Glasses-free large-screen three-dimensional display and super multiview camera for highly realistic communication

Masahiro Kawakita
Shoichro Iwasawa
Roberto Lopez-Gulliver
Naomi Inoue
Glasses-free large-screen three-dimensional display and super multiview camera for highly realistic communication

Masahiro Kawakita,a,* Shoichiro Iwasawa,b Roberto Lopez-Gulliver,c and Naomi Inoued

aJapan Broadcasting Corporation, Science and Technology Research Laboratories, Tokyo, Japan
bNational Institute of Information and Communications Technology, Universal Communication Research Institute, Tokyo, Japan
cRitsumeikan University, Department of Information Systems Science and Engineering, Shiga, Japan
dKDDI Research Inc., Saitama, Japan

Abstract. We studied highly realistic communication systems using a super multiview three-dimensional (3-D) video system. We propose the developed glasses-free 200-in. 3-D display using ∼200 high-definition (HD) projector units to reconstruct natural life-size 3-D moving objects, such as cars and humans. We also analyzed the optimal arrangement of the multiview camera system for the 3-D display. In the experiments, we developed a prototype super multiview 3-D camera system using compact HD video cameras with real-time convergence compensation circuits that correct the captured images via image processing optimized for the display. We also performed demonstration experiments using the developed 3-D display and camera system in a public area. As a result, we highlight several possibilities for actual applications of the glasses-free super multiview 3-D video system. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.OE.57.6.061610]

Keywords: glasses-free; three-dimensional display; large-screen; three-dimensional camera; multiview camera.

1 Introduction

We conducted research on highly realistic communication systems using natural three-dimensional (3-D) images that do not require special glasses for viewing. Highly realistic communication systems will enable communication with people in remote locations as if one were actually in their presence. Such systems would allow us to present, appreciate, and learn from valuable cultural assets as if they were actually present, greatly enriching our lives.

In recent years, augmented reality (AR) and virtual reality (VR), which are highly realistic technologies, have spread throughout games, entertainment, and other applications, and new sensors, displays, and other equipment for general consumer use are now being commercialized. At the same time, studies are also underway on next-generation highly realistic 3-D images that do not require special glasses to reproduce multiple parallax images and groups of light rays in an attempt to achieve 3-D image displays for practical use. In addition to the technology for capturing and displaying light fields, comprehensive efforts are now under way to proceed with the standardization of image formats, compression coding, and evaluation methods.

Natural 3-D image reconstruction technology without special glasses is undergoing steady progress. An integral 3-D system based on integral photography is classified as a 3-D system of the spatial image reconstruction type. An integral 3-D system that features horizontal as well as vertical parallax is being developed for future TV broadcasting. An integral 3-D system requires many pixels to achieve high image quality. Therefore, ultrahigh-definition (UHD) video with 8 K resolution is used to capture and display integral 3-D images consisting of ∼100,000 pixels. Furthermore, multiple display devices are used to enhance the resolution of integral 3-D images. Lee et al.12 studied an optimal configuration design of a simple multiprojection 3-D display using only a vertical diffuser film for the screen. However, uneven brightness in 3-D images is likely to occur, which is reduced by suitable arrangements of the projectors. Efrat et al.13 developed a 3-D display for the viewing environment similar to that at a movie theater. Here, a 3-D image can be displayed to observers at their respective seat positions of different viewing distances using slanted barriers in front of a flat panel display. However, the viewing area is limited to the width of the seat, and it is difficult to display 3-D images with a large viewing angle.

In this study, to realize a highly-realistic communication system based on multiview 3-D video, we proposed a projection-type 3-D display method that can display glasses-free 3-D images on a large screen using multiple projectors. We have designed and fabricated a 200-in. 3-D display system using ∼200 HD projector units to enlarge the viewing area and to enable many observers to simultaneously see 3-D images from various angles. The resolution and color reproducibility of the 3-D images displayed are of HD image quality.

We have also developed a super multiview 3-D capture system with compact HD camera units that include
real-time video signal correction circuits for camera calibration and image rectification. First, we analyzed the optimal camera arrangements for multiprojection-type 3-D displays. Second, based on experiments, we developed a super multiview camera system using ~200 HD cameras to capture moving objects. The registrations of captured images were arranged by high-accuracy camera adjustment and geometrically compensated by image processing for the multiprojection 3-D display. Experimental results showed that we successfully captured and displayed 3-D videos of moving objects using our system.

