Feature

Integral 3-D Television Using a 2000-Scanning Line Video System

We have developed an integral three-dimensional (3-D) television that uses a 2000-scanning line video system. An integral 3-D television system enables capture and display of 3-D color moving images in real time. We previously developed a system that uses a high-definition (HD) television system and reconstructs images using 54 (horizontal) x 59 (vertical) elemental images. To further improve the picture quality, our new system typically uses 6 times as many elemental images, of size 160 (horizontal) by 118 (vertical) arranged at 1.5 times the density. To evaluate the resolution and viewing area characteristics we conducted a test to compare the system with a conventional TV system. In the resolution test, taking the Nyquist frequency response of the image reconstructed on the lens array as a reference, we measured the spatial frequency formed at a response equivalent to this at arbitrary depths. Also, to assess the viewing area, we measured the angular range over which a viewer can move relative to a display device. We confirmed that an image near the lens array can be reconstructed at approximately 1.9 times (283 cpr) the spatial frequency of the conventional system, with a viewing angle that is 1.5 times (12°) wider.

1. Introduction

Integral photography (IP) is a technique for shooting and displaying three-dimensional (3-D) images. The viewer of the reconstructed 3-D images does not have to wear any special viewing glasses, and the appearance of the image changes naturally as the viewer shifts position. Due to these advantages, IP has been studied extensively since it was devised in 1908 by M.G. Lippmann, with the aim of improving image quality. The resolution of the image reconstructed by IP is determined by the pitch of the elemental lenses that make up the lens array and the resolution of image capture and of display media (e.g. film, charge-coupled devices, LCD panels). If film in IP is replaced by electronic media, a real-time 3-D TV system can be constructed (hereinafter "integral 3-D TV"). Previously we examined an integral 3-D TV based on a high-definition-television (HDTV) system and developed a first prototype based on the HDTV system. Here, to evaluate the performance of an integral 3-D TV system featuring a 2000-scanning line video system for enhanced image quality (hereinafter "second prototype"), we measured the resolution characteristics and viewing area characteristics of the system, and compared the results with those of the earlier prototype. We report the results herein.

2. Reconstructed Image Characteristics

2.1 Spatial Frequency and nyquist frequency of reconstructed images

Upon examining the general characteristics of the resolution of the reconstructed images produced by an integral 3-D television, we considered a 3-D image as a depthwise stack of planar images. We thus use MTF (Modulation Transfer Function) at any depthwise position. MTF expresses the response of planar images for each spatial frequency. Also, as the spatial frequency of reconstructed image we use the spatial frequency β (measured in units of cycles/radian, hereinafter "cpr") when the image is viewed by an observer.

Figures 1 and 2 show the arrangement for image capture and display in IP. An object having a spatial frequency of v_c cycles/mm (frequency x_c mm/cycle) is captured through a lens array. In Fig. 1 the spatial frequency per radian when the object is viewed through an elemental lens is expressed as α_c cpr; the frequency of the reconstructed elemental image as e_c mm/cycle; the distance between the object and the lens array as z_c mm; the distance between the lens array and the image capture device as g_c mm; and the pitch of the elemental lenses as p_c mm. Figure 2 shows how the elemental image captured by the capture device is shown on the display device and how the image is reconstructed through the lens array at the spatial frequency of v_d cycles/mm (frequency x_d mm/cycle). In Fig. 2 the frequency of the displayed elemental image is expressed as e_d mm/cycle; the spatial frequency per radian when the reconstructed image is viewed through the elemental lens as α_d cpr; the distance between the lens array and the reconstructed image as z_d mm; the distance between the display panel and the lens array as g_d mm; the pitch of the elemental lens as p_d mm; the distance between the lens array and the observer as L mm; and the spatial frequency per radian when the reconstructed image is viewed by the observer is expressed a β cpr. The higher the spatial frequency β, the greater the accuracy with which the reconstructed image can be observed. In both Figs. 1 and 2 the direction to the left of the lens array is negative and the direction to the right is positive.

In the settings in Figs. 1 and 2 the geometric relationship between the object and the reconstructed image of the object is expressed by the equations below.

