

Video Quality Assessment through PSNR Estimation for Different Compression Standards

Renuka. G. Deshpande¹, Lata L Ragha², Satyendra Kumar Sharma³

¹Pacific academy of Higher Education and Research, Udaipur, India

²Terna College of Engineering, Navi Mumbai, India

³Modern Institute of Technology Research Centre, Alwar, India

Article Info

Article history:

Received Sep 17, 2017

Revised Nov 14, 2017

Accepted Apr 27, 2018

Keywords:

H.264

Motion-compensation

MSE

PSNR

SSIM

ABSTRACT

There is a threefold increase in video traffic over internet. Due to this video compression has become important. Compression of video signals is quiet an interesting task but comes at the cost of video quality. After compression, two methods are scientifically applied to evaluate the quality of video; Subjective and objective analysis. In subjective approach the compressed video is shown to a group of viewers and their feedback is recorded. Objective approach aims to set up a mathematical model which can approximate the results of subjective analysis. One such approach is based on the measurement of PSNR. When a signal is applied to the encoder for compression, too much of compression results in a signal with a smaller size but at the same time quality of the signal degrades. In this paper we will compare the quality of compressed video signals produced by H.264, Mpeg2 and Mpeg4 encoder based on the values of MSE and PSNR. Lower the value of MSE, higher will be the PSNR. Comparative plots of MSE, PSNR, SSIM and images for subjective analysis have been added at the end of this paper.

Copyright © 2018 Institute of Advanced Engineering and Science.

All rights reserved.

Corresponding Author:

Renuka. G. Deshpande,
Pacific academy of Higher Education and Research,
Udaipur, India.
Email: renukadshpande@gmail.com

1. INTRODUCTION

Video Streaming is gaining popularity over Internet. It is estimated that video streaming will increase threefold by 2020. With the advances in mobile handsets and wireless networks, video passes through different types of networks before it reaches the client. Video encoding /Video Compression plays an important role in video streaming. In this paper we compare the different compression (encoding) standards based on video quality measurements.

1.1 Block-oriented Motion Based Compensation

A video stream is a sequence of still pictures (frames) as shown in Figure 1. When this sequence is streamed at a rate of 25-30 frames per second it gives the viewer an illusion of motion. Each Frame can be divided into number of slices, where each slice is further made up of macro-blocks. The size of a Macro-block for H.264/AVC is 16x16. Each macro-block holds smaller units called as blocks. Blocks in turn are made up of pixels. The size of a block can either be 16x16, 16x8, 8x16, 8x8, 8x4, 4x8 or 4x4 [1]. Compression of a video is achieved by working around the macro-blocks. If we closely observe a pixel in a block we will see that if a pixel is "blue" then adjacent pixels will also be blue or lighter blue or dark blue. So if we encode each of these pixels independently, number of bits will be high. To achieve Compression it is necessary to identify redundancy in a block or frames. Spatial and temporal redundancies are the

techniques that can be applied to achieve compression. Where Spatial redundancy is applied to pixels, temporal redundancy is applied to frames [2].

In case of spatial redundancy, pixels that are close to each other spatially are compared. It is found that neighbouring pixels have nearly the same value, therefore due to redundancy they can be compressed using fewer bits. Consecutive frames also exhibit redundancy. Near-by frames are mostly identical but just shifted due to motion therefore by using temporal redundancy two consecutive frames are compared and the difference or error between them relative to the amount of shift is calculated. As this difference is relatively small therefore it can be coded and stored using fewer number of bits. This is termed as “motion compensation” and the motion vector corresponds to the amount of shift. H.264 encoder compresses the video signals based on motion compensation, hence H.264 is called as block-oriented motion based compensation technique [3].

