

Subjective Evaluation of Ultra-high Definition (UHD) Videos

Tariq Rahim¹, Soo Young Shin^{*}

^{1, *} Kumoh National Institute of Technology, Gumi, South Korea,
[E-mail : tariqrahim@kumoh.ac.kr]

^{*} Corresponding Author : Soo Young Shin
[E-mail : wdragon@kumoh.ac.kr]

*Received December 31, 2019; revised April 8, 2020; accepted April 25, 2020;
published June 30, 2020*

Abstract

This paper presents a detailed subjective quality assessment for the ultra-high definition (UHD) videos having frame rates of 30fps and 60 fps. The subjective assessment is based on the ITU-R BT-500 recommendations, where double stimulus continuous quality scale (DSCQS-type II) test is performed for the evaluation of the perceived quality of the user's in terms of differential mean opinion score (DMOS). Encoding of the UHD videos by opting encoders i.e. H.264/AVC, H.265/HEVC, and VP9 at five different quantization parameter (QP) levels is done to investigate the perceived user's quality of experience (QoE) given as DMOS. Moreover, the encoding efficiency as the encoding time for each encoder and qualitative performance by employing full-reference (FR) quality metrics are presented in this work.

Keywords: Encoders, H.264/AVC, H.265/HEVC, quality of experience, quantization parameter, and VP9.

1. Introduction

The extensive use of digital video for entertainment and communication motivates researchers to acquire a better perception of the content of videos. The fast and robust quality of the video's measurements is necessary to maximize the quality of experience (QoE) and resource allocation control within networks [1]. Nowadays, most online video web sites and broadcasters produce video content at high-definition (HD) resolutions. The success of HD contents has led to the development of 4K ultra-high definition (UHD) contents regarded as the future standard in video applications. The film industry and display manufacturers are paying more attention to this 4K UHD format, and video streaming services such as Amazon, Netflix, and YouTube already are supporting 4K UHD resolution videos [2].

In perfect conditions, UHD is supposed to accommodate viewers with improved visual experience by a wide field of view (FOV) in both vertical and horizontal directions of the screen [3]. 4K UHD has a spatial resolution four times higher than the full high definition (FHD) resolution, which has a format of 1920×1080 and can, eventually process a higher amount of visual information. The extension of the resolution is the initial step towards a naturalistic and immersive visual experience [4].

Nevertheless, the 4K UHD format yields a new challenge, Managing the amount of additional data in 4K video services demands increased bandwidths and storage capacity. To overcome this, the compression of videos is an indispensable and crucial step. High-efficiency video coding (HEVC) [2] is a standard video compression commonly used for 4K UHD applications. Nevertheless, the previous standard, i.e., H.264/AVC [2], is still the most prevalent in real-world applications. Recently, VP9 [2] was introduced as a part of the WebM project. VP9 is an open-source technique for compression that is adopted for research. For any compression system, a higher compression rate leads to more visible distortions. Hence, it is necessary to examine the trade-off between the compression efficiency and the perceived quality. The main improvement introduced by the UHD resolution concerning over previously achieved resolutions is video quality. The quality of videos can be measured using subjective or objective techniques [5]. Subjective techniques involved subjective tests, i.e. directly asking viewers their opinion on the perceived quality. Objective techniques are classified into three main classes: full Reference (FR), reduced reference (RR) and no reference (NR) quality assessments. All of these classes model the quality of the video based on mathematical equations after feature extraction from the video source [1].

Higher frame rates have a direct relationship with the quality of the videos. The perceived quality of the human as a subjective analysis is more accurate when it comes to high frame rates and high-resolution contents. The development of modern video encoders appeals to an investigation for extensive subjective analysis. This analysis is based on different cases such as content-based, multiple frame rates (from low to high) or compression artifacts as quantization parameters (QP). Furthermore, the subjective analysis is either limited to specific encoders for different artifacts and cases mentioned above. For any compression system, a higher compression rate leads to more visible distortions. Hence, it is necessary to examine the trade-off between compression efficiency and the perceived quality.

This work presents an extensive subjective quality assessment of 4K UHD videos having frame rates of 30fps and 60 fps. All videos were encoded with three different encoders (H.264/AVC, H.265/HEVC, and VP9) at five different quantization parameter (QP) levels to evaluate the performance of each encoder and perceived user's QoE given as DMOS values.

In comparison to previous work, our work is opting for three different encoders at five different QP levels for different frame rates making an extensive subjective analysis reflected as DMOS. Furthermore, the encoding efficiency as the encoding time for each encoder and qualitative performance by employing full-reference (FR) quality metrics are presented in this work. A qualitative result for finding a correlation for the subjective quality assessment model and FR quality metrics are also provided.

The principal contributions of this paper can be summarized as follows:

- An extensive subjective analysis of 4KUHD videos at frame rate of 30 and 60 fps. The dataset only consisted of 15 representative videos at a frame rate of 30 fps up-sampled to 60 fps to extensively analyze 4K UHD contents.
- Encoding both frame rate contents via H.264/AVC, H.265/HEVC, and VP9 encoders under five different QP (22, 27, 32, 39, 44) levels.
- A comprehensive result for all opted encoders for both frame rates is provided in terms of DMOS against each QP level, along with the encoding efficiency as the encoding time for each encoder and qualitative performance by employing FR quality metrics.

3. Related Work

The UHD was recommended as a standard by the ITU recommendation (ITU-R BT.1769) covering the spatial resolutions, of 7680×4320 (8K UHD) and 3840×2160 (4K UHD) [3]. The UHD is an attempt to provide a better visual experience by allowing a wide FOV that virtually incorporates the entire human visual FOV, which is mostly covered by the spatial resolution [6]. A subjective and objective analysis for UHD contents at 120fps where video contents were encoded by HEVC proving that an increase in perceived quality is ensured by high frame rates [7]. A subjective quality assessment for the real-time application (video conferencing) HD video contents at a frame rate of 50fps resulted in an average bit savings of 42.40% for HEVC in comparison to VP9 [8].

