Ultrahigh-Definition Video System with 4000 Scanning Lines

M. Kanazawa[1], K. Mitani[1], K. Hamasaki[1], M. Sugawara[1], F. Okano[1], K. Doi[2], M. Seino[3]

[1]Japan Broadcasting Corporation, Japan; [2]Victor Company of Japan, Japan; [3]Ikegami Tsushinki, Japan

Abstract

The authors are investigating an ultrahigh-definition video system that can provide viewers with a greater sensation of reality than HDTV. The present target is to develop a system with 4000 scanning lines. This system should play a major role in any application that requires high resolution such as digital cinema. A video camera and projection display together with a disc recorder system have been developed as experimental devices for the 4000 scanning line system. At the present time, the number of panel pixels is limited to 2k×4k for both CCD and LCD. Due to this resolution constraint, four panels (two greens, one red and one blue) are combined to produce a resolution of 4k×8k pixels (16 times that of HDTV) in both the camera and display. The two green panels are arranged by the diagonal-pixel-offset method to achieve the above resolution. After describing the development of these devices, the paper will discuss issues to be addressed for realizing an ultrahigh-definition video system.

Introduction

The HDTV system has been developed to provide realistic images on a display, but in some applications, including digital cinema, video systems with more improved resolution than HDTV are often required. From the previous research, the followings are clear.

	HDTV	4k × 8k video
Pixel Number	1,080 × 1,920	4,320 × 7,680
Viewing angle (pixel invisible)	30 deg. horizontally	More than 100 deg. horizontally
Comparison with movie	Equivalent to 35 mm motion film	More than twice of 70 mm motion film
Displayable character number*	3,600	58,000 (as many as two sheets of newspaper)

** Pixels of 24 × 24 are assigned to one character.*

Table 1: Specifications of 4k × 8k video system

The preferable viewing angle for viewers depends on the picture size and resolution. Our project has shown that the viewers prefer larger angle with wide-screen and high-resolution images than that with HDTV where the angle of 30 degrees is recommended (1). The sensation of reality induced to an observer by watching a display screen increases with the viewing angle and saturates at an angle around 100 degrees horizontally (2). From human vision characteristics, the spatial resolution of a display is desired to have 60 pixels per degree (3). To meet the above conditions, a display with approximately 4k × 8k pixels is necessary.

At the present time, the most feasible way to develop a display with such a huge pixel number is by combining projection displays. Recently, a LCD panel of 2k × 4k pixels for projection display was introduced to produce a display with 2k × 4k pixels (4), and it became possible to realize a 4k × 8k pixel image.

Table 1 shows the specifications of the 4k × 8k video system, comparing with those of HDTV. Figure 1 shows the relationship between human visual field (2) and display areas of HDTV and the 4k×8k system. While the display area of HDTV is covering the effective field (3), that of the 4k × 8k system covers the induced field, which is related to the observer's posture judgment and therefore the sensation of reality. At a viewing angle of 100 degrees, which corresponds to a viewing distance of 0.75 H (H: picture height), viewers can enjoy much higher sensation of reality with 4k × 8k video images than with HDTV. The resolution of the 4k × 8k system is 16 times of that of HDTV and is more than twice of that of 70 mm motion film. Not only for a wide-screen application, but an application for a high information terminal is possible with a display area of two sheets of newspaper.

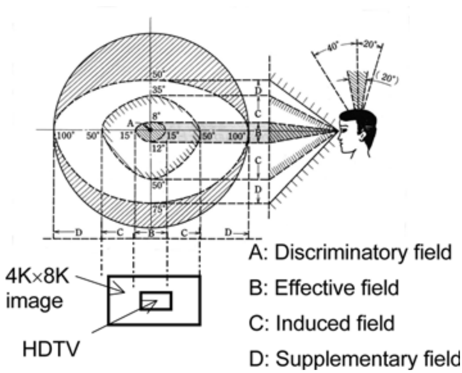

Figure 1: Visual field and display areas of 4k × 8k system

Experimental devices of ultrahigh-definition video system

Configuration of experimental devices

The following measures were incorporated in the actual development of the equipment and devices to enable parallel processing and to simplify hardware implementation.

