Performance Analysis and Energy Consumption of H.264 and H.265 Video Coding Standards

Hardi Hajar, Imade Fahd Eddine Fatani
Sciences and Techniques for the Engineer Laboratory
National School of Applied Sciences, Sultan Moulay Slimane University, Morocco

Abstract

A codec is a software that encodes and compresses a video for storing or before transporting through networks. The process of video encoding has become a major part of information technologies. Hence, a variety of codecs has been developed through the years. in this paper we provide an analysis of the two wildly used video coding standards H.264 and H.265. In order to elaborate this study, for each one of the codecs, we calculated four essential metrics: PSNR, SSIM, VIF and VMAF. We report a higher performance of H.264 for most of the metrics however in terms of size the H.265 presented a higher compression rate. Moreover, in terms of Energy consumption, H.265 showed lower energy consumption than H.264.

Keywords: Video coding, H.264/AVC, H.265/HEVC, Energy Optimization

1. Introduction

Streaming of videos, be it music, movies or simply just speech videos has been growing rapidly in the recent past. It is not unusual for one to stream a video only to find that the video quality is low or so much bandwidth is wasted since the quality of the output video is not measured, hence higher bitrates are used than is necessary. Users are nowadays more interested in the quality of the video and the cost of content delivery network. Video codecs go along the way in determining the visual quality of a video, therefore it becomes important to compare the different video codecs in terms of their performance metrics. Through such comparison video encoding channels are optimized to lower the bitrate as low as possible without negatively affecting the video quality. In this article we will compare the performance of two video codecs namely: H.264 and H.265.

2. From AVC to HEVC

H.264/MPEG 4 Part 10, referred to as AVC Advanced Video Coding is a codec that was developed in 2003 thanks to the joined efforts of the two standards organizations ISO/IEC MPEG and ITU-T VCEG under the name Joint Video Team (JVT). H.264/MPEG 4 Part 10 aims at resolving the problems raised by its precedents: H.261, H.262, MPEG-2, H.263 (Baseline), H.262+ (Profile 3), H.263++ (Profile 5), H.26L, MPEG-1, MPEG-2, MPEG-4 version 2, MPEG-7, MPEG-21. The main advantages of H.264 compared to the previous codecs is its efficient coding, resilience to errors and its network friendliness [1]. Since then, H.264 has been the most popular video codec. Hence most of devices have H.264 built on in order to decode videos without compromising the device processor.

In 2010, a new collaboration between the two standards organizations ITU-T VCEG and ISO/IEC MPEG under the name Joint Collaborative Team on Video Coding (JCT-VC) Created a new codec, H.265 or also referred to as HEVC High efficiency Video coding to counterbalance the limits of H.264/MPEG 10 standard [2]. HEVC can be considered as the future of codecs since it has showed great results with 4k resolution videos with almost the same bitrate as AVC.

In this work we weigh up the performance of the two codecs, for this purpose, there are several metrics to be considered when analyzing the codec performance, we will consider the following metrics:

- PSNR
- SSIM
- VIF
- VMAF

3. Methodology

In order to achieve the comparison, we encoded the original video into H.264 and H.265 codec using Python and ffmpeg. The resulting video from each of the codec will then be compared to the original video by calculating the metrics mentioned above. Finally, these metrics will be compared by use of graphs for the two codecs in order to determine which codec is the best.

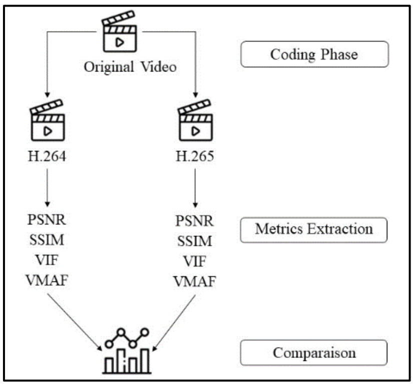

Figure 1. Methodology

For this comparison, we created a 6 seconds video, it's composed of 13 parts: bars of shades of gray, a black and white hypnotic spiral, white noise, a countdown timer, bars of gradient colors fading, two tv screen tests, circles of colors moving, rectangles of shades of gray, horizontal rainbow shades, loading animation, moving colored circles and a text ticker. This variety of primary and secondary colors aims to test the color range and the use of static and dynamic clips of shades of gray is to test the grayscale range. Moreover, the purpose of using different shapes and animations is to test the liability to details and see if the encoder can handle smooth curves as well as angular shapes. Finally, the use of a text ticker is to inspect if the encoder is well adapted to text.

