An Ultrahigh-Definition Color Video Camera With 1.25-inch Optics and 8k x 4k Pixels

Hiroshi Shimamoto[1], Takayuki Yamashita[1], Noriyuki Koga[1], Kohji Mitani[1], Masayuki Sugawara[1], Fumio Okano[1], Masato Matsuoka[2], Jiro Shimura[2], Isao Yamamoto[3], Taku Tsukamoto[3], and Satoshi Yahagi[4]

[1]NHK (Japan Broadcasting Corporation) 1-10-11 Kinuta Setagaya-ku Tokyo 157-8510 Japan
[2]Ikegami Tsushinki Co.,Ltd. 4-6-15 Higashikoujiya Ohta-ku Tokyo 144-0033 Japan
[3]Astrodesign,Inc. 2-22-12 Kamikodanaka Nakahara-ku Kawasaki Kanagawa 211-0053 Japan
[4]Fuji Photo Optical Co.,Ltd. 1-324 Uetake, Kita-ku Saitama-city Saitama 331-9624 Japan

Abstract

We have developed an ultrahigh-definition color video camera that uses new 1.25-inch 8M-pixel CMOS digital imagers and diagonal pixel shifting in a 4-pickup system. The camera head weighs less than 40 kg. We also have developed a 5x zoom lens and a signal processing system incorporating a function for real-time lateral chromatic aberration correction. This function makes it possible to suppress optically generated color shifts across the entire zooming range. Moreover, we have developed optical-multiplexing transmission equipment and a camera control unit. This camera has a limiting resolution of more than 3200 TV lines. The SN ratio on an HDTV video, which is extracted 1/16 from 8k x 4k pixels video for the reason of SNR measurement device, was about 45 dB. The sensitivity was 2000 Lux at F2.8 with a dynamic range of 200%.

1. Introduction

We have actively been developing a 4000-scanning-line ultrahigh-definition video system for future high-reality broadcast services [1]. We already have developed an 8k x 4k pixel ultrahigh-definition color video camera ('8k-camera') system [2]. This camera employed four 60-frame/sec progressive-scanning 8M-pixel CCD imagers with a 2.5-inch optical system [3] and featured diagonal pixel shifting in a 4-pickup system [4]. Last year, the camera was used to shoot video of various sites including Yakushima, an island off the coast of Kyushu, Japan that has been registered a World Heritage Site. A theater in NHK Science & Technical Research Laboratories (STRL) has been showing these highly realistic 8k x 4k pixel video programs since 2002, and they have been favorably received by many visitors.

The 8 million pixels making up each of the CCD devices used in the camera must be driven at high speed, and to this end, the internal imager circuitry is divided into 16 channels corresponding to 16 image-sensing areas. The analog output-amplifier characteristics of these areas do not match, however, and this requires special effort to correct. In addition, the

Figure 1. External view of two 8k x 4k pixels camera heads. Previous prototype (right:2.5-inch CCD) and newly developed (left:1.25-inch CMOS).

camera has been designed for use with a 2.5-inch optical system resulting in relatively large optical lenses and optical prisms, and consequently, the prototype camera head of the first prototype weighs about 80 kg (Fig. 1 Right). This state of affairs has generated a desire for a smaller and lighter 8k-camera.

To meet this need, we developed a new 8M-pixel CMOS imager with a 1.25-inch optical system [5], and developed a new 8k x 4k pixel, ultrahigh-definition color video camera with a camera-head weight half that of the previous camera (Fig. 1 Left). We also developed a new camera system featuring 5x zoom lens, optical-multiplexing transmission equipment, a camera control unit (CCU), and a signal processing system.

2. Camera system

2-1. Outline of the system

Table 1 lists the specifications of the camera, and Fig. 2 shows a block diagram of the camera system.

System	7680 x 4320 pixels 60 frames per sec. progressive scanning
Optical format	approx. 1.25 inch
Lens	5x zoom f = 12 ~ 60 mm F2.1
Color imaging method	Four-panel imaging (GGBR) (Diagonal pixel offset between green imagers)
Image sensor	8M- pixel CMOS
SNR	Approx. 45dB (on HDTV format)
Sensitivity	2000 lux, F2.8
Dynamic range	200%
Power consumption*	300 W
Weight*	40 kg

Table 1. Specifications of the camera system.

