J. Vis. Commun. Image R. 57 (2018) 76–83

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

HDR video quality assessment: Perceptual evaluation of compressed HDR video ☆

Xiaofei Pan [a,f], Jiaqi Zhang [b], Shanshe Wang [c,*], Shiqi Wang [d], Yun Zhou [e], Wenhua Ding [f], Yahui Yang [a]

[a] School of Software and Microelectronics, Peking University, Beijing, China
[b] Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[c] School of Electronics Engineering and Computer Science, Peking University, Beijing, China
[d] Department of Computer Science, City University of Hong Kong, Hong Kong, China
[e] Academy of Broadcasting Science, SAPPRFT, Beijing, China
[f] China Central Television, Beijing, China

ARTICLE INFO

Article history:
Received 5 July 2018
Revised 25 September 2018
Accepted 16 October 2018
Available online 22 October 2018

Keywords:
High dynamic range (HDR)
Subjective quality assessment
Video compression

ABSTRACT

Compared with standard dynamic range (SDR) video, the high dynamic range (HDR) video can provide us significantly enhanced viewing experience. In particular, compared to SDR video, the HDR video has better contrast and preserves more details for the same scene. With the rapid development of HDR video compression technology, there is a lack of trusted quality measure of HDR video compression. In order to facilitate the future development of objective HDR quality assessment, we build a HDR video quality assessment database, in which the bitstream is created by compressing a series of HDR video sequences. In the compression, the quantization parameters (QP) are set to 12 levels according to the configuration of the codec. The subjective quality of each bitstream is rated by 22 viewers. It is revealed that the subject viewers have arrived at a reasonable agreement on the subjective quality of different QP levels. This paper presents the results of subjective assessment of HDR compressed video, which also exhibits that there is significant room to further improve the objective HDR video quality assessment algorithms.

1. Introduction

Recently, there has been a dramatic increase of the demand for digital video services. With the advancement of video technology, the video quality can be improved through higher video resolution (SD, HD, UHD) and frame rate (e.g., from 30 to 120 fps). The high dynamic range (HDR) video, which renders a greater dynamic range of luminosity, is believed to be a big leap in TV technology [1]. Different from the 4 k and 8 k technologies which improve the video quality in terms of resolution, HDR advances the video technology from the perspective of the ratio of luminance. As such, the dynamic range of video rendered by display can be increased, and the visual scene that is more close to the real scene perceived by human visual system (HVS) can be experienced.

Besides increasing the perception of sharpness [2], higher dynamic range can also support more saturated colors than low dynamic range. Fig. 1 shows the luminance the HVS can view, and apparently HDR video covers a wider luminance range than the traditional SDR video.

The concept of dynamic range lies in the ratio between the maximum and minimum luminance perceived from a scene in a natural environment or rendered by a display. In digital cameras, f-stop is the most commonly used unit for measuring the dynamic range, which describes the luminance by power of 2. Moreover, here the total amount of all visible light that passes through the specific space is termed as luminance, and it is usually measured in candela per unit area which is equivalent to nits [4]. For example, in the real world, the luminance of the sun is around 6×10^8 nits, and the starlight in the night has a luminance around 10^{-4} nits or lower [5]. Within a room, a human face has a luminance about 50 nits, and a dark surface may be 1 nit [6].

As a matter of fact, although the HVS can perceive all the objects in the dynamic range from 10^{-6} nits to 10^8 nits, which is equivalent to the total 14 log units, viewing all of the luminance over this range simultaneously is not possible. Dynamic range from 10^{-1} to 10 nits is called the scotopic range and the range from 0.01 to 10^8 nits is called photopic range. The overleap range is termed as the mesopic range [3]. The HVS is capable of perceiving the

☆ This paper has been recommended for acceptance by Zicheng Liu.

* Corresponding author.
E-mail addresses: panxiaofei@cctv.com (X. Pan), zhangjiaqi17@mails.ucas.edu.cn (J. Zhang), sswang@pku.edu.cn (S. Wang).

https://doi.org/10.1016/j.jvcir.2018.10.016
1047-3203/

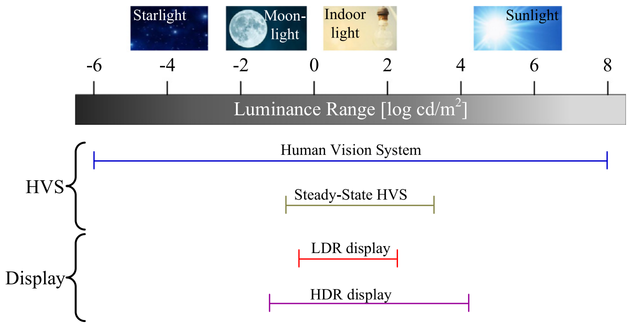


Fig. 1. Dynamic range of human vision and display [3].

dynamic range around 3.7 log unit under specific viewing conditions [3,7]. As such, an adaption process of HVS is performed in a new environment, where the dynamic range of light level has a dramatic change.

Wide color gamut (WCG) moves forward to improve the viewing experience, as the range of colors that is rendered can be further extended. In HDTV standards [9], BT.709, the color gamut of which corresponds to the triangle area labeled with BT.709 in Fig. 2, was adopted. UHDTV systems employ a wider color gamut, which was known as BT.2020 [10]. To give an intuition about the color gamut, based on the research of [8], Fig. 2 also shows the Pointer's gamut, which describes an approximate of the gamut of real scene as can be perceived by HVS, and every color can be reflected through the material of Pointer's gamut [11]. Basically, BT.2020 covers a much wider area of the color space than BT.709. The BT.2020 convers 63.4% of the chromaticities while BT.709 covers only 33.5% of the charomaticties [11]. Pointer's Gamut is significantly bigger than the BT.709 color space while BT.2020 color space is designed to include Pointer's gamut. Moreover, it is clarified here that the concept of HDR and WCG is independent, and there is no direct relationship between WCG and HDR. In other words, SDR videos may possess a wider color gamut, BT.2020, but still have a dynamic range lower than HDR video. Alternatively, a HDR video can own higher dynamic range and a wider color gamut simultaneously.

