A NEWLY-DEVELOPED 8K/HLG BROADCASTING CAMERA SYSTEM WITH THREE CMOS IMAGE SENSORS

Tsuyoshi Sakiyama, Takayuki Yamashita,

Toshihiro Hanatani, and Seiji Mitsuhashi

NHK (Japan Broadcasting Corporation), JAPAN

ABSTRACT

To date, we have developed various types of 8K cameras while considering the trade-off between cost and performance at each period of development. Our first 8K camera employed four 4K image sensors. This was followed by a second model, which used an 8K single-chip colour image sensor with the Super 35 optical format, thus offering compatibility with digital cinema lenses. These cameras have been used effectively in a wide range of applications, such as covering and recording international sport events and entertainment programs as well as producing programs for 4K/8K test broadcasting, beginning in 2016. In preparation for both the 4K/8K broadcasting starting in 2018 as well as the 2020 Olympic Games in Tokyo, we have reconsidered which parameters should be seen as the most important ones for 8K broadcast cameras, and we have also reviewed the appropriate elemental technologies for these cameras. In light of the technological advancements that have been made, we decided to develop a three-chip 8K camera with 5/4-in image sensors. The camera we developed supports both high dynamic ranges and high frame rates, and we succeeded in reducing its size to that of an HDTV camera. Both its signal-to-noise ratio and its sensitivity achieved the best performances of all 8K cameras. This paper discusses both the parameters and the elemental technologies of all of the 8K broadcast cameras that we have examined. It then continues to describe the characteristics of a new camera that we have developed and demonstrates its superior practical performance.

INTRODUCTION

In August of 2016, we began test broadcasting based on the “8K Super Hi-Vision” format (referred to as “8K” hereafter) via a broadcasting satellite. Starting with broadcasts from the 2016 Summer Olympics in Rio de Janeiro, a variety of fascinating events, unique to 8K, were broadcast each day.

In the course of developing various cameras for handling the production of 8K programs in a wide range of genres, we have been developing both cameras that can output live streams as well as cameras with built-in recording devices that can handle on-location filming while also emphasising mobility. In addition to achieving the results of our efforts to date, we switched from using the dual green (DG) format [1] which is the same as Bayer pixel arrangement to using the YCbCr422 format in pursuit of higher picture quality. Furthermore, we converted the optical design from the single-chip style to the three-chip style, and we also succeeded in raising the sensitivity by increasing the camera's light use efficiency.

8K FORMAT

8K signal interfaces

There are 16 times the number of pixels in 8K video signals than in conventional HDTV(2K) signals, and this represents 32 times the amount of information in those signals at 60 fps (the standard frame rate). In order to realise this rate with the hardware that was currently available, we designed and utilized the Dual Green (DG) format, which could reduce the amount of information transmitted. The bit-rate of DG is 24 Gbps, and this interface bundles eight links for the 3-Gbps serial digital interface (3G-SDI) [2].

We considered both improved picture quality and signal compatibility with 4K devices (whose rapid development is remarkable) and concluded that it was necessary to convert 8K devices to the colour difference signals of Y, Cb and Cr and subsampled signals into 4:2:2 ratio as well. However, since it is also necessary to bundle 16 links for the 3G-SDI, it is not realistic to establish this as the mainstream standard for future development.

A mode in which both four links for the 12G-SDI are bundled together and an 8K-422 signal is transmitted is described in SMPTE ST2082-12 [3].

High dynamic range with Hybrid log-gamma

The type of high dynamic range (HDR) prescribed for the 8K test broadcasts in Japan is Hybrid log-gamma (HLG) which is standardised in Recommendation ITU-R BT.2100 [4]. Therefore, 8K production devices are being developed to be compatible with HLG. HLG is characterised both by having the same γ curve as standard dynamic range (SDR) has in the region where the video level is less than 50% and by treating the brightness of the object as the relative video level. Thus it is highly compatible with conventional SDR equipment, so HLG is handled easier by engineers than Perceptual Quantisation (PQ). Given the above, this camera system is also compatible with HLG mode.

Wide Colour Gamut (WCG)

In the UHDTV standardised in ITU-R BT. 2020 (Rec. 2020), it is overwhelmingly larger than the conventional 709 colour gamut and the P3 colour gamut for D-cinema. In single-chip colour cameras, CFA changes and linear matrices are used to realise the correspondence. On the other hand, we have completed examining the design of a colour separation prism for the three-chip optical system with spectral characteristics covering WCG in Rec. 2020. Further details of the design are described in reference. [5]

High Frame Rate(HFR)

Features of UHDTV includes not only ultra-high definition, but also higher frame rate. From the results of the past research [6], it is also known that the improvement of the subjective evaluation value with respect to the frame frequency is saturated at 240 Hz, but the increase rate at the corresponding 100 or 120 Hz which is twice the conventional rate is the largest. Considering the future of the broadcasting system, this camera system also supports 120 Hz shooting and outputting.