The 5x zoom lens was developed for this 8k-camera. The lens has a remote controller for zooming, focusing and iris movement. Focus can be controlled from the CCU in order to adjust the focus precisely with a high-resolution monitor. The lens iris is also controlled by the CCU.

The color separation prism divides an optical image into two green (G1, G2), blue (B), and red (R) images with one image incident on one 4k x 2k pixel CMOS image sensor ('four-pickup imaging method'). Each imager has on-chip analog-to-digital converters (ADCs) and outputs 10-bit digital signals. Signals from the four imagers are converted into HD-SDI format within the camera head. The 16 channels of HD-SDI electrical signals are converted into eight 3-Gbps optical signals (one optical signal for every two channels) that are then grouped together in one optical fiber through 8-wavelength multiplexing, allocating from 1310 nm to 1450 nm at 20nm intervals. The resulting multiplexed signal can be transmitted up to 1 kilometer on an SMPTE 311M optical fiber cable. On the receiving side, the signal is converted back into 16 channels of HD-SDI electrical signals.

The CCU performs a variety of tasks including fixed-pattern-noise cancellation (FPN cancel) and pixel and/or column defect correction. The electrical power of the camera head is transmitted through the optical fiber cable from the CCU; consequently, one-cable operation is possible.

The signal processing system synthesizes and interpolates 4-four streams of pixel-shifted 8M-pixel video into 32M-pixel video (8k x 4k pixels; '8k-video'). This system also performs lateral chromatic aberration correction processing, a contour compensation processing, a gamma compensation processing, and a focus-aid function, and is capable of configuring and outputting monitor video simultaneously. Since there are presently no CRT or LCD monitors that can directly display 8k-video, we use an LCD monitor that can display 4k x 2k pixels video ('4k-video') and apply filter processing to