HDR video is defined based on the signal that has a dynamic range higher than SDR video, as SDR video only supports the luminance in the range of approximate 0.1 to a few hundred nits [12]. Regarding the definition of dynamic range, SDR has a range less than 10 f-stops, and HDR has a range greater than 16 f-stops. There also exists an intermediate signal called enhanced dynamic range (EDR), which usually has the range between 10 and 16 f-stops. To generate HDR video, typically there are two transfer functions, PQ and HLG [13]. Due to the emerging of HDR video and the advances of HDR display, the corresponding HDR video compression algorithms have been developed based on the existing video coding standards such as HEVC [14,15] and AVS2 [16–18]. AVS2 is a new generation of video coding standard developed by the IEEE 1857 Working Group under project 1857.4. AVS2 is also the second generation of AVS video coding standard developed by Audio and Video Coding Standard(AVS) Working Group of China, which is designed to achieve significant coding efficiency improvements relative to the preceding AVS1 [19] standards. In these standards, several methods have been developed for HDR video coding. In particular, to support high bit-depth coding for HDR video, the HDR video format is signalled in the Video Usability Information (VUI) and the Supplement enhancement information (SEI) is also defined to help the conversion from the HDR signal to the display at the receiver side.

The VUI supports the convey of the transfer function and the color space. With the transfer function, the original linear video signal can be converted into a nonlinear signal based on the PQ and HLG functions. The nonlinear video signal can be derived with the color primaries and the transfer function. The supplement enhancement information (SEI), which includes mastering display color volume SEI message, tone mapping information SEI message, provide the conversion mechanisms to support SDR-HDR interoperability. In [20], a method that respectively adjusts the chroma and luma QP value was proposed to improve the perceptual quality improvement, which brings the bit-rate saving 27% in terms of MOS [21] based on HEVC. In [22], the authors proposed a residual highway convolutional neural network (RHCNN) for in-loop filtering in HEVC, which can improve the reconstruction video quality. A content adaptive image subsampling method is also proposed in [23] in stereo interleaving video coding.

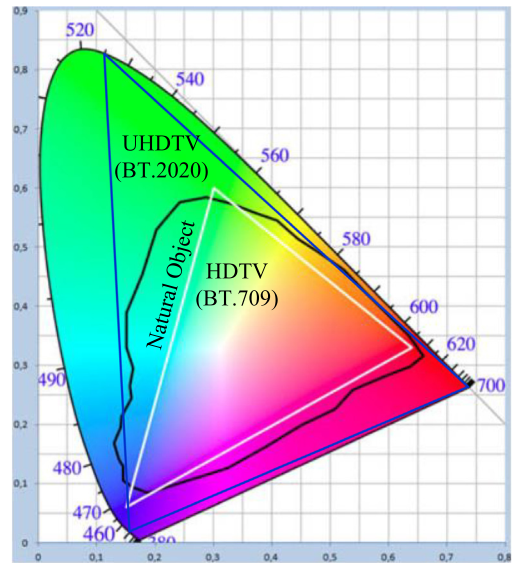


Fig. 2. The pointer's Gamut and BT.2020, BT709 [8].

In view of the importance of HDR visual quality, objective quality measures were also proposed for quanlitative quality assessment of the HDR video content [24,25]. In [24], the authors proposed a novel objective quality assessment method on the basis of sparse-domain representation. In [26], efforts have been devoted to objective evaluation of the performance of HDR video compressed content. The PSNR-based HDR metrics which modify the formulation of Mean Square Error (MSE) for computing higher dynamic range, have also been proposed in the literature. In these metrics, the bias of color transfer function has been removed with tPSNR [27,28]. Another PSNR-based metric is deltaE2000, which assesses the color difference according to the CIE DE 2000 standard [27,28]. In [29], the authors proposed puPSNR as the quality assessment method which is capable of handling luminance from 10^{-5} to 10^{8} nits. Moreover, puSSIM is an extension to Structural Similarity Index Metric (SSIM) which can measure the structural similarity between the HDR source video and the decoded video [30]. There also exist some metrics which are independent of the dynamic range. In particular, HDR-VDP considers physical luminance while the color information is not taken into account [31,32], and HDR-VQM tackles this problem and takes color information into consideration [25].

Although HDR video has many desired advantages, the lack of HDR dispalys and video content have hindered the research of HDR video quality, which plays important roles in quantifying and monitoring the quality of HDR video system. Inspired by this, we conduct subjective study on the quality of HDR compressed videos and construct a database including 12 pristine videos and 144 compressed versions with the corresponding mean opinion score (MOS). In particular, a series of HDR videos with AVS2 encoder [16–18] are compressed and 22 subjects were invited to rate the compressed HDR video with the SONY BVM-X300 Monitor monitor. The subjective evaluations can greatly facilitate the further investigation of the HDR video quality and the development of new objective evaluation methods.

The rest of this paper is organized as follows. Section 2 introduces two formats of HDR video, including PQ as well as HLQ. The HDR source videos and compression configurations are described in Section 3. Section 4 details the subjective test and data analysis. Section 5 concludes this paper.

