https://doi.org/10.2352/ISSN.2470-1173.2017.12.IQSP-233

Development and Image Quality Evaluation of 8K High Dynamic Range Cameras with Hybrid Log-Gamma

Ryohei Funatsu, Kazuya Kitamura, Toshio Yasue, Daiichi Koide, Hiroshi Shimamoto; NHK Science & Technology Research Laboratories; Tokyo, Japan

Abstract

The 8K ultra-high-definition television (UHDTV) is a next generation television system with a high realistic sensation. High dynamic range (HDR) is a new standard for television systems, defined in Recommendation ITU-R BT. 2100. Hybrid log-gamma (HLG) is one of the HDR standards, jointly suggested by NHK (Japan Broadcasting Corporation) and BBC (British Broadcasting Corporation), and is highly compatible with the conventional standard dynamic range system. Although a "full-featured" 8K camera with HLG has already been developed, most existing 8K cameras do not comply with the HLG standard. In this paper, we describe a method for adapting existing 8K cameras to HLG, thus enhancing their dynamic range. Based on subjective image quality evaluation results, we propose a guideline to choose the dynamic range setting for each shooting scene, considering the noise performance of the particular used 8K camera.

Introduction

The 8K ultra-high-definition television (UHDTV) is a next generation high sensation television system [1]. An 8K test broadcast via satellite has been launched on August 1, 2016 in Japan [2], and a regular broadcasting service will be started in 2018. With the upcoming launch of the broadcast service, research and development of practical 8K cameras has been accelerated, to accommodate the increased demand for 8K UHDTV content.

High dynamic range (HDR) was defined as an image standard in ITU-R Rec. BT.2100 [3]. Hybrid log-gamma (HLG) is one of the HDR standards, jointly suggested by NHK (Japan Broadcasting Corporation) and the BBC (British Broadcasting Corporation); it is highly compatible with the conventional standard dynamic range (SDR) system standardized in ITU-R Rec. 709 [4]. The above-mentioned 8K test broadcast in Japan includes the HDR standards. However, most existing 8K cameras—except for the newly developed full-featured 8K camera[5]—were developed prior to the establishment of the HLG standard, and they consequently only meet the requirements defined for SDR systems. Therefore, the HLG processing method and its effects on the image quality of conventional 8K cameras have not been sufficiently discussed.

In this paper, we describe a method for adapting existing 8K cameras to the HLG standard, and subsequently evaluate the image quality of the HLG images captured by such cameras. Based on the evaluation results, we propose a guideline to decide on the dynamic range setting for each shooting scene, taking into consideration the noise performance of the used 8K camera.

High-dynamic-range Standard Overview

High dynamic range television can deliver images with high brightness accuracy on brighter displays. Two HDR methods have been standardized in ITU-R Rec. BT.2100. One methods is the so-called perceptual quantization (PQ) and the other is the hybrid log-gamma (HLG) method. The PQ specification defines a very wide range of display brightness levels, using an electro-optical transfer function (EOTF) finely tuned to match the human visual system. In contrast, the HLG specification defines an opto-electronic transfer function (OETF) based on the scene linear light intensity, as is also done in conventional SDR systems. Therefore, HLG is more suitable for HDR/SDR simultaneous broadcasting and live productions such as sport contents.

The parameter values for HLG are listed in Table 1. In this table, E is a signal—for each color component—proportional to the scene linear light, and E' is the resulting non-linear signal; the opto-optical transfer function (OOTF) maps the relative scene linear light level onto the display linear light level; Y_S is the normalized linear scene luminance, L_W and L_B are the nominal peak luminance and black luminance of the display in units of cd/m^2, respectively, and γ is the system gamma value derived from the peak luminance L_W.