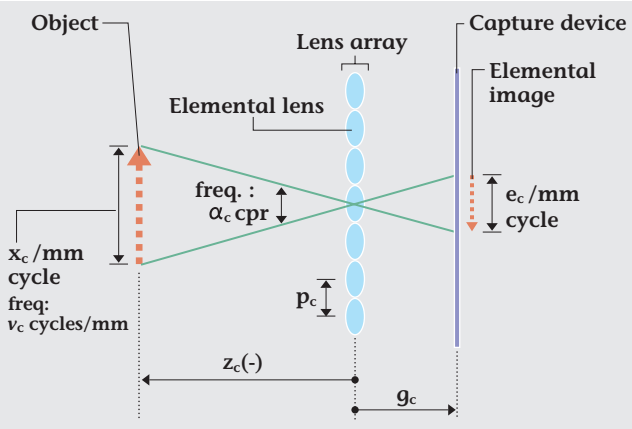

Figure 1: Schematic of IP for capture

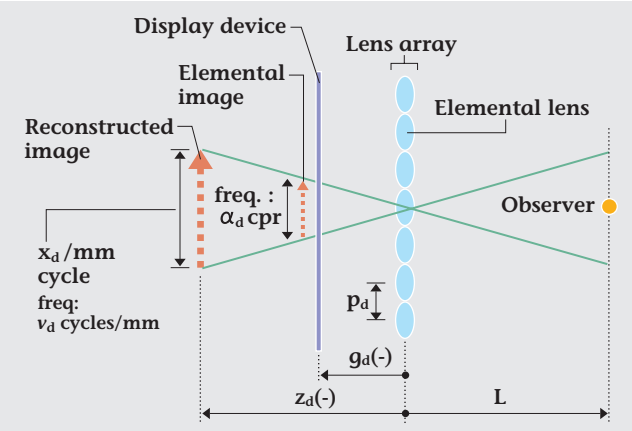

Figure 2: Schematic of IP for display

$$\xi = \frac{X_d}{X_c} = \frac{(p_d / p_c)(e_d / e_c)g_c}{Z_c\{(e_d / e_c) - (p_d / p_c)\} + (e_d / e_c)g_c} \quad (1)$$

$$\eta = \frac{Z_d}{Z_c} = \frac{(p_d / p_c)(-g_d / g_c)g_c}{Z_c\{(e_d / e_c) - (p_d / p_c)\} + (e_d / e_c)g_c} \quad (2)$$

Here, given the following equations

$$a = p_d / p_c \quad (3)$$
$$b = -g_d / g_c \quad (4)$$
$$d = e_d / e_c \quad (5)$$

Eq. (1) and (2) can be expressed as follows:

$$\xi = \frac{X_d}{X_c} = \frac{abg_c}{Z_c(d-a) + dg_c} \quad (6)$$

$$\eta = \frac{Z_d}{Z_c} = \frac{abg_c}{Z_c(d-a) + dg_c} \quad (7)$$

Here, a = d. That is, if the ratios of the elemental lens interval and elemental image between the display side and image capture side are equal, then the relationship between v_c and v_d can be expressed by the equation below, since from Eq. (6) ξ =a.

$$v_d = v_c / a \quad (8)$$

In this case, the spatial frequency β when the reconstructed image is viewed by an observer can therefore be expressed as follows, using the spatial frequencies of the object and the reconstructed image:

$$\beta = v_d(L - Z_d) = v_c(L - Z_d)a \quad (9)$$

Using the value of β expressed in Eq. (9), the MTF (MTF_T (β)) of the reconstructed image at the spatial frequency β cpr in the position zd from the lens array can be expressed by the equation below.

$$MTF_T(\beta) = [MTF_{LC}(\beta) \cdot MTF_{dC}(\beta)] \cdot [MTF_{Ld}(\beta) \cdot MTF_{dd}(\beta)] \quad (10)$$

Here, MTF_{Lc} and MTF_{dc} indicate the MTF of the elemental lens and capture device, respectively, on the capture side, and MTF_{Ld} and MTF_{dd} indicate the MTF of the elemental lens and display device, respectively, on the display side. Now, when focused close to the elemental lens, the value of MTF of the reconstructed image, as expressed by Eq. (10), decreased abruptly at positions away from the focal point. Also, when the capture device and the display device are focused at infinity relative to the elemental lens, the MTF of the reconstructed image has a high value close to the lens array, and decreases gradually over a wide range in the depth direction. For this reason, in this paper we discuss settings for focusing the capture devices and display device at infinity relative to the elemental lenses, assuming that the object exists over a wide range in the depth direction.