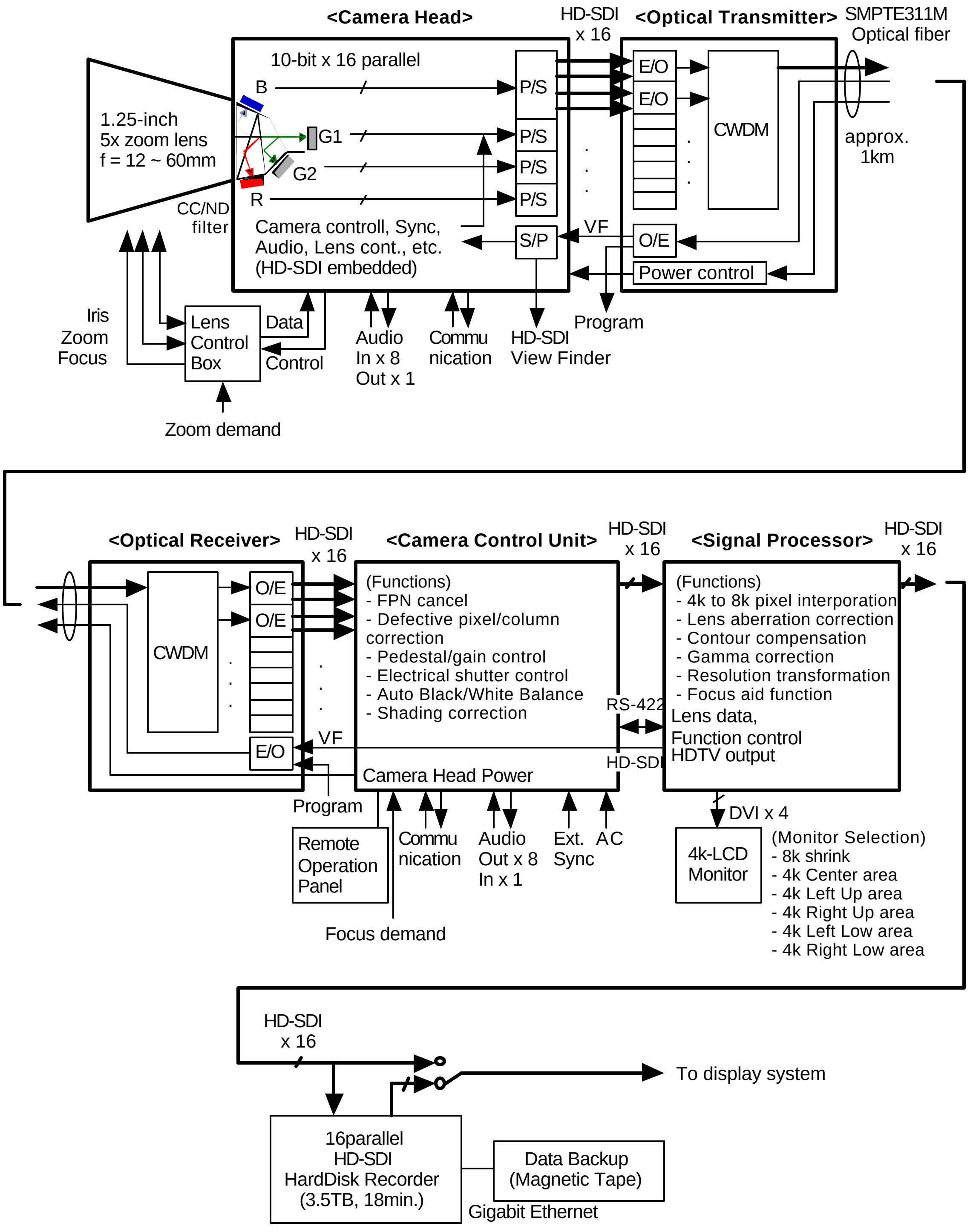

Figure 2. Block diagram of the 8k x 4k camera system.

reduce the 8k-video to 4k-video for display in order to see the whole picture at once. It is also possible to extract part of 8k-video as 4k-video, because monitoring picture without reduction is indispensable to bring a camera into focus. The monitoring area is selectable for five area such as center, left-upper, right-lower, etc. The signal processing system can also prepare HDTV output, and it is displayed on a viewfinder and used for compositing of a camera frame. A built-in focus-aid function detects the signal components of high-resolution video from the 8k-video and displays them on a low-resolution monitor optionally.

The 8k-video adjusted by the signal processing system can be reconverted back into 16 channels of HD-SDI signals for recording on a 16-parallel HDTV hard-disk recorder. Alternatively, the video can be directly output to liquid-crystal projector display equipment designed especially for 8k-video.

2-2. Camera head

Zoom lens

Table 2 lists the zoom lens specifications, and Fig. 3 shows depth-of-field calculations for focal lengths f = 12 mm and 60 mm at f-stops from F2.8 to F11. These calculations were performed for a circle of least confusion of 2.1 µm on the G signal with the pixel shift. When viewing angle and f-stop are same, depth of field increases as pixel size decreases. This means that the depth of field with this smaller 8k-camera will be deeper than with the former 2.5-inch 8k-camera. Though dependent on the material content and production, a feeling of immersion can be obtained from video shown on a large screen as long as the entire screen is in focus. Likewise, video extracted from 8k-video can be easily used as video material provided that this focusing condition holds. The camera achieves pan focus for a shooting distance of about 6 meters at a focal length of 12 mm, for wide-angle image shots of 67 degrees to horizontal. Conversely, when a shallow depth of field is needed, depending on material content and production requirements, zoom and macro, close-up lens, etc. can be used.

Focus length	12mm 60mm
Zoom	5x
Iris	F2.1
M.O.D. (from front lens)	1 m (0.07 m macro)
Shooting angle	67 ~ 15 deg. (Horizontal) 41 ~ 8 deg. (Vertical)
Weight	Approx. 8 kg (lens only)
System equipments	Servo module System controller Zoom demand Focus demand

Table 2. Specifications of the 5x zoom lens.

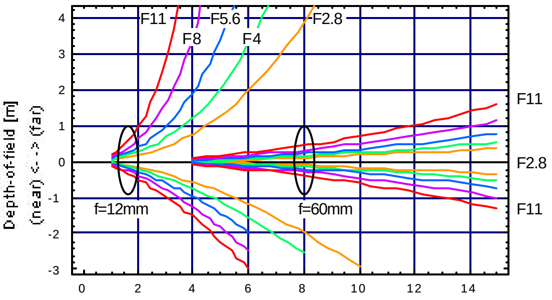

Figure 3. Depth-of-field calculations for focal length f = 12 mm and 60 mm at f-stops from F2.8 to F11.

8M-pixel CMOS imager [5]

Table 3 lists the specifications of newly developed imager. The effective number of pixels is exactly four times the number of pixels of the HDTV 1080/60P format, which makes it easy to employ existing HDTV electronic devices in the signal-processing circuitry and equipment. The principal reasons for adopting CMOS, not CCD, are the need for low power consumption and a compact configuration.