2. HDR standard

For the production of HDR, the Perceptual Quantizer (PQ) [33] and Hybrid log-Gamma (HLG) [34] are two emerging systems. These two approaches are designed to convert images from the light of the original scene to the representation for the final display, and in this process the artistic intention of the content creator is also desired to be retained. The Recommendation ITU-R BT.2100 [35] has provided the parameters when using the Perceptual Quantization (PQ) and Hybrid Log-Gamma (HLG) approaches. Before introducing these two approaches, there are three terms that must be used to explain the difference between these approaches, and Fig. 3 shows the end-to-end video chain. Generally, the scene light and the display light can be characterized by the three functions:

- OETF: Opto-Electronic Transfer Function, which is the non-linear function coverting linear scene light falling on the sensor into the video digital electronic signal.
- EOTF: Electro-Optical Transfer Function, which is the non-linear function coverting video digital electronic signal into the linear light emitted by the display.
- OOTF: Optical-Optical Transfer Function, which is the function describing the relationship between light scene falling on the image sensor and the light emitted by the display.

As such, a key difference between the HLG and PQ approaches is that HLG is "scene referred", while PQ is "display referred". The "scene referred" approach, HLG, codes the signal from the camera. The video signal contains the image information based on the brightness and color of the original scene at each pixel, and the brightness of the scene and the brightness of the end user display would be compared by the final display. As such, the final display will adjust the OOTF gamma without any metadata that implies the brightness or viewing environment of the final display. On the other hand, PQ is "displayed referred", as it codes the signal intended to be shown on a display. The video signal contains the image information on how bright the scene is. The electronic signal reflects the luminance and color that appears on the final display. When the PQ signal does not match the final display, the signal will need to be adjusted for the specific monitor, and this adjustment is called "display mapping" [35]. In the HLG approach, HDR video signal encoded by the HLG is automatically backward compatible with SDR TVs, but additional metadata to provide the HDR

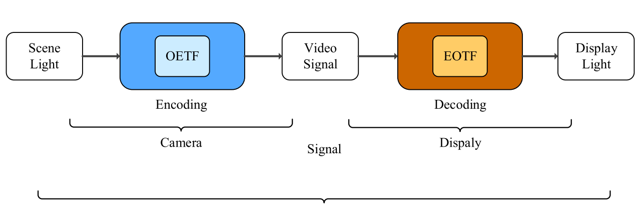


Fig. 3. End-to-end video chain [13].

enhancement is also required through the PQ approach. Fig. 4 shows the video process chain of these two approaches.

3. Video dataset for subjective evaluations

3.1. Motivation

The quality measure peak-signal-to-noise ratio(PSNR), which has been widely adopted in video coding as the objective video quality measure, has been criticized due to its low correlation with the HVS. As such, developing trustful objective quality measures for HDR video is very critical, which requires the establishment of the subjective HDR video quality assessment database. To our best knowledge, there is no available HDR compressed video database with subjective quality ratings. As such, here we build the dataset from the following motivations.

(1) The HDR video dataset can greatly facilitate the development of HDR video quality assessment algorithms. The database can also provide valuable information regarding the critical factors that influence the quality of HDR video.
(2) As the compression artifacts play the dominate role in the HDR video quality, the trust objective measures developed from the database can be further applied in the optimization of the HDR video compression. Previous works also indicate that the optimization of the perceptual quality measures can significantly improve the video coding performance [36,37].
(3) The database also provides a straightforward way to compare the two approaches, PQ and HLG, that are used to generate the HDR video from the natural scene.
From our experimental results, we can observe the differences between the two approaches, which also motivate us to further improve them.

3.2. Description of video content

There are 12 HDR video sequences to be compressed which are captured in 6 scenes. In particular, six HDR sequences are generated by HLG approach and the other six sequences are generated by the PQ approach. The formats of the sequences are summarized in Table 1, where the first column shows the names of the sequences, the second, third, fourth columns describe the frame rate, the spatial resolution and the pixel format, and the fifth column is the duration of the sequence.

The gamut of these sequence is in the ITU-R BT.2020 color gamut. In Fig. 5, we provide the content of HDR video sequences, and the characteristics are as follows:

(1) National Dance
It is a typical sing and dance scene in the evening party. Actors were dressed in bright costumes and quickly rotate and dance on the stage. The major factors that influence the video quality are the accuracy of highly saturated color reproduced after decoding, the performance of motion compensation, as well as the accuracy of the reproduced color generated by pre- and post-processing systems.
(2) The dancing girl under the tree
It is an outdoor scene that a girl is dancing under the tree captured by a fixed camera. The scene is full of ripples and branches with bright details. The resolution and the bright details are the major concerns when performing quality evaluation.
(3) Lady's skin tone
It is a woman dressed in dancing costume under the sunshine captured by a fixed camera. The reflective point on the Women's clothes has high brightness. The skin tone