Recently, a subjective quality assessment was conducted for the perceptual quality of 4K UHD contents compressed by HEVC and AVC [9]. The evaluation showed that, on average the content compressed by HEVC had a bit rate 66.50% lower and a visual quality 44.40% higher in terms peak signal-to-noise ratio (PSNR) compared to the content compressed by AVC. Similarly, the superiority of HEVC over AVC for 4K UHD contents was demonstrated in [2], where the compression efficiency was compared based on a subjective evaluation among HEVC, VP9, and AVC; in this evaluation HEVC outperformed both VP9 and AVC in terms of bitrate savings [2].

In [5], a subjective assessment for HD and UHD contents compressed by HEVC at the same bit rate was performed. Furthermore, objective and subjective assessments were presented for the comparison analysis of AVC and HEVC encoders for four classes of video resolutions, namely UHD, 1080p, 720p, and 480p; the study showed the superiority of HEVC over its competitor [9]. Moreover, in [10] a comparison for compression efficiency among AVI, HEVC, and VP9 using subjective quality evaluation are done for HD videos having a frame rate of 30fps employing high to low bit rates. Similarly, an attempt to evaluate the compression efficiency comparison between H.264/AVC, HEVC/H.265, VP9, and aomenc (AV1) are done for HD videos for multiple QP levels providing results in terms of PSNR and

VMAF [11]. All of the previous work is attempting to do subjective quality assessment focusing either on HD or low frame rates that lack the user's QoE in relationship with high frame rates. The presented work is employing different encoders for different QP levels having multiple frame rates to investigate user's QoE reflected in the form of DMOS. Furthermore, the encoding efficiency as the encoding time for each encoder and qualitative performance by employing full-reference (FR) quality metrics are presented in this work.

3. Video Database and Encoder Settings

This section details about the database specification and subjective test in this study.

3.1 Video Database Specification

Owing to the limitation of the 4K contents, we used an openly available database containing 15 videos of 4K UHD sequences. These videos are available both in 4:4:4 and 4:2:0 YUV color sampling having 10 and 8 bits per sample respectively [12]. The UHD video contents are available at the frame rate of 30 fps but were up-sampled to 60 fps to perform a more extensive subjective quality assessment to reflect the relationship of frame rate and DMOS values for different encoders at different QP levels. Fig. 1 depicts the frame sample obtained from the SJTU 4K video sequence dataset [12] with the associated indexes and names (available for research purpose at medialab.sjtu.edu.cn/HDR/index.html).

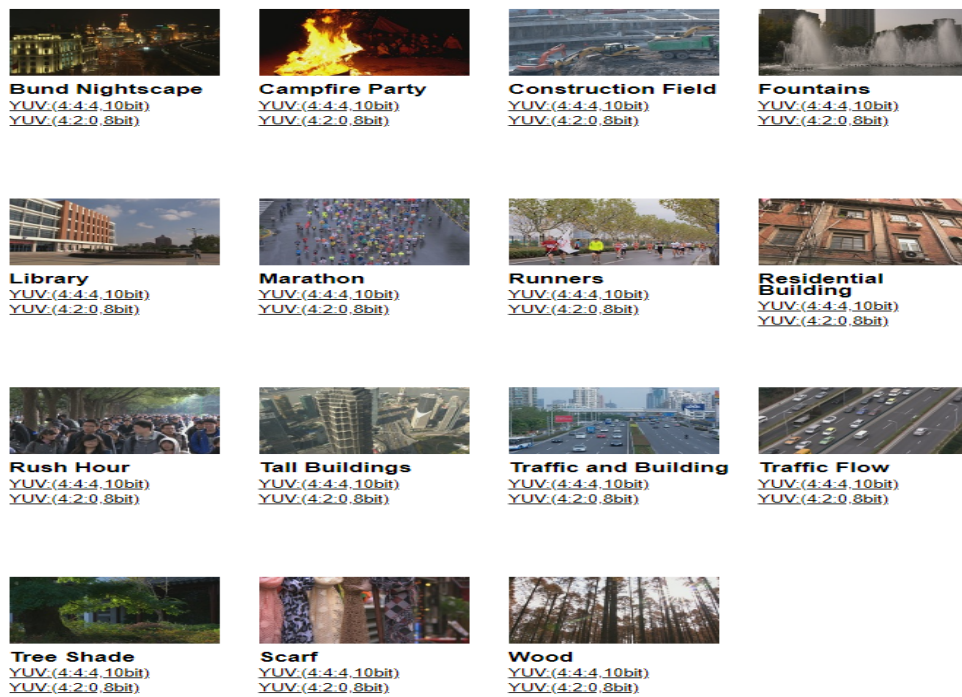


Fig. 1. A Frame sample of each 15 videos in the SJTU 4K video sequence dataset.

3.2 Description of the Content

The spatial information (SI) and temporal information (TI) of the scenes are important parameters and features for the representation of any video sequences. These parameters play a significant role in determining the amount of video compression and exercise a vital impact on the corresponding researches to some degree. The SI is a space-dependent feature of a video that is based on the output of the Sobel filter featuring the areas having a high spatial frequency i.e. edge regions and can be measured with the following equation as:

$$SI = \text{Max}_{time}\{\text{std}_{space}[\text{Sobel}(f_n)]\}, \quad (1)$$

where "*std*" is the measured standard deviation of each frame pixels that is filtered out by Sobel operator. This filtering process is repetitive for all the f_n frames resulting in a time series of spatial information of the video clip. After this, a "*Max_{time}*" is selected for the estimated time series that characterizes the SI of the video clip.