Figure 1 shows the OETFs of the HLG and conventional SDR systems. When $E \leq 1$, the HLG OETF curve is compatible with the SDR gamma curve. In contrast, when $E > 1$, the HLG OETF becomes logarithmic, and larger scene linear light intensities can

Table 1: Parameter values for the HLG system

OETF	$E' = \begin{cases} \sqrt{E}/2 & 0 \leq E \leq 1 \\ a \cdot \ln(E-b)+c & 1 < E \end{cases}$ a = 0.17883277, b = 0.28466892, c = 0.55991073	
EOTF	$F_D = OOTF[E] = OOTF[OETF^{-1}[E']]$ $= \frac{L_W - L_B}{12^\gamma} Y_S^{\gamma-1} E + L_B$ $\gamma = \begin{cases} 1.2 & L_W \leq 1000 \\ 1.2 + 0.42\log_{10}(L_W/1000) & L_W > 1000 \end{cases}$	
Signal Values	Video Data (E')	4–1019 (10 bits) 16–4079 (12 bits)
	Black (E = E' = 0)	64 (10 bits) 256 (12 bits)
	Nominal Peak (E = 12, E' = 1)	940 (10 bits) 3760 (12 bits)

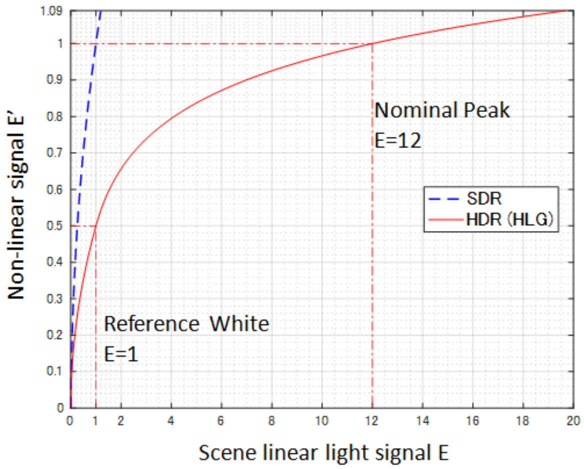

Figure 1. OETFs of HLG and conventional SDR (Rec. ITU-R BT.709) system.

therefore be expressed. $E = 1$ ($E' = 0.5$) is the reference white level, which is equal to the signal level of 100 % of the conventional SDR television system. The maximum range of display brightness in the HLG system is approximately 20 times that of the SDR system.

In this paper, "dynamic range" is defined as the percentage of the achievable maximum scene linear light level to the reference white level ($E = 1$). For example, a camera with a "1200 % dynamic range setting" can shoot scenes up to a linear light signal of $E = 12$. In this case, signals with $E > 12$ will be clipped, and the maximum non-linear signal is $E' = 1$.

Adapting existing 8K cameras to HLG

Overview of 8K cameras

Most 8K cameras had been developed before the standardization of HDR television systems. Therefore, we adapted

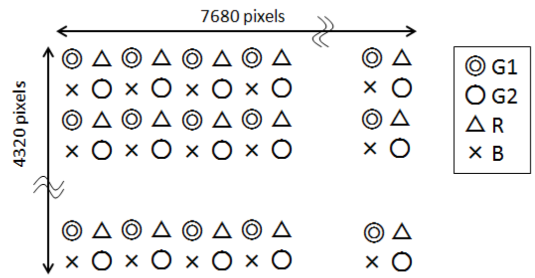

Figure 2. Sampling structure of Bayer pattern video signal.

those cameras to HDR standard, by changing their signal processing circuitry. Table 2 lists the existing 8K cameras [5–8] that were adapted to HLG.

A Full-featured camera is the newest type of cameras meeting the maximum parameter requirements of UHDTV, as standardized in Rec. ITU-R BT.2020 and BT.2100, with RGB 4:4:4, 120 fps, wide color gamut, 12-bit depth, and HDR. This camera employs three 33-Megapixel, 120-fps CMOS image sensors, and a wide color gamut color separation prism.

Single-chip cameras were developed in 2015 as practical 8K cameras with high sensitivity and high mobility. Those cameras are similar in appearance to conventional broadcast cameras, and are used in 8K program production for the 8K test broadcast. The color sampling structure of the single-chip camera is Bayer pattern, as shown in Figure 2; color interpolation signal processing is required to generate a video signal for 8K displays.