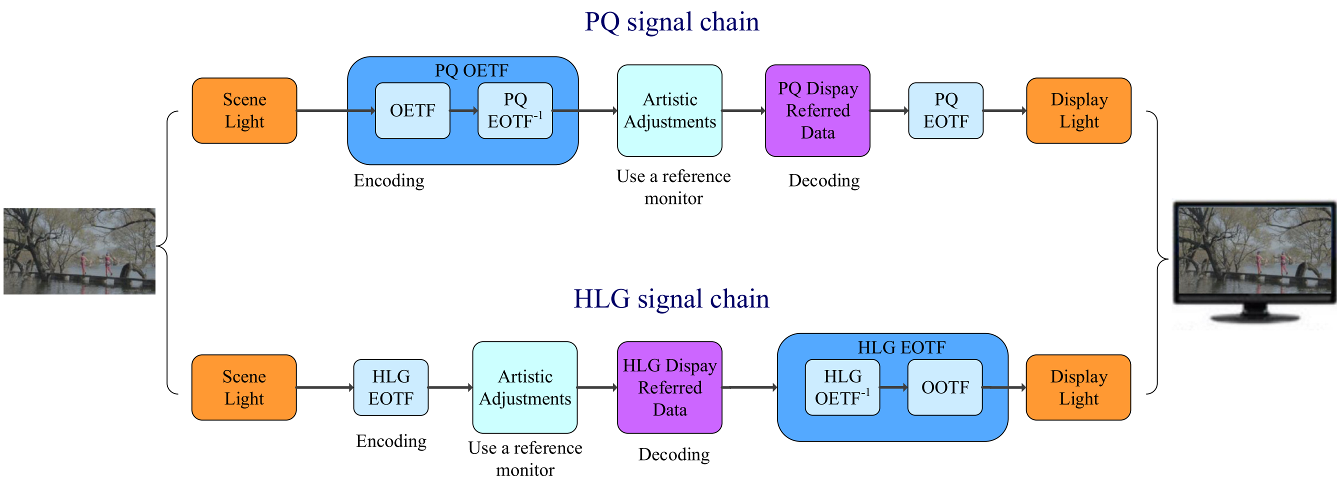


Fig. 4. PQ and HLG process chain [13].

Table 1
The information of the test sequence.

Seq Name	Framerate	Resolution	Pixel format	Duration
National Dance	50	3840 × 2160	YUV420p	10 s
The Dancing girl under the tree	50	3840 × 2160	YUV420p	10 s
Lady's skin tone	50	3840 × 2160	YUV420p	10 s
Night scene in Huangpu River	50	3840 × 2160	YUV420p	10 s
Flowers	50	3840 × 2160	YUV420p	10 s
Bamboo leaves	50	3840 × 2160	YUV420p	10 s

Fig. 5. Illustration of the source HDR video content.

reduction and processing capabilities of the pre and post processing system are the major concerns in quality evaluation.

(4) Night scene in Huangpu River
It is a night scene of the Huangpu River captured by a panning camera. The foreground is the river and cruise ships in the night, while the background is the standing tall buildings. The random motions observed in the water waves, as well as the dark areas are the major concerns.

(5) Flowers
It is a scene of flowers under natural light. Flowers are colorful with high saturation. As such, how the color influences the visual quality can be investigated in this scenario.

(6) Bamboo leaves
It is a scene of bamboo leaves swaying in the breeze captured by a fixed camera. The bright details as well as the tree structures may significantly influence the visual quality.

3.3. Video compression configurations

We use AVS2 Main-10bit profile [38,39] to encode the HDR sequences, and the constant quantization parameter is chosen as the coding strategy. The random access configuration is used, such that frame-level QP can be further adjusted by the encoder according to the frame type and picture order count (POC). Based on the frame rate, the intra-period is set to be 48, and the group of picture

(GOP) size is set to be 8. Other encoding configurations are also included in the appendix. The maximum QP in AVS2 codec standard is 79, and the selected QP values are between [15, 70]. We select 12 values as the encoding QPs from this range, including 15, 20, 25, 30, 35, 39, 43, 48, 54, 61, 65 and 70. The six HDR video scenes, two different approaches including PQ and HLG, and 12 QP values produce 144 compressed videos for subjective evaluations in total.

4. Subjective testing and data analysis

4.1. Subjective testing

Based on the video set, subjective test was conducted to obtain the subjective scores. In total, 22 subjects from China Central Television(CCTV), Academy of Broadcasting Science, Academy Of Broadcasting Planning, Peking University and No.3 Research Institute of China Electronics Technology Group Corporation attended the subjective test. The details of the viewing conditions are shown in Table 2, and the SONY BVM-X300 HDR monitor, SONY KD-75Z9D television, LG OLED65G6P television were used for display. The parameters of these equipments are shown in Tables 3–5. The subjective quality evaluation method called Subjective Assessment of Multimedia Video Quality (SAMVIQ) in Recommendation ITU-R BT.1788 [40] was used, this method can compare the quality of several versions which processed with the same sequence. The range of the score rated by viewer is between 0 and 100, and the majority elements which affect the final score consist of visible distortion, contrast, color reproduction, sharpness.

Table 2
Subjective evaluation conditions.

No	Project	UHDTV Parameter value
1	The watching distance	1.5 times of image height
2	The peak brightness of the screen display	Shown in Tables 3–5
3	Indoor ambient lighting	Low
4	Viewers seating layout	Horizontal direction in the vertical line within $\pm 30^{\circ}$

4.2. Data analysis

To ensure the consistency of the MOS of each subject, the data analysis is further conducted. In particular, the decision of whether each individual score should be reserved or eliminated, is based on the deviation of the individual score and the corresponding mean scores from all the subjects of the test sequence. According to the Recommendation ITU-R BT.1788, the relationship between the video quality and the score of the subjects are regarded as linearity, such that Pearson correlation coefficient is used to eliminate the outlier. In particular, this coefficient can reveal that whether the individual score of one subject is consistent with the mean scores from all subjects in this subjective test. The correlation coefficients between the subjects will be decreased when the consistency is broken by the outlier subject. As such, we compute the correlation coefficient between the *jth* subject and the MOS of a specific video,

$$r_j = \frac{\sum_{i=1}^{n} x_i y_{ji} - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_{ji}\right)}{n}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} y_{ji}^2 - \frac{\left(\sum_{i=1}^{n} y_{ji}\right)^2}{n}\right)}} \tag{1}$$

where r_j is the correlation coefficient between the *jth* viewer and MOS, and the range of j is 1–22, the number of viewers. Moreover, the range of i is 1 to the number of test video n, and x_i is the mean score of all subjects for the video, y_{ji} indicates the score of i_{th} test video of the subject j, and n is the total number of videos to be tested. The computed correlation coefficient can be used to remove the unreliable subjective score of the observer.