For the TI that represents temporal changes of the video clip led to compute the motion difference features of successive frames by maintaining the same space location and is given as:

$$M_n = f_n(i, j) - f_{n-1}(i, j) \quad (2)$$

$$TI = \text{Max}_{time}\{\text{std}_{space}[M_n(i, j)]\}; \quad (3)$$

where $f_n(i, j)$ and $f_{n-1}(i, j)$ indicates the i_{th} and j_{th} column of the respective n_{th} frame of the clip. The motion difference feature is represented by M_n . So, TI can be demarcated as the maximum value *Max_{time}* of the time series of a set of calculated *std_{space}*. Here it should be noted that the above equation (1, 2 & 3) applies for grayscale images recommended by ITU-R BT.500-13 [13], but as our video contents are colored having three channels i.e. y, u, and v, SI and TI were computed for each channel and then averaged to get a single value. A pseudo-code implemented in MATLAB (R2018b) for the computation of SI and TI for the colored contents is given in Algorithm 1 where the Sobel filter value for computing edges was kept at 0.0904 after hit a trial [13].

Fig. 2 represents the SI and TI plot for the 15 representative videos from the SJTU 4K database. Four videos based on the SI and TI plot were selected for the subjective evaluation at a frame rate of 30 fps and 60 fps. The selected four representative video clips from SI and TI plotting are based on the higher values as information in the form of SI and TI values as shown in **Fig. 2**.

Algorithm 1 Computation of SI-TI values for SJTU 4K videos

Input: SJTU 4K video \longrightarrow Original video content
Output: SI & TI \longrightarrow SI and TI values

1: SET $start = 1$ \longrightarrow Starting point initialization
2: SET $finish = 4$ \longrightarrow Ending point initialization
3: F_N \longrightarrow Frames within video content
4: **for** $i = 1$ to $F_N - 1$ \longrightarrow For all videos, run this loop
5: $M = \text{READ}(\text{SJTU 4K video}, i)$
6: $N = \text{READ}(\text{SJTU 4K video}, i+1)$
7: **for** $j = start$ to $finish$ \longrightarrow Runs 4 times for each frame
8: $U = \text{FILTER}(M(:, :, j), \text{SOBEL})$ \longrightarrow Filter the j_{th} color channel from the yuv video content frame M.
9: $\text{Spatial}(j) = \text{Max}(\text{STD}(U))$ \longrightarrow computes the SI of the j_{th} color channel
10: $V = N(:, :, j) - M(:, :, j)$
11: $\text{Temporal}(j) = \text{Max}(\text{STD}(V))$ \longrightarrow computes the TI j_{th} color channel
12: **end for**
13: $SI(i) = \text{MEAN}(\text{Spatial})$ \longrightarrow SI of the i_{th} video frame as an average
14: $TI(i) = \text{MEAN}(\text{Temporal})$ \longrightarrow TI of the i_{th} video frame as an average
15: **end for**

16: $SI_{video} = \text{SUM}(SI) / F_N$ \longrightarrow SI of the input SJTU video content
17: $TI_{video} = \text{SUM}(TI) / F_N$ \longrightarrow TI of the input SJTU video content

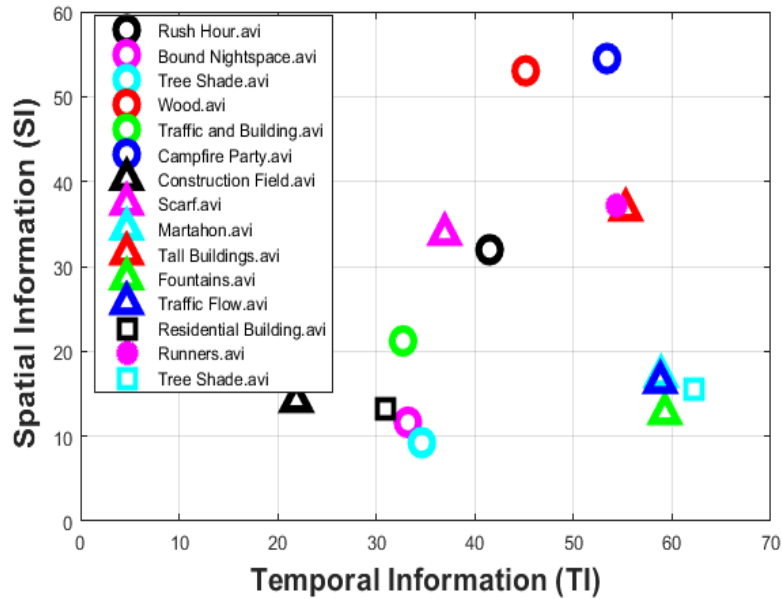


Fig. 2. Spatial and Temporal Information for the SJTU 4K Video database.

3.3 Encoder Settings and Quantization Parameters

From the SI and TI plotting of the SJTU 4K (UHD) database as shown in Fig. 2, four video sequences are selected based on their SI and TI values for the encoding. The four videos are encoded by H.264/AVC, H.265/HEVC, and VP9 at five different QP levels (22, 27, 32, 39, 44). The QP values were chosen to reflect the better performance of the opted encoders: low and high QP indicate low and high compression, respectively. The higher the QP levels, the higher is compression resulting in degradation in quality [14]. When encoded with five different QP levels, videos result in twenty encoded video sequences; thus, all three encoders generate 60 encoded sequences. An extensive data is generated that investigate the QoE against each QP level for both 30fps and 60fps reflected as DMOS values.