A silent and high-sensitivity camera was developed in 2013 for theater content production requiring silence and high sensitivity. Although this camera has three 33-Megapixel CMOS image sensors, the color sampling structure of the output video is still Bayer pattern, because of a four-pixel addition scheme used to achieve higher sensitivities.

A cube-type, ultra-compact camera is the smallest 8K camera, which was developed in 2013. It weighs only 2 kg, whereas the full-featured 8K camera weighs over 50kg.

Table 2: Existing 8K cameras adapted to HLG

Type	Full-featured [5]	Single-Chip [6]	Silent and high-sensitivity [7]	Cube-type [8]
Imaging Method	3-chip	Single-chip	3-chip	Single-chip
Optical Size	1.7 in	Super 35	2.5 in	Super 35
Sampling Structure	RGB 4:4:4	Bayer pattern	Bayer pattern	Bayer pattern
Frame Rate	120 fps	60 fps	60 fps	120 fps
Development Year	2016	2015	2013	2013
Appearance				

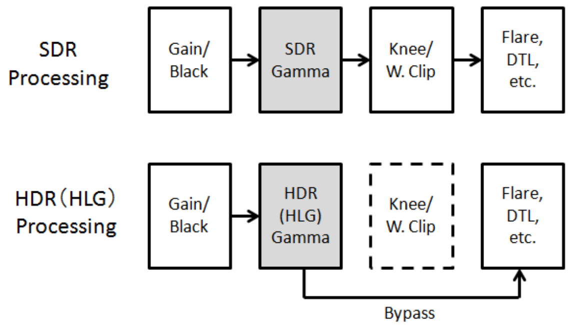

Figure 3. Difference in the signal processing flow between conventional SDR cameras and 8K HDR cameras.

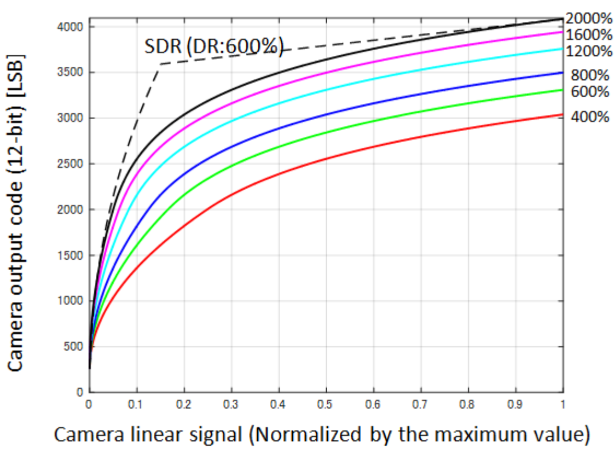

Figure 4. HLG OETF curves for the dynamic range settings of 400 %, 600 %, 800 %, 1200 %, 1600 %, and 2000 % used in 8K HDR cameras. The dotted line illustrates a conventional OETF curve for the 600 % dynamic range, using knee and white clip processing.

Method for adapting existing 8K cameras to HLG:

Existing 8K cameras can be made to adopt the HLG OETF by modifying the conventional gamma correction circuit of SDR cameras with the addition of HLG OETF. Figure 3 shows the difference in the signal processing flow between conventional SDR cameras and HLG cameras. In HDR processing, the knee and white clip circuits used for dynamic range extension are bypassed, because the dynamic range can be extended by the HLG OETF curve itself.

Figure 4 shows the HLG OETF curves used in 8K HDR cameras. The horizontal axis is the linear signal level, normalized by its maximum value. The maximum level is usually determined by the saturation level of the image sensor, which means that the horizontal axis is equivalent to the full range of the image sensor. The vertical axis is the actual camera output signal code after HLG processing. The camera has a variable HLG setting, with dynamic range values of 400 %, 600 %, 800 %, 1200 %, 1600 %, and 2000 %. An SDR curve with a dynamic range of 600 % using knee and white clip circuits is also shown, for reference. The OETF curve for the dynamic range *DR* [%] is given by adapting formula (1) to the OETF in Table 1.

$$E = \frac{DR}{100} I , \tag{1}$$

where *I* is the linear signal level normalized by its maximum level.