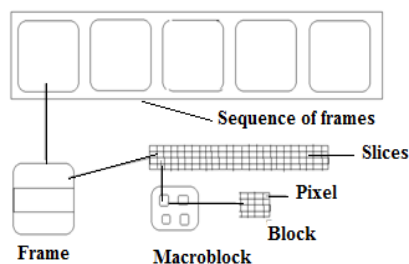


Figure 1. Contents of a Video Stream

1.2 Lossy and Lossless Compression

Every signal (data or image) contains two parts: information and redundancy. Information contains the vital part of the signal whereas redundancy holds the repetitive portion. Reducing the repetitive part gives a compressed signal. Signals can be compressed by applying lossy or lossless compression. Lossless compression (sometimes also termed as a reversible compression) technique uses some general purpose compression algorithms based on Run length coding, Lempel-ziv compression, Huffman coding etc [4]. Lossless compression technique results in decompression of signals without any loss of information (decompressed signal is a replica of original signal). The signal is compressed without affecting the video quality but it also requires a larger space on the disk as redundancy is not reduced. Lossy compression is an irreversible process. It not only removes redundancy but also some of the non-redundant part of the signal which results in loss of information. The compressed signal does not reconstruct all of the original information (it is just an approximation of the original signal) but this loss is undetectable in case of video compression due to perception of human eye. Lossy compression occupies a smaller space on the disk. H.264/AVC, Mpeg2 and Mpeg-4 are some of the lossy video compression encoders [5].

From earlier results [6] we know that H.264 encoder produces a compressed signal at a lower bit rate in comparison to other encoders. To extend our work further we now look forward to estimate the video quality of compressed signal. In communication, Noise plays a very important role and is defined as an unwanted random phenomenon which affects the performance of all systems over the entire range of frequency spectrum. Its effect on image or video signals is observed in the form of blurring or fading of signal. Noise performance of any system (processing, encoding, decoding, transmitting, receiving system etc) can be evaluated by measuring the value of signal-to-noise ratio [7].

Popularity of high definition devices and systems in the field of communication, entertainment, analysis etc has recently increased the demand for high resolution videos [13] therefore video compression is a field demanding great amount of work today. Apart from Video Compression lossless audio compression is also gaining popularity due to use of IEEE 1857.2 audio compression tools. [15]

This paper is organized as: Section 2 which covers the encoding issues, Section 3 which covers the Video Quality estimation based on PSNR measurement, Section 4 discusses the experimental set-up and results, and Section 5 concludes the work.

2. ENCODING ISSUES

In [1] the main features of H.264 are compared with MPEG-2 and MPEG-4 compression techniques. The comparative table highlights the advantage of using H.264 as it allows motion compensation for various available blocks. The RGB color system used for color image is expressed in terms of YCbCr

representation in [2], it has been shown that chrominance signal Cb, Cr, Cg can be expressed with lower resolution in comparison to luminance Y without affecting visual quality. Features and applications of H.264 are discussed in [3] and average bit rate achieved is compared to previous coding schemes along with bandwidth requirement. It has been highlighted that H.264 video needs lower B.W, storage space and download time in comparison to MPEG-2 and MPEG-4 (ASP). Signal analysis based on S/N ratio estimation by applying Taguchi method has been discussed in [7]. Three general standard S/N equations for classification are presented as "larger the best", "lower the better", and "nominal the best". Taguchi method discusses the estimation of video quality in presence of background noise by calculating the Power signal-to-noise ratio. A no-reference PSNR estimation discussed in [12] uses the maximum likelihood with linear prediction estimates to achieve a better PSNR estimation. In [13], the proposed system was implemented to compress video using modified HEVC technique based on saliency features & motion vector entropy was used to relate variations in the scene. In [14] the proposed system combines spatial saliency & temporal saliency features together for different macroblocks in association with transformed residuals for real time video compression. The PSNR values (more than 30 db) for all three channels Y, U, V are compared with existing techniques.

3. PROPOSED METHOD FOR VIDEO QUALITY MEASUREMENT

PSNR and SSIM are used to measure quality of video signals. PSNR is based on the estimation of signal-to-noise ratio. As we know that a video sequence is a collection of still images and a color image has three values per pixel which are obtained as additive colors by mixing red, green and blue lights with several different intensities. If intensities of all three colors are made zero it produces the darkest color, black and if the intensity is highest it produces a white. An illusion of motion is sensed when these still images are streamed at the rate of 25-30 frames per second. When moving from one frame to another, human eye cannot detect the variation in the depth of different color components (if a light blue has become a bit dark in the next frame) but it can clearly make out the difference whenever the intensity of the image changes. Each pixel thus holds the information about the intensity and chrominance of image [8]. If the color signals were to be directly transmitted, three carriers would have been required for three components R, G, B. During the earlier days of color television it was necessary to make them backwards compatible with the B/W TV and also to conserve bandwidth required to transmit a color video [9]. To meet the Television standards of transmission and reception, the RGB color model was therefore converted to Y'UV color space. Most commonly it is represented as YUV. Y is the luma which defines the overall electronic brightness of a pixel and U, V are the chrominance component. The values of YUV can be obtained from the RGB components as follows [10].