In IP, when a reconstructed image with focus at infinity is observed, the resolution of the image is limited not only by the response expressed by Eq. (10) but also by the Nyquist frequency shown below, due to the sampling structure of the lens array.

$$\beta_{nyq} = L / 2p_d \quad (11)$$

Considering this, in this paper, using the MTF of the reconstructed image generated at the Nyquist frequency on the lens array (=MTF_T (β_{nyq})) as a reference, we examine the relationship between spatial frequencies with equivalent MTF and the depthwise positions of the reconstructed image.

2.2 Response on the lens array

If an object exists on the lens array, its image is also reconstructed on the lens array. In this case the individual elemental lenses of the image capture device and the display device do not form elemental images, meaning that the MTF of each lens is zero, regardless of the spatial frequency. As for the entire lens array, however, the MTF

can be assumed to exert the same aperture effects as ordinary television, and the MTF at the Nyquist frequency can be expressed by the following equations:

$$\lim_{Z_c \to 0} MTF_{Lc}(\beta_{nyq}(Z_c)) = \frac{2J_1[(2/\pi)(W_c/P_c)]}{(2/\pi)(W_c/P_c)}$$
$$\lim_{Z_d \to 0} MTF_{Ld}(\beta_{nyq}(Z_d)) = \frac{2J_1[(2/\pi)(W_d/P_d)]}{(2/\pi)(W_d/P_d)} \quad (12)$$

Here, J is the vessel function; w_c and w_d are the aperture diameter of the elemental lenses of the image capture device and display device, respectively.

2.3 Viewing angle

As Fig. 3 shows, in IP the range within which an observer can move (viewing area V) depends on the distance between the elemental image and the elemental lens (g_d) and also the area of the elemental image (w_{el}). The range is largest when the center of each elemental image is viewed by the observer through each corresponding elemental lens. In this case the angle (viewing angle θ) when the image in the viewing area (V) is observed through the center of the lens array can be obtained by the following equation:

$$\theta = 2\arctan[w_{el}/(-2g_d)] \quad (13)$$

The experimental systems discussed in this paper are set up in such a way that the positional relationship between elemental images and elemental lenses satisfies Eq. (13).

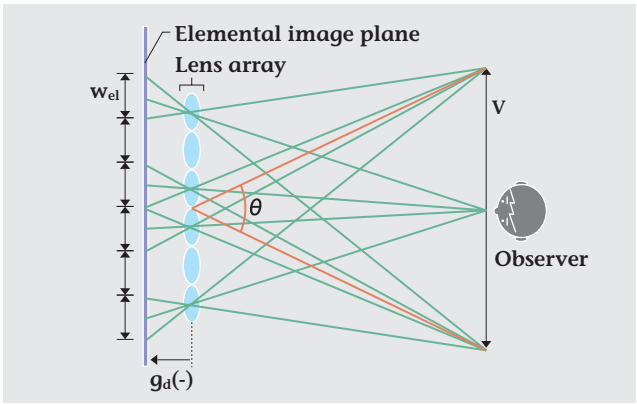

Figure 3: Viewing area of IP

3. Prototype

3.1 Construction and specifications

As shown in Fig. 4, in integral 3-D TV, a television camera is used as the image capture device, and an LCD panel is used as a display device for the real-time capture and display of 3-D images. To avoid the problem of pseudoscopic image formation (where the reconstructed image is inverted depth-wise relative to the object) the lens array in the image capture system consists of gradient-index lenses, and a depth-control lens is inserted between the object and the lens array so that 3-D images can be formed both in front of and behind the lens array on the display side.

Our second prototype, featuring a 2000-scanning line video system, was constructed and uses a 3-CCD (approx. 8 million pixels per image capture device) camera for image capture. In Table 1 we show the specifications of

Table 1: Specification of the camera system

Parameters	3-CMOS[a] Camera System
System	3840×2160 pixels/frame 60 frames/s, progressive scanning
Image senser	1.25-in.[b] 3840×2160 pixels CMOS
Imaging method	Three-panel imaging (GBR[c])
Lens	f = 63 mm, F5.6