2 System Design of Multiview Three-Dimensional Video

2.1 Basic System Configuration

Figure 1 shows the basic configuration of our multiview 3-D video capture and display system. A projector array is used as an image display device and consists of several small projectors called “projector units” arranged horizontally and vertically. From each projector unit, images with horizontal parallax are projected and superimposed on the plane for image display that combines an optical film that has anisotropic diffusion characteristics with a condenser lens. The diffuser film is a rear screen that has a small diffusion angle in the horizontal direction relative to the incident light, and a wide diffusion angle in the vertical direction. These diffusion characteristics enable the system to produce different images at various horizontal angles, thereby allowing the observer to see parallax images based on the observer’s position. Conversely, in the vertical direction, the incident light becomes widely diffused, thereby eliminating the effects of the projection angle in that direction. One can therefore arrange the projector units vertically, so that increasing the number of units increases the image density, which is defined as the horizontal number of parallax images per unit viewing angle. Increasing the parallax image density can be expected to produce a high-quality 3-D display at high resolution with smooth motion parallax over a wide viewing zone.

A 3-D camera system consists of numerous camera units arranged in a horizontal row, unlike the projector unit, so as not to cause vertical parallax. It is difficult to arrange the camera pitch closely by several cm or less due to limitations imposed by the size of the cameras. Therefore, we generate interpolated images to increase the parallax. Image quality correction using high-accuracy adjustment of camera calibration and geometrical compensation are common problems in the multiview 3-D camera method. In this work, we study suitable 3-D camera conditions for a 3-D display system using multiprojectors and an optical screen consisting of an anisotropic diffusion film and a Fresnel lens.

2.2 Optimal Arrangement of Projector Array and Camera Array

First, we study the camera and display system arrangement to reconstruct the correct viewing angle and visual field. Figure 2 shows the viewing angle and visual angle in our multiview 3-D display system. The optical axes of images projected from the projector units converge at one point in the center of the display screen as shown by the dashed lines in the figure. All projector units have a projection lens that moves laterally to shift the optical axes of the light sources. In the figure, the viewing angle \( \theta_d \) is expressed as

\[
\theta_d = 2 \arctan\left( \frac{L_d}{2a_d} \right).
\]

where \( L_d \) is the width of the projector array, and \( a_d \) is the distance between the projector array and the condenser lens.

The width of viewing area \( V \) is expressed as

\[
V = \frac{L_d}{a_d} z_0.
\]

where \( z_0 \) is the distance between the display screen and an observer. In this method, the interval between parallax images at the observer’s position should be less than the distance between human eyes, \( e \approx 65 \text{ mm} \). Therefore, the number of parallax images \( N_p \) required is

\[
N_p \geq \frac{V}{e} = \frac{L_d z_0}{a_d e}.
\]

For example, in the case of \( L_d = 2 \text{ m}, a_d = 2.5 \text{ m}, \) and \( z_0 = 2.5 \text{ m} \), we require a minimum of 30 parallax images.

On the other hand, the visual angle \( \Omega_d \) in Fig. 2 is expressed as

\[
\Omega_d = 2 \arctan\left( \frac{W_d}{2z_0} \right) \approx \frac{a_d}{z_0} \varphi_d,
\]

where \( W_d \) is the total width of the projected images and \( \varphi_d \) is the projection angle. From Eq. (4), the visual angle \( \Omega_d \) is proportional to the projection angle \( \varphi_d \) and projection distance \( a_d \), and is inversely proportional to the viewing distance \( z_0 \).

Next, we consider the capture angle in the 3-D camera system. In Fig. 3, the dashed lines show the optical axes.
of each camera that converge at point $O$ in plane $P$. The convergence angle $\theta_c$ is thus expressed as

$$\theta_c = 2 \arctan \left( \frac{L_c}{2a_c} \right),$$  \hspace{1cm} (5)$$

where $L_c$ is the width of the camera array, and $a_c$ is the distance from the camera array to convergence plane $P$. For the condition of when the capture angle $\phi_c$ is equal to the viewing angle $\theta_d$, we obtain

$$\frac{L_c}{a_c} = \frac{L_d}{a_d}.$$  \hspace{1cm} (6)$$

For the condition of when capture angle $\phi_c$ is equal to visual angle $\Omega_d$, we obtain

$$\frac{W_c}{a_c} = \frac{W_d}{z_d}.$$  \hspace{1cm} (7)$$

From Eq. (6) to reconstruct the correct viewing angle in the display system, i.e., $L_d/L_c$ (ratio of the width of the projector array and one of the camera arrays) must be equal to $a_d/a_c$ (ratio of the projector distance and camera-to-object distance). From Eq. (7), to reconstruct the correct visual angle in the display system, i.e., $W_d/W_c$ (ratio of the size of the projected image and one of the capture images) must be equal to $z_d/a_c$ (ratio of the viewing distance and one of the capture distances).