$$Y = 0.299 * R + 0.587 * G + 0.114 B \quad (1)$$

$$U = 0.492 (B - Y) \quad (2)$$

$$V = 0.877 (R - Y) \quad (3)$$

Rearranging the above equations converts YUV to RGB color space back again.

As the human perception is more prone to changes in the intensity of image therefore the luma (Y) component is the most important parameter in defining the quality of a video signal. Hence to test the PSNR performance of different encoders the luma component is selected as the judging parameter. As the value of PSNR is defined by considering the mean squared error i.e. MSE, the plots of MSE and PSNR will be obtained for the Y component. Another method which can be used to compare two signals is based on the measurement of structural similarity (SSIM). Both these methods are employed in this paper for comparison of codec performance.

4. EXPERIMENTAL SETUP AND RESULTS

Figure 2 shows the steps involved to evaluate the quality of compressed video. A video of 9.8 MB and 2.08 min duration was considered for carrying out the experiment. In three independent trials the video was encoded using H264, Mpeg2 & Mpeg-4 encoders. To achieve compression, the size of output file was decided and set to equal 80% the size of input file, the bit rate was accordingly set. As bit rate = $\frac{\text{Size (MB)}}{\text{Duration (s)}}$ therefore in each case the bit rate used for compression was same (393b/s and size of output file obtained was 8MB).

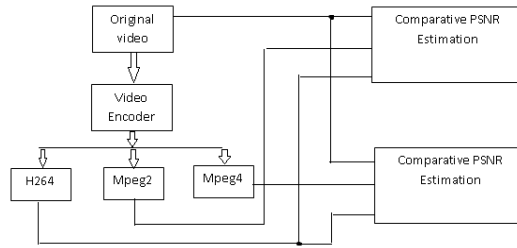


Figure 2. Experimental Setup

As shown in Figure 2, PSNR [11] is calculated by evaluating the two signals (original & compressed video) according to Equation (4)

$$PSNR = 10\log_{10} \left(\frac{R^2}{MSE} \right) \tag{4}$$

Where R is the maximum variation in input signal and MSE is the mean squared error [12] given by Equation (5)

$$MSE = \sum_{M,N} \frac{[I_1(m,n) - I_2(m,n)]^2}{M * N} \tag{5}$$

Lower the value of MSE, higher will be the value of PSNR.

Case I: In the first case, the upper Comparative PSNR estimator is analyzed. The inputs to this estimator are the original signal, H264 video and Mpeg2 video. The Comparative PSNR estimator performs two independent PSNR evaluations where first the original video and Mpeg2 video are compared for calculating PSNR. Simultaneously the second trial involves estimating PSNR of original signal and H264 video. Figure 3 shows the plot of MSE where the original signal is taken as the reference signal and H264 and Mpeg2 videos are taken as the compressed signals. A comparative plot of the two compressed signals against the referenced signal shows the variation in the values of MSE obtained for different frames.

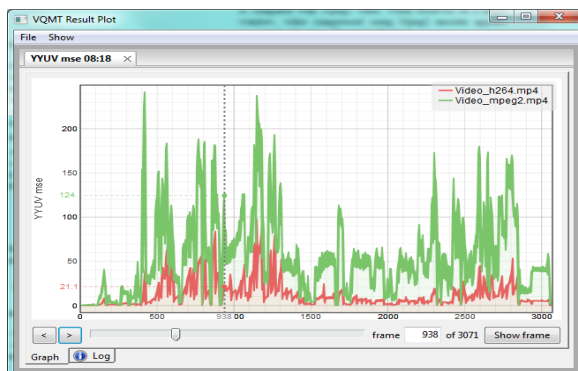


Figure 3. MSE (Comparison of H264 and Mpeg2 Video)