Figure 4 shows the block diagram of this 8M-pixel CMOS imager. In this configuration, pixel signals read out from photo diodes in the photosensitive area are stored in sample hold circuits for each column and digitized by 10-bit column-parallel ADCs prepared for every two columns for a total of about 2000 converters. Digital data is then temporarily stored in an SRAM bank and 16 pixels are output in parallel every clock cycle.

Pixel number (Horizontal x Vertical)	Effective: 3840 x 2160 Total: 3936 x 2196
Pixel size	4.2 µm x 4.2 µm
Scanning	60 frames per sec. Progressive
ADC	1968 Column parallel
Output	10-bit digital 16 parallel
Frequency	49.5 MHz
Pixel aperture	68%

Table 3. Specifications of the 1.25-inch 8M-pixel CMOS imager.

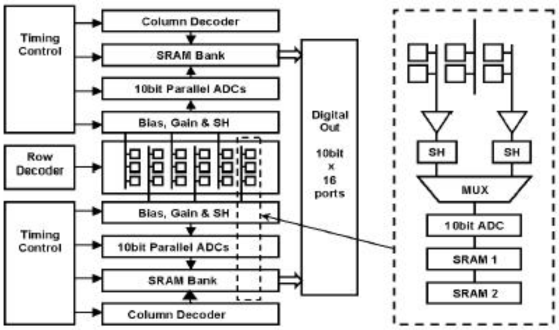

Figure 4. Block diagram of the imager.

Four-pickup Imaging

The 8k x 4k ultrahigh-definition video camera requires an imager having about 32 million pixels. However, a device having such a high number of pixels and capable of capturing moving images at 60 frames per second has not yet been reported. We therefore decided to take the same approach as we did with the previous CCD camera and use 8M-pixel CMOS imagers in a 4-pickup system with pixel shifting to obtain 8k x 4k video.

Figure 5 shows the pixel-spatial-sampling arrangement in this 4-pickup system. The imaging system uses two sensors for the green channel and one sensor for each of the red and blue channels. As for the optical image, by shifting the two green-channel sensors diagonally opposite each other at half a pixel apart, the imaging system's Nyquist frequencies in the horizontal and vertical directions become double that of one sensor. As a result, since the green signal contributes most to the luminance signal of the images, this sampling pattern can effectively increase the system resolution by using imagers with a relatively small number of pixels. Image sampling pattern in this method is equivalent to that of a 32-megapixel single-chip color device.

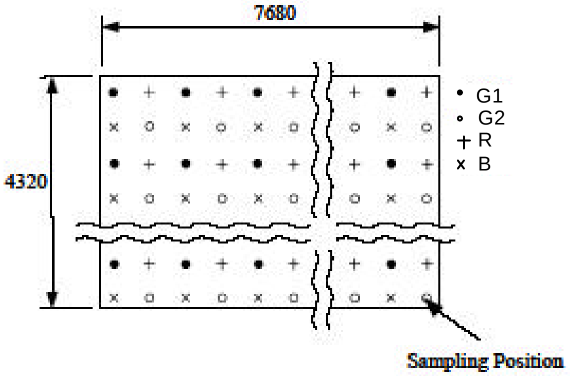

Figure 5. Pixel-spatial-sampling arrangement in the 4-pickup system.

When increasing the resolution by pixel shifting in this manner, inaccurate imager alignment might degrade the spatial resolution of the synthesized image. To deal with this problem, we devised a technique for attaching the prism to each imager [6]. With this technique, we were able to keep the attachment error to within 0.5 um (within 1/10th of a pixel) and achieve a limiting resolution of 3200 TV lines or more.

2-3. Lateral Chromatic Aberration Correction

For fixed focus, lateral chromatic aberration caused by the optical characteristics of a lens is generally easy to correct. This is not true, however, for a zoom lens in which the amount of aberration changes. In particular, the ultrahigh-definition video of an 8k-camera is easily affected by chromatic aberration even if a high-precision lens is used. For this reason, we incorporated a function for real-time correction of lateral chromatic aberration into the signal-processing system. Figure 6 shows a block diagram of this function. The lateral chromatic aberration of a lens changes according to the f-stop value, focal length, and focus position. It is therefore desirable to obtain data beforehand for correcting the lateral chromatic aberration for each lens parameter. This is done by taking video of patterns and then record the data for correcting these patterns. In the current prototype system, however, we made corrections only in relation to the focal length because of limited data capacity.