In this step, r is a vector consisting of the correlation coefficients of all subjects, $mean(r)$ indicates the average of the correlation of the observers of this video, and $std(r)$ indicates the standard deviation of all the correlations coefficients. Here, the value of Max Correlation Threshold (MCT) is set to be 0.85 based on [40]. After

Table 3
SONY BVM-X300 Monitor(HDR Mode) Parameters.

No	Project		Parameter value	
			Monitor1	Monitor2
1	Display size (diagonal)		30 inches	
2	Display type		OLED	
3	Measured peak brightness		1053 cd/m^2	1036 cd/m^2
4	Measured black level brightness		0.0005 cd/m^2	0.0005 cd/m^2
5	Measured white point coordinates (CIE 1931 xy)		(0.3081, 0.3207)	(0.3082, 0.3183)
6	Base color chromaticity coordinates(CIE 1931 xy)	Red	(0.6788, 0.3187)	(0.6784, 0.3188)
		Green	(0.2507, 0.7047)	(0.2496, 0.7050)
		Blue	(0.1396, 0.0509)	(0.1397, 0.0509)

Table 4
SONY BVM-X300 television(HDR Mode) Parameters.

No	Project		Parameter value	
			Television1	Television2
1	Display size (diagonal)		75 inches	
2	Display type		LCD	
3	Measured peak brightness		1904 cd/m^2	1854 cd/m^2
4	Measured black level brightness		0.001 cd/m^2	0.001 cd/m^2
5	Measured white point coordinates (CIE 1931 xy)		(0.3096, 0.3249)	(0.3108, 0.3266)
6	Base color chromaticity coordinates(CIE 1931 xy)	Red	(0.6784, 0.3132)	(0.6782, 0.3136)
		Green	(0.2795, 0.6602)	(0.2812, 0.6586)
		Blue	(0.1528, 0.0527)	(0.1525, 0.0529)

Table 5
SONY BVM-X300 television Parameters.

No	Project		Parameter value	
			Television1	Television2
1	Display size (diagonal)		65 inches	
2	Display type		OLED	
3	Measured peak brightness		577 cd/m^2	618 cd/m^2
4	Measured black level brightness		0.0005 cd/m^2	0.0005 cd/m^2
5	Measured white point coordinates (CIE 1931 xy)		(0.3049, 0.3111)	(0.3034, 0.3115)
6	Base color chromaticity coordinates(CIE 1931 xy)	Red	(0.6734, 0.3212)	(0.6701, 0.3215)
		Green	(0.2671, 0.6612)	(0.2679, 0.6681)
		Blue	(0.1405, 0.0510)	(0.1410, 0.0513)

removing the unreliable subject, Fig. 6 shows the relationship between the bitrate of HDR video compressed with different QP and MOS. We can observe that the MOS value is increased with the bitrate, and converges to a constant value when the QP is less than a specific value.

5. Objective evaluation

5.1. Correlation between subjective and objective evaluations

Due to the larger dynamic range of HDR video compared to the SDR video, new objective measures have been specifically developed for HDR. In [41,42], the authors showed the difference between some objective measurements and their correlations with perceived video quality, these objective measurements can be further divided into the following categories.

(a) Metrics computed in linear domain:
  – PSNR
  – PSNR-Lx: represent the distortion in the lightness domain of the CIELab color space [43].
  – deltaE100 [27,28]
  – HDR-VQM [25]
(b) Metrics computed in PO-TF domain [27,28]:
  – tPSNR-x: PSNR computed on x component.
(c) Metrics computed using multi-exposure [44]:
  – mPSNR: multi-exposure peak-signal-to-noise ratio.

Here, we use the HDRTools(V0.17) [45] to obtain the objective quality based on these quality measures. The computation process of all these objective can also be found in [43].

The subjective value can be treated as groundtruth to evaluate the accuracy of objective metrics which estimate the perceived

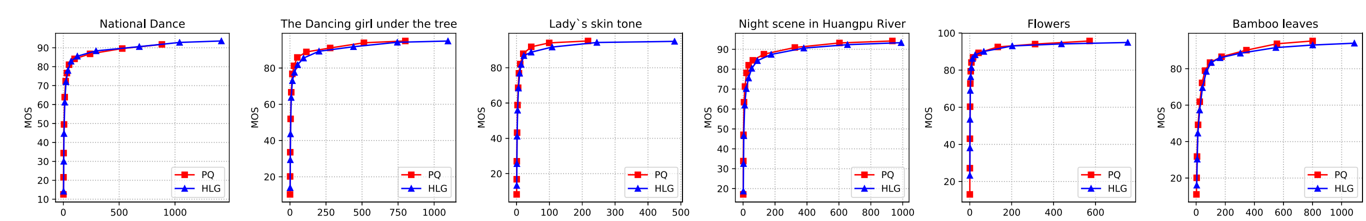


Fig. 6. The relationship between MOS and bitrate.

Table 6
Performance of different objective metrics based on the proposed database.

