Available online at www.ejournals.uofk.edu

UofKEJ Vol. 10 Issue 1 pp. 15- 18 (February2020)

UNIVERSITY OF KHARTOUM ENGINEERING JOURNAL (UofKEJ)

Comparative Objective Analysis of video quality Between H.265/HEVC and H.264/AVC

Jaafar A. A. Alnoor, Amin B. A. N. Mustafa

Dept of Electrical and Electronic Eng, Faculty of Eng.

University of Khartoum sudan , AL- Neleen University of Khartoum sudan

(E_Mail: jaawish@gmail.com)

Abstract: This paper shows the difference in video quality between two compressed videos using H.264 AVC (Advanced Video Coding) and H.265 HEVC (High Efficiency Video Coding) encoders. To evaluate video completely it should be prepared video files that have a variety of bit rates and content. There are many video quality assessment methods. We can divide the min to subjective and objective methods. Subjective are conducted by a human perception and objective are conducted by a computer software which is calculating the video quality. All of these methods have theirs advantages and disadvantages. To generate compressed videos from the original video FFmpeg (Fast Forward-moving picture experts group) converter has been used. MSU-VQMT (Moscow State University's Video Quality Measure- mentTool) was used to perform comparative objective analysis of video quality. Delta, MSE (mean square Error), MSAD (Mean Sum of Absolute Difference), PSNR (Peak Signal-to-Noise Ratio), and SSIM (Structural Similarity Index Measure) metrics were measured. The result from FFmpeg shows that the size of the compressed video using the H.265 codec has been decreased by 50% compared to the compressed video using the H.264 codec. The comparison of metrics shows that delta, MSAD, PSNR, and SSIM values of H.265 encoded video was decreased, while Delta and MSE value was increased compared to H.264 encoded Video. That's mean the overall video quality was decreased but the video size was enhanced.

***Keyword*:** *Delta, MSE, MSAD, PSNR, SSIM.*

1. INTERDUCTION

IDEO quality is a measure of perceived video degrada- tion (typically, compared to the original video). For many stakeholders such as content providers, service providers, and network operators, the assurance of video quality is an important task. Because of that, there is always a need to improve the quality of Video. In this case, quality means the highest perceptual quality with the smallest size of the video file. The video Codecs are being improved because of a huge improvement in technologies.

Compression methods such as and H.265 use prediction of moving objects between certain frames as a way to compress videos and reduce bandwidth. H.265 uses Adaptive Motion Vector Prediction for inter-frame prediction [6]. It should be noted that H.265 is a more computationally expensive compression method, which is likely the reason that it has not yet become as prevalent as H.264. There are different compression technologies and different standards are available. Standardization is necessarily to guarantee compatibility.

They are a lot of video compression formats for example: H.120, H.261, MPEG-1, MPEG-2 Part 2, H.263, Motion JPEG, MPEG-4, H.264 and HEVC. HEVC is the latest and most efficient video compression standard. These standards are developed by two organizations ISO/IEC and ITU-T [5]. Another type of compression is JPEG 2000. This is a wavlet based compression that uses the Discrete Cosine Transformation (DCT).

This type of compression works on individual frames rather than predicting frames. This is a good option for high-bandwidth situations, but may not provide ac- ceptable quality at low bandwidths. VP8 and VP9 are alternate encoding algorithms which are competitors with H.264 and respectively.

A. H.264

Also known as MPEG-4 Part 10/AVC.H.264. In this stan- dard we can achieve reduced video file for about 80% com- Paring to Motion JPEG and for 50% comparing to MPEG-4 maintaining the same image quality. In the same bit rate.

AVC ENCODER

An H.264 video encoder carries out prediction, transform and encoding processes to produce a compressed H.264 bit stream. An H.264 video decoder carries out the complement- tary processes of decoding, inverse transform and reconstruct- tion to produce a decoded video sequence. A block of residual samples is transformed using a 4x4 or 8x8 integer transform, an approximate form of the Discrete Cosine Transform (DCT). The transform outputs a set of coefficients, each of which is a weighting value for a standard basis pattern. When combined, the weighted basis patterns re-create the block of residual samples. Fig.1. shows how the inverse DCT creates an image block by weighting each basis pattern according to a coefficient value and combining the weighted basis patterns.