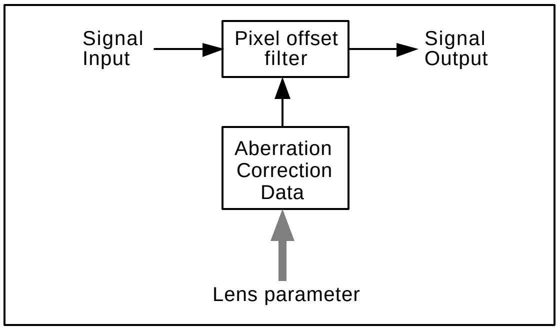

Figure 6. Block diagram of the lateral chromatic aberration correction.

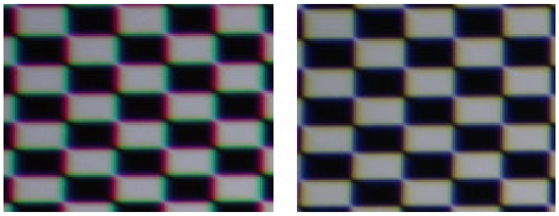

Figure 7-a (left) and 7-b (right).
An example of the lateral chromatic aberration correction.
Before (a) and after (b).

To measure lateral chromatic aberration, we prepared a checkered pattern chart. It was easier to detect grid points than a simple crosshatch pattern with fine grid lines. For each of the G, B, and R signals, we determined the spatial coordinates of the checkered pattern's grid points and used the resulting shifts to compute the aberration for each color channel. It was found that B and R had a maximum shift of about five pixels on each imager with respect to G due to zooming. Figures 7-a and 7-b show a picture before and after lateral-chromatic-aberration correction.

In the above way, we have succeeded in correcting lateral chromatic aberration over the entire range of a 5x zoom lens in real time. In addition, contour compensation processing, which might emphasize pseudo colors if lateral chromatic aberration were left unattended, can now be performed while shooting video since lateral chromatic aberration is removed. This significantly reduces the time involved in checking picture quality and the amount of post-processing required after recording.

3. Camera characteristics

3-1.Resolution

The resolution of the 8k-camera is becoming close to theoretical limitation. There are three main factors for determining a camera's spatial frequency characteristics (modulation transfer function (MTF) characteristics).

(1) Optical diffraction phenomenon in the lens

This is a physical phenomenon that occurs even in an ideal lens having no aberration. The MTF is calculated by the following equation.

$$MTF_{dif}(m) = \frac{2\boldsymbol{b} - \sin 2\boldsymbol{b}}{\boldsymbol{p}}$$
$$\left[\boldsymbol{b} = \cos^{-1}\left(lF_{no}\frac{m}{2V}\right)\right] \tag{1}$$

Here, l is a wavelength, F_{no} is an f-stop value, m is a spatial frequency [TV lines], and V is a vertical length of the imager.

(2) Optical low-pass filter (LPF)

An optical LPF is inserted to suppress pseudo signals caused by spatial sampling in a solid-state imager. It is calculated by the following equation.

$$MTF_{LPF}(m) = \left|\cos\left(\frac{\boldsymbol{p}m}{2f_s}\right)\right| \tag{2}$$

Here, f_s is a spatial sampling frequency [TV lines]. The equation takes on cosine characteristics, therefore the LPF not only cuts the target frequency but also drops within the effective band. For G channel in 4-pickup system, two optical LPF are used for diagonal directions.

(3) Pixel aperture ratio

When pixels of an image sensors are arranged with a pixel sampling interval of d_0 in both a horizontal- and vertical- direction and the relative positions of two image sensors are shifted a half of the pixel pitch to a diagonal direction, the spectrum on a two-dimensional frequency plane (m,n) is expressed as follows.

$$G_S(m,\boldsymbol{n}) = G(m,\boldsymbol{n}) \otimes \frac{1}{d_0^2}\sum_m\sum_n \boldsymbol{d}\left(m - \frac{m}{d_0}, \boldsymbol{n} - \frac{n}{d_0}\right)\left\{1 + e^{-pj(m+n)}\right\} \tag{3}$$