[a] Complementary metal-oxide semiconductor
[b] 1in.=2.54cm
[c] Green, blue, red

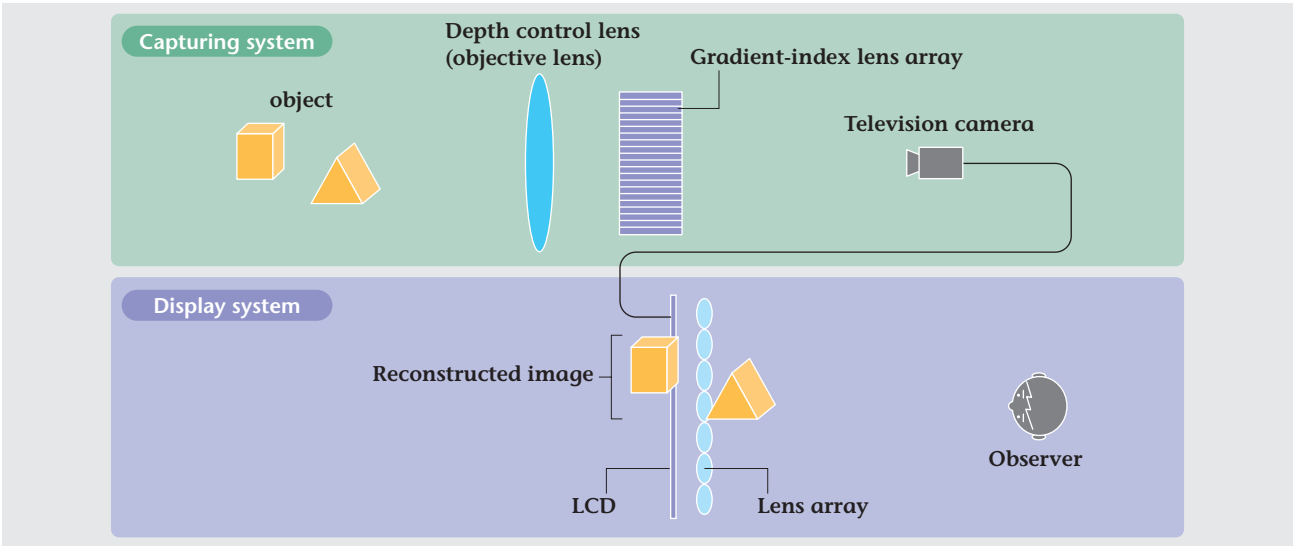

Figure 4: Schematic diagram of the integral 3-D TV system

Table 2: Specification of the integral 3-D television

Capturing system					
System	Television camera (Active pixels)	Gradient-index lens array			
		Diameter (mm)	Number of lenses	Focal length (mm)	Arrangement
First	Approx. 1000 (H) × 1000 (V)	1.085	54 (H) × 59 (V)	-2.65	Offset (delta)
Second	Approx. 3200 (H) × 2160 (V)	1.085	160 (H) × 118 (V)	-2.65	Offset (delta)

Display system								
System	LCD			Lens array				
	Pixel width (mm)	Active pixels	Picture hight (mm)	Diameter/ pitch (mm)	Number of lenses	Focal length (mm)	Arrangement	Nyquist frequency (cpr)
First	0.21	Approx. 1000 (H) × 1000 (V)	201	3.5/4.02	54 (H) × 59 (V)	19.8	Offset (delta)	150
Second	0.1245	Approx. 3200 (H) × 2160 (V)	249	2.64/2.64	160 (H) × 118 (V)	8.58	Offset (delta)	280

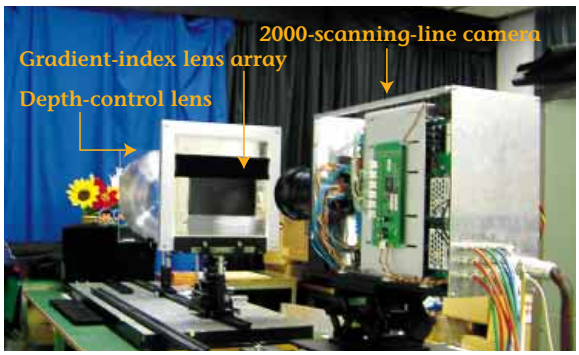

Figure 5: (Color online) Experimental setup for capture (second system)

(a) An example of reconstructed image (first system)

(b) An example of reconstructed image (second system)

Figure 6: (Color online) Example of reconstructed images

the three-CCD camera, and in Fig. 5 we give an outline of the image capture system of the second prototype. In Table 2 we show the specifications of our first and second prototypes. The Nyquist frequency in Table 2 was calculated assuming that the viewing distance is six times the display height. In order to improve the resolution, the second prototype has 6 times more effective pixels and elemental lenses than the first system, and a narrower pixel pitch and elemental lens pitch in the display device. Also, in order to increase the viewing angle, as expressed by Eq. (13), in the second prototype we reduced the focal length of the elemental lenses used in the display device. Figure 6 shows reconstructed images produced by the first and second prototypes. Figure 7 shows how images produced by the second prototype change according to the positions of the observer.