We implemented the open-source library FFmpeg; in which libx264, libx265, and libvpx are used as wrapper encoders for H.264/AVC, H.265/HEVC, and VP9, respectively. To lessen the impact of rate control mechanisms for encoder performance, we have implemented a constant rate factor for libx264 and libx265, while constant quantization parameter is selected for libvpx. The details of encoder settings are shown in Table 1, where the preset was chosen as “*veryfast*” for balancing the encoding efficiency and obtain higher quality.

Table 1. Encoder settings with wrapper used.

Encoder	Wrapper	Settings
H.264/AVC	libx264	ffmpeg-i wood.avi -c:v libx264-params qp=22-44-c:a aac- preset veryfast-vbr 3-an output.mp4
H.265/HEVC	libx265	ffmpeg-i wood.avi -c:v libx265-params qp=22-44-c:a aac- preset veryfast-vbr 3-an output.mp4
VP9/WebM	libvpx	ffmpeg-i wood.avi -c:v libvpx-vp9-b:v 2000k-pass 2-c:a opus -b:a -qmin 21 -qmax 23 -f webm output.webm

4. Subjective Quality Assessment

This section explains the experimental investigation for the determination of the association between the QP and user's QoE in terms of DMOS and encoding efficiency as the encoding time and qualitative performance in terms of PSNR and SSIM.

4.1 Video Sequences

This subsection explains the method and materials chosen for the subjective quality assessment test. For the subjective quality assessment, the number of participants was selected based on the International Telecommunication Union (ITU)'s ITU-R BT.500-13 recommendations [13]. For the subjective quality assessment, four representative video clips are chosen from the SI and TI plotting as shown in Fig. 2 that are encoded by H.264/AVC, H.265/HEVC, and VP9 encoders at five different QP levels.

In the rest of the paper, the terms will be use as for ease usage.

- Source video (SV): unimpaired/original video sequence.
- Encoded video (EV): impaired/encoded video sequence.
- Clip: either SV or EV.

4.2 Subjective Experiment Setup

The subjective quality assessment was carried out in a specially designed room based on the ITU-R BT.500-13 recommendations [13]. The room contains only materials related to the subjective quality assessment test with addition to a high-performance display of UHD 4K SAMSUNG series NU7100-7. The size of the display is 75 inches having an aspect ratio of 16:9 and a screen resolution of 3480×2160 . A tool (MSU-PVQT) for subjective quality assessment (VQA) provided by the Moscow State University is employed for the test [15].

A training session for the acquainting of the participants was organized for the subjective quality assessment. The participants were selected from the electronics department, Kumoh National Institute of Technology. After the outlier criterion, a total of 18 participants based on ITU-R BT.500-13 recommendations of diverse age from 22 to 30 was selected. To generate a diverse investigation and QoE of users, 6 doctoral, 6 masters, and 6 undergraduate participants were selected. For the scoring as mean opinion score (MOS), Double Stimulus Continuous Quality Scale Type II (DSCQ-II) is used. In DSCQS-II testing side-by-side clips are played concurrently i.e. encoded and reference clip, but participants are unaware of that which one is reference one. This type of testing ensures an unbiased subjective evaluation reflecting the user's QoE as a perceptual quality in the form of DMOS values.

Table 2. Subjective opinion scores scale

Visual quality	Opinion scores
Excellent	80-100
Good	60-80
Fair	40-60
Poor	20-40
Bad	0-20

4.3 Analysis of Opinion Scores

The quality evaluation of the clip is done by taking the opinion scores (Oss) of the participants on a scale of 0-100. The scale is categorized as that 0 represents bad while 100 represents the excellent score on a continuous rating scale as shown in **Table 2**. Based on the ITU-R BT.500-13 recommendations [13] the participants were shown both SVs and EVs, unaware of the fact that which one is SVs or EVs. After completion of each test, the participants are asked to record their OSs individually on a continuous rating scale of DSCQS-II.

4.4 Method for Scoring

As mentioned earlier, we have used the DSCQ type II method for subjective quality assessment. The reported opinion scores (OS) on the five-point rating scale are changed to a normalized scale ranging within 0 and 100. For assessing FR VQA metrics, it is recommended to use differential mean opinion score (DMOS) [14] and can be calculated as eq. 4 and eq. 5. MOS_{ith}^{SV} and MOS_{ith}^{EV} are the measured MOS of SV and EV sequences of the *ith* video of 4K video, respectively.

$$DMOS_i = MOS_{ith}^{SV} - MOS_{ith}^{EV}, \quad (4)$$

Where; MOS for each clip frame rate was computed for each participant and encoder as OS in the generalized form as follows:

$$MOS_{ith} = \frac{1}{K} \sum_{i=1}^K OS_{ith}^{jth}, \quad (5)$$

where, OS_{ith}^{jth} is the recorded OS for all $ith = 1.2.3...K$ for the jth clips, while K represents the total number of participants number involved in the subjective quality assessment. A representation of the DSCQS type-II rating scale can be viewed in Fig. 3. In Fig. 3 the video quality differential is referring to the estimation of differential opinion score (DOS) of a clip [14].