Metric	PQ			
	PLCC	SROCC	KROCC	RMSE
PSNR	0.8952	0.9155	0.7768	12.3370
mPSNR	0.8050	0.8271	0.6395	16.4206
tPSNR-XYZ	0.8901	0.9104	0.7705	12.6124
tPSNR-X	0.8874	0.9102	0.7768	12.7568
tPSNR-Y	0.8949	0.9158	0.7760	12.3529
tPSNR-Z	0.8858	0.9087	0.7627	12.8457
DelatE100	0.9452	0.9491	0.8088	9.0355
PSNRL100	0.9199	0.9359	0.7836	10.8529
HDR-VQM	0.9196	0.9177	0.7517	10.8718
Metric	**HLG**			
	PLCC	SROCC	KRCC	RMSE
PSNR	0.9137	0.9381	0.8086	10.1997
mPSNR	0.7167	0.7742	0.5892	17.5025
tPSNR-XYZ	0.8017	0.8308	0.6648	14.9989
tPSNR-X	0.7959	0.8352	0.6607	15.1928
tPSNR-Y	0.8139	0.8613	0.6909	14.5804
tPSNR-Z	0.7899	0.8116	0.6450	15.3897
DelatE100	0.8443	0.8880	0.7108	13.4460
PSNRL100	0.9243	0.9495	0.8290	9.5784
HDR-VQM	0.8859	0.8984	0.7162	11.6436

quality. To demonstrate the accuracy and monotonicity of the existing objective measures and the subjective MOS, the Spearman rank order correlation coefficient (SROCC), Pearson linear correlation coefficient (PLCC), Kendallrank-order correlation coefficient (KROCC) and root mean square error (RMSE) have been computed for each metric. The results are shown in Table 6. Results show that PSNR-DeltaE100, PSNR-L100 and HDR-VQM outperform other metrics, with the PLCC and SROCC all above 0.84 and 0.88. PSNR is also a reliable metrics, with a high correlation with subjective scores. In general, mPSNR, tPSNR-XYZ are also promising metrics, with a PLCC and SROCC above 0.7.

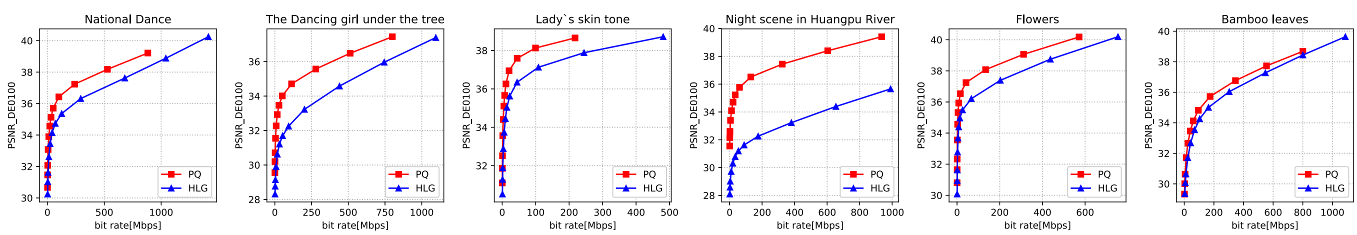
5.2. Comparisons between PQ and HLG approaches

The HDR test videos were generated from natural scenes, using PQ approach and HLG approach, respectively. The differences between two approaches have been discussed in Section 2. Fig. 7 compares the rate-distortion performance of these two approaches, and the relationship between MOS and bitrate were also provided in Fig. 6.

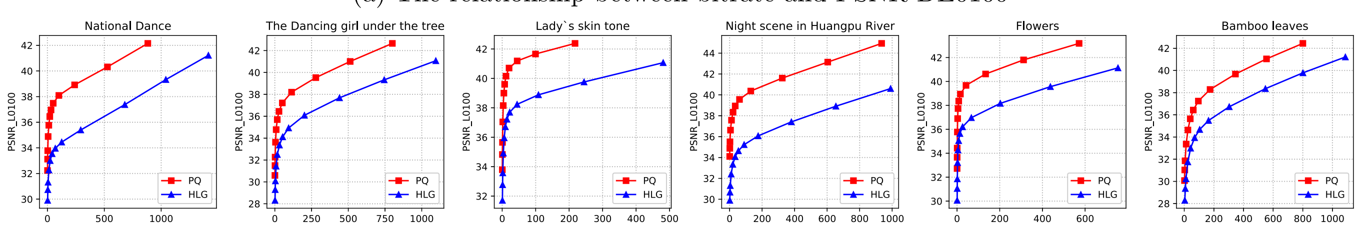
From Fig. 6, we can see that the bitrate value of the video generated by HLG approach is higher than the video generated by PQ approach in general. From Fig. 7,we can see the performance in terms of objective metrics for PQ approach is higher than HLG approach. The reason for this phenomenon is that PQ approach is a constant standard without adaptation, which implies for each input level there is a fixed output luminance value. HLG approach is the "scene referred" approach, implying that the brightness of the scene and the brightness of the end user will be compared.