Although the nominal peak level defined in the standard is equivalent to a 1200 % dynamic range, some upper and lower dynamic range settings relative to 1200 % are applied to the camera by considering the “super-white” level and the signal-to-noise ratio (SNR). The signal level range from 100 % to 109 %, which is also called “super white,” can be used for the video signal. The camera’s dynamic range of approximately 2000 % corresponds to 109%, which means that the camera system has indeed a dynamic range up to 2000 %.

HDR captured image

Figure 5 shows lamp images captured with a full-featured 8K camera using each one of the possible dynamic range settings. As shown, images with a higher dynamic range setting have enhanced reproducibility of the brightness gradation without saturation.

SNR and Sensitivity Measurement

Measurement results

The SNR and sensitivity are very important to evaluate the image quality of an HDR video. The SNRs and sensitivities of all 8K HDR cameras were measured at 60-fps operation. The SNR and sensitivity results obtained for each dynamic range setting are shown in Figures 6 and 7, respectively. The sensitivities were measured with different light sources suitable for each of 8K cameras, because the color reproduction of each 8K camera was designed for a specific light source, and those light sources were used in past sensitivity measurements. A 3200 K Tungsten light source was used for the single-chip camera and silent and high-sensitivity camera, a 5600K LED light source was used for the cube-type camera, and an artificial solar light was used for the full-

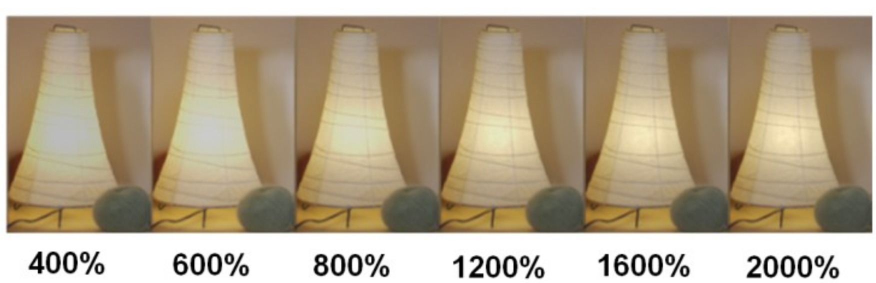

Figure 5. Lamp images captured by a full-featured camera with different dynamic range settings.

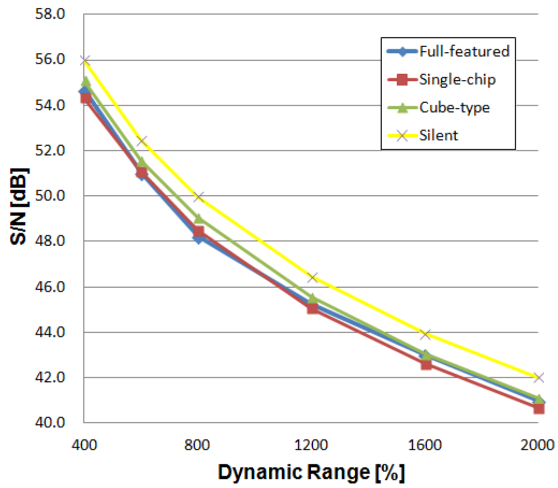

Figure 6. SNRs measured at the different dynamic range settings.

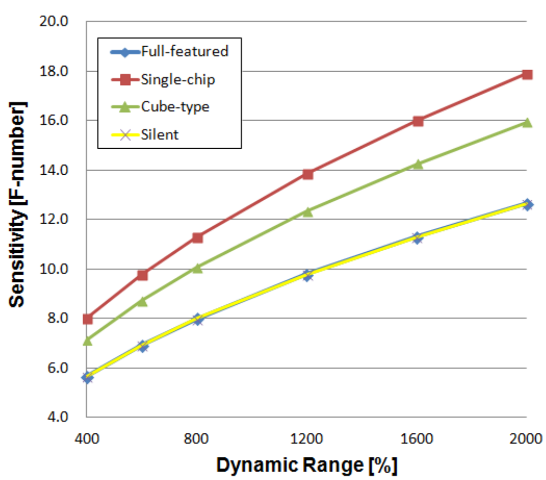

Figure 7. Sensitivities measured at the different dynamic range settings.

featured camera, respectively.