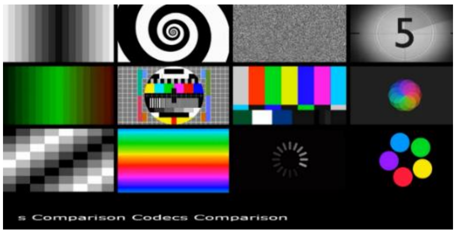

Figure 2. Screenshot for the video used for testing

4. Performance Analysis

This is one of the most popular video performance metrics and is majorly used for reconstructing lossy compression video codec. It is a measure of human perception of the codec quality. It quantitatively evaluates the error introduced by the compression codec. The higher the PSNR value measured in dB, the better the codec and the value can be up to infinity. However, PSNR sometimes performs poorly on videos encoded with different codecs hence should not be the sole metric relied upon to make a decision.

PSNR is calculated by the ratio between the maximum possible power of the luminance and the mean squared error (MSE) [3].

$$MSE = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[f(i,j) - F(i,j)]^2}{M.N}$$

$$PSNR = 20.log_{10}\left(\frac{2^n - 1}{\sqrt{MSE}}\right)$$

Where f is the original signal and F is the reconstructed signal and MxN the frame size and $2^n - 1$ is the maximum possible of luminance.

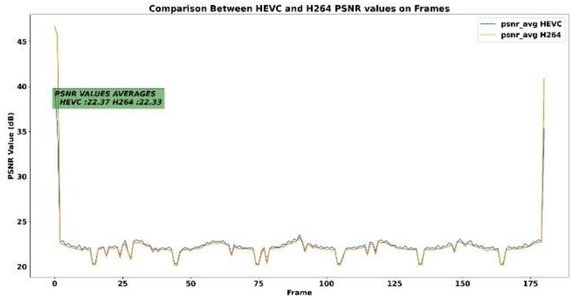

Figure 3. PSNR values comparison for H.264 and H.265

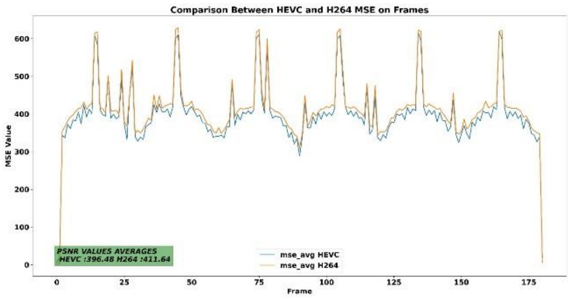

Figure 4. Mean Square Error (MSE) for PSNR

From Figure 3 shown above, it is observed that PSNR values for H.265 are mostly higher than those of H.264 signifying the possibility of H.265 being a better video codec in terms of human perception. This means that a human observer is likely to conclude that H.265 has the best quality. The average PSNR value HEVC is 22.37 while that of H.264 is 22.33 further buttressing our assumption. Figure 4 on the MSE also shows that the H.265 codec has the lowest MSE on PSNR comparison.

Calculating just the PSNR average value comparison is not enough. As already stated, PSNR measures human perception of the video codec. Human perception is in 3 color modes Y, U and V representing luminance (brightness), blue color projection and red color projection respectively. From the analysis it is evident that H.264 outperforms H.265 in luminance and blue color projection, however H.265 outperforms H.264 in red color projection.

 | DOI: 10.20533/ijmip.2042.4647.2023.0068

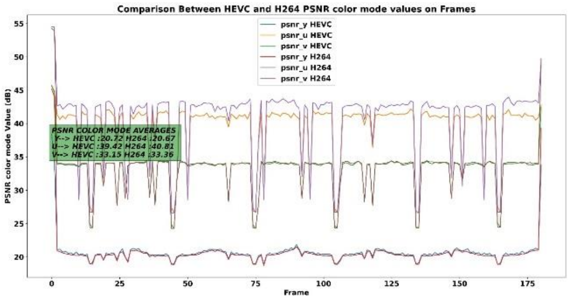

Figure 5. PSNR values per color mode for both H.264 and H.265

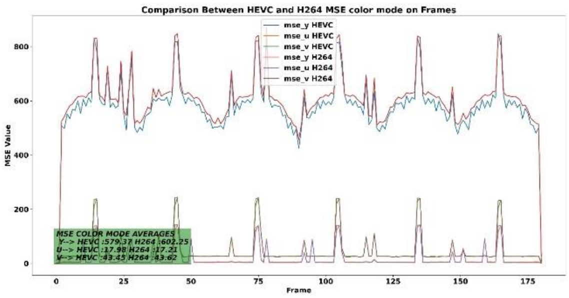

Figure 6. MSE values per color mode for both H.264 and H.265

5. Structural Similarity Index (SSIM)

SSIM is a human perception measure that calculates the degradation of video image quality caused by encoding the video into different codecs. Compared with PSNR, SSIM is concerned about the visible structure of the video and it is often better than PSNR in analyzing video quality.