6. Conclusions

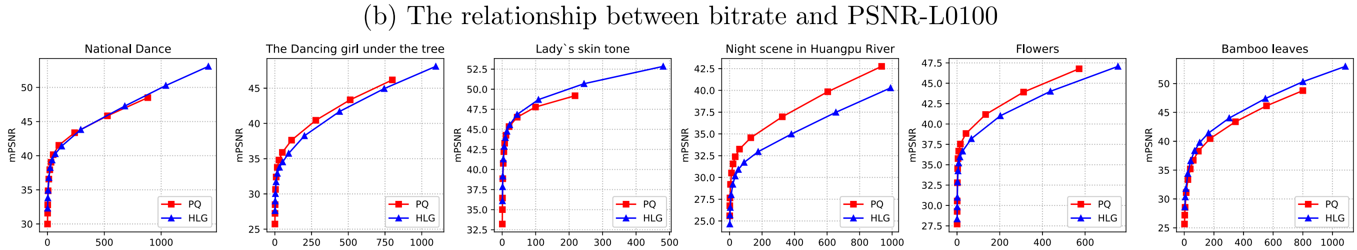
In this paper, we have established the database which contains the subjective ratings of multiple compressed HDR videos. The database is expected to greatly facilitate the HDR visual quality assessment and compression, and play important roles in the future HDR research. In particular, the HDR video content, subjective test environment, procedures, outlier processing method and objective evaluation results have also been provided and analyzed. The subjective and objective results can also exhibit the difference of HDR videos generated by the PQ and HLG approaches, respectively, such that the production chain can be further improved to deliver videos of better quality.



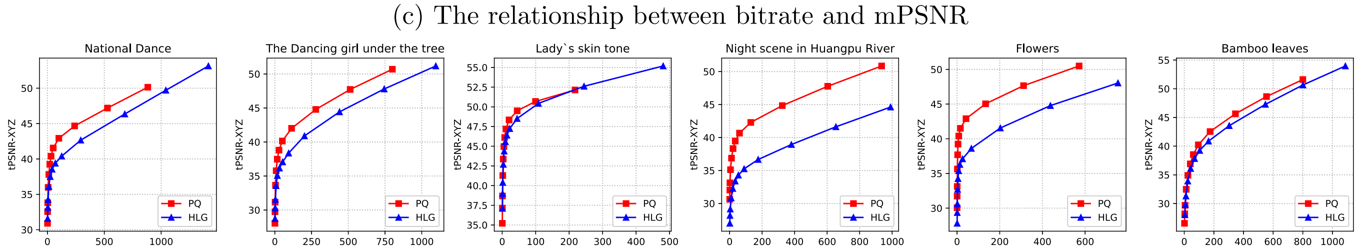
(a) The relationship between bitrate and PSNR-DE0100



(b) The relationship between bitrate and PSNR-L0100



(c) The relationship between bitrate and mPSNR



(d) The relationship between bitrate and tPSNR-XYZ

Fig. 7. The relationship between bitrate and some objective. In each subfigure, the x axis represents bit rate, and the y axis represents objective metric.

Conflict of interest

There is no conflict of interest.