The SNRs were measured using a normalized linear scene luminance (Y_s) signal with an 8K full-resolution (7680×4320 pixels). The SNR can be calculated by the following equation,

$$SNR = 20\log_{10}\left(\frac{S_Y}{N_Y}\right), \tag{2}$$

where S_Y is the linear signal level corresponding to a scene linear light level of $E = 1$, and N_Y is the dark noise level calculated from the standard deviation of a dark image. 8K cameras with Bayer pattern color sampling structures require color interpolation to generate a full-resolution 8K image. Therefore, the following 3×3 linear color interpolation filters were used in this measurement to generate the full-resolution 8K image.

$$f_G = \begin{bmatrix} 0 & 0.25 & 0 \\ 0.25 & 1 & 0.25 \\ 0 & 0.25 & 0 \end{bmatrix}, \tag{3}$$

$$f_{B,R} = \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}. \tag{4}$$

In (3) and (4), f_G is the color interpolation filter for the green signal, and $f_{B,R}$ is the equivalent filter for the red and blue signals. The full-resolution 8K color images were converted to a luminance signal by the following equation:

$$Y_S = 0.2627R_S + 0.6780G_S + 0.0593B_S, \tag{5}$$

where R_S, G_S, and B_S denote the scene linear light signals for the red, green, and blue color components, respectively.

Sensitivity was measured using a gray scale chart with a white patch of 83.0 % reflection ratio. The sensitivity is defined by the F-number that brings the signal level of the white patch to 50 % ($E' = 0.5$) when the patch is illuminated with a 2000-lux light intensity. The sensitivities of the 120-fps cameras (the full-featured and cube-type cameras) were converted to 60-fps operation.

The obtained results show that the SNR decreases for higher dynamic ranges because the reference white level becomes closer to a dark level. In contrast, the sensitivity of higher dynamic ranges increases for the same reason. For example, a dynamic range two times higher decreases the SNR by -6 dB.

SNR and sensitivity conversion

In general, if SNR and sensitivity measurements are obtained for a certain dynamic range setting, the SNR and sensitivity values for other dynamic range settings can be derived from the known measurements. Designating the SNR and sensitivity measurements at the dynamic range setting D_1 [%] by SNR_1 and F_1, the SNR and sensitivity values at the D_2 [%] setting (SNR_2 and F_2) can be calculated as follows:

$$SNR_2 = SNR_1 + 20\log_{10}\left(\frac{D_1}{D_2}\right) \tag{6}$$

$$F_2 = F_1\sqrt{\frac{D_2}{D_1}} \tag{7}$$

Image Quality Evaluation

Overview

Generally, if the noise performance of HDR video is good enough as an 8K HDR content, a higher dynamic range setting is preferred. Therefore, we implemented a subjective assessment to evaluate the noise level of HDR video acceptable in 8K contents. Based on the obtained results, we propose a guideline to choose the dynamic range setting, taking into consideration the noise performance of the particular used 8K HDR camera.

Subjective noise assessment setup

The overall setup for the mentioned subjective noise assessment is illustrated in Figure 8. It consists of a 4K HDR organic light-emitting diode (OLED) display and a 4K uncompressed recorder. The viewing distance from the 4K display is fixed at 1.5H (H is the display height), which is equivalent to 0.75 H for an 8K display. The viewing distance of 0.75H for 8K