Here, $G(m,n)$ is a frequency response of the pixel, m and n are the harmonic degree. For simplicity, we assume a square pixel aperture and consider a one-dimensional case. The MTF characteristics are given by the following equation.

$$MTF_{apt}(m) = \frac{\sin\left(\boldsymbol{p}\frac{d}{d_0}\frac{m}{f_s}\right)}{\left(\boldsymbol{p}\frac{d}{d_0}\frac{m}{f_s}\right)} \tag{4}$$

Here, d is a pixel aperture width, therefore d / d_0 [x100%] denotes the aperture ratio along one dimension. In a 4-pickup system, the pixel-sampling interval under pixel shifting is short, which means that the aperture ratio becomes high relative to a 3-pickup system and may even exceed 100%.

Total MTF characteristics are the product of these three factors. Figures 8 and 9 shows the results of the calculation for G and B/R, respectively. The parameters for the calculations are, l = 550 nm, V = 9 mm, f-stops of F2.8, 4, 5.6, and 8. The optical LPF is to be given a null response to the sampling frequencies of f_{s_G} = 8640 TV lines for G and

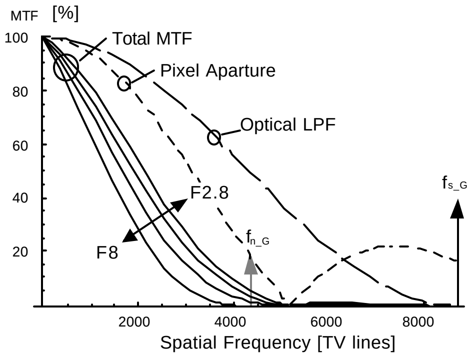

Figure 8. A MTF calculation for the total characteristics of this camera (green channel).

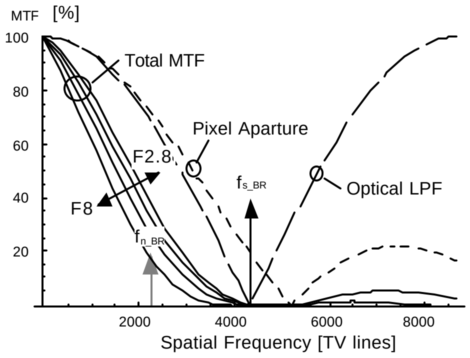

Figure 9. A MTF calculation for the total characteristics of this camera (blue and red channel).

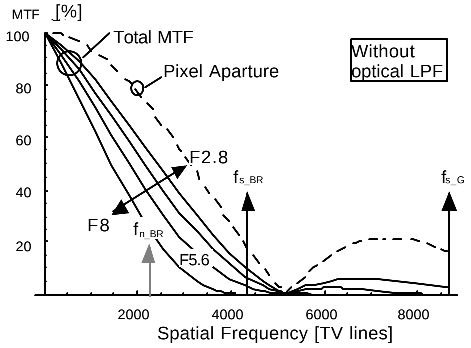

Figure 10. A MTF calculation for the total characteristics of this camera (without optical LPF).

f_{s_BR} = 4320 TV lines for B and R, so that aliasing distortion does not occur at low frequencies where it is easily visible. The pixel aperture ratio in horizontal or vertical direction is 83 % for B and R, and 167 % for G.

Figure 10 shows a calculated MTF characteristics without inserting an optical LPF for G, B, and R. As can be seen, MTF characteristics are greatly affected by the pixel aperture ratio, and response is low at the sampling frequency of each color channel. The drop in MTF characteristics could be electronically corrected by enhancing high-frequency components at the aperture compensation circuit, but doing this would increase noise. Therefore, we decided to omit the optical LPF for G and to place priority on maintaining MTF response as high as possible. For B and R signal, the amount of pseudo signals aliasing will be more excessive than for G as the spatial sampling frequency is half that of G. Nonetheless, we omitted the optical LPF for B and R because (1) the degree of aliasing that does occur at visually noticeable low frequencies is low especially when the f-stop value is over F5.6; (2) for B and R, we also placed priority on maintaining MTF response and avoiding any increase in noise caused by aperture compensation; (3) several methods to improve the resolution of a single-chip color imager have been reported [7] and we would like to examine them with an eye to using them with this camera in the future.