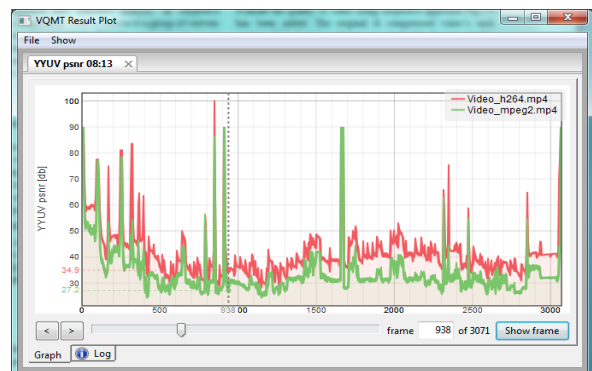


Figure 4. PSNR (Comparison of H264 and Mpeg2 Video)

It can be seen from the plot (Figure 3) that the value of MSE for Mpeg2 video is quiet larger then H264 (shown with green color in Figure 3). After estimating the values of MSE, PSNR is estimated. Figure 4 shows the comparative plot of PSNR for H264 and Mpeg2 video. Since the value of MSE for Mpeg2Video is high therefore value of PSNR is low (represented by green color in Figure 4). To evaluate the quality of video using subjective approach Figure 5 has been added. The original & compressed video's each contain 3071 frames, a frame was randomly chosen (in this case frame no. 938) for which the values of MSE for H264 and Mpeg2 video were 22 and 124, and the values of PSNR were 34.9db and 27.2db respectively.

In Figure 5 H264 video is compared with Mpeg2 video. When observed on a larger window, video compressed using Mpeg2 encoder appears corrupted with noise (Encircled Region).

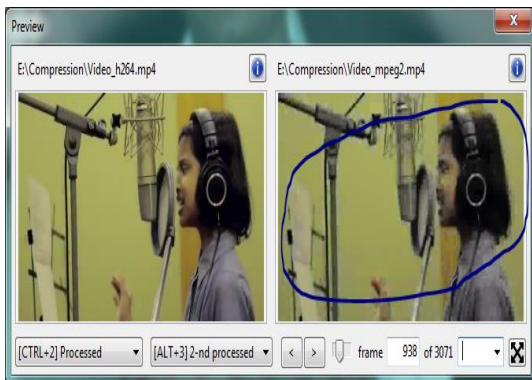


Figure 5. Subjective Analysis (Comparison of H264 and Mpeg2 Video)

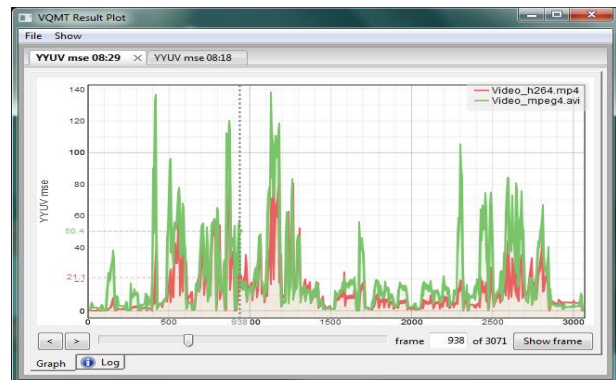


Figure 6. MSE (Comparison of H264 and Mpeg4 Video)

Case II: In the second case, the lower Comparative PSNR estimator is analyzed. The inputs to this estimator are the original signal, H264 video and Mpeg4 video. In the first trial, original video and Mpeg4 video are compared for calculating PSNR. Simultaneously the second trial involves estimating PSNR of original signal and H264 video. Figure 6 shows the comparative plot of MSE for H264 and Mpeg4 video signals. The measurement of MSE at frame no. 938 gave the values as 21.1 and 50.4 for H264 and Mpeg4 video's respectively. The Plot of PSNR shown in Figure 7 shows the value as 34.9db and 31.1db respectively. Figure 8 has been added for subjective analysis where distortions in the frame can be clearly observed in the encircled region.