- Given that the ultra-high-definition video system is to have 4 × 4 times the number of HDTV pixels (for a total of 4320 × 7680 pixels), the authors simplified the devices design by using and synchronising multiple devices based on the HDTV format.
- At present, only 2048 × 3840 pixel panels are available for use in both cameras and display equipment. Therefore to achieve ultrahigh-definition with the above panels, we increased the number of pixels in both the horizontal and vertical directions for both camera and display by combining two panels to color G in the

diagonal-pixel-offset method (5). In addition, we applied one panel of 2048 × 3840 pixels to R and B each. This scheme makes use of human visual characteristics. Specifically, while the human vision features high spatial frequency response with respect to brightness, it suffers from relatively low response with respect to hue and chroma. Because G makes a bigger contribution than R and B to brightness components, achieving high resolution even in only G can still have a great effect.

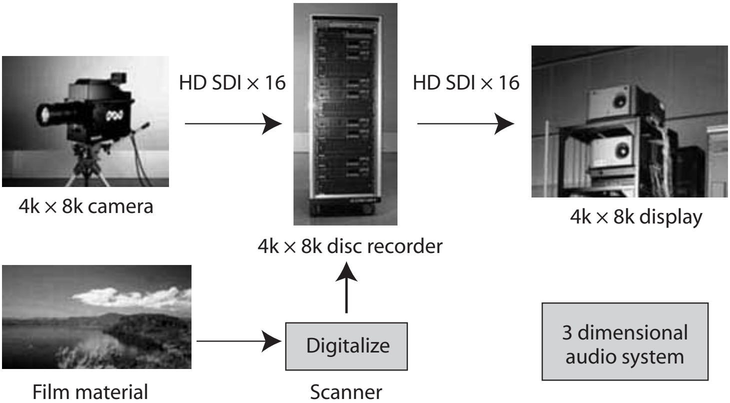

Figure 2: Experimental devices of ultrahigh-definition video system

Figure 2 shows the configuration of developed equipment and devices. The camera and display are developed based on the diagonal-pixel-offset method. Figure 3 shows how images are reproduced by the pixel-offset method. The picture quality is discussed later. The video signal is transmitted through 16 channels of HDTV SDI (Serial Digital Interface) format signal (in total, approximately 24 Gbps).

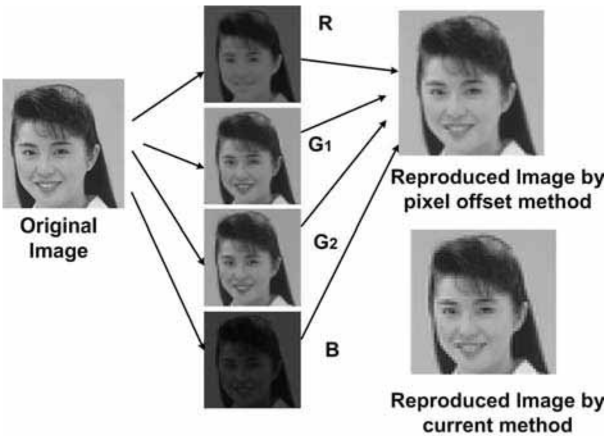

Figure 3: Experimental devices of ultrahigh-definition video system

Camera system

Combining four 2.5 inch CCD (Charge Coupled Device) panels with 2048 × 3840 pixels, a video camera, which has equivalently 4096 × 7680 pixels for green while 2048 × 3840 pixels for red and blue was developed. The image area of the CCD panel is partitioned into 16 regions. The multiple-output structure with its 16 outputs enables imaging with a high data rate of 594 mega-pixels per second for ultrahigh-definition moving images of 60 Hz progress scanning. Figure 4 shows the optical structure of the camera and the spatial position of each CCD (6). Table 2 shows the specifications of the camera.