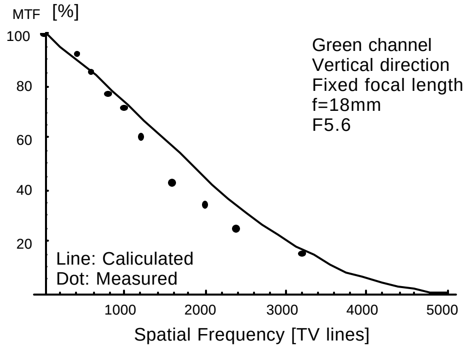

Figure 11. Measured MTF characteristics of this camera.

Figure 11 shows the MTF characteristics measured in the vertical direction for the G signal in an actual video camera at an f-stop of F5.6. A lens with fixed focal length of 18 mm was used for the measurement in order to minimize the lens' influence. The MTF was obtained by converting measured amplitude response (AR) according to Coltman's equation [8]. The measurement result almost matched the calculated one.

3-2.SN ratio

The signal-to-noise ratio (SNR) was measured for HDTV video, because there is presently no device that can measure SNR for 8k-video. An HDTV video was extracted from the 8k-video (comprising 1 / 16 of the signal) and converted from progressive scanning to interlaced scanning. The result was about 45 dB. Although this SNR was 5 dB less than that of our previously developed 2.5-inch CCD camera, the SNR per unit of photosensitive area was improved considering that the CMOS imager had only 1/4 the area of the CCD. The sensitivity of the camera was 2000 Lux at F2.8 with a dynamic range of 200%.

4. Conclusion

We have developed an 8M-pixel CMOS digital imager with a 1.25-inch optical system, and have developed a 8k x 4k pixel, ultrahigh-definition video camera featuring a 4-pickup imaging system and diagonal pixel shifting. The camera head weighs less than 40 kg and the system achieves a limiting resolution of 3200 TV lines or more. Although this camera has a photosensitive area only 1/4 the size of our previously developed 2.5-inch CCD camera, its SNR when extracting 1/16 of the video for HDTV is still about 45 dB when shooting under 2000 Lux at F2.8 and a dynamic range of 200%. In relation to an optical LPF, we showed in the section 3 that almost no aliasing signals occur in the G of our camera without inserting an optical LPF and that, for B and R, the degree of aliasing that does occur at visually noticeable low frequencies is minimal. For these reasons, we chose not to insert an optical LPF in this camera and to place priority on suppressing any increase in noise through contour compensation.

This camera system is the first that we have equipped with a 5x zoom lens. We have also incorporated a lateral chromatic aberration correction function into the signal processing system that enables real-time modification of the aberration correction value according to the focal length of the zoom lens. This function makes it possible to suppress optically generated color shifts across the entire zooming range. It also enables contour compensation processing to be performed while shooting thereby reducing the amount of post-processing after recording. In addition to the above, we have also developed optical-multiplexing transmission equipment, the CCU, and signal processing system, which makes for a highly practical camera. This camera system is currently being used to record video material (location shots) to be shown at the 2005 World Exposition to be held in Aichi, Japan from March 2005.

References

[1] F. Okano et al., "Ultrahigh-Definition Television System with 4000 Scanning Lines," NAB BEC Proceedings, pp.437-440, 2004

[2] K. Mitani et al., "Ultrahigh-definition color video camera system with 4K-scanning lines," Proceedings of SPIE, Vol.5017, pp.159-166, 2003

[3] C. Smith et al., “An 8M-CCD for an Ultra High Definition TV Camera,” IEEE Workshop on Charge Coupled Devices and Advanced Image Sensors, Nagano, Japan, June 1999
[4] M. Sugawara et al., “Four-chip CCD Camera for HDTV,” SPIE Proc., Vol. 2137, pp. 122-129, 1994
[5] I. Takayanagi et al., “A 1-1/4 inch 8.3M Pixel Digital Output CMOS APS for UDTV Application,” Digest of Technical Papers. ISSCC. 2003, vol.1, pp.216-217, 2003
[6] T. Yamashita et al., “A New Alignment Method for an 8k x 4k-pixel ultrahigh-definition Camera with Four Imagers,” Proceedings of SPIE, 5301, pp.76-86, 2004
[7] S. C. Park et al., “Super-Resolution Image Reconstruction -A Technical Overview-,” Signal Processing Magazine, vol.20, no.3, pp.21-36, May 2003
[8] J. W. Coltman, “The Specification of Imaging Properties by Response to a Sine Wave Input,” J. Opt. Soc. Am., 44(6): 468-471, 1954