stream, but those details are not important to understand the basic coding techniques. What follows is a brief description of MPEG-2 compression techniques without focusing on the exact specification of the bit stream, standard format used for satellite TV, digital cable TV, DVD movies, and HDTV. In addition, MPEG-2 is a commonly used format to distribute video files on the internet.

INTRODUCTION
MPEG-2 is an evolution of MPEG-1, an earlier MPEG coding standard finalized in 1991. In fact, MPEG-2 is a superset of the MPEG-1 standards in that any MPEG-2 decoder can decode an MPEG-1 video[2]. The additions to MPEG-2, therefore, are what make it a separate standard. The major additions are:

- Support for higher resolution video
- Support for interlaced video (as used on standard definition TV (SDTV))
- Optimized for higher bit rates (typically 4 Mb/s and above, versus 1.5 Mb/s and below for MPEG-1)
- Scalability via layered encoding to support a variety of quality levels/transmission bandwidths from one coded source These features plus the base technology of MPEG-1 make MPEG-2 very attractive for video storage and transmission. The adjustable compression level allows for a significant reduction in bit rate with a just noticeable difference (jnd) in video quality or higher quality video[4] for a fixed bit rate. Coding a frame in MPEG-2 format always begins by representing the original color frame in YCbCr format. The Y component represents luminance, while Cb and Cr represent chrominance differences. The three components of this color model are mostly uncorrelated so this transformation is a useful first step in reducing redundant information in the frame. Another way to reduce the amount of information to be encoded is by taking advantage of human vision characteristics. Human eyes are much more sensitive to luminance than chrominance, so it is common to subsample both chrominance channels. The most commonly used subsampling format is denoted by 4:2:0, which means that chrominance sampling is decimated [5] by 2 in the horizontal and vertical direction. Both chrominance channels are reduced to one quarter the original data rate in this way and the net effect is for total frame data rate to be cut in half with hardly any perceptual effect on image quality.

I, P, and B frames
The next encoding step can vary from frame to frame. There are actually three possible types of frames, called I, P, and B frames. Each will be discussed separately below.

I frame coding
An I frame is intra or spatially coded so that all the information necessary to reconstruct it is encoded within that frame's segment[7]of the MPEG-2 bit stream. It is a self-contained image compressed in a manner similar to a JPEG image.

Each block of the frame is processed independently with an 8x8 discrete cosine transform (DCT). This transform generates a representation of each 8x8 block in the frequency domain instead of the spatial domain. Since, as noted above, the pixels in each block of a natural video are likely to be correlated, the resulting DCT coefficients typically consist of a few large values and many small values. The relative sizes of these coefficients represent how important each one is for reconstructing the block accurately.

The coefficients are then quantized using

$$QDCT = \text{round}\left(\frac{8 \times DCT}{scale \times Q_i}\right)$$

IJESMR
International Journal OF Engineering Sciences & Management Research

where DCT and QDCT represent the coefficient matrices before and after quantization, scale is used to determine quality, and Qi is a perceptually weighted quantization table. The default value of Qi,

Qi = [8 16 19 22 26 27 29 34]
[16 16 22 24 27 29 34 37]
[19 22 26 27 29 34 34 38]
[22 22 26 27 29 34 37 40]
[22 26 27 29 32 35 40 48]
[26 27 29 32 35 40 48 58]
[26 27 29 34 38 46 56 69]
[27 29 35 38 46 56 69 83]

weights low frequency content higher that high frequency content, mimicking the response characteristics of a human eye. The quality scalar, scale, can be any value from 1 to 31. If desired, MPEG-2 allows the value of scale to be mapped non-linearly to higher compression levels, with a maximum of 112.

This quantization process is lossy because the true floating point values of the DCT coefficients are not preserved. Instead, the quantized coefficients are just an approximation of the true coefficients and the quality of the approximation determines [7] the quality of the frame reconstructed from this bit stream.