The output of the transform, a block of transform coefficients, is quantized, i.e. each coefficient is divided by an integer value Quantization reduces the precision of the transform coeffi- cients according to a quantization parameter (QP). Typically, the result is a block in which most or all of the coefficients are zero, with a few non-zero coefficients. Setting QP to a high value means that more coefficients are set to zero, resulting in high compression at the expense of poor decoded image quality. Setting QP to a low value means that more non- zero coefficients remain after quantization, resulting in better decoded image quality but lower compression [7].

B. H.265/HEVC/ MPEG-H PART2

High Efficiency Video Coding (HEVC) is the newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The most important feature of this standard is that compression is Improved compared to previous standards -50% bitrate reduc- tion for equal perceptual video quality.This standard includes text specifications as well as reference software source code, example of encoding and decoding HEVC video file [9].

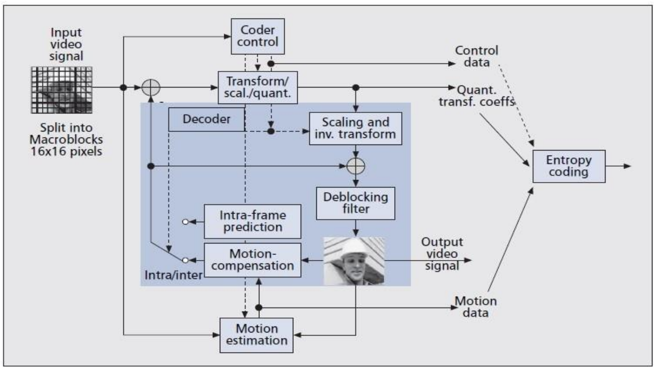

Fig.1. AVC video encoder

HEVC ENCODER

For encode the video in this standard first picture should be divided into macro blocks. Next step is using intra-frame compression for reducing the spatial redundancy. Then using interframe compression the temporal redundancy is reduced. In the next step transformation and quantization for reduce data compression is used. In the last step using entropy coding the final redundancy and motion vectors transmission are reduced [3]. HEVC encoder generates a valid sequence of bits and it could be divided into few steps. Fig. 2.showst h e block diagram of the HEVC encoder. HEVC as well as previous standards since H.261 follows the classical block- based video coding approach. The coding algorithm is a joint of inter-picture prediction to use spatial and temporal statistical dependencies. For future use of spatial statistical dependencies code of the prediction of residual signals is transformed [9].

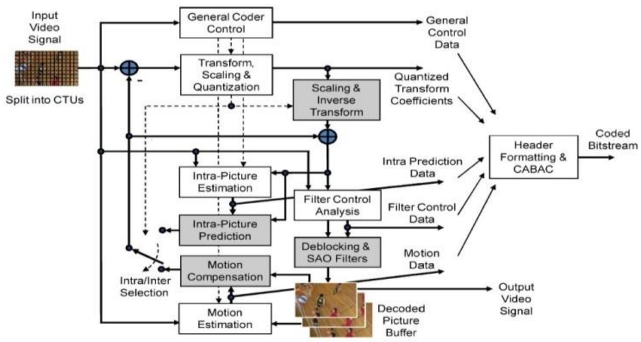

Fig.2. HEVC video encoder

HEVC uses tristimulusYCbCr color space with 4:2:0 sampling for representing colors in a video. It divides a color representation into Y, Cb and Cr elements. The Y is responsible for brightness and it is also called luma. Cband Cr are called chroma and are responsible for in which way color goes from gray to blue and red. In sampling structure, each chroma element has one fourth of the luma element. It is a typical used structure because the human visual system is more sensitive for luma than for chroma elements. Luma CB and two chroma CBs are creating a coding unit (CU). Each CU can be divided into Transform Units (TUs). Transform sizes are 3232, 1616, 88 and 44. Larger TUs are applicable for encoding stationary signals, smaller TUs are applicable for encoding impulsive signals. This transforms are based on Discrete Cosine and Discrete Sine Transforms (DCT and DST) [9].