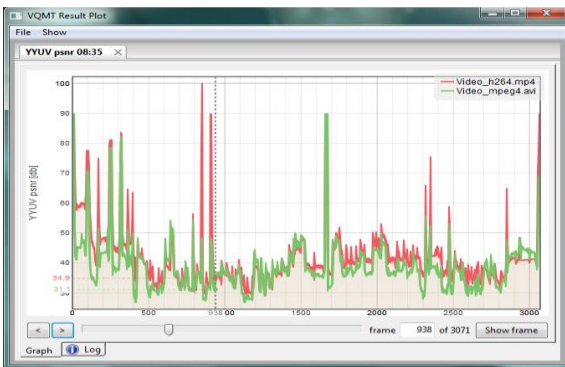


Figure 7. PSNR (Comparison of H264 and Mpeg4 Video)

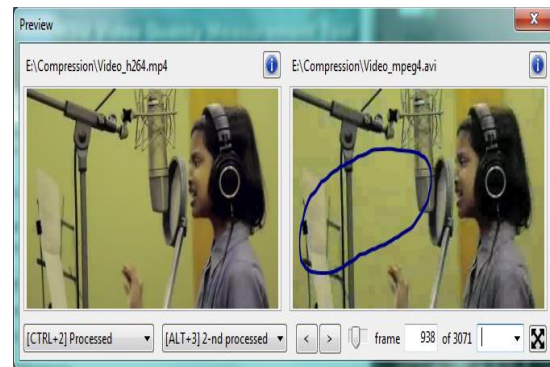


Figure 8. Subjective Analysis (Comparison of H264 and Mpeg4 Video)

To estimate video quality, it won't be fair to compare the performance of different encoders for a single frame of reference. To attain a truly acceptable result it is necessary that the PSNR of entire video (3071 frames) must be taken into consideration. In case of images involving RGB, the average value of PSNR is calculated by taking the mean of PSNR for all frames. VQMT presents a limitation at this point as it is best used for frame-frame analysis but does not return the average value of PSNR Therefore after performing frame-frame analysis on the signals, average values of PSNR were calculated for the signal by comparing each of them independently with the original signal. For this Ffmpeg was used in cmd and the following observations were made.



Figure 9. SSIM (Comparison of H264 and Mpeg4 Video)

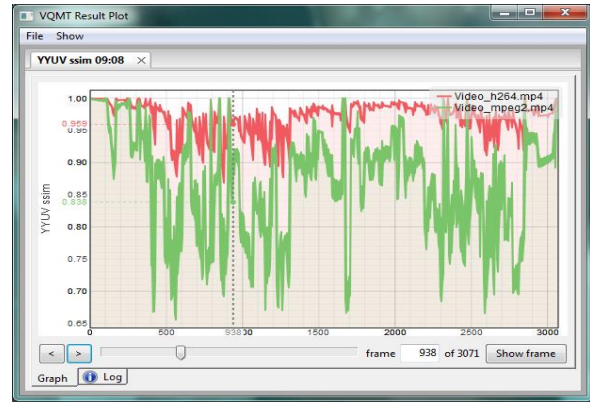


Figure 10. SSIM (Comparison of H264 and Mpeg2 Video)

For 8 bit data the attainable values of PSNR ranges between 30-50db whereas it moves up to a value of 60-80db for a 16bit data. The average PSNR for H264 video is quiet high therefore it gives the perception of a good quality video to the human eye whereas Mpeg4 video with the lowest PSNR clearly appears in the form of a blurred image and highly distorted by noise (Table 1). These observations also bring forth the difference in performance of different lossy encoders.

Table 1. Encoders with Average PSNR Values

Encoder	Average PSNR
H264	38.78db
Mpeg2 Video	32.21db
Mpeg4 Video	23.07db

Moving a step ahead the compressed video signals are compared on the basis of structural similarity (SSIM) with the original signal. Figure 9 and Figure 10 have been added for SSIM. The overall structural similarity between the original and various compressed signals are obtained using ffmpeg and the results are depicted in Table 2.