### 2.3 Resolution Characteristics of Three-Dimensional Video System

We estimate the required resolution of the 3-D system by correlating the Nyquist frequency of the sampling and the resolution of the camera and display.$^{14,15}$

Camera units are arranged so that the optical axes converge at one point $O$, as shown in Fig. 4. A grating pattern is placed at a distance $z_i$ from the convergence plane. With one cycle of the grating being $(a_c - z_i)/f_c$, the spatial frequency of the grating from each camera is $f_c$ [cycle/rad]. The spatial frequency $f_o$ [cycle/rd] of the grating viewed from point $O$ on the convergence plane is then expressed as

$$f_o = \frac{f_c}{a_c - z_i}.$$  \hspace{1cm} (8)$$

When the camera units are spaced at a pitch $p_c$, the angular pitch of sampling is $p_c/a_c$ rad, and the Nyquist frequency $f_{\text{cqyq}}$ [cycle/rd] of the sampling is $a_c/2p_c$. To prevent aliasing with this sampling, the following condition must be met:

$$|f_o| < f_{\text{cqyq}} = \frac{a_c}{2p_c}.$$  \hspace{1cm} (9)$$

From Eqs. (8) and (9), the condition to prevent aliasing is expressed as

![Fig. 2 Configuration of 3-D display system.](image-url)

![Fig. 3 Configuration of 3-D capture system.](image-url)
The resolution \( f_c \) is also limited by the camera’s maximum frequency \( f_{c\text{ max}} \) that depends on the pixel pitch of the capture device.

\[
f_c \leq \frac{a_c}{2p_c} \left| \frac{a_c - z_i}{z_i} \right|. \tag{10}
\]

The resolution \( f_c \) is also limited by the camera’s maximum frequency \( f_{c\text{ max}} \) that depends on the pixel pitch of the capture device.

Next, the sampling effect of the display system, as well as the camera system, is examined. Projector units are arranged to converge and the optical axes converge at the center of the display screen, as shown in Fig. 4. A grating pattern is displayed at a distance \( z_i \) from the display screen. The spatial frequency of the grating viewed from the observer’s position is \( f_i \) [cycle/\( \text{rad} \)]. The spatial frequency \( f_d \) of the grating viewed from the display screen is expressed as follows:

\[
f_d = f_i \frac{z_i}{z_o - z_i}. \tag{11}
\]

When the projector units are spaced at a pitch \( p_d \), the angular pitch of sampling is \( p_d/a_d \text{ rad} \), and the Nyquist frequency \( f_{\text{d Nyq}} \) of the sampling is \( a_d/2p_d \). To prevent aliasing with this sampling, the following condition must be met:

\[
|f_d| < f_{\text{d Nyq}} = \frac{a_d}{2p_d}. \tag{12}
\]

From Eqs. (11) and (12), the condition to prevent aliasing is expressed as

\[
f_d < \frac{a_d}{2p_d} \frac{z_o - z_i}{z_i}. \tag{13}
\]

The resolution \( f_d \) is also limited by the display’s maximum frequency \( f_{d\text{ max}} \), which depends on the pixel pitch of the display device.

Figure 5 shows the relationships between the depth position of the object and the capture spatial frequency and depth position of the reconstructed image and the display spatial frequency. From Fig. 5, we examine the required camera conditions.

The maximum spatial frequency of the camera system must be higher than that of the display system

\[
f_{c\text{ max}} \geq f_{d\text{ max}}. \tag{14}
\]

The depth positions of \( z_{cf}, z_{df}, z_{cn}, \) and \( z_{dn} \) shown in Fig. 5 must meet the following conditions:

\[
z_{cf} \leq z_{df}. \tag{15}
\]

3 Characteristics of Three-Dimensional Display Using Projector Array

3.1 Gap of Parallax Images

Figure 6(a) shows an overhead view of the basic arrangement of the 3-D system using projector units, where the number of projector units is almost the same as the number of horizontal parallax images. Thus, the horizontal pitch \( p_d \) between the projector units is set to a value that matches the number of parallax images per horizontal unit angle. The viewing angle \( \phi_d \) of the 3-D image displayed at the center of the screen in Eq. (1) will become

\[
\theta_d = 2 \arctan \left( \frac{Np_d}{a_d} \right), \tag{19}
\]

where \( N \) denotes a natural number, and the total number of projector units is \( 2N + 1 \). Distance \( z_o \) between the condenser lens and a suitable viewing position has the relationship
\[
\frac{1}{a_d} + \frac{1}{z_o} = \frac{1}{f}. \tag{20}
\]

where \(f\) is focal length of condenser lens. Therefore, the viewing area \(V\) in Eq. (2) is expressed as

\[
V = 2 \left| \frac{f}{a_d - f} \right| N p_d. \tag{21}
\]

Equations (19) to (21) make it possible to determine the number of projector units \(2N + 1\), arrangement pitch \(p_d\), distance \(L_d\) between the projector array and display screen, and other parameters based on the viewing angle \(a_d\) for the 3-D image, viewing area \(V\), and other factors that are required for the 3-D image. These parameters are used as guidelines during system design.