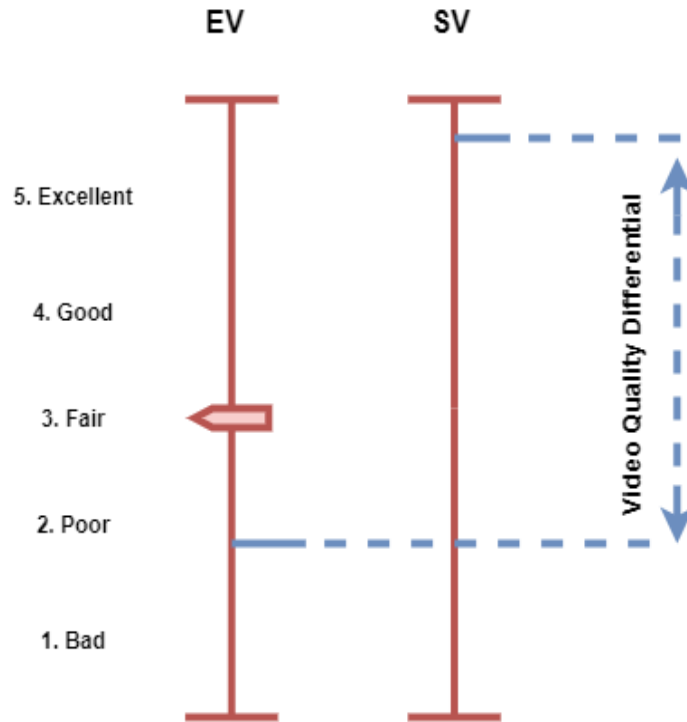


Fig. 3. DSCQS type-II illustrating video quality differential

5. Experimental Results and Discussion

Unlike previous works, an extensive subjective quality assessment of 4K UHD videos having frame rates of 30fps and 60fps using H.264/AVC, H.265/HEVC, and VP9 encoders at five different QP levels (22, 27, 32, 39, 44) is done. Fig. 4 shows the subjective quality assessment results that evaluate the performance of each encoder and perceived user's QoE given as DMOS values for both 30fps and 60fps. The results show a vivid demonstration of the decrease in the degradation of the perceived quality of the video, i.e. the QoE of the users in terms of DMOS values as the QP level increase. Fig. 4(a) and Fig. 4(b) shows the results of DMOS against QP for 30fps and 60fps respectively, where it can be seen that for 60fps clips the QoE is lesser than 30fps for high QP levels for all opted encoders. Moreover, the performance of H.265/HEVC encoder outperforms H.264/AVC and VP9 encoders for both frame rates at low to high QP levels in terms of DMOS values.

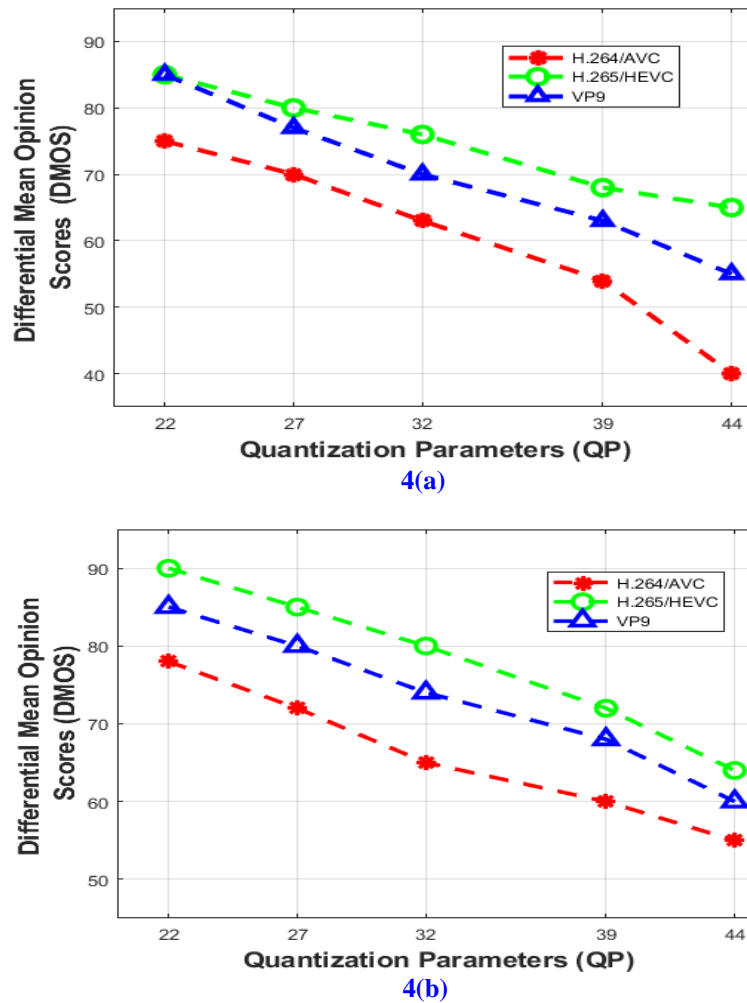


Fig. 4. Subjective evaluation in terms of DMOS at (a) 30 fps and (b) 60 fps.