SSIM extracts 3 main features for video images such as Luminance $l(x, y)$, Contrast $c(x, y)$ and structure $s(x, y)$ and therefore analysis is performed on the 3 features. The SSIM values range from 0 – 1 and the greater the value the better the codec.

It's calculated as follows [4]:

$$SSIM(x, y) = l(x, y).\, c(x, y).\, s(x, y)$$

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

The SSIM value is calculated using the average of all the SSIM values.

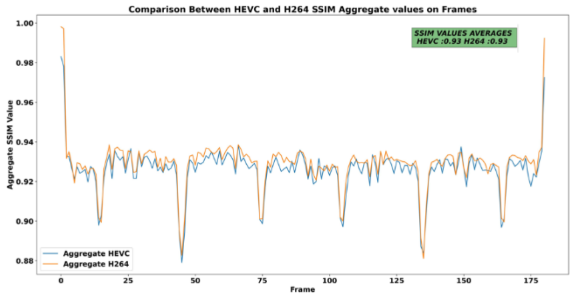

Figure 7. The aggregate SSIM value on color modes for H.265 and H.264

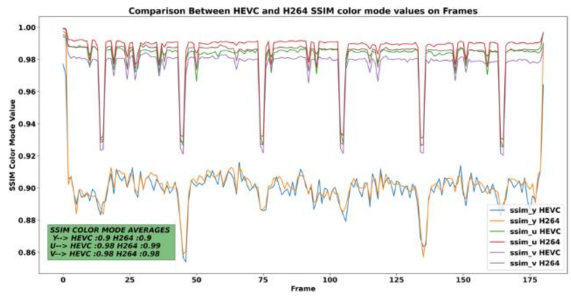

Figure 8. SSIM values for each of the color modes for H.264 and H.265

From Figure 7, the aggregate SSIM value for H.264 seems slightly greater than that of H.265 however the average values are the same (0.93 for H.265 and 0.93 for H.264) and are almost equal to the optimum value of 1 indicating both the of two codecs perform particularly well in terms of SSIM.

When we focus on the SSIM values for the color modes as in Figure 8, the Y channel (Brightness) for H.265 and H.264 have an average value of 0.9. The U (Blue projection) is also almost equal to 1 with 0.99 for H.264 but slightly lower for H.265 with 0.98, as for the V red color projection the two codecs have the same average of 0.98.

We therefore conclude that H.264 and H.265 comparison in terms of SSIM values yield almost the same results, however H.264 results are slightly better in terms of the blue color projection.

6. Visual Information Fidelity (VIF)

VIF is a type of quality metric performance measure which takes into account that the quality of a video is directly proportional to the loss of information fidelity. It is a combination of four scales. VIF values range from 0 to 1 and the higher the value, the better the codec.

VIF is calculated using $I(\vec{C}^N; \vec{E}^N \mid s^N)$ and $I(\vec{C}^N; \vec{F}^N \mid s^N)$ which indicates the optimal information picked up by the brain from a particular channel [5].

$$I\left(\vec{C}^N; \vec{E}^N \middle| s^N\right) = \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{M} log_2\left(1 + \frac{s_i^2 \lambda_k}{\sigma_n^2}\right)$$

$$I\left(\vec{C}^N; \vec{F}^N \middle| s^N\right) = \frac{1}{2}\sum_{i=1}^{N}\sum_{k=1}^{M} log_2\left(1 + \frac{g_i^2 s_i^2 \lambda_k}{\sigma_v^2 + \sigma_n^2}\right)$$

$$VIF = \frac{\sum_{j\epsilon channels} I\left(\vec{C}^{N,j}; \vec{F}^{N,j} \middle| s^{N,j}\right)}{\sum_{j\epsilon channels} I\left(\vec{C}^{N,j}; \vec{E}^{N,j} \middle| s^{N,j}\right)}$$

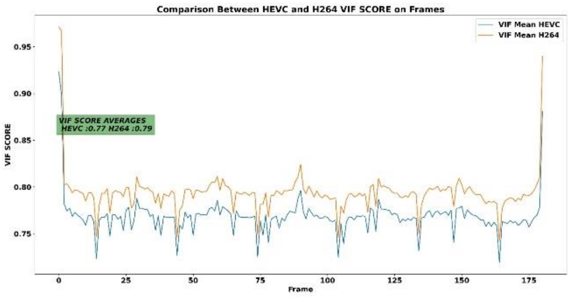

Figure 9. VIF score for H.264 and H.265

Figure 9 indicates the VIF score for H.265 and H.264. VIF Mean refers to the average score for the 4 scales in each frame. The VIF mean line plot for H.264 is higher than that of H.265 indicating that it is the best in terms of VIF. This is confirmed by the average VIF score of 0.77 for H.265 and 0.79 for H.264.

7. Video Multi-method Assessment Fusion (VMAF)

VMAF is a metric developed by Netflix which infuses human vision modeling together with machine learning and has become more popular for its success in automating subjective testing of video image quality which requires users to watch and give a score to videos. VMAF values range from 0 to 100 and the higher the VMAF value the better.