Table 2. Encoders with SSIM Values

Encoder	SSIM
H264	0.978116
Mpeg2 Video	0.900587
Mpeg4 Video	0.863064

5. CONCLUSION

In this paper we have evaluated the quality of video signals based on the performance of encoders. It was found that some lossy encoders result in degradation and poor quality of compressed signals. In real world applications the signals after compression have to travel from the source to destination using wired or wireless media. At each point of communication noise is added to the signal owing to internal (noise added by the encoding and transmitting system) and external (noise added by the transmitting medium) factors. This further degrades the video quality and results in low PSNR for received signal. Therefore by taking all these factors into account and the amount of loss in the value of PSNR during transmission, it is necessary that the encoded signal should have a high value of PSNR prior to transmission. From Table 1 it can be seen that PSNR value achieved with H264 is highest, approximately 39db (range should be between 30-50db). When the signals were compared on the basis of SSIM it was found that signal compressed with H264 encoder resembles the original signal by 97.8 %. Therefore H264 is the widely used video compression standard.

REFERENCES

- [1] Juurlink.b, Alvarez mesa, Chi C.C , Azevado A, Meenderinck C, Ramirez A, “Understanding the Application : An Overview of the H.264 Standard”, Springer, Scalable Parallel Programming applied to H264/AVC decoding Chapter 2, 2012 XIII, 5-15.
- [2] Kusuma.H.R , Dr.Mahesh Rao, “Video Compression Using Spatial and Temporal Redundancy –A Comparative Study” *International Journal of Innovative Research in Science, Engineering and Technology*, June 2015, 4(6): 4801-4808.
- [3] Mamatha .R .B, Dr. Keshaveni N., “Comparative Study of Video Compression Techniques H.264/AVC”, *International Journal of Advanced Research in Computer Science and Software Engineering*, November 2014, 4(11): 874-877.
- [4] S.R. Kodituwakku , U. S.Amarasinghe , “Comparison Of Lossless Data Compression Algorithms For Text Data”, *Indian Journal of Computer Science and Engineering*, 1(4), 416-425.
- [5] “An Explanation of Video Compression Techniques”, White Paper, Axis Communication.
- [6] Renuka G. Deshpande, Lata L. Raha, “Performance Analysis of Various Video Compression Standards”, *International Conference on Global Trends in Signal Processing, Information Computing and Communication*, (ICSPICC- 2016), Jalgaon, Maharashtra, India, December 2016.
- [7] V. Roshan Joseph and C. F. Jeff Wu, “Performance Measures in Dynamic Parameter Design”, *Journal of Japanese Quality Engineering Society*, 2002, 82-86.
- [8] Douglas A, Kerr P.E. “Chromaticity and Chrominance in color definition- UT Mathematics” ma.utexas.edu, Aug 2003, (3), 1-4.
- [9] William F Schreiber, Robert R Buckley “ *Introduction to color Television-Part I*” Proceedings of the IEEE, Jan 1999, 87(1), 173-179.
- [10] Nishad.P.M, Dr. R.Manicka Chezian, “Various Colour Spaces and Colour Space Conversion Algorithms” *Journal of Global Research in computer science*, Jan 2013, 4(1), 44-48.
- [11] Tom´as Brand´ao and Maria Paula Queluz, “No-Reference Psnr Estimation Algorithm For H.264 Encoded Video Sequences” EURASIP, 16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 2008.
- [12] Mahesh Kamat, Roopa Kulkarni, Dr. Prashant Patavardhan, “Analysis of Lossy Video Compression Technique”, *International Research Journal of Engineering and Technology*, September 2015, 2(6), 388-395.
- [13] K Bhanu Rekha , Ravi Kumar A.V “High Definition Video Compression Using Saliency Features”, *Indonesian Journal of Electrical Engineering and Computer Science*. Vol. 7, No. 3, September 2017, pp. 708 ~ 717 DOI: 10.11591/ijeecs.v7.i3.pp708-717.
- [14] K Bhanu Rekha , Ravi Kumar A.V “High Definition Video Assessment Using Saliency Features”, *Indonesian Journal of Electrical Engineering and Computer Science* Vol. 7, No. 3, September 2017, pp. 761 ~ 772 DOI: 10.11591/ijeecs.v7.i3.pp761-772.
- [15] Teddy Surya Gunawan, M. Khalif Mat Zain, Fathiah Abdul Muin, Mira Kartiwi , “Investigation of Lossless Audio Compression using IEEE 1857.2 Advanced Audio Coding”, *Indonesian Journal of Electrical Engineering and Computer Science* Vol. 6, No. 2, May 2017, pp. 422 ~ 430 DOI: 10.11591/ijeecs.v6.i2.pp422-430.