References

[1] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, K. Myszkowski, High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting, Morgan Kaufmann, 2010.
[2] J. Stevens, S.S. Stevens, Brightness function: effects of adaptation, JOSA 53 (3) (1963) 375–385.
[3] T. Kunkel, E. Reinhard, A reassessment of the simultaneous dynamic range of the human visual system, in: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization, ACM, 2010, pp. 17–24.
[4] Wikipedia, candela per unit area, <https://en.wikipedia.org/wiki/Candela_per_square_metre>, online, 2018.
[5] E. François, C. Fogg, Y. He, X. Li, A. Luthra, A. Segall, High dynamic range and wide color gamut video coding in hevc: status and potential future enhancements, IEEE Trans. Circuits Syst. Video Technol. 26 (1) (2016) 63–75.
[6] F. Xiao, J.M. DiCarlo, P.B. Catrysse, B.A. Wandell, High dynamic range imaging of natural scenes, Color and Imaging Conference, vol. 2002, Society for Imaging Science and Technology, 2002, pp. 337–342.
[7] J.A. Ferwerda, S.N. Pattanaik, P. Shirley, D.P. Greenberg, A model of visual adaptation for realistic image synthesis, in: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, ACM, 1996, pp. 249–258.
[8] M. Pointer, The gamut of real surface colours, Color Res. Appl. 5 (3) (1980) 145–155.
[9] I. Recommendation, 709-5, parameter values for the hdtv standards for production and international programme exchange, ITU Radiocommunication.
[10] I. Recommendation, 2020:parameter values for ultra-high definition television systems for production and international programme exchange, International Telecommunication Union, Geneva.
[11] K. Jansen, The pointers gamut-the coverage of real surface colors by rgb color spaces and wide gamut displays, TFT Central, Feb.
[12] A. Luthra, E. Francois, W. Husak, Draft requirements and explorations for hdr/wcg content distribution and storage, ISO/IEC JTC1/SC29/WG11 MPEG2014 N 14510.
[13] T. Borer, A. Cotton, M. Pindoria, S. Thompson, Approaches to high dynamic range video, in: Digital Media Industry & Academic Forum (DMIAF), IEEE, 2016, pp. 71–76.
[14] G.J. Sullivan, J. Ohm, W.-J. Han, T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Trans. Circuits Syst. Video Technol. 22 (12) (2012) 1649–1668.
[15] F. Bossen, B. Bross, K. Suhring, D. Flynn, Hevc complexity and implementation analysis, IEEE Trans. Circuits Syst. Video Technol. 22 (12) (2012) 1685–1696.
[16] S. Ma, S. Wang, W. Gao, Overview of ieee 1857 video coding standard, in: 2013 20th IEEE International Conference on Image Processing (ICIP), IEEE, 2013, pp. 1500–1504.
[17] W. Gao, S. Ma, An overview of AVS2 standard, in: Advanced Video Coding Systems, Springer, 2014, pp. 35–49.
[18] S. Ma, T. Huang, C. Reader, W. Gao, AVS2? Making video coding smarter [standards in a nutshell], IEEE Signal Process. Mag. 32 (2) (2015) 172–183.
[19] L. Yu, F. Yi, J. Dong, C. Zhang, Overview of avs-video: tools, performance and complexity, Proceedings of SPIE - The International Society for Optical Engineering 5960 (5960) (2005) 679–690.
[20] J. Ström, K. Andersson, M. Pettersson, P. Hermansson, J. Samuelsson, A. Segall, J. Zhao, S.-H. Kim, K. Misra, A.M. Tourapis, et al., High quality hdr video compression using HEVC main 10 profile, in: Picture Coding Symposium (PCS), 2016, IEEE, 2016, pp. 1–5.
[21] V. Baroncini, K. Andersson, A. Ramasubramonian, G. Sullivan, Verification test report for hdr/wcg video coding using HEVC main 10 profile, in: JCTVC-X1018, 24th JCT-VC meeting, Geneva, Switzerland, 2016.
[22] Y. Zhang, T. Shen, X. Ji, Y. Zhang, R. Xiong, Q. Dai, Residual highway convolutional neural networks for in-loop filtering in hevc, IEEE Trans. Image Process. A Publ. IEEE Signal Process. Soc. 27 (8) (2018) 3827.
[23] Y. Zhang, X. Ji, H. Wang, Q. Dai, Stereo interleaving video coding with content adaptive image subsampling, IEEE Trans. Circuits Syst. Video Technol. 23 (7) (2013) 1097–1108.
[24] L. Xie, X. Zhang, S. Wang, X. Zhang, S. Ma, Quality assessment of tone-mapped images based on sparse representation, in: 2016 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE, 2016, pp. 2218–2221.
[25] M. Narwaria, M.P. Da Silva, P. Le Callet, HDR-VQM: an objective quality measure for high dynamic range video, Sig. Proces.: Image Commun. 35 (2015) 46–60.
[26] A. Luthra, E. Franois, W. Husak, A. Tourapis, T. Ebrahimi, Exploration experiments for hdr and wide gamut content distribution, ISO/IEC JTC1/SC29/WG11 MPEG2014 N 14549.
[27] M. Rerabek, P. Korshunov, P. Hanhart, T. Ebrahimi, Correlation of subjective scores and objective metrics for hdr video quality assessment, ISO/IEC JTC1/SC29/WG11 MPEG2014/m35273.
[28] P. Hanhart, M. Řeřábek, T. Ebrahimi, Towards high dynamic range extensions of hevc: subjective evaluation of potential coding technologies, Applications of Digital Image Processing XXXVIII, vol. 9599, International Society for Optics and Photonics, 2015, p. 95990G.
[29] T.O. Aydın, R. Mantiuk, H.-P. Seidel, Extending quality metrics to full luminance range images, Human Vision and Electronic Imaging XIII, vol. 6806, International Society for Optics and Photonics, 2008, p. 68060B.
[30] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[31] R. Mantiuk, K.J. Kim, A.G. Rempel, W. Heidrich, HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions, ACM Trans. Graphics (TOG), vol. 30, ACM, 2011, p. 40.
[32] M. Narwaria, R. Mantiuk, M.P. Da Silva, P. Le Callet, HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images, J. Electron. Imaging 24 (1) (2015) 010501.
[33] High dynamic range electro-optical transfer function of mastering reference displays, in: Smpte St, 2015, pp. 1–14.
[34] T. Borer, A. Cotton, A display-independent high dynamic range television system, Smpte Motion Imag. J. 125 (4) (2016) 50–56.
[35] R. Sector, Image parameter values for high dynamic range television for use in production and international programme exchange, International Telecommunication Union.
[36] S. Wang, A. Rehman, Z. Wang, S. Ma, W. Gao, Perceptual video coding based on SSIM-inspired divisive normalization, IEEE Trans. Image Process. 22 (4) (2013) 1418–1429.
[37] S. Wang, A. Rehman, Z. Wang, S. Ma, W. Gao, Ssim-motivated rate-distortion optimization for video coding, IEEE Trans. Circuits Syst. Video Technol. 22 (4) (2012) 516–529.
[38] Y. Zhao, W. Zeng, E. Chai, X. Zheng, J. Zheng, On a 10-bit coding profile for AVS2 standard, in: 2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW), IEEE, 2014, pp. 1–5.
[39] FaleiLuo, xavs2, https://github.com/pkuvcl/xavs2, online, 2018.
[40] B. ITU-Recommendation, Methodology for the subjective assessment of video quality in multimedia applications, International Telecommunications Union, Radio-communication Assembly.
[41] P. Hanhart, M. ebek, T. Ebrahimi, Subjective and objective evaluation of HDR video coding technologies, in: Eighth International Conference on Quality of Multimedia Experience, 2016, pp. 1–6.
[42] R. Mukherjee, K. Debattista, T. Bashford-Rogers, P. Vangorp, R. Mantiuk, M. Bessa, B. Waterfield, A. Chalmers, Objective and subjective evaluation of high dynamic range video compression, Signal Process. Image Commun. 47 (C) (2016) 426–437.
[43] A. Segall, V. Baroncini, J. Boyce, J. Chen, T. Suzuki, Joint call for proposals on video compression with capability beyond hevc, Joint Video Exploration Team (JVET) of ITU-T SG 16 2017.
[44] J. Munkberg, P. Clarberg, J. Hasselgren, T. Akenine-Möller, High dynamic range texture compression for graphics hardware, ACM Trans. Graphics (TOG), vol. 25, ACM, 2006, pp. 698–706.
[45] A.M. Tourapis, Hdrtools: Software status, ISO/IEC JTC1/SC29/WG11 MPEG2016 M 38422.