CONVENTIONAL AND CURRENT 8K CAMERAS

We have developed a variety of 8K cameras (Figure 1). A first-generation camera system, weighing approximately 80 kg, used four 8M-pixel charge-coupled device (CCD) sensors for 4K along with a 2.5-in optical system. Second-generation camera system, with its head weighing 32 kg, employed a 5/4-in optical format that was smaller than that of the first generation and also used four complementary metal-oxide-semiconductor (CMOS) sensors with 8M pixels. The third-generation camera systems were smaller and lighter, at 20 kg each, while still employing the same optical format and sensors as the second-generation cameras. In 2014, we began developing two kinds of practical single-chip cameras, each adopting a Super 35 (S35) sensor to further reduce its size and weight. The first camera developed was a “portable style” camera with a shape well suited to shoulder-loaded operation and a high level of development aiming at providing both the same ease of operation and the same low noise level as current 2K cameras provide. The second camera was a “multipurpose style” camera that was both small and capable of handling integrated recording device operations. Its ability to shoot in various styles was enabled by combining several units together on its small camera head. The miniature NAB2015 camera was unveiled and exhibited in Las Vegas in April 2015, and footage of on-location reporting, taken in both Japan and Las Vegas by an 8K camera, was shown simultaneously. In addition, we carried out live 8K production at both the Red and White Singing Festival, which was held at the end of the same year, and the NFL Super Bowl, which was held in February of the following year. Moreover, we undertook on-location 8K production at the Louvre in Paris, on “Shiretoko” which is the one of the world's natural heritages in Japan, and in other locations, and we applied a new HDR-based shooting method that achieved dynamic video expression.

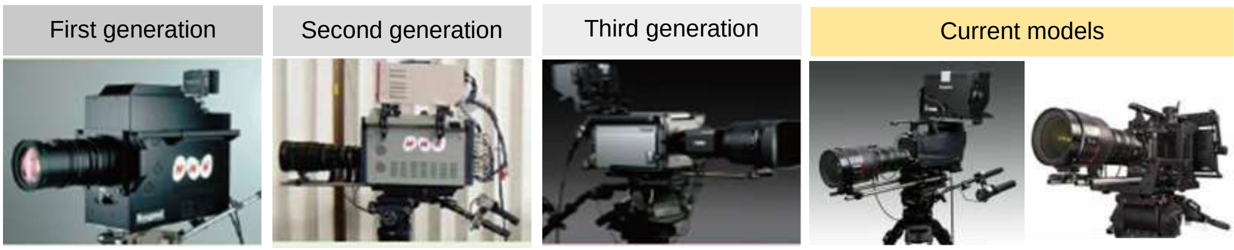

Figure 1 - 8K cameras we developed with Japanese manufacturers

REQUIREMENTS FOR 8K-BROADCASTING CAMERAS.

The followings are requirements for 8K broadcasting cameras we consider:

#1 – potential to use a zoom lens: a zoom lens is essential to meet the production requirements of television programs. Live sport feeds, in particular, require a high-magnification zoom of at least 20 times. An optical system that both has a reasonably sized zoom lens and can guarantee 8K image quality must be chosen.

#2 - the assurance of sufficient dynamic range (DR) and signal-to-noise ratio (S/N): although this is already a crucial camera characteristic, because it is tied directly to the fact that the expansion of HDR production provides the contents with a richer tone, a lens providing a magnification of at least 1,200%, as required by HLG, is necessary. If the integrated production of 2K, 4K, and 8K video is considered in conjunction with the signal-to-noise ratio (S/N), it is necessary to ensure an S/N that is equivalent to that of 2K video when down-converting or cropping 8K video to 2K video. Because of this, an S/N of 8K must be achieved as 60 dB at least, based on the static characteristics of a standard 2K camera.

#3 - a deep depth of field: in either television production employing a studio set or outdoor sport events, a depth of field in focus with both all of the performers and the set, or athletes on the same fields: it is important not only that the depth of field be deep, but also that 8K resolution is assured.