The arrangement of the projector array is limited by the size of each unit, thereby making it difficult to arrange the projectors densely in the horizontal direction alone. For this reason, the system is composed not only horizontally but also vertically, in the 2-D manner as shown in Fig. 6(b).

The gap between displayed parallax images \(p'\) is expressed as

\[
p' = \left| \frac{f}{a_d - f} \right| p_d. \tag{22}
\]

The ideal gap of parallax images \(p'\) is less than the size of a pupil, but it is difficult to create a practical 3-D system based on this requirement. The gap of parallax images in our previous system was 29.4 mm, which could be used to display 3-D images with natural motion parallax.\(^6\) We designed the system parameters in the proposed 200-in. 3-D display system to be \(p' = 22.8\) mm, which is approximately one-third of the interpupillary distance, to enhance smooth motion parallax in 3-D images. The number of required projectors was estimated from the gaps of the parallax, the optimal viewing distance, and the required viewing angle.

3.2 Uniformity of Three-Dimensional Image’s Brightness

We experienced problems related to the appearance of stripe noise, i.e., nonuniform brightness and color areas in 3-D images, reduced 3-D image resolution, and unnatural images owing to observer movements, which made it difficult to increase the screen size using conventional methods. Stripe noise occurred because of the angle pitch of the parallax images, uniform brightness and color in projected images, and diffusion angle of the diffuser film.

As shown in Fig. 7(a), the incident light from a projector unit is emitted at a horizontal angle \(\theta_m\) relative to the optical axis of the condenser lens. Let the horizontal diffusion be a Gaussian distribution; the optical distribution \(i_m\) of the output beams from the \(m\)'th projector unit can then be expressed as

\[
i_m = \frac{1}{\sqrt{2\pi}\sigma_x} \exp \left[ -\frac{(\theta - \theta_m)^2}{2\sigma_x^2} \right]. \tag{23}
\]

where \(\sigma_x\) denotes the standard deviation of the distribution of optical diffusion angles, and it represents the size of the diffusion angle. This is known as the diffusion angle in this paper. Incident light from the group of projector units is emitted and superimposed as shown in Fig. 7(b), where the optical distribution \(I(\theta)\) is given by

\[
I(\theta) = \frac{N}{\sqrt{2\pi}\sigma_x} \sum_{m=1}^{N} \exp \left[ -\frac{(\theta - \theta_m)^2}{2\sigma_x^2} \right]. \tag{24}
\]

The flatness and crosstalk in the brightness distribution between parallax images will be affected by the screen’s horizontal diffusion characteristics in a 3-D image synthesized from parallax images. A small diffusion angle causes irregular brightness between parallax images in terms of the brightness value of the synthesized light shown in Fig. 7(b), resulting in uneven brightness in the image. In contrast, increasing the diffusion angle increases the proportion of the crosstalk area in Fig. 7(b).

For example, let the number of parallax images be 50. First, we evaluate the degree of synthesized light brightness modulation using \(M = (I_{\text{max}} - I_{\text{min}})/(I_{\text{max}} + I_{\text{min}}) \times 100\%\) (where, as shown in Fig. 7(b), \(I_{\text{max}}\) denotes the maximum brightness and \(I_{\text{min}}\) the minimum brightness), before estimating the crosstalk rate as the proportion of the crosstalk relative to the entire output beam. The degree of modulation...
$M$ [%] of uneven brightness relative to the diffusion angle $\sigma_x$ and crosstalk rate $R$ [%] is shown in Fig. 8.

When the horizontal diffusion angle $\sigma_x$ is small, the degree of modulation ($M$) of the synthesized beam will be greater, thereby resulting in more uneven brightness. In contrast, when $\sigma_x$ is high, the crosstalk ($R$) of the output beam will be higher between parallax images. Based on Fig. 8, we can balance the uneven brightness and crosstalk, where the appropriate diffusion angle $\sigma_x$ can be projected in a range from 0.25 to 0.3. This corresponds to a half-value angle of 0.5 deg to 0.7 deg in the diffusion angle distribution of the beams. Another method involves selecting the optimal diffusion characteristics from these relationships to improve the screen material and structure, thereby optimizing the diffusion angle distribution characteristics of the beams.