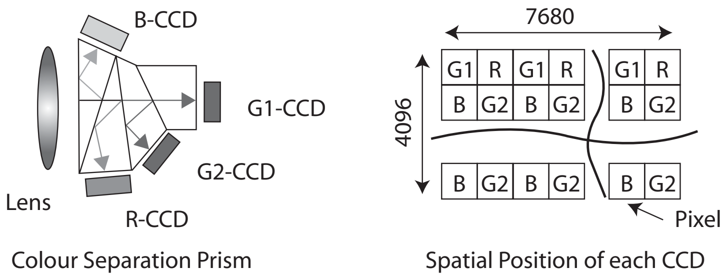

Figure 4: Color separation prism for 4-CCD pickup method and spatial position of each CCD

Item	Content
Format	4 k lines progressive scanning
Sensor	2.5 inch CCD × 4
Lens	Fixed focal-length lens (f: 50 mm)
S/N on HDTV format	Approximately 50 dB
Sensitivity	2,000 lux, F2.8
Weight	76 Kg (including Lens)

Table 2: Specifications of camera system

The resolution and depth of field are examined. In order to increase the pixel number of an image sensor panel for ultrahigh-definition, the pixel size should be smaller or the panel size must be bigger. Assuming that the pixel size is equal to the Airy disc diameter, the maximum F-number of lens, Fmax is shown in equation (1), where ε is the pixel size and λ is the wavelength of the incident light. When the lens is used out of this condition, the frequency response at the Nyquist frequency would be less than 50%.

$$F_{max} = \varepsilon / (1.22 \cdot \lambda) \quad (1)$$

Depth of field is defined as the area within which a point light source is focused in one pixel area and is shown in equation (2). The maximum depth of field, L_{max} can be obtained by substituting equation (1) into (2).

Figure 5 shows the relationship between the pixel size and L_{max}, F_{max}. When the pixel size is reduced to increase the number of pixels, the depth of field can be made constant by using a wide-angle lens with a shorter focal length.

$$L = \varepsilon \cdot s^2 \cdot F / f2 \qquad (2)$$

where L: depth of field, s: the distance from the lens to the object, F: F-number of the lens, f: the focal length of the lens

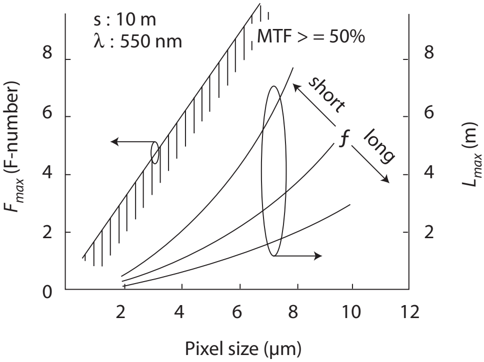

Figure 5: Pixel size, maximum depth of field and Maximum F-number

Display System

Combining four 1.7 inch LCoS (Liquid Crystal on Silicon) panels with 2048 × 3840 pixels, a display consisting of two projection units, one for green and the other for red/blue was developed (7). Figure 6 shows the optical structure of the display. In order to combine two green images, four PBS (Polarizing Beam Splitter) and four-λ/2 phase shifting plates are used in the green projector, while for red and blue images, two PBSs and one cross-dichroic prism are used. Table 3 shows the specifications of the display.

Item	Content
Panel	1.7 inch LCoS × 4
Type	Front projection
Screen	320 inch (4m × 7m) Gain : 0.85
Light output	> 5,000 lumen
Contrast ratio	> 700:1

Table 3: Specifications of display system

The two green panels must be located apart from each other with exactly 0.5 pixels horizontally and vertically to obtain ultrahigh-definition. If there is misalignment between the two panels, the resolution would be degraded and artefact would be observed. The output of the display can be shown in equation (3) in one dimension for the sake of simplicity.

$$g(x) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ h(2 \cdot n) \cdot \delta\left(y - 2 \cdot n - \frac{q}{2}\right) + h(2n+1) \cdot \delta\left(y - (2 \cdot n + 1) + \frac{q}{2}\right) \right\} \cdot a(x-y)\, dy \qquad (3)$$

Where x: position on display, q: misalignment, g(x): output, a(x): aperture, h(x): input to display, δ(x): Dirac's delta function

Equation (4) is the Fourier transform of equation (3) and shows the frequency characteristics due to the misalignment.