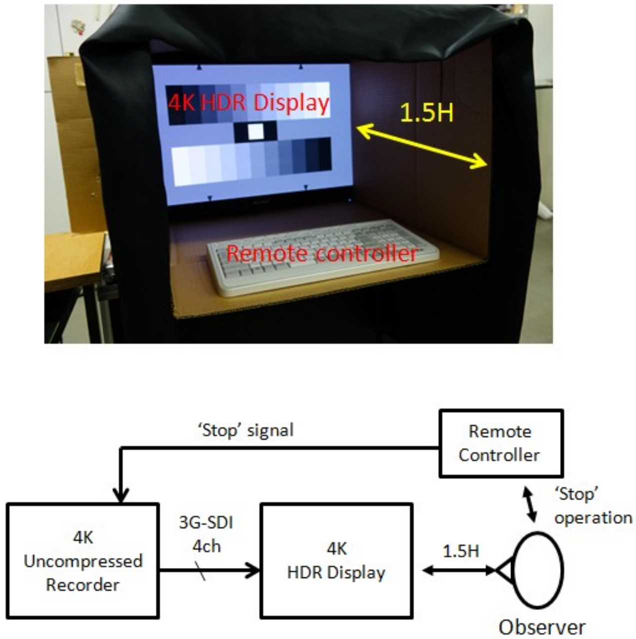

Figure 8. Setup for the subjective HDR noise assessment.

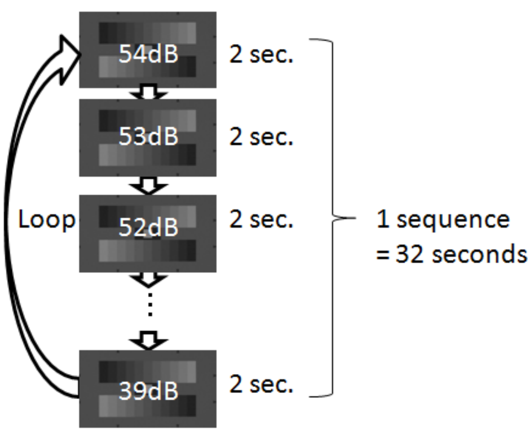

Figure 9. Loop playback video sequence for subjective noise assessment. The sequence is 32 seconds long, and the SNR degrades 1 dB every two seconds. The video is monochrome, without color components, and the noise is added to the luminance (Y) signal.

displays corresponds to the viewing angle of 100 degrees that maximizes the sense of presence, according to past assessments [9].

Observers determine their "acceptable" noise level by looking at a loop playback video sequence of a gray scale chart shot by the single-chip 8K camera [6]. The camera is fixed and the gray scale chart does not move. The video SNR changes from 54 dB to 39 dB, degrading 1 dB every two seconds, as shown in Figure 9. The video is monochrome, without color components, and the noise is added only to the luminance (Y) signal. The observer stops the video when the noise of the video reaches an "unacceptable" level. For example, when an observer stops the video at a 45 dB SNR it means that her or his acceptable SNR is 46 dB. Every observer repeats the same assessment five times; the five obtained results of each observer are averaged.

Table 3: Subjective assessment setup specifications

Display	4K OLED (Sony BVM-X300)
Display size	29.5 in (663.6 mm × 349.9 mm)
Peak luminance	1000 cd/m^2
System gamma	1.2
Viewing distance	1.5 H (H: display height)
Video format	2160 / 59.94P, YCbCr 4:2:2, 10-bit, Uncompressed
Contents	Gray-scale chart captured by a single-chip camera.
Maximum non-linear signal level	E' = 0.5, at the top of gray scale
Noise level	39–54 dB (Y signal)
Observers	16 video specialists
Environment	Dark room
Assessment	"Acceptable" noise level as an 8K video content.

The specifications of the subjective noise assessment setup are listed in Table 3. The peak luminance of the display is 1000 cd/m^2, and the system gamma is set at 1.2, which complies with the HLG EOTF standard shown in Table 1. The maximum non-linear signal level at the top step of the gray scale is $E' = 0.5$, and the video sequence does not contain signals brighter than the reference white level. It is assumed that when the noise level at a lower signal level is acceptable, the same noise level is also acceptable with higher signal levels, because the contrast perception in low luminance conditions is higher than in brighter conditions. Sixteen observers are used; all of them are video specialists. The observation environment is a dark room, and the display is covered with a shading fabric, as shown in Figure 8.