In terms of the vertical direction shown in Fig. 9(a), the angle of the beam that enters the diffuser screen varies depending on which projector unit is involved, i.e., upper or lower. Consequently, a difference also arises in the diffusion direction of the beam depending on whether an upper or lower projector unit is involved. When observed from position $y_0$, as shown in Fig. 9(a), the distribution $I(y)$ of the brightness values in the $y$-axis direction of the screen (as viewed by the observer) is as follows:

$$I(y) = \frac{1}{\sqrt{2\pi}\sigma_y} \sum_{m=-N_y}^{N_y} \exp \left[ -\frac{(y - y_0)^2}{2\sigma_y^2} \right]$$

$$\approx \frac{1}{\sqrt{2\pi}\sigma_y} \sum_{m=-N_y}^{N_y} \exp \left[ -\frac{(y - y_0)^2}{2\sigma_y^2} \right],$$

where $2N_y + 1$ is the number of lines in the projector array in the vertical direction, $\theta_{me}$ denotes the irradiation angle of the beam in the vertical direction, $\theta_o$ is the angle of the beam from each image location to the observer as shown in Fig. 9(a), and $\sigma_y$ is the diffusion angle in the vertical direction. For example, if we assume that $N_y = 3$, $a_d = 2.5$ m, $p_x = 0.15$ m, $z_0 = 3$ m, and $y_0 = 0$ m, then the distribution of brightness values in the $y$-axis direction of the display image will be as shown in Fig. 9(b).

As indicated in Eq. (25), the image is shaded depending on location $y$ of the image, location $y_0$ of the observer, and location $mp_x$ of the projector unit. Note that a comparatively large diffusion angle characteristic is needed to obtain an image with a brightness distribution that provides a certain degree of uniformity in the vertical direction.

4 Fabrication of 200-in. Three-Dimensional Display

4.1 Projector Unit

We developed a compact projector unit measuring 180 (W) × 180 (H) mm² to pack the projector units densely, as shown in Fig. 10. One of the main impediments to image quality is the stripe noise between parallax images, which is influenced by the uniformity of brightness and color balance in the parallax images. To solve these problems, the brightness and color balance between projector units are compensated for using power-controlled LED light sources in the projector units. The uniformity of brightness in each projected image is also compensated for electrically by using a shading compensation circuit in each projector unit. The specifications of the projector unit are summarized in Table 1.

4.2 Prototyped Three-Dimensional Display System

The accuracy of the light control on the display screen affects the resolution and natural motion parallax of 3-D images.
We selected a suitable diffuser film with a 0.88 deg horizontal diffusion angle and 35 deg vertical diffusion angle and we combined it with an optimally designed 200-in. Fresnel lens with an aspherical surface (Fig. 11). As a result, approximately 200 parallax images could be displayed at a viewing distance of 5.517 m. The stripe noise in the displayed 3-D images, which reduced the image quality in our previous prototype system,16 as shown in Fig. 12(a), was reduced in our newly developed 200-in. display system, as shown in Fig. 12(b).

The specifications and performance of our newly developed 3-D display are listed in Table 2. The on-screen surface resolution of the 3-D images is 1920 pixels horizontally and 1080 pixels vertically. The viewing angle is 40 deg, and the width of the viewing area is 4 m at a viewing distance of 5.517 m. It is possible to display moving images with a frame rate of 60 fps. Figure 13 shows a reconstructed 3-D image of a life-size computer graphics (CG) car, and parallax images observed from the left, center, and right views. Motion parallax is present owing to the difference in the positional relationships of the car door and the interior depending on the observer’s location. We can also confirm that the reflections on the door change according to the viewing position just as for a real object as shown in Fig. 14.

---

**Table 1 Specifications of projector unit.**

<table>
<thead>
<tr>
<th>Specification</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display device</td>
<td>0.7 in., LCOS × 3</td>
</tr>
<tr>
<td>Number of pixel</td>
<td>1920 (H) × 1080 (V)</td>
</tr>
<tr>
<td>Projection lens</td>
<td>Focal length: 21.4 to 42.8 mm</td>
</tr>
<tr>
<td></td>
<td>F number: 3.2 to 4</td>
</tr>
<tr>
<td>Light source</td>
<td>RGB LEDs</td>
</tr>
<tr>
<td>Center wavelength</td>
<td>R: 617 nm, G: 525 nm, B: 465 nm</td>
</tr>
<tr>
<td>Luminous flux</td>
<td>20 lm</td>
</tr>
<tr>
<td>Size</td>
<td>180 (W) × 1300 (D) × 180 (H) mm</td>
</tr>
</tbody>
</table>