$$
\begin{aligned}
G(\omega) &= A(\omega)\cdot\sum_{n=-\infty}^{\infty} H(\omega-n\cdot\pi)\cdot\left\{\exp\left(-j\cdot\frac{q}{2}\omega\right)+\exp\left(-j\cdot n\cdot\pi+j\cdot\frac{q}{2}\omega\right)\right\} \\
&= A(\omega)\cdot\left\{\cos\left(\frac{q}{2}\omega\right)\cdot H(\omega)-j\cdot\sin\left(\frac{q}{2}\omega\right)\cdot H(\omega-\pi)+j\cdot\sin\left(\frac{q}{2}\omega\right)\cdot F(\omega+\pi)\right\} \\
&\cong A(\omega)\cdot\left\{H(\omega)-j\cdot\frac{q}{2}\omega H(\omega-\pi)+j\cdot\frac{q}{2}\omega H(\omega+\pi)\right\} \qquad q << 1
\end{aligned}
\tag{4}
$$

Where ω: spatial frequency (Nyquist frequency is π), q: misalignment,
G(ω): Fourier transform of g(x), H(ω): Fourier transform of h(x),
H(ω) = 0 for |ω| >π, A(ω): Fourier transform of a(x)

From equation (4), it is shown that the misalignment causes a component of which frequency is (ω-π) and amplitude is nearly proportional to the misalignment and the frequency.

Considering human vision characteristics, it is found that the misalignment of 0.02 pixels can cause visible artefacts (8). However, the conclusion is derived by calculation with a test pattern and more generous misalignment would be allowed for general images. An actuating tool is attached to one of the green panels and can move the panel with sufficient precision to prevent from the artefact described above.

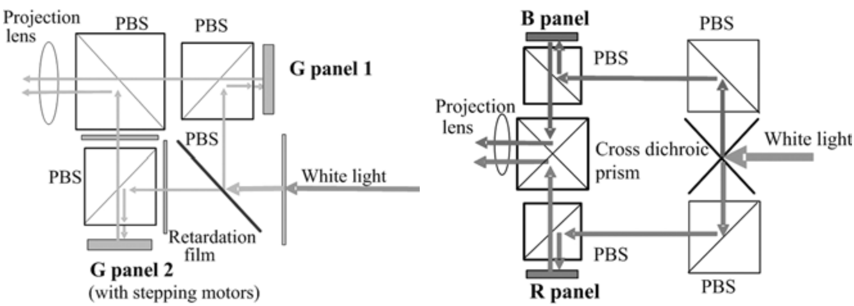

Figure 6: Optical structure of display

Other devices

A disc recorder system was developed. The device consists of 16 HDTV disc recorder units and can record approximately 18 minutes of the ultrahigh-definition video signal in real time with a capacity of nearly 3.5 TB. The input/output of the device is 16 channels of HDTV SDI format signal.

Method	Resolution	
	Green	Red/Blue
A (original)	4k × 8k	
B	2k × 4k × 2 (spatial offset)	
C	2k × 4k × 2 (spatial offset)	2k × 4k
D	2k × 4k	

Table 4: Evaluated methods

In the current audio systems, loud speakers are placed on a horizontal plane to produce two-dimensional audio. However, in the ultrahigh-definition video system, the screen is large and the audio system is required to match with the size. Therefore, a great number of loud speakers (in total, 22 loud speakers plus 2 bass speakers and two speaker arrays) are installed three-dimensionally and fed independently to make a three-dimensional audio system.

Results and discussion

Subjective evaluation of resolution

The authors conducted a subjective assessment test in order to confirm the picture quality of the pixel-offset method. The picture quality related to pixel structures of the four methods shown in Table 4 was investigated. The image from the Method-A has full 4k×8k pixels for red, green and blue. The Method-B and Method-C are the spatial pixel-ffset method, where the image from Method-B has 2k×4k×2 pixels for red, green and blue, and that from Method-C has 2k×4k×2 pixels for only green. The image from Method-D has 2k×4k pixels for red, green and blue (half resolution of Method-A). Table 5 shows the condition of the evaluation.

Item	Content
Image size	480 × 640 pixels
Display	20 inch CRT
Viewing distance	6 H (H: picture height)
Peak luminance of display	75 cd/m^2
Evaluation method	EBU method variant II (9)
No. of assessors	5 experts
No. of images	3

Table 5: Evaluation conditions

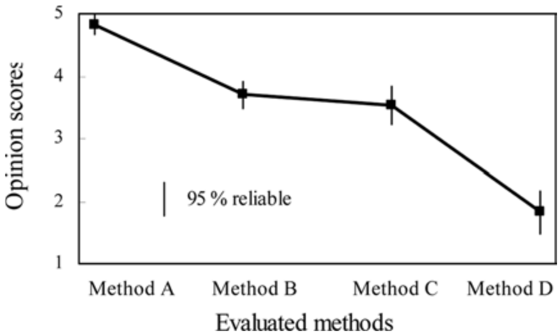

Figure 7: Result of subjective evaluation

Figure 7 shows the result of evaluation. Because the scores of the Method-B and Method-C were as high as that of Method-A (Original) and the score of the Method-D (half resolution) was very low, the spatial pixel-offset method was proved to be effective. Because the Method-C is the most cost effective, this method was adopted to develop the ultrahigh-definition video devices.