(a) Objects

(c) Upper viewpoint

(e) Left viewpoint

(b) Center viewpoint

(f) Right viewpoint

(d) Lower viewpoint

Figure 7: (Color online) Changes in reconstructed images viewed from different positions (second system)

3.2 Geometric Relationship Between Object and Reconstructed Image

In Table 3 we show the geometric relationship between the object and reconstructed image for the first and second prototypes, as calculated from Table 2 and Eqs. (1) and (2).

Table 3: Geometrical relation between an object and a reconstructed image

Magnification	First system	Second system
ξ : Lateral	3.71	2.32
η : Depth	7.47	3.24

4. Comparison Experiments

4.1 Resolution characteristics

As described above, when the image capture system and display system are focused at infinity, the spatial frequency of the reconstructed image is limited by the Nyquist frequency, which is determined by the lens pitch. In this experiment the MTF of the Nyquist frequency of the reconstructed image formed on the lens array (=MTF_T (β_{nyq})) was used as the reference level. We then measured the depth-wise changes of the spatial frequency having the same MTF and compared the results of the first and second prototypes. Figure 8 shows the experimental setup used to measure resolution characteristics. In the figure, the distance L between the lens array and the measurement camera is set at six times the display height (H), since the display height is assumed at the same angle. Using the equipment shown in Fig. 8 we measured the resolution characteristics as follows. Firstly, a sine-wave pattern of the Nyquist frequency is formed on the lens array, and its degree of modulation is measured using a measurement camera and oscilloscope positioned at a point 6H from the lens array. Next, spatial frequencies having a sine-wave pattern with the same degree of modulation are measured at intervals of 0.25 H, between image depth positions of -1.5 H (behind the lens array) to 1.5H (in front of the lens array). The relationship of the reproduced pattern and the object arrangement pattern is shown in Table 3. Measurement was carried out using the central part of the display screen.

Considering only the lens array, the MTF values of the reconstructed images at the Nyquist frequency, as given by Eq. (12), were 0.57 and 0.52 for the first and the second prototypes, respectively. The values plotted in Fig. 9 (black dots) indicate the measurements of the spatial frequencies of equivalent MTF. The result shows that the second prototype exhibits higher spatial frequencies than the first prototype within the range of $0.78H_2$ in front of