---

We selected a suitable diffuser film with a 0.88 deg horizontal diffusion angle and 35 deg vertical diffusion angle and we combined it with an optimally designed 200-in. Fresnel lens with an aspherical surface (Fig. 11). As a result, approximately 200 parallax images could be displayed at a viewing distance of 5.517 m. The stripe noise in the displayed 3-D images, which reduced the image quality in our previous prototype system,16 as shown in Fig. 12(a), was reduced in our newly developed 200-in. display system, as shown in Fig. 12(b).
5 Super Multiview Camera for Three-Dimensional Display

5.1 Camera Units and Field Test of Multiview Capture

We developed a compact camera unit to capture dense multiview video as shown in Fig. 15. The specifications of the camera unit are listed in Table 3. The width of the

![Fig. 12 Displayed 3-D images with (a) our previous 70-in. 3-D display and (b) our newly developed 200-in. 3-D display. Stripe noise was reduced by optimization of the optics and development of a new projector unit.]

<table>
<thead>
<tr>
<th>Table 2 Specifications and performance of 200-in. 3-D display.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screen</strong></td>
</tr>
<tr>
<td>Diffuser film</td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td>Fresnel lens</td>
</tr>
<tr>
<td><strong>Projector units</strong></td>
</tr>
<tr>
<td>Number of units</td>
</tr>
<tr>
<td><strong>3-D image</strong></td>
</tr>
<tr>
<td>Resolution</td>
</tr>
<tr>
<td>Frame rate</td>
</tr>
<tr>
<td>Number of parallax images</td>
</tr>
<tr>
<td>Interval of parallax images</td>
</tr>
<tr>
<td>Viewing angle</td>
</tr>
</tbody>
</table>

![Fig. 13 Displayed 3-D image of life-size CG car and left, center and right view images.]

![Fig. 14 Motion parallax of the displayed 3-D image of life-size CG car (Video 1, MPEG, 9 MB [URL: https://doi.org/10.1117/1.OE.57.6.061610.1])].

5 Super Multiview Camera for Three-Dimensional Display

5.1 Camera Units and Field Test of Multiview Capture

We developed a compact camera unit to capture dense multiview video as shown in Fig. 15. The specifications of the camera unit are listed in Table 3. The width of the

![Fig. 12 Displayed 3-D images with (a) our previous 70-in. 3-D display and (b) our newly developed 200-in. 3-D display. Stripe noise was reduced by optimization of the optics and development of a new projector unit.]

<table>
<thead>
<tr>
<th>Table 2 Specifications and performance of 200-in. 3-D display.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screen</strong></td>
</tr>
<tr>
<td>Diffuser film</td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td>Fresnel lens</td>
</tr>
<tr>
<td><strong>Projector units</strong></td>
</tr>
<tr>
<td>Number of units</td>
</tr>
<tr>
<td><strong>3-D image</strong></td>
</tr>
<tr>
<td>Resolution</td>
</tr>
<tr>
<td>Frame rate</td>
</tr>
<tr>
<td>Number of parallax images</td>
</tr>
<tr>
<td>Interval of parallax images</td>
</tr>
<tr>
<td>Viewing angle</td>
</tr>
</tbody>
</table>

![Fig. 13 Displayed 3-D image of life-size CG car and left, center and right view images.]

![Fig. 14 Motion parallax of the displayed 3-D image of life-size CG car (Video 1, MPEG, 9 MB [URL: https://doi.org/10.1117/1.OE.57.6.061610.1])].

5 Super Multiview Camera for Three-Dimensional Display

5.1 Camera Units and Field Test of Multiview Capture

We developed a compact camera unit to capture dense multiview video as shown in Fig. 15. The specifications of the camera unit are listed in Table 3. The width of the

![Fig. 12 Displayed 3-D images with (a) our previous 70-in. 3-D display and (b) our newly developed 200-in. 3-D display. Stripe noise was reduced by optimization of the optics and development of a new projector unit.]

<table>
<thead>
<tr>
<th>Table 2 Specifications and performance of 200-in. 3-D display.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Screen</strong></td>
</tr>
<tr>
<td>Diffuser film</td>
</tr>
<tr>
<td>Size</td>
</tr>
<tr>
<td>Fresnel lens</td>
</tr>
<tr>
<td><strong>Projector units</strong></td>
</tr>
<tr>
<td>Number of units</td>
</tr>
<tr>
<td><strong>3-D image</strong></td>
</tr>
<tr>
<td>Resolution</td>
</tr>
<tr>
<td>Frame rate</td>
</tr>
<tr>
<td>Number of parallax images</td>
</tr>
<tr>
<td>Interval of parallax images</td>
</tr>
<tr>
<td>Viewing angle</td>
</tr>
</tbody>
</table>

![Fig. 13 Displayed 3-D image of life-size CG car and left, center and right view images.]