C. VIDEOCOMPRESSIONS

Video compression is a technology where the size of a video file is minimized but the quality is kept high, preferably with no noticeable distortion. Compression allows more efficient storage and transmission of the data. Reduction in the file size can be accomplished with little effect on the visual quality or has no effect on visual quality. When the size of the file is reduced by raising the compression level for a compression technique, the quality of the video file can be a little bit affected. Removing and reducing unwanted data is necessarily for sending a digital video file over network more effectively [5].

There are two types of compression: lossy and lossless. Lossy compression does not allow recovering the whole amount of the original data. It is used for data which includes a lot of redundancies and which is insensitive to losses. It is used in images, videos or sound files. Compared to lossless compression, lossy compression provides higher compression ratios. Lossless compression provides a whole recovery of the original data. It is used in files where loss causes a lot of damage such as text and executable files [4].

D. VIDEO QUALITYASSESSMENTS

Generally video quality assessment can be divided into two main categories: objective video quality metrics and subjective video quality metric. Subjective metric is based on the tests conducted on a group of people who are judging the quality of the video by watching it. There are a few important factors to perform a subjective experiment for example careful plan- ning, assessment method, selection of test material, viewing conditions or timing of the material. Subjective method is hard to conduct in real time that is why objective metrics are created. Objective metrics are calculated by the computer. Some metrics are more and some are less similar to how the human perceives quality of the video.

Objective video quality metrics can be divided into three categories: full reference (FR), reduced reference (RR), and no-reference (NR). With FR method original image or video is available as a reference when distorted image or video is compared with the original one. In RR method we provide features about texture or other characteristic of the original image or video. The input in this method is the comparison between reduced information from original file and information from distorted file. NR method does not require access to original image or video but it use information in bitstream or search for information in pixel domain [8].

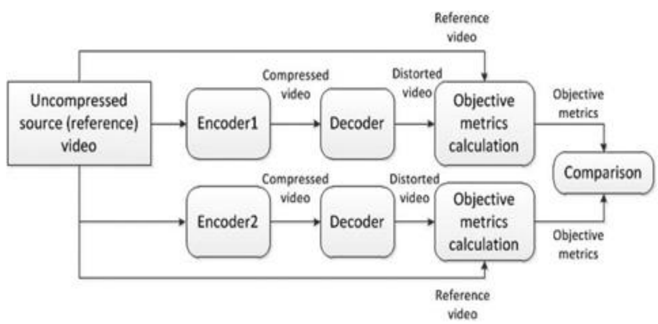

Fig.3. A typical encoder comparison setup using FR objective quality metrics.

1) Universal reference objective quality metrics:

Delta: The value of this metric is the mean difference of the color components in the correspondent points of image. This metric is used for checking codecs/filters for errors like losses or growths of luminance, not for quality comparisons.

$$d(X,Y) = \frac{\sum_{i=1,j=1}^{m,n} (X_{i,j} - Y_{i,j})}{mn} \quad (1)$$

The values are in -255...255. 0 for identical frames.

a) MEAN SQUARED ERROR: MSE is the average of the squared differences between the luminance values of corresponding pixels in two frames. It allows evaluating the degree of image reconstruction by a decoder. Values are from 0 (no difference) to 65025 (maximum difference at 8 bit color depth). This factor should be as small as possible [2]. It is defined by:

$$MSE = \frac{\sum_{i=1,j=1}^{m,n} (X_{i,j} - Y_{i,j})^2}{mn} \quad (2)$$

MEAN SUM of ABSOLUTE DIFFERENCE (MSAD):

The value of this metric is the mean absolute difference of the color components in the correspondent points of image.

$$MSAD = \frac{\sum_{i=1,j=1}^{m,n} |X_{i,j} - Y_{i,j}|}{mn} \quad (3)$$

The values are in 0...255. 0 for id entical frames.

a) PEAK SIGNAL-TO-NOISE-RATIO(PSNR):

This met- ric, which is used often in practice, called peak-to-peak signal- to-noise ratio. Accepted values for 8 bit color: from 0 (max difference - red) to infinitive dB - (no difference - black) however for practical situations PSRN value is not bigger than 50dB...The bigger the PSNR, the better.

$$PSNR = 10 \cdot \log_{10} \frac{MaxErr^2 \cdot w \cdot h}{\sum_{i=0,j=0}^{w,h} (X_{i,j} - Y_{i,j})^2} \quad (4)$$

Where Max Err maximum possible absolute value of color components difference, w video width, h video height.