Subjective noise assessment results

The acceptable SNR levels obtained from the 16 video specialists are shown in Figure 10. In this figure, the horizontal axis is the individual observer's ID. Each dot shows the mean acceptable SNR value obtained from that observer's five trials; the error bar indicates the corresponding standard deviation. The global average of the mean acceptable SNRs for all observers is 48.2 dB. Figure 11 shows the histogram of all the mean acceptable SNRs. The standard deviation in this histogram is approximately 1.9 dB.

The results of a questionnaire about the attention points in the assessment reveal that all observers focused on the dark areas in

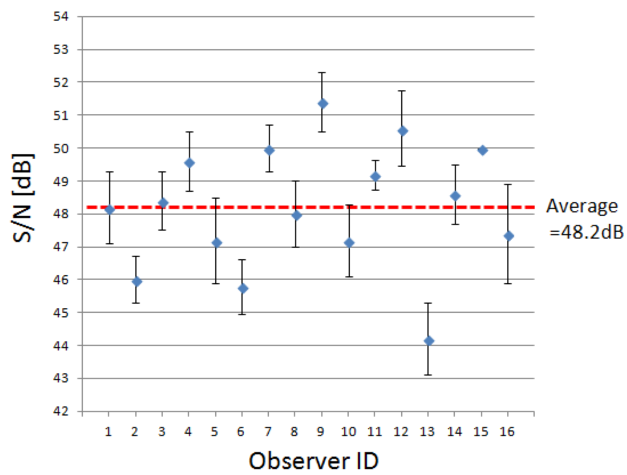

Figure 10. Acceptable SNR results obtained from the 16 video specialists. The horizontal axis is the observer's ID. Each dot is the mean acceptable SNR value obtained from the five trials of an observer, and the error bar is the corresponding standard deviation. The global average of the mean acceptable SNRs is 48.2 dB.

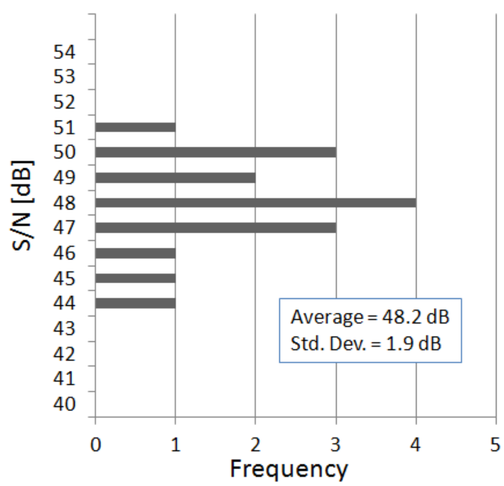

Figure 11. Histogram of the 16 mean acceptable SNRs. The standard deviation of the histogram is approximately 1.9 dB.

the gray scale chart. In particular, observers with a higher acceptable SNR tended to pay attention to only one dark location. In contrast, observers with lower acceptable SNR values tended to look both at the entire chart and at specific dark locations alternately. This indicates that the acceptable SNR values will change with image brightness.

Real-life shooting scenes will generally contain more complex textures and movements than the video used in this assessment. Moreover, the actual environmental illuminance on the screen is higher than that used on this assessment, and the preferred viewing distance from the 8K display is larger than 0.75H [10]. Therefore, the acceptable SNR requirements will be relaxed in realistic conditions; we therefore consider that the average value of the acceptable SNR values obtained in this assessment can be used as a practical benchmark in realistic HDR productions.

HDR production guideline

As mentioned, one of the objectives of this study was the proposal of a guideline to choose the dynamic range setting, considering the noise performance of the particular used 8K HDR camera. The proposed procedure can be described as follows:

First, we set the illuminance (L_S) and F-number (F_S) for the shooting scene, and choose the preferred SNR (SNR_S). Next, we convert the F-number at the shooting scene to the reference 2000 lux illuminance using the following equation:

$$F_{2000} = F_S \sqrt{\frac{2000}{L_S}} \,. \tag{8}$$

From the sensitivity characteristics of the HDR camera, we then determine the candidate dynamic range settings with a higher sensitivity than F_{2000}. Finally, from this set of candidate dynamic range settings we select one with a higher SNR than SNR_S. If there are no dynamic range settings meeting the SNR condition, we need to increase the scene illuminance L_S, select a lower F-number, or set the preferred SNR to a lower value, thus relaxing the requirement.