![Fig. 14 Motion parallax of the displayed 3-D image of life-size CG car (Video 1, MPEG, 9 MB [URL: https://doi.org/10.1117/1.OE.57.6.061610.1])].
camera units is 30 mm to allow for a dense camera array. The output video has an HD resolution of 1920 (H) × 1080 (V) pixels. All camera units can operate with an external synchronization video signal. Each camera has a compact oblique stage to adjust the camera position accuracy as shown in Fig. 15.

Figure 16 shows the field test setup for capturing actual live moving objects using a camera array of 64 camera units at the Koyasan Kongobuji Temple in Japan. Each camera requires only one cable to control camera gain and white balance, and to transmit the captured video signal to an image processor and camera control unit. Moreover, camera unit pairs are connected by one cable in a daisy-chain fashion. Therefore, we can reduce the number of cables between the camera control units and the camera units to half of the number of camera units. To record 3-D sound, we also place small microphone arrays below and above the camera array, as shown in Fig. 16. The camera array system is mounted on a tripod, which enables changing the shooting angle and direction. The number of camera units can be changed according to the distance from the camera to the objects, as well as the size of the objects.

5.2 Real-Time Capture and Display System

Figure 17 shows the system configuration of the real-time capture and 3-D video display system. To capture and display 3-D images in real-time, we developed a dynamic convergence compensation circuit for each camera control unit. The registrations of captured images were arranged by high-accuracy camera adjustment geometrically compensated for using image processing for the multiprojection 3-D display. The output 3G-serial digital interface (SDI) signals from the camera control units are connected to the signal sources of the 3-D display system. Experimental results are shown in Fig. 18, and we successfully captured 3-D images by developing multiview camera array, and displayed 3-D videos of life-size moving objects in real-time. In this demonstration, we connected the camera side and display side with multiple metallic video cables. It will be possible to transmit these 3-D videos between distant places using a highly efficient compression method for multiview images.

<table>
<thead>
<tr>
<th>Camera unit</th>
<th>Number of pixel</th>
<th>1920 (H) × 1080 (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame rate</td>
<td></td>
<td>60 fps</td>
</tr>
<tr>
<td>Output video signal</td>
<td>3G-SDI</td>
<td></td>
</tr>
<tr>
<td>Image sensor size</td>
<td>1/3 in.</td>
<td></td>
</tr>
<tr>
<td>Camera lens</td>
<td>Focal length</td>
<td>6 mm</td>
</tr>
<tr>
<td>Angular field of view</td>
<td>43 deg (H) × 33 deg (V)</td>
<td></td>
</tr>
<tr>
<td>Iris range</td>
<td></td>
<td>F1.2 to F16</td>
</tr>
</tbody>
</table>
Next, we designed a real-time capture system using the above-mentioned compact camera units for the dense multi-view 3-D display described in Sec. 4. First, we consider the parameters of the camera pitch and projector pitch for the reconstruction required to correct the 3D images. From a geometrical analysis using Fig. 19, to conserve the depth-to-distance ratio of the objects so that $a_c/w_c = a_d/w_d$, we obtain the following relationship:  

$$\frac{g_d}{g_c} = \frac{C_d}{C_c},$$  

(26)

where $g_d$ is the focal length of the projection lens, $g_c$ is the focal length of the camera lens, $C_d$ is the size of the display device, and $C_c$ is the size of the capture device. On the other hand, the condition to conserve the depth-to-distance ratio of an object so that $a_c/b_c = a_d/b_d$ is

$$\frac{p_d}{p_c} = \frac{a_d}{a_c}.$$  

(27)

In our developed 3-D display, the pitch of the projector array $p_d$ is 33 mm, the projection distance $a_d$ is 8 m, the distance between the camera and objects, $a_c$, is 5.517 m, and is equal to the viewing distance. As mentioned previously, a camera pitch of 22.8 mm is required. The 22.8-mm value is less than the width of the camera units, so we developed a synthesized optical system consisting of a mirror and half mirror to combine a two-array camera system as shown in Fig. 20(a). A two-array system consisting of two sets of 96 camera units with 45.6-mm intervals [Fig. 20(b)] was combined with the synthesized optical system, and the pitch of the multiview camera became effectively 22.8 mm.