Next the 8x8 block of quantized coefficients is arranged into a vector by indexing the matrix in zigzag order given by

[1 2 6 7 15 16 28 29]
[3 5 8 14 17 27 30 43]
[4 9 13 18 26 31 42 44]
[10 12 19 25 32 41 45 54]
[11 20 24 33 40 46 53 55]
[21 23 34 39 47 52 56 61]
[22 35 38 48 51 57 60 62]
[36 37 49 50 58 59 63 64]

This ordering is approximately low frequency to high frequency for a very important reason: small blocks (8x8) of natural video are likely to contain mostly low frequency content. This ordering tends to group non-zero terms at the front of the vector and zero terms at the end. This type of distribution is beneficial to the coding that follows.

The final steps for coding an I frame are lossless with respect to the quantized DCT coefficients. For each block, the DC coefficient is differentially coded with the last block's DC term. The AC coefficients are run-length encoded (RLE) and then Huffman coded. The result is a compact bit stream from which the quantized DCT coefficients can be reconstructed perfectly.

P frame coding

A P frame is inter or temporally coded. That means it uses correlation between the current frame and a past frame to achieve compression. The MPEG-2 standard[5] dictates that the past frame must be an I or P frame, but not a B frame.

Temporal coding is achieved using motion vectors. The basic idea is to match each macroblock in the current frame with a 16x16 pixel area in the past reference frame as closely as possible. Closeness here can be computed in many ways, but a simple and common measure is sum of absolute differences (SAD). The offset from the current macroblock position to the closest matching 16x16 pixel area in the reference frame is recorded as two values: horizontal and vertical motion vectors.

The search to find the best motion vectors is performed in the luminance channel only. Whatever motion vector is found then applies to all three channels of that macroblock.

There are a few common algorithms for finding motion vectors, such as:

1. Sequential search -- This is a brute force method where every possible match is tested within a given search window.
2. Logarithmic search -- Search 9 locations centered around the original macroblock, then keep repeating the search centered on the best match of the last iteration. At each iteration the search window gets smaller until the desired fidelity is reached.
3. Hierarchical search -- The macroblock and search window are decimated once or more, and searching takes place in order of lowest to highest resolution so the motion vectors can be refined at every step.

These motion vectors are a simple way to convey a lot of information, but are not always a perfect match. To give better quality reconstruction, error between the actual macroblock and the predicted macroblock is then

IJESMR
International Journal OF Engineering Sciences & Management Research

encoded. This coding of the residual is almost exactly the same as I frame coding. In fact the only difference is that the quantization equation becomes

$$QDCT = floor\left(\frac{8 \times DCT}{scale \times Qp}\right)$$

where floor is used in place of round, and Qp is a different weighted quantization table. The default for Qp is

Qp = [16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]
[16 16 16 16 16 16 16 16]

During reconstruction, the residual is decoded and added to the motion vector predicted macroblock.

B frame coding

A B frame is simply a more general version of a P frame. Motion vectors can refer not only to a past frame, but to a future frame, or both a past and future frame. Using future frames is exactly like a P frame[7] except for referencing the future. Using past and future frames together works by averaging the predicted past macroblock with the predicted future macroblock. The residual is coded like a P frame in either case.

Frame sequencing of MPEG-2

The ordering of I, P, and B frames is fairly flexible in MPEG-2. The first frame must always be an I frame because there is no other data for a P or B frame to reference. After that, I, P, and B frames can be mixed in any order.

Experimentation has shown that a repeating sequence such as

I B B P B B P B B P B B

yields good quality and compression. Remember that the P frames reference the last I or P frame and the B frames reference the closest past and future I or P frames. Due to those requirements, the frames must be coded and transmitted out of display order. The ordering would be

I P B B P B B P B B I B B

where the second I frame is actually from the next sequence of frames.

Code structure and features

The code for this project is organized into many functions that mostly reside within a single m-file. The top-level function performs four major steps:

1. Load the movie to encode.
2. Encode the movie and return MPEG data.
3. Decode the MPEG data and return the reconstructed movie.
4. Save both movie versions and MPEG data to disk for later viewing and/or analysis.