#4 - Modulation Transfer Function (MTF) characteristics: with respect to resolution characteristics, a 30% of modulation at 800 TV lines is the standard for an 2K camera. 3,200 TV lines is 80% of the Nyquist frequency for an 8K camera. So, a 30% of modulation at 3,200 TV lines is required for 8K.

#5 - the same power consumption, weight, and dimensions as current 2K and 4K cameras: it is necessary to achieve these while still satisfying the performance requirements of 8K cameras.

CAMERA DESIGN

After examining the following characteristics of a camera that meets the above-mentioned requirements, we decided to choose a 5/4-inch optical format, three-chip style camera. For your reference, the sensor sizes in each optical format are shown in Figure 2.

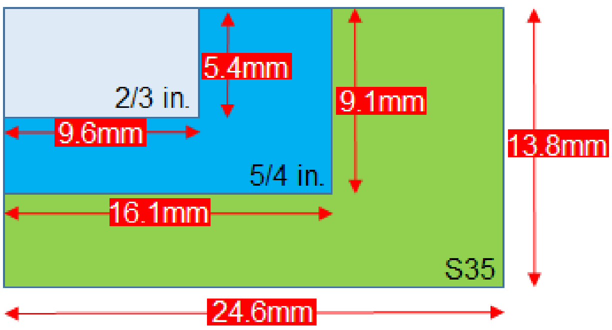

Figure 2 – Comparison of sensor sizes in each optical format.

Design Possibilities for the Zoom Lens

We examined the zoom magnifications, F-stops and dimensions of the zoom lenses for each optical format of 8K based on them for a 2/3-inch optical format available on the market. According to a reference [7], the lens covers focal lengths ranging from 9 mm in wide angle to 720 mm in telephoto and F-number in telephoto is 3.5. 205 mm of the front lens element diameter is introduced by these parameters and following equation:

$$\frac{f}{D} = F_{no} \tag{1}$$

where $f[mm]$, D[mm] and F_{no} are the focal length, the front lens diameter and F-number, respectively. There are restrictions on the front lens element diameter because it is difficult to produce a diameter that was greater than that of the reference 2/3-in lens. Furthermore, an F-stop needs to be smaller than that of the reference 2/3-in lens because of a low sensitivity due to smaller pixel sizes of 8K image sensors. When a lens is designed with a same front lens diameter and an F-stop of 2.8, the focal length in telephoto becomes 574 mm according to Eq. 1.

In the case of the S35 lens, the focal length in wide-angle is 22.5 mm with the same viewing angle as the reference 2/3-in lens, and the magnification level becomes 25.5 times. In the case of the 5/4-in lens, the magnification level attained is 38 times because its focal length in wide-angle is 15 mm with the same viewing angle as the reference 2/3-in lens.

As a result, the optical format of 5/4-in is suitable for live sport production and compares with that of the S35.

Examination Optical Format Based on Resolution

Figure 3 shows the relationships between the spatial frequencies and the degrees of modulation for ideal lenses. The respective graphs correspond to different F-stops. This figure shows that an approximately 20% degree of modulation can be obtained at an F-stop of F8 at 156 lp/mm, which is equivalent to the Nyquist frequency of the S35 8K sensor. In contrast, in the case of 2/3-in 8K sensor, no significant degree of modulation can be obtained when it does not have an F-stop of at most F4 at 400 lp/mm, which is equivalent to the Nyquist frequency. In addition, it is both too difficult and too expensive to design a lens that operates at 400 lp/mm, so the sensor with larger size should be chosen as possible.

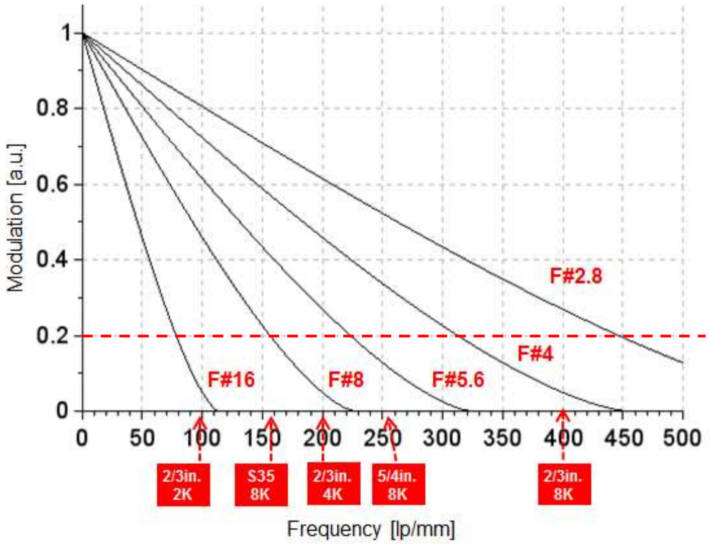

Figure 3 – Relationships between frequencies and degrees of modulation for ideal lenses.