5.4 Demonstration Experiments

We carried out social demonstration experiments for the highly realistic communication system by introducing...
the developed multiview 3-D camera and display systems in a new development area north of Osaka station in Japan from April 2013 to November 2015. In the demonstration experiments, we used the developed 200-in. 3-D display system and 3-D contents captured by the camera array.

Figure 21(a) shows the installed 200-in. 3-D display system, called ray emergent imaging (REI), and the super multiview camera array system. The height of the screen center is 2 m, to enable many people to observe 3-D images simultaneously, as shown in Fig. 21(b). The super multiview 3-D camera system described in Sec. 5.3 is installed under the

Optical Engineering 061610-11 June 2018 • Vol. 57(6)
screen, so that the appearance of visitors can be captured and displayed on the screen in real-time.

There were up to 2500 visitors per day to the exhibition during the summer vacation periods. We have examined the reactions and behavior of visitors when viewing several 3-D contents as shown in Fig. 22, and surveyed the visitors about the impressions of the observed 3-D images.21

In the results, the visitors stopped at and showed the most interest in the content of the colorful real objects spread over a wide area, as shown in Fig. 22(a). The visitors also felt the most motion parallax with the content of Fig. 22(b).22 We believe that the observers experienced a high degree of realism owing to the reproduction of natural glasses-free 3-D images on the large screen in life-size dimensions, as if the objects were really in front of them.

6 Conclusion

To achieve a highly realistic communication system based on a natural 3-D video, we proposed a projection-type 3-D display method that can display glasses-free 3-D images on a large screen using multiple projectors. We designed and fabricated the 200-in. 3-D display system to allow 40 deg viewing angles. The gap between the parallax images was 22.8 mm, which helped us to produce natural HD 3-D images with smooth motion parallax. We also developed the super multiview 3-D capture system with compact HD camera units that have real-time video signal correction circuits for camera calibration and image rectification.

We performed social demonstration experiments by introducing the developed 3-D video system in a public area. The large 3-D image, which does not require observers to wear special glasses, enables several people to observe large 3-D images and share a 3-D space and environment. By reproducing people, vehicles, and other familiar objects and environments in life-size dimensions, the observer can experience a high degree of realism. Such images are expected to find many applications in the verification of industrial designs, publicity, exhibitions of cultural heritage, works of art, and other areas, in addition to being used for digital cinema.

Acknowledgments

This research was conducted when the authors were employed by Universal Communication Research Labs in the National Institute of Information and Communications Technology. Part of this research was supported by the “Research and Development of Glasses-Free 3-Dimensional Image Technology, 3-Dimensional Image Support Technology” research program of the Ministry of Internal Affairs and Communications, Japan. The authors would like to extend special thanks to Masahisa Sakai, Yasuyuki Haino, and Masahito Sato of JVC KENWOOD, Inc. for their technical support in developing the 3-D display system. Our gratitude is also extended to Kongobuji Temple and NHK Enterprises, Inc. for their support in the capture field test at Koyasan and to ORIX Real Estate Corporation for their support in the capture test at Kyoto Aquarium. The authors would like to thank Knowledge Capital Association and KMO Corporation for their support of the experimental demonstration in GRAND FRONT OSAKA.

References


Masahiro Kawakita received his BS and MS degrees in physics from Kyushu University and his PhD in electronic engineering from the University of Tokyo in 1988, 1990, and 2005, respectively. In 1990, he joined NHK (Japan Broadcasting Corporation), Tokyo. Since 1993, he has been at the Science and Technical Research Laboratories of NHK, where he has been researching applications of liquid crystal devices and optically addressed spatial modulators, 3-D TV cameras, and display systems.

Shoichiro Iwasawa received his PhD in electrical engineering and electronics from Seikei University, Japan. He is a senior researcher at the National Institute of Information and Communications Technology, Japan, and a member of ACM and IEEE. His fields of interest include image-based social listening, 3-D display, pervasive computing, computer graphics, and computer vision.

Roberto Lopez-Gulliver received his PhD of engineering in information and media science from Kobe University. He worked as a senior researcher at ATR and NICT research labs from 1994 to 2014, and he is currently an associate professor at Ritsumeikan University in Kyoto, Japan. His research interests include interactive autostereoscopic 3-D imaging and displays, virtual reality, and human computer interaction.

Naomi Inoue received his DEng, ME, and BE degrees from Kyoto University in 1998, 1984, and 1982, respectively. From 1987 to 1991, he was a researcher at ATR Interpreting Telephony Research Laboratories. In 1991, he joined KDD R&D Laboratories. He joined NICT Universal Media Research Laboratory in 2006, and became its director in 2010. He promoted research on ultra-realistic communication for 10 years. He rejoined KDDI Research, Inc. in 2016.