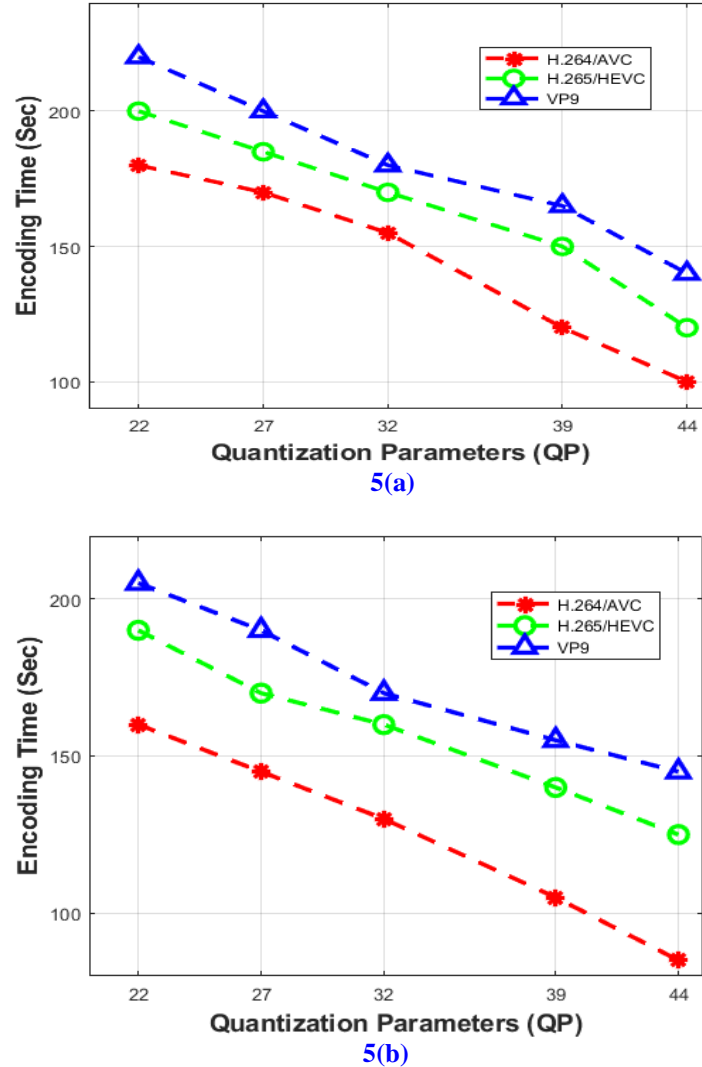


Fig. 5. Encoding efficiency of the selected encoders at (a) 30 fps and (b) 60 fps.

Moreover, the encoding time of the implemented encoders for different QP levels is done to evaluate the efficiency of each encoder. Encoding time is also an important feature of an encoder for contents-based encoding. **Fig. 5(a)** and **Fig. 5(b)** shows the encoding time for 30fps and 60fps UHD contents, respectively. The encoding time for both 30fps and 60fps for H.264/AVC encoder requires the least encoding time (sec) for UHD video sequences from low to high QP levels. **Fig. 5** also illustrates that though VP9 is an open-source encoder and readily use nowadays, still taking more encoding time than H.265/HEVC encoder. Here it should be noted that the better performance of H.264/AVC for UHD content is at the cost of QoE of end users making it a tradeoff between speed and quality.

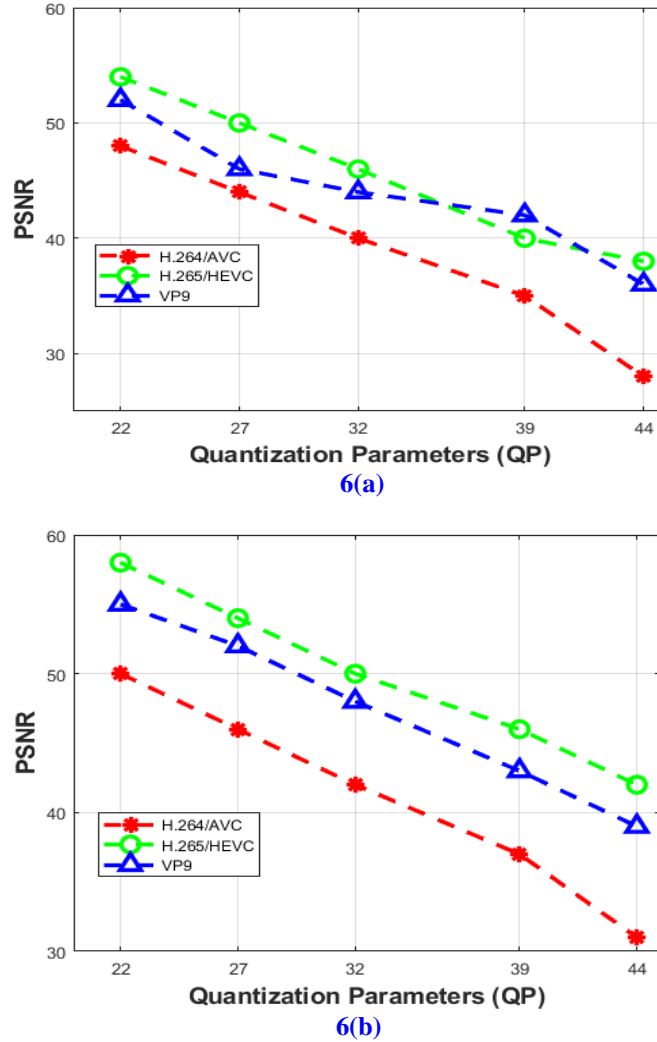


Fig. 6. PSNR for quality evaluation among SV and EV at (a) 30 fps and (b) 60 fps.

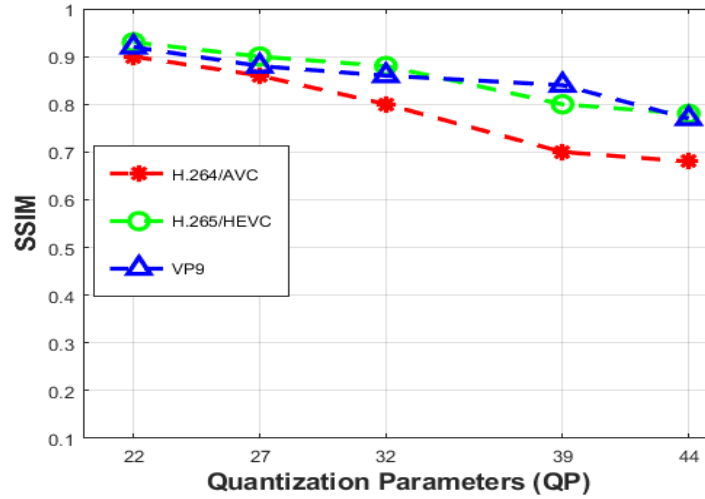
Furthermore, a comprehensive investigation of the performance of the selected encoders under each QP level is performed by using FR quality metrics i.e. of PSNR [16] and structural similarity index (SSIM) [17]. The PSNR is an FR metric for the calculation of the perceived quality in decibel dB. **Fig. 6(a) and Fig. 6(b)** depicts the generated PSNR for UHD contents at 30fps and 60fps. The higher the value, the better the quality and can be calculated as:

$$MSE = \sum_i^M \sum_j^N [(SV_s) - (EV_s)]^2, \quad (6)$$

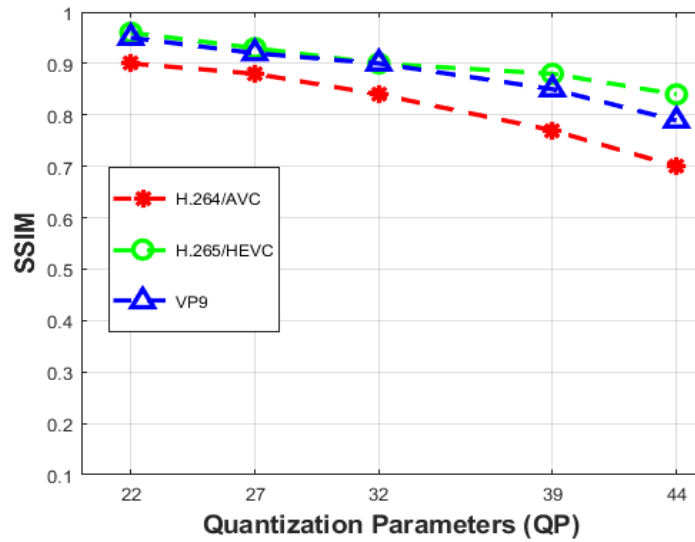
Where SV_s and EV_s represents the source and encoded clip respectively. So PSNR can be is given as:

$$PSNR = 20 \log_{10} \left(\frac{Max_{SV}}{\sqrt{MSE}} \right) \quad (7)$$

Fig. 6 depicts the PSNR generated for 30fps and 60fps of the encoders implemented. It can be seen clearly in **Fig. 6(a)** and **Fig. 6(b)** that for low to high levels of QP, encoders H.265/HEVC and VP9 is outperforming encoder H.264/AVC. **Fig. 6** also validates subjective assessment done by the users for both frame rates as shown in **Fig. 4** that QoE is high for low QP levels (DMOS values).



7(a)



7(b)

Fig. 7. SSIM for quality evaluation among SV and EV (a) 30 fps and (b) 60 fps.

We also computed the SSIM [17] that represents the structural degradation between SVs and EVs for each level of QP to evaluate the performance of each encoder. **Fig. 7** depicts the SSIM generated for 30fps and 60fps of the encoders implemented. Shown in **Fig. 7(a)** and **Fig. 7(b)**, a similar behavior of the encoders is reported that for low to high levels of QP, encoders H.265/HEVC and VP9 is outperforming encoder H.264/AVC. **Fig. 7** also validates subjective

assessment done by the users for both frame rates as shown in Fig. 4 that QoE is high for low QP levels (DMOS values). Therefore, for UHD video sequences, the H.265/HEVC is superior to all compared encoders at every QP level in terms of DMOS values, encoding efficiency, PSNR, and SSIM. The subjective evaluation reflected the video quality perceived by users, while the FR assessment reflected the amount of degradation in the quality of SV sequences compared to the EV sequences caused by different encoders.

Moreover, a comparative qualitative analysis is done to reflect the subjective results for finding the correlation between the DMOS scores and FR quality metrics. Table 3 shows qualitative analysis benchmarked with the convolutional neural network (CNN) based subjective analysis methods. Clearly, it can be seen that subjective analysis done using ITU-R BT-500 recommendations using (DSCQS-type II) is outperforming the CNN based subjective quality assessment in terms of PSNR and SSIM, and correlation coefficients (CC) such as Pearson linear correlation coefficient (PLCC) and Spearman rank-order correlation coefficient (SROCC). The results shown in Table 3 are only for HEVC and 30fps as its performance was good in our findings. Albeit, the CNN based subjective quality assessment is lesser time consuming and without human opinion for QoE of users, it is suitable for no reference-quality assessment [18]. In addition to that, the extraction of various spatial cues is done for training phase making CNN based models as hit and trial model [19]. Moreover, handcrafted features are still required for higher-level feature extraction making the CNN models deep.

Table 3. PLCC and SROCC for HEVC at 30fps

Encoder HEVC (30fps)	PSNR			SSIM			PLCC			SROCC		
	Our	CNN based		Our	CNN based		Our	CNN based		Our	CNN based	
		[18]	[19]		[18]	[19]		[18]	[19]		[18]	[19]
QP 22	55dB	42dB	No	0.92	No	No	0.93	No	No	0.88	No	No
QP 27	50dB	39dB	42dB	0.90	No	0.86	0.91	No	0.85	0.86	No	0.85
QP 32	46dB	37dB	40dB	0.86	No	0.83	0.85	No	0.82	0.85	No	0.84
QP 39	40dB	35dB	No	0.80	No	No	0.82	No	No	0.82	No	No
QP 44	39dB	29dB	No	0.79	No	No	0.77	No	No	0.79	No	No

We also calculated the CC as PLCC and SROCC of 60fps video contents when encoded with H.265/HEVC. Shown in Table 4, a high correlation was found between FR quality metrics and CC findings. Table 4 also shows a similar behavior of high CC values as observed by users as DMOS values as shown in Fig. 4.