Examination Based on Both DR and S/N of Pixel Pitch of Image Sensor

The DR of a camera is essentially determined by the ratio between the saturation change of the sensor and the noise. The DR of the HLG is 1,200% based on this standard. Moreover, in order to realise a 10-bit tone for a video level of 100%, assuming that the noise level is set at one electron, the full well capacity (FWC) must be a minimum of 12,000 electrons. FWC of the image sensor usually depends on the pixel pitch. Figure 4 shows pitch dependency of the FWC for the CMOS image sensor with conventional CMOS image sensors reported in the past [8]. In reference [8] it was described that the logarithms of FWC increase linearly with the pixel pitch under 5 μm.

This figure gives the following equation:

$$\log_{10} c = 0.25x + 0.5 \qquad (2)$$

where x [μm] is the pixel pitch, and c [ke^-] is the full well capacity of the pixel.

Solving this equation, we find that x = 2.32 when c = 12000. In other words, we should choose a pixel size of 2.32 μm or more.

Similarly, if the noise is under one electron, the S/N achieves to 60 dB (= $20\log(2^{10})$). If we assume the use of a filter that is equivalent to a four-pixel addition to the down-conversion output to 2K, an S/N improvement of about 12 dB is attained, so the required conditions are met.

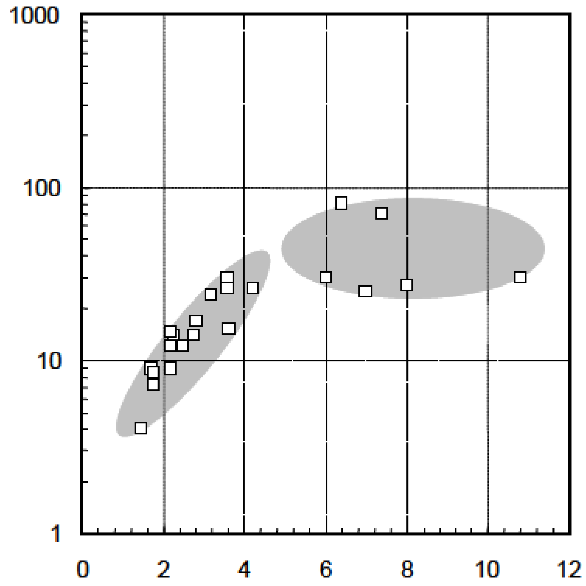

Figure 4 - Pixel pitch vs full well capacity for conventional CMOS image sensors.

Examination of Optical Format Based on Depth of Field

The sensor size of an 8K camera is larger than that of a 2K camera, and the number of pixels of the former is larger than that of the latter. The following equation shows that the depth of field becomes shallower for both the same angle of view and the same F-stop.

$$d_n = d_{\mathrm{pof}} {f^2}/{f^2 + F_{no} c_{coc} (d_{pof} - f)}$$
$$d_f = d_{\mathrm{pof}} {f^2}/{f^2 - F_{no} c_{coc} (d_{pof} - f)} \qquad (3)$$

where d_n : near distance, d_f : far distance, d_{pof}: point of focus
c_{coc}: circle of confusion, F_{no}: F-stop, relative aperture
f: focal length.

A comprehensive examination of both the sensitivity and the depth of field of a camera taken as an example is shown in Table 1. For studio production, we set the distance between a camera and an object to 6 m. In addition, we assumed that the angle of view was wider than it was in 2K production. This was because we assumed that the 8K videos would be displayed on a larger screen than 2K videos.

Optical format	2/3 in.	5/4 in.	S35	2/3 in. 2K

				(Reference)
Pixel pitch (μm)	1.25	2.4	3.2	5
Circle of confusion (μm)	2.5	4.8	6.4	10
F-stops	4	8	13	8
Object distance (mm)	6,000			
Focal length (mm)	10	19	25	14
Horizontal angle of view (°)	51.3	51.8	52.4	37.8
Depth of front field (mm)	2,250	2,338	2,664	4,260
Depth of rear field (mm)	9,000	10,585	23,809	-10,140.8

Table 1 – Comparison of depth of fields in each optical format.

SPECIFICATON AND EVALUATION OF THE PROTOTYPED CAMERA

Figure 5 - New 8K camera with three CMOSs.