The VMAF is based on three other quality metrics: Visual Information Fidelity (VIF) [6], Detail Loss Metric (DLM) [7] and Mean Co-Located Pixel Difference (MCPD) [8].

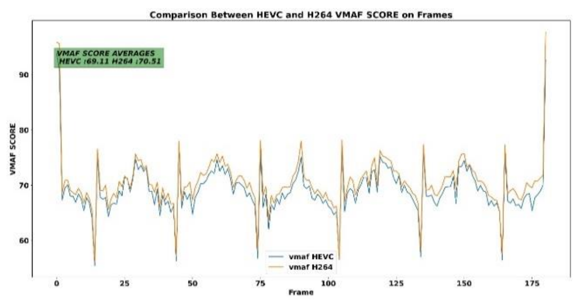

Figure 10. VMAF score for H.264 and H.265

Figure 10 above shows the line plots for H.264 being higher than those of H.265 signifying that H.264 performs better than H.256 when analyzed on the VMAF score. This is confirmed by the average VMAF score which indicates 69.11 for H.265 and 70.51 for H.264.

8. Energy Consumption

In order to use these video codecs in a sustainable way with the current high demand, we need to find which codec present less energy consumption. For this purpose, we used Intel Power Gadget to calculate the average of consumption of power of the processor. The log results show a higher consumption of energy with the H.264 codec compared to H.265, thus the latest is a more energy saving codec.

Processor Energy =IA Energy + GT Energy + Others

IA Energy is the Energy of the CPU and processor cores

GT Energy is Energy of the processor graphics

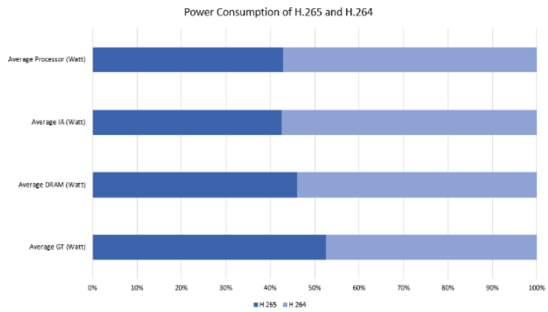

Figure 11. Power Consumption of H.265 and H.264

9. Conclusion

We have analyzed the performance of the quality for the two codecs focusing on 4 main areas. In the first metric PSNR, we have noticed that H.265 is better than H.264 in encoding this particular video. In the second metric SSIM the performance is identical for the two codecs, for the third metric VIF H.264 slightly outperform H.265 however for the last metric VMAF H.264 achieves better results. Nonetheless it’s worth noting the high performance of H.265 in terms of the size of the video after compression with a reduction of 62% compared to H.264 with only 30.42%.

10. References

[1] Tamhankar, A., Rao, K. R. (2003). An overview of H. 264/MPEG-4 part 10. In Proceedings EC-VIP-MC 2003.

4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications (IEEE Cat. No. 03EX667) (Vol. 1, pp. 1-51). IEEE.

[2] Sullivan, G.J. and Ohm, J-R. (2010). "Recent developments in standardization of high efficiency video coding (HEVC)", Proc. SPIE 7798, Applications of Digital Image Processing XXXIII, 77980V. DOI: 10.1117/12.863486.

[3] Wang, Z., Lu, L, Bovic, A. C. (2004). "Video quality assessment using structural distortion measurement," Signal Processing: Image Communication, special issue on "Objective video quality metrics", vol. 19, no. 2, pp. 121-132.

[4] Ou, T., Huang, Y., Chen, H.H. (2011). "SSIM-Based Perceptual Rate Control for Video Coding," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 5, pp. 682-691. DOI: 10.1109/TCSVT.2011.2129890.

[5] Sheikh, H. R., and Bovik, A. C. (2005). A visual information fidelity approach to video quality assessment. In The first international workshop on video processing and quality metrics for consumer electronics (Vol. 7, No. 2, pp. 2117-2128). sn.

[6] Sheikh, H.R., Bovik, A.C. (2006). "Image information and visual quality," in IEEE Transactions on Image Processing, vol. 15, no. 2, pp. 430-444, DOI: 10.1109/TIP.2005.859378.

[7] Li, S., Zhang, F., Ma, L., Ngan, K.N. (2011). "Image Quality Assessment by Separately Evaluating Detail Losses and Additive Impairments," in IEEE Transactions on Multimedia, vol. 13, no. 5, pp. 935-949, DOI: 10.1109/TMM.2011.2152382.

[8] Rassool, R. (2017). "VMAF reproducibility: Validating a perceptual practical video quality metric," 2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1-2, DOI: 10.1109/BMSB.2017.7986143.