Table 4. PLCC and SROCC for HEVC at 60fps

Encoder HEVC	PSNR	SSIM	PLCC	SROCC
QP 22	58dB	0.95	0.94	0.94
QP 27	55dB	0.91	0.90	0.92
QP 32	50dB	0.90	0.91	0.90
QP 39	47dB	0.88	0.87	0.85
QP 44	32dB	0.84	0.85	0.81

7. Conclusions and Future Works

We have presented an extended subjective quality assessment based on the ITU-R BT-500 recommendations for 4K UHD videos having frame rates of 30fps and 60 fps. All videos were encoded with three different encoders (H.264/AVC, H.265/HEVC, and VP9) at five different quantization parameter (QP) levels to evaluate the performance of each encoder and perceived user's QoE given as DMOS values. In comparison to previous work, our work is opting for three different encoders at five different QP levels for different frame rates making an extensive subjective analysis reflected as DMOS. Furthermore, the encoding efficiency as the encoding time for each encoder and qualitative performance by employing full-reference (FR) quality metrics are presented in this work.

The future work comprised of designing a detailed objective model that can reflect the subjective results in terms of video quality and metrics and correlation coefficient models.

References

- [1] Alizadeh, M., and M. Sharifkhani, "Subjective video quality prediction based on objective video quality metrics," in *Proc. of 2018 4th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. IEEE, 2018. [Article \(CrossRef Link\)](#)
- [2] Cheon, Manri, and Jong-Seok Lee, "Subjective and objective quality assessment of compressed 4K UHD videos for immersive experience," *IEEE Transactions on Circuits and Systems for Video Technology*, 28.7, 1467-1480, 2017. [Article \(CrossRef Link\)](#)
- [3] BT2020, I. T. U. R, "Parameter values for ultra-high definition television systems for production and international program exchange," 2012, available at, <https://www.itu.int/md/R12-WP6C.AR-C-0250/en>
- [4] Yurek, Jeff, "Ultra-High Definition: Beyond pixel count," *IEEE Consumer Electronics Magazine*, 4.4, 89-9, 2015. [Article \(CrossRef Link\)](#)
- [5] Sotelo, Rafael, et al, "Subjective video quality assessments for 4K UHD TV," in *Proc. of 2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. IEEE, 2017. [Article \(CrossRef Link\)](#)
- [6] The Present State of Ultra-high Definition Television, document Rec. BT.2246, ITU-R, 2017, available at, <http://www.itu.int/pub/R-REP-BT.2246>
- [7] Mackin, Alex, Fan Zhang, and David R. Bull, "A study of high frame rate video formats," *IEEE Transactions on Multimedia*, 21.6, 1499-1512, 2018. [Article \(CrossRef Link\)](#)
- [8] Řeřábek, Martin, et al, "Quality evaluation of HEVC and VP9 video compression in real-time applications," in *Proc. of 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2015. [Article \(CrossRef Link\)](#)

- [9] Akyazi, Pinar, and Touradj Ebrahimi, "Comparison of compression efficiency between HEVC/H.265, VP9 and AV1 based on subjective quality assessments," in *Proc. of 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2018. [Article \(CrossRef Link\)](#)
- [10] Guo, Liwei, Jan De Cock, and Anne Aaron, "Compression performance comparison of x264, x265, libvpx and aomenc for on-demand adaptive streaming applications," in *Proc. of 2018 Picture Coding Symposium (PCS)*. IEEE, 2018. [Article \(CrossRef Link\)](#)
- [11] Song, Li, et al, "The SJTU 4K video sequence dataset," in *Proc. of 2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 2013. [Article \(CrossRef Link\)](#)
- [12] Series, B.T, "Methodology for the subjective assessment of the quality of television pictures," *Recommendation ITU-R BT*, 500-513, 2012. [Article \(CrossRef Link\)](#)
- [13] Usman, M. A., Usman, M. R., & Shin, S. Y, "Quality assessment for wireless capsule endoscopy videos compressed via HEVC: From diagnostic quality to visual perception," *Computers in biology and medicine*, 91, 112-134, 2017. [Article \(CrossRef Link\)](#)
- [14] MSU, "Msu perceptual video quality tool," available at, https://compression.ru/video/quality_measure/perceptual_video_quality_tool.html.
- [15] Yang, Shu, et al, "An objective assessment method based on multi-level factors for panoramic videos," in *Proc. of 2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 2017. [Article \(CrossRef Link\)](#)
- [16] Wang, Zhou, Ligang Lu, and Alan C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal processing: Image communication*, 19.2, 121-132, 2004. [Article \(CrossRef Link\)](#)
- [17] Alizadeh, M., A. Mohammadi, and M. Sharifkhani, "No-Reference Deep Compressed-Based Video Quality Assessment," in *Proc. of 2018 8th International Conference on Computer and Knowledge Engineering (ICCCKE)*. IEEE, 2018. [Article \(CrossRef Link\)](#)
- [18] Ahn, Sewoong, and Sanghoon Lee, "Deep blind video quality assessment based on temporal human perception," in *Proc. of 2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018. [Article \(CrossRef Link\)](#)



Tariq Rahim is a Ph.D. Wireless and Emerging Network System Laboratory (WENS Lab.) in Department of IT Convergence Engineering, Kumoh National Institute of Technology, Republic of Korea. He has completed Master in Information and Communication Engineering from Beijing Institute of Technology, PR. China 2017. His research interests include image and video processing and quality of experience for high resolution videos. Email: tariqrahim@kumoh.ac.kr



Soo Young Shin was born in 1975. He received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from Seoul National University, Seoul, South Korea, in 1999, 2001, and 2006, respectively. From July 2006 to June 2007, he was a Visiting Scholar with the FUNLab, University of Washington, Seattle, WA, USA. For three years, he was with the WiMAX Design Lab, Samsung Electronics, Suwon, South Korea. He currently holds the position of Associate Professor with the School of Electronics, Kumoh National Institute of Technology, Gumi, South Korea. His research interests include next generation mobile wireless broadband networks, cognitive radio networks, non-orthogonal multiple access, cooperative communications, wireless local area networks and personal area networks, IoT, wireless sensor networks, and industrial and military networks. Email: wdragon@kumoh.ac.kr