Items	Specification and parameters	
	8K three-chip camera	2K three-chip camera
Optical format	5/4 in.	2/3 in.
Frame rate	59.94, 119.88 frames/s	59.94 fields/s
Resolution limit	Approx. 4,000 TVL	Approx. 1,000 TVL
S/N	62 dB (SDR)/ 56 dB (HLG)	60 dB
Sensitivity *1	F#8.0 (59.94p)	F#11.0 (59.94i)
Dynamic range	About 650%/1300% (HLG)	About 600%
Dimensions W × D × H	Head: 164 mm × 395 mm × 216 mm CCU: 440 mm × 500 mm × 221 mm	Head: 135 mm × 363 mm × 197 mm CCU: 424 mm × 410 mm × 133 mm
Weight	Head: 8.0 kg CCU: 25 kg	Head: 6.7 kg CCU: 17 kg

*1 Aperture for 700mV, 2000 lux 89.9% reflectance

Table 2 – Camera specification and parameters

The camera was prototyped based on the examinations and design specifications discussed in the preceding section. Figure 5 shows the appearance of the prototyped camera. We measured and evaluated each parameter of this camera; the specification and results are shown, with comparison to a 2K camera as a reference, in Table 2.

Regarding the picture quality, the resolution limit achieves approximately 4,000 TV lines using the lens with a 5/4-in optical format which is made for conventional 8K cameras. The S/N of 8K in SDR is better than that of 2K. Unfortunately, the S/N in HDR does not achieve our target, 60dB. On the other hand, the dynamic range obtained in HDR mode is above our target of 1200%. As for physical characteristics, the head of camera weighs about 1.2 times as that of 2K in spite of complicated signal processing and the larger sensors in it. Dimensions of 8K camera are closely similar to those of 2K, so we can carry it on the shoulder more easily than conventional 8K cameras.

CONCLUSIONS

We presented the specifications required for 8K cameras used for broadcasting and undertook the trial production of a three-chip 8K camera by applying a design that satisfied the required conditions. Based on the measured characteristics of the camera, we verified that its specifications make it useful as a practical 8K camera. By ensuring a DR wider than any other 8K cameras, a richer and clearer image expression was achieved.

On the other hand, the conventional single-chip camera still has other advantages. Therefore, we will continue to promote the production of numerous attractive 8K programs by making full use of the two kinds of 8K cameras.

In 2018, 8K broadcasts in Japan will move from their test phase into the practical phase. Looking towards the upcoming 2020 Olympic Games in Tokyo, we are ready to shoot the highest quality 8K you have ever seen!

REFERENCES

1. Sakiyama, T., Ichikawa, K., Abe, M., Mitsuhashi, S., Yamashita, T. and Miyazaki. M., 2016. 8K-UHDTV Production Equipment and Workflow which Realize an Unprecedented Video Experience, SMPTE 2016 Annual Technical Conference and Exhibition, pp. 1 to 11.
2. SMPTE ST 425-1:2011, Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface.
3. SMPTE ST 2082-12:2016, 4320-line and 2160-line Source Image and Ancillary Data Mapping for Quad-link 12G-SDI.
4. Recommendation ITU-R BT.2100-0, 2016. Image parameter values for high dynamic range television for use in production and international programme exchange.
5. Masaoka, K., Yamashita, T., Nishida, Y. and Sugawara, M., 2015. Color Management for Wide-Color-Gamut UHDTV Production, SMPTE Motion Imaging Journal, Vol. 124, no. 3, pp. 19 to 27.
6. Nishida, Y., Masaoka, K., Sugawara, M., Ohmura, M., Emoto, M. and Nakasu, E., 2011. Super Hi-Vision System Offering Enhanced Sense of Presence and New Visual Experience. The Best of IET and IBC. Vol. 3, pp. 5 to 10.
7. http://www.fujifilm.com/products/optical_devices/tv_cine/studio_field/ua80x9_12xext/
8. Mizobuchi, K., Adachi, S, Akahane, N., Sawada, H., Ohta, K., Oshikubo, H. and Sugawa, S., 2009. 4.5 μm Pixel Pitch 154 ke- Full Well Capacity CMOS Image Sensor, Proc. of Int'l Image Sensor Workshop, pp. 235 to 238.

ACKNOWLEDGEMENTS

The authors would like to offer our special thanks to Sony corporation for jointly developing the 8K camera. The authors also would like to thank Dr. Mizobuchi for giving permission for use of the figure from his paper [8].