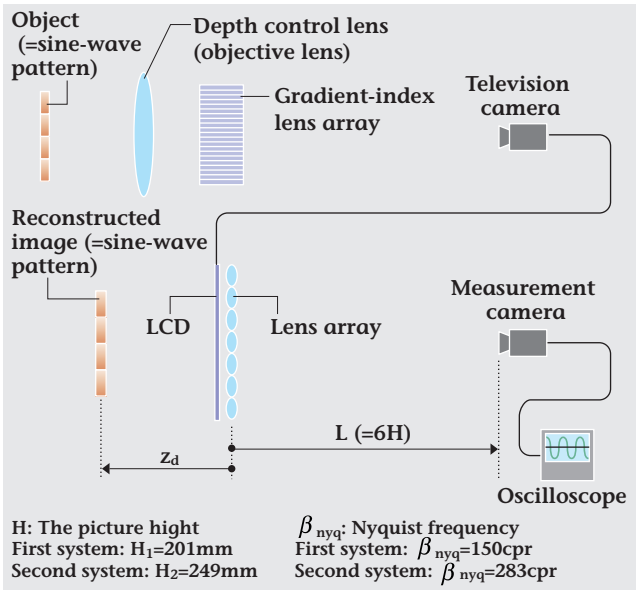

Figure 8: Experimental setup for measuring the resolution characteristics

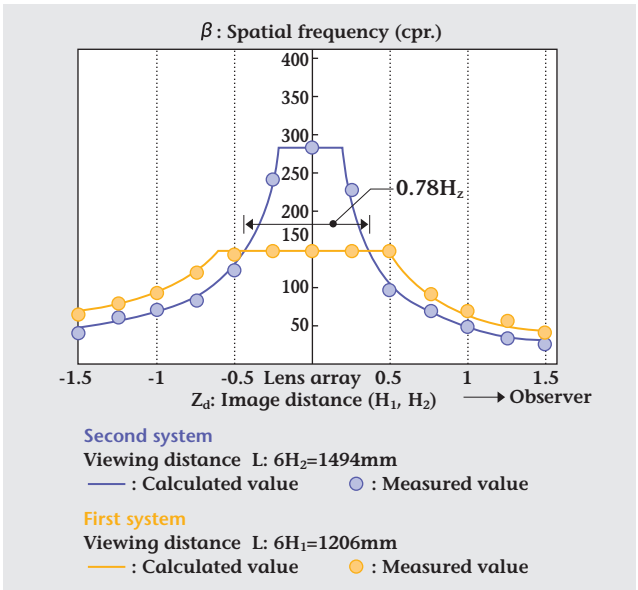

Figure 9: Experimental result of the resolution characteristics

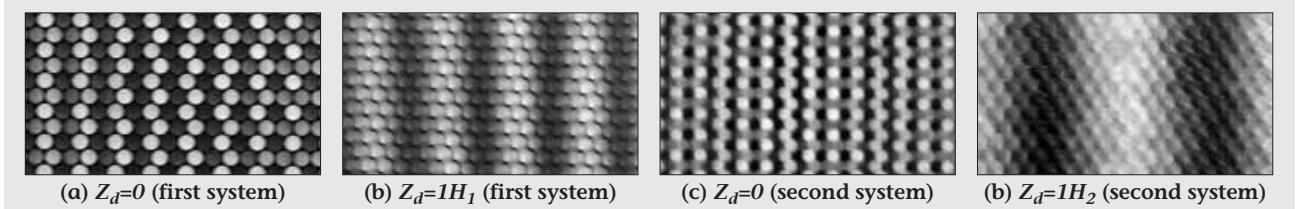

Figure 10: Example of the reconstructed images of test patterns

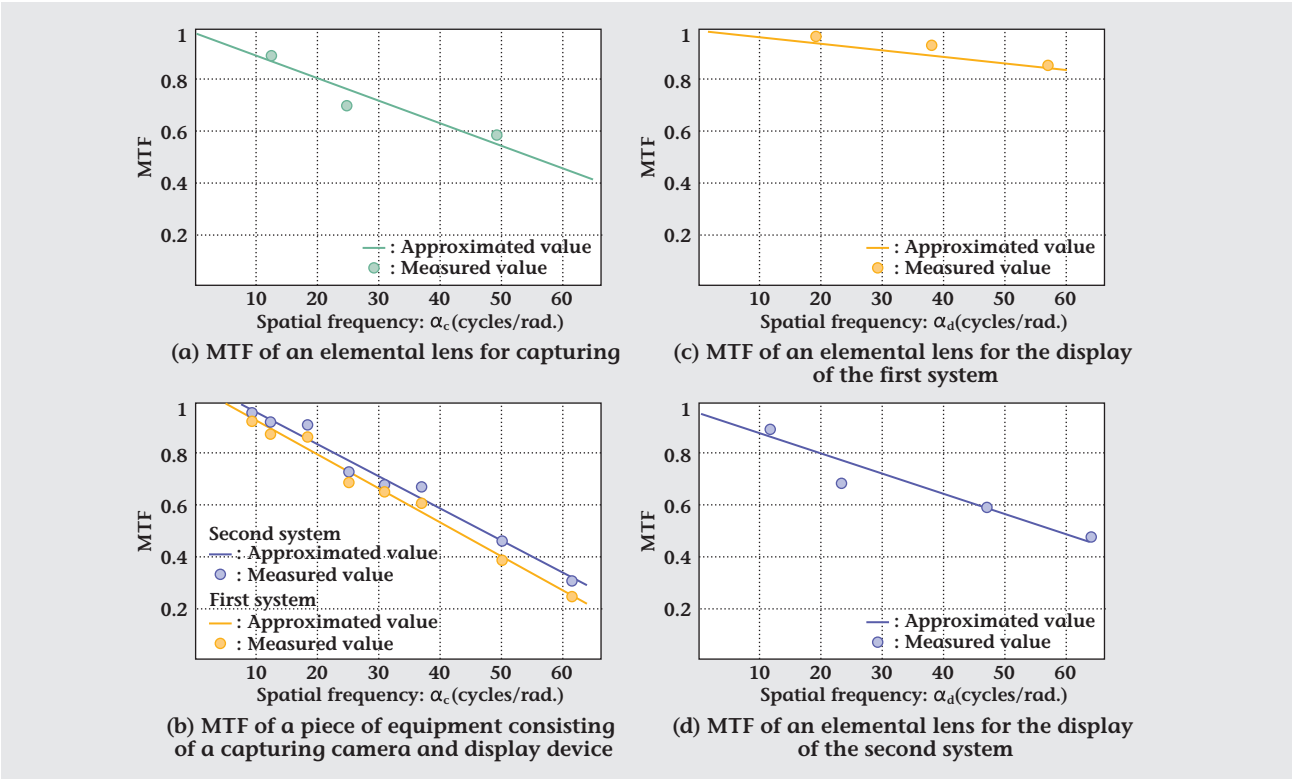

Figure 11: MTFs of an elemental lens, capturing camera, and display device

and behind the lens array (H_2 denotes the effective display height of the second prototype). Close to the lens array, notably, it is possible to reconstruct an image at 1.9 times the spatial frequency. Also, in both the first and the second prototypes, the spatial frequency tends to decreases as an image moves further away from the lens array. Figure 10 shows how the pattern used for the measurement is reconstructed. The reconstructed images formed on the lens array and at a point 1H from the lens array are shown in (a) and (b) for the first prototype, and (c) and (d) for the second prototype. Comparing the images reconstructed on the lens array in Fig. 10 too, it is clear that for both the first and second prototypes, the images formed at position 1H from the lens array have a lower spatial frequency.

The values in Fig. 9 (solid line) are calculated as the product of each of the individually measured MTFs of the elemental lenses of the image capture device, the system comprising the capture camera and display panels, and the elemental lenses of the display device. The calculation method is described below. Figure 11 shows the measured MTFs for the elemental lenses of the image capture device, the system made