Specifications of system

Figure 8 shows the floor plan of an experimental room where the devices are installed. Ultrahigh-definition images are projected onto a screen of 320 inch (4m×7m) and the luminance on the screen is 50 cd/m^2, which is similar to that of a movie theater. Viewers watching images 3 meters away from the screen would enjoy a very high sensation of reality with a viewing angle of 110 degrees. Figure 9 shows the external view of the room and an example of the screen image.

A demonstration video program of 3 minutes was produced. Combined with the three dimensional audio system, the program was exhibited to visitors to NHK Science and Technical Research Laboratories.

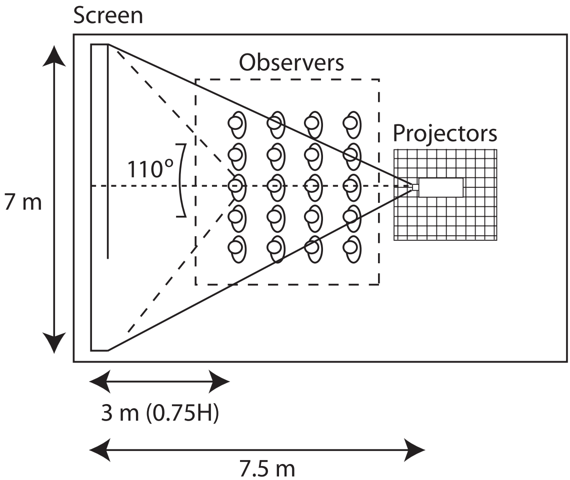

Figure 8: Floor plan

Figure 9: External view

Conclusions

Ultrahigh-definition video systems have a lot of opportunities for various potential applications.

A camera and display with 4k×8k pixels (16 times that of HDTV) was developed to realize nearly the maximum of sensation of reality. There are 4k×8k pixels equivalently for green, while there are 2k×4k pixels for red and blue. A spatial pixel-offset method was used for green.

The ultrahigh-definition video system is still at a basic stage. The performances of the devices must be improved. In particular the camera system should be smaller and be equipped with a zoom lens to be used in program production.

In order to investigate the possibilities for future broadcast services, the authors will make a quantitative evaluation of sensation of reality with the experimental system described in this paper.

Reference

1 N.Narita, et al, 2002. Optimum screen parameters for ultrahigh definition and wide screen image systems, Study of screen aspect ratios and maximum pixels with still image, Journal of ITEJ Vol.56, No.3, pp.437-446, 2002

2 T.Hatada, et al., 1980. Psychophysical analysis of the sensation of reality induced by a visual wide-field display, SMPTE Journal, Vol.89 pp.560-569 August 1980

3 1988. Hi-Vision Technology; Japan Broadcast Publishing Co., Ltd. 1988, pp.19

4 K.Tanaka, M.Satoh and et al., 2001. Q-HDTV projector, ITE Annual Convention 2001 10-7

5 M.Sugawara, et al., 1995. A study on the advantages of spatial-pixel-offset method in four-chip CCD imaging, Journal of ITEJ Vol.49, No.2, pp.212-218 1995

6 K.Mitani, et al., 2002. Image Capturing for Ultrahigh-Definition and Ewide-Screen Video System, Proceeding of IDW'02 IA-1, pp.7-10

7 M.Kanazawa, et al., 2002. A wide-screen projector of 4k×8k pixels, Proceeding of IDW'02 3D3-1, pp.

8 M.Kanazawa, et al., 2003. An Ultrahigh-Definition Display using the pixel offset method, (under reviewed)

9 Methodology for the Subjective Assessment of the Quality of Television Pictures, ITU–R Recommendations, BT.500