This simulation is built with variables controlling much of the coding process to allow experimentation. Some of the parameters that can be easily adjusted are the choice of movie to encode, frame type pattern (ex: I P P I P P), quantization scale factor, sequential or logarithmic motion vector search, and motion vector search window size.

Experiments and observations

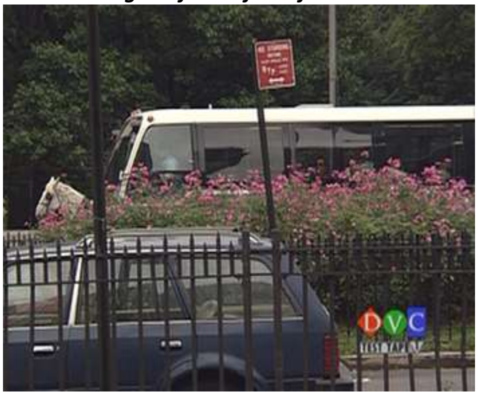
For this test, I encoded the same frame as an I frame at many quality levels.

ISSN 2349-6193
Impact Factor (PIF): 2.243

IJESMR

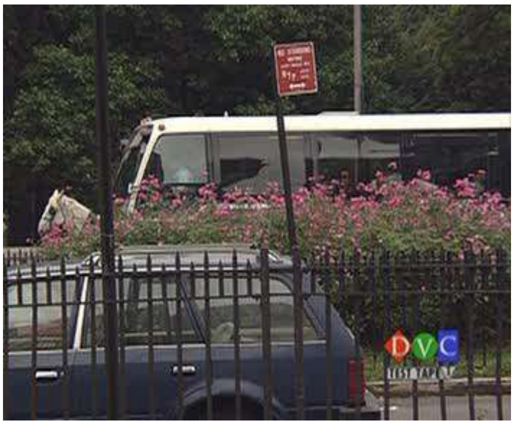
International Journal OF Engineering Sciences & Management Research

Original frame for reference:

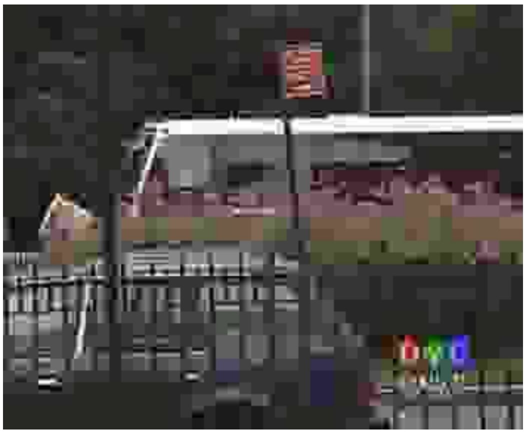


Results:

Scale = 8



Scale = 112

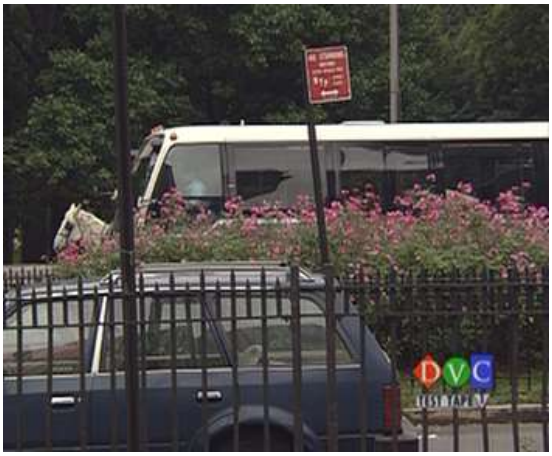


Original frame

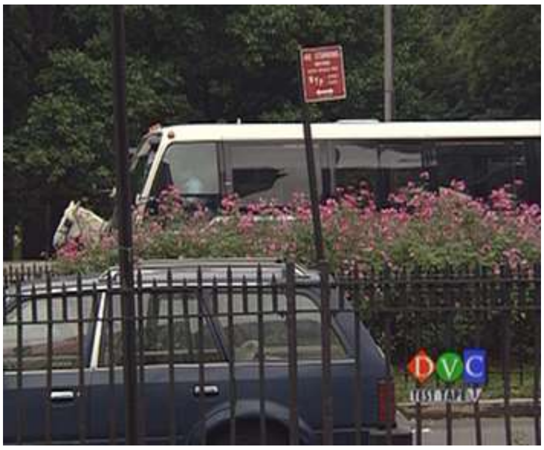
ISSN 2349-6193
Impact Factor (PIF): 2.243