up of a capture camera and display panels, and the elemental lenses of the display device. In Fig. 11 the MTF is expressed on the vertical axis, the spatial frequency α_c during image capture on the horizontal axis in (a) and (b), and the spatial frequency α_d during display on the horizontal axis in (c) and (d). The measurements are plotted and approximated by a straight-line function.

1st prototype

$$MTF_{Lc}(\alpha_c) = -0.0086\alpha_c + 0.9794$$
$$MTF_{dev}(\alpha_c) = -0.066\alpha_c + 1.0675$$
$$MTF_{Ld}(\alpha_d) = -0.0827\alpha_d + 1.0116$$

2nd prototype

$$MTF_{Lc}(\alpha_c) = -0.0086\alpha_c + 0.9794$$
$$MTF_{dev}(\alpha_c) = -0.062\alpha_c + 1.0888$$
$$MTF_{Ld}(\alpha_d) = -0.0879\alpha_d + 0.973 \quad (14)$$

In Fig. 11(b), as the MTF of the capture camera and display panel, the resolution pattern is positioned where an elemental image is formed during capture; the image captured by the camera is shown on the display panel; and the degree of modulation of the image is measured. The results therefore reflect the resolution characteristics of both the capture camera and the display panel. In Figs. 11(a), 11(c), and 11(d), the results of measuring the MTF of elemental lenses in the focused state are shown. Next, using the measurement results in Fig. 11, we calculated the spatial frequency that satisfies the following equation.

$$MTF_T(\beta_{nyq}) = MTF_{Lc}(\alpha_c) \cdot MTF_{dev}(\alpha_c) \cdot MTF_{Ld}(\alpha_d) \quad (15)$$

Note that MTF_{Lc}, MTF_{dev} and MTF_{Ld} express the MTFs of the elemental lenses for capture, the system comprising the capture camera and display panel, and the elemental lenses for display, respectively.