For example, let us consider shooting a 500 lux illuminance scene with an F-number of 4.0, using the single-chip camera in Table 2. The preferred SNR is assumed to be 48.2 dB, in accordance with the obtained subjective assessment results. In this case, F_{2000} is 8.0, and all dynamic range settings from 400 % to 2000 % belong to the set of candidate dynamic settings (see Figure 7). Using this result, and from Figure 6, we conclude that the dynamic range settings from 400 % to 800 % can be chosen.

Based on the subjective assessment results, we recommend a preferred SNR of at least 48.2 dB, when the shooting scene contains many objects with different brightness and various lighting conditions.

Conclusion

We developed 8K HDR cameras using the HLG standard, based on Rec. ITU-R BT.2100. First, we illustrated the method to adapt existing 8K cameras to HLG, in order to utilize the native dynamic range of 8K image sensors efficiently. Next, we measured the SNRs and sensitivities of the 8K HDR cameras. We then implemented a subjective, observer-based, experiment to assess the SNR levels required for 8K contents. Finally we proposed a guideline to decide the appropriate dynamic range setting for a given camera and shooting scene.

The HDR standard of Rec. ITU-R BT.2100 assumes that the minimum peak luminance of an HDR display is 1000 cd/m^2, which corresponds to the 1200 % dynamic range setting in HLG. However, the measured SNRs of 8K HDR cameras at the 1200 % dynamic range are lower than the averaged acceptable SNR (48.2 dB). Therefore, the SNR at the higher dynamic range settings should be improved, by improving the image sensor's dynamic range and the used signal processing techniques.

We hope this work contributes to support the choice of an appropriate dynamic range setting for each shooting scene when an existing 8K camera is used for HDR acquisition. Moreover, this work is expected to serve as a basis for discussion of the required noise performances of HDR 8K cameras.

References

[1] Recommendation ITU-R BT.2020, "Parameter values for ultra-high definition television systems for production and international programme exchange."

[2] Y. Nishida et al., "Launching 8K/4K High Dynamic Range Television (HDR-TV) Broadcasting," 2016 NAB BEC proceedings, pp.259-265, 2016.

[3] Recommendation ITU-R BT.2100, "Image parameter values for high dynamic range television for use in production and international programme exchange."

[4] Recommendation ITU-R BT.709, "Parameter values for the HDTV standards for production and international programme exchange."

[5] K. Kitamura et al., "Full-Specification 8K Camera System," 2016 NAB BEC proceedings, pp.266-271, 2016.

[6] Y. Miki et al., "Ready for 8K UHDTV broadcasting in Japan, " Proc. IBC 2015 Conference, Amsterdam Netherlands, 2015.

[7] T. Nakamura et al., "Development of a Super Hi-Vision camera toward shooting theater contents," Proc. ITE Winter Annual Convention, 11-12, 2013.

[8] H. Shimamoto et al. "A Compact 120 Frames/sec UHDTV2 Camera with35mm PL Mount Lens," SMPTE Motion Imaging J. vol. 123, no.4, 2014, pp.21-28, 2014.

[9] K. Masaoka et al., "Contrast Effect in Evaluating the Sense of Presence for Wide Displays," J. SID, 14, pp.785-791, 2006.

[10] Y. Kusakabe et al., "Subjective evaluations of preferred viewing distance and psychophysical effects of extremely high resolution images using Super Hi-Vision 85-inch LCD, " ITE Technical Report vol. 36, no.9, pp.245-250, 2012.

Author Biography

Ryohei Funatsu joined NHK in 2002. Since 2004, he has been engaged in research of ultrahigh-definition TV camera systems at NHK Science and Technology Research Laboratories. His research interests include the development of high-resolution image sensor and the high-performance signal processing. Funatsu received B.E. and M.E. degrees in electrical and electronic engineering from Tokyo Institute of Technology in 2000 and 2002, respectively. He is a member of the Institute of Image Information and Television Engineers of Japan (ITE).