IJESMR

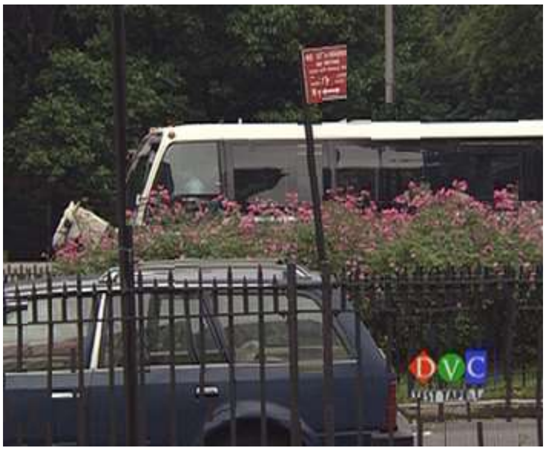
International Journal OF Engineering Sciences & Management Research



Decoded with residual



Decoded without residual



The predicted frame without residual is a good match except for some very obvious blocking artifacts. For example, see the red sign above the bus or iron bars below the 'DVC' logo. This type of blocking error makes sense considering that without residual added, the frame is made up completely of 16x16 blocks from the last frame. This video, with slow and consistent panning, is probably an ideal choice for this experiment since there is little relative motion within the frame.

ISSN 2349-6193
Impact Factor (PIF): 2.243

IJESMR

International Journal OF Engineering Sciences & Management Research

CONCLUSION

Our DCT&QDCT methods are works very efficiently compresses the video without loss of data using Huffman encoding and Run-length-codingtechnique. Compare to other compression technology Huffman coding is more efficient and also it uses minimum resource to compress the video at low cost. In video compression technique Huff Man Based on the results obtained, the proposed method Huffman encoding will give better performance.

REFERENCES

1. Muzhir Shaban Al-Ani and Talal Ali Hammouri, Video Compression Algorithm Based on Frame Difference Approaches International Journal on Soft Computing (IJSC) Vol.2, No.4, November 2011.
2. J. Mohanalin Rajarathnam, A novel Fuzzy based Medical video compression using H.264, Georgian Electronic Scientific Journal: Computer Science and Telecommunications 2008, No.3(17).
3. Mostafa Mofarreh-Bonab, Mohamad Mofarreh-Bonab, Adaptive Video Compression using PCA Method International Journal of Computer Applications (0975 – 8887)Volume 44– No.21, April 2012.
4. Detlev Marpe, Heiko Schwarz and Tomas Wiegand, “Context-Based Adaptive Binary Arithmetic Coding in the H.264/AVC Video Compression Standard”, in IEEE Transactions on Circuits and Systems for Video Technology. Vol 13, No.7 July 2003.
5. Mitesh Shah, Hetal Patel, "Design of a New Video Compression Algorithm Using Accordion Function ", International Journal of Science and Modern Engineering (IJISME) ISSN: 2319-6386, Volume-1, Issue-6, May 2013.
6. Sangeeta Mishra, Sudhir Savarkar, “Video Compression Using EZW and FSBM”, International Journal of Scientific and Research Publications, Volume 2, Issue 10, October 2012, ISSN 2250-3153.
7. Charu Pandey, Satish Kumar, Rajinder Tiwari, "An Innovative Approach towards the Video Compression Methodology of the H.264 Codec: Using SPIHT Algorithms “, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-2, Issue-5, November 2012.
8. Mohammad Hossein Bisjerdi, Alireza Behrad, "Video Compression USING a New Active Mesh Based Motion Compensation Algorithm in Wavelet Sub-Bands”, Journal of Signal and Information Processing, 2012, 3, 368-376.