From Eqs. (14) and (15) the spatial frequencies of the elemental images required in order to have same MTF as that of the reconstructed image formed on the lens array

at the Nyquist frequency are α_c =6.67cpr and α_d =13.45cpr for the first prototype and α_c =7.35cpr and α_d =9.78cpr for the second prototype. That is, if the spatial frequency of the elemental image is higher than these values, it is not possible to reconstruct an image with the same MTF as that of the Nyquist frequency image on the lens array. The solid lines in Fig. 9 represent the spatial frequency β of the image reconstructed by the elemental image with the spatial frequency that satisfies Eqs. (14) and (15). However, for spatial frequencies higher than the Nyquist frequency, the spatial frequency is substituted by the Nyquist frequency. If an image is reconstructed at a distance from the lens array, its spatial frequency is lower for the second prototype than for the first one. This is because the upper limit of the spatial frequency passing through the elemental lenses of the display system is α_d =13.45 cpr for the first prototype and α_d =9.78cpr for the second prototype. To remedy this situation, it is necessary to increase the resolution of the elemental lenses of display device in the second prototype.

4.2 Viewing area characteristics

The viewing angle is calculated from Eq. (13) and Table 2 to be 11.6° and 17.5°, respectively, for the first and the second prototypes. The actual measurement results were approximately 8° for the first prototype and 12° for the second, which correspond to 70% of the calculated values. With the observer at a distance of 6 times the effective display height from the lens array, the viewing area (*V* in Fig. 3) was 170 mm for the first prototype and 320 mm for the second. The reason why the measured viewing angles are 70% of the calculated values is due to displacement in the lens position. There is no such discrepancy between the measured and calculated values of resolution, because resolution is less sensitive to the precision of lens position, since it is measured in the central part of the display area. At the same time, the viewing area deteriorates if positional accuracy is not maintained over the whole lens array. Considering, however, that the percentage difference between the measured and calculated values is approximately the same for the first and second prototypes, we can conclude that the larger lens array used in the second prototype has approximately the same positional accuracy as that used in the first system.

4.3 Discussion

In the experiment described in this chapter, the response of the reconstructed image formed on the lens array at the Nyquist frequency, which is determined by the pitch of elemental lenses, is used as a reference, and the spatial frequency with the same response is measured at different depth-wise positions. In an experiment to compare our first and second prototypes, we found that the second protype can reconstruct an image near the lens array at a spatial frequency 1.9 times that of the first system. When the image was at a significant distance from the lens array, however, the spatial frequency was lower in the case of the second prototype than the first. In order for the second prototype to reconstruct images at the same spatial frequency as that of the first prototype even at a distance from the lens array, it is necessary to improve the spatial frequency characteristics of the elemental lenses used in the display device. At present the MTF is 0.9 at a spatial frequency of 9.78 cpr. We need to achieve the same MTF, 0.9, when the spatial frequency is 13.45 cpr.

As for viewing angle, the second prototype offers a view that is 1.5 times wider than the first prototype. The actual performance, however, is not exactly as designed. For better viewing area characteristics, the positional accuracy of the elemental lenses that make up the lens array needs to be improved.

5. Conclusion

In this report we discussed our latest integral 3-D TV system, featuring a 2000-scanning line video system. We conducted an experiment to measure the resolution and viewing area of this latest prototype and compared the results with those of a previous 3-D TV system based on a HDTV system. As a result, we confirmed that the new system is superior to the previous one in terms of the resolution of the reconstructed image and the viewing area, due to the higher resolution of the lens array, the display panel, and the image capture camera. However, the resolution of the system is not high enough for the receivers used for NTSC broadcasting. Nevertheless, our latest prototype can reconstruct an image near the lens array with 1.9 times the spatial frequency (283 cpr) of the previous prototype, and a viewing angle that is 1.5 times wider (12°). To increase the resolution of images reconstructed at a distance from the lens array, it is necessary to improve the MTF of the elemental lenses of the display system. The measured viewing angle was 70% of the theoretical calculated value. This was due to some positional inaccuracy in the elemental lenses that make up the lens array. Thus, the positional accuracy of lens array needs to be improved further.

3-D video systems need to present more information to observers than normal 2-D video systems. They therefore need more pixels and higher resolution for both image capture and display. Optical components also need to have higher resolution at higher spatial frequencies. We intend to continue our efforts to make overall improvements to 3-D video systems, in order to produce superior 3-D motion images.

(Jun ARAI, Takayuki YAMASHITA, Makoto OKUI and Fumio OKANO)