Generally, this metric is equivalent to Mean Square Error, but it is more convenient to use because of logarithmic scale. . It has the same disadvantages as the MSE metric.

e) STRUCTURAL SIMILARITY INDEX MEASUR:

Main idea of (SSIM) is to compare distortion of three image components:

- Luminance comparison
- Contrast comparison
- Structure comparison

Final formula after combination of these comparisons is the following:

$$SSIM(X,Y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x + \mu_y + c_1)(\sigma_x + \sigma_y + c_2)} \quad (5)$$

Acceptable values from -1 (maximum difference) to 1 (no difference). Higher value means better quality. One of the advantages of the SSIM metric is that it better represents human visual perception than does PSNR. SSIM is more complex, however, and takes more time to calculate.

2. METHODOLOGY

To respectively perform the video quality prediction dif- ferent video files should be tested. Sequences with different Amount of motion, animations, color, frame rate, bitrates were tested. The original video file is compared to the distorted.
Video file.

Fig. 3. Shows encoder comparison setup using FR objective quality metrics.

In the table below there are parameters of video which were used in tests.

Table 1. Parameters of video

Video	
Length	00:04:48
Frame width	1280
Frame height	720
Data rate	2077kbps
Total bitrate	2222kbps
Frame rate	34 frames/second

Fig.4. H.265 vs. AH.264 Video frame

A comparative quality evaluation between two compressed videos and original video, using objective measures of assess- ment. FFmpeg was used in compression of the original video to generate two compressed video, the first one using H.264 codec and the second using H.265 codec. For the purpose of comparison MSU VQMT was used to calculate metrics values.

1) FFmpeg: FFmpeg is a very fast video and audio con- verter. It can also grab from a live audio/video source. The command line interface is designed to be intuitive, in the sense that FFmpeg tries to figure out all parameters that can possibly be derived automatically. You usually only have to specify the target bitrate you want. FFmpeg can also convert from any sample rate to any other, and resize video on the fly with a high quality poly phase filter. As a general rule, options are applied to the next specified file [1].

2) MSU VQMT: MSU Video Quality Measurement Tool is professional software that is used to perform deep comparative objective analysis of video quality. The main functionality of this software is to calculate objective quality metrics for digital multimedia content (video or image) using reference (when comparing several processed/compressed/distorted video se- quences to original one) or non-reference (when analyzing content and getting mark of its quality) types of analysis [2].

3. RESULTS AND DISCUSSION

This section summaries the findings and contributions made. To respectively perform the video quality prediction, video file Sequence with different amount of motion, animations, color, frame rate, bitrates should be tested.

- Color component: YYUV
- Files: raw video compared to: H.264 video and H.265 video

Table *2.* THE AVERAGE VALUES OF METRICS

Metric	H.264	H.265
Delta	-0.021556	-0.023710
MSE	38.015209	60.119526
MSAD	0.426976	0.691311
PSNR	40.539188	36.396862
SSIM	0.991614	0.983565

Table 2. and Fig.5.show that H.265 delta value is greater than H.264 delta value. It is clear that the value of the delta has increased by 10%. This increase has little effect on video quality because it is close to zero.

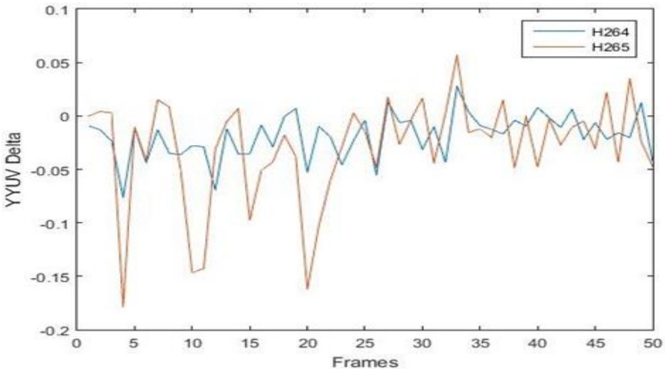

Fig.5. Delta Metric Value of H.264 and H.265

Fig. 6.Shows MSE values. From table II clearly it can be seen that there has been an increase of about 58%. Increasing this metric means a decrease in video quality.

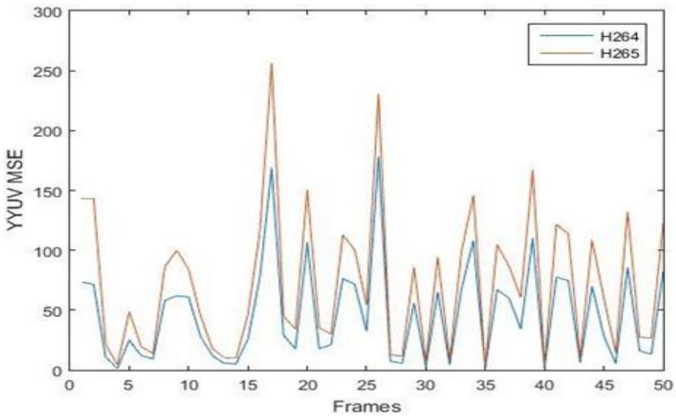

Fig.6. MSE Metric Value of H.264 and H.265

Fig. 7. Shows MSAD values. It can clearly be seen that there has been a small increase in the MSAD value. From table II we found that MSAD increased by 58%, which mean the video quality decreased. But the value (0.691311) is near zero so the effect on video quality will be small.

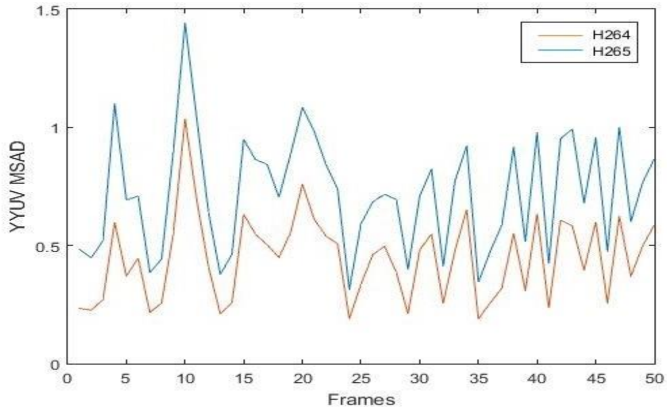

Fig.9. SSIM Metric Value of H.264 and H.265

4. CONCLUSION

Video quality is evaluated for specification of system requirements, comparison of competing service offerings, transmission planning, network maintenance, client-based quality measurement, and so on. From the comparison of raw video, H.264 compressed video and H.265 compressed video, the new H.265 standard reduces video files by 50% compared to the previous H.264 standard. Through the metrics used, metric values were changed. The video quality has decreased meanwhile video file volume reduced and this will help in video transmission over the networks.

ACKNOWLEDGMENT

The author would like to thank Prof. Ameen Babkir Ab- dalnabi for his guidance and help.

REFERENCES

[1] Fmpeg.org. Ffmpeg documentation. 07/01/2011.

[2] CS MSU Graphics and Media Lab Video Group. Msu vqmt documenta- tion. 2007.

[3] https://sonnati.wordpress.com/2014/06/20/h265- part-i-technical overview/. Page visited 02.02.2016

[4] Page visited 05.07.2018 https://www.linkedin.com/pulse/brief-history- video-codecs-yoav nativ.

[5] http://www.axis.com/global/en/learning/web-articles/technical-guide-to- networkvideo/ Page visited 01.07.2018.

[6] Loren Merritt and Rahul Vanam. X264: A high performance h.264/avc encoder. 10 2018.

[7] Iain E. Richardson. The h.264 advanced video compression standard.

[8] M. Shahid. No-reference image and video quality assessment: a classifi- cation and review of recent approaches.

[9] G. J. Sullivan and T. Wiegand. Overview of the high efficiency video coding (hevc) standard.