High-Definition Equivalent 3D Imaging and Display System

Masanori Kano[1], Takuya Omura[1], Hayato Watanabe[1], Jun Arai[1] and Masahiro Kawakita[1,2]

Corresponding author's e-mail address: kanou.m-gc@nhk.or.jp
[1]Japan Broadcasting Corporation (NHK), Tokyo 157-8510, Japan
[2]Osaka Institute of Technology, Osaka 573-0196, Japan
Keywords: 3D, Imaging, Display, High-definition equivalent.

ABSTRACT

We have been developing 3D video systems for future video media. High-quality 3D imaging and display technology is required to realize highly realistic 3D videos. In this paper, we report the development of a 3D imaging and display system with high-definition equivalent resolution.

1 Introduction

We have been advancing the research and development of 3D video systems that can provide highly realistic and immersive experiences, which cannot be provided by conventional 2D video systems. The features of these 3D video systems are that they do not require special glasses, have full parallax in the horizontal and vertical directions, can be viewed by multiple people, and are less likely to cause visual fatigue owing to binocular vision with vergence-accommodation conflict. These features can be realized by imaging rays emitted from a real subject and reconstructing equivalent rays in space. However, it is challenging to image subjects and display high-quality 3D videos. The indexes for evaluating the quality of a 3D video system include resolution, the depth range of a reconstructed 3D video (ray density), and viewing zone, where 3D videos can be viewed correctly. It is therefore necessary to acquire and reconstruct a large amount of ray information to improve these indexes. We previously developed a 3D video system with standard-definition (SD) resolution [1-3]. This system was able to image a subject and display a 3D video with a viewing angle of 35.1° × 4.7° at SD resolution. However, as the resolution of this system is not sufficient, it is difficult to image and display a subject with fine patterns.

In this paper, we report the development of a 3D video system with high-definition (HD) equivalent resolution [4]. This system consists of an imaging system and a display system, as shown in Fig. 1. The imaging system captures videos of the subject. Based on the captured videos, it generates a multi-view video, hereafter referred to as a displayed multi-view video. As the rays emitted by the subject are reconstructed by displaying the multi-view video on the display system, viewers can watch the 3D video.

2 Imaging System

A displayed multi-view video consists of the videos of a subject taken from various directions. The number of viewpoints and resolution are determined by the specifications of the display system. Generally, an extremely large number of viewpoints are required. However, it is difficult to capture videos from all viewpoints using real cameras because of system scale

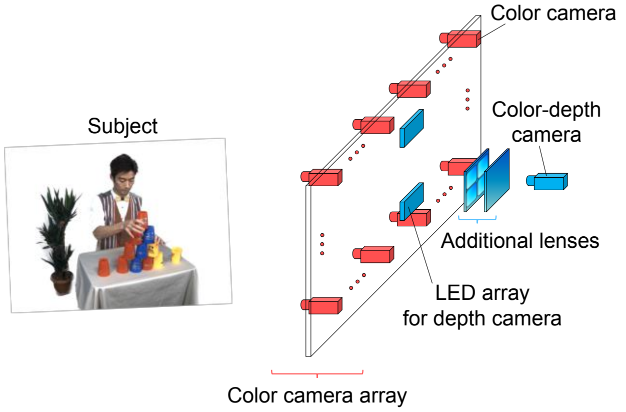
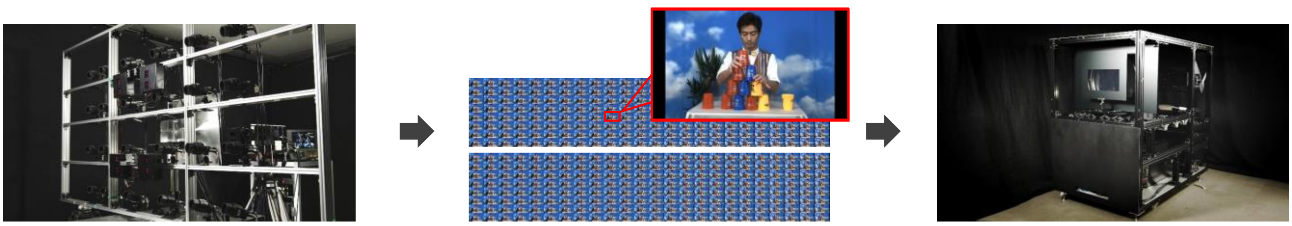


Fig. 2 Imaging system configuration



(a) Imaging system

(b) Displayed multi-view video

(c) Display system

Fig. 1 Overview from imaging to display

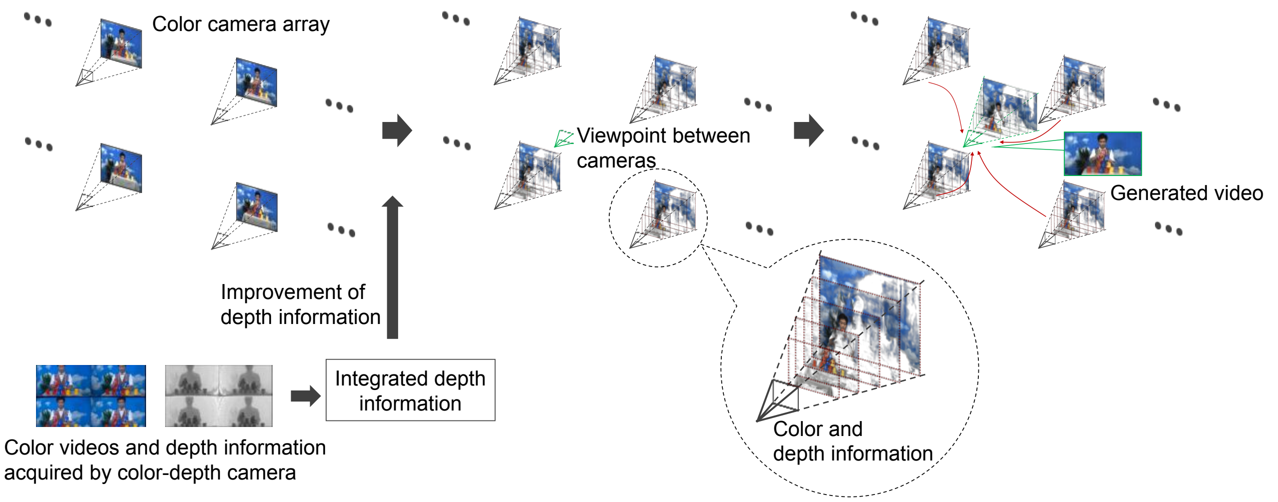


Fig. 3 View interpolation processing

and cost. Hence, to generate the displayed multi-view video using fewer cameras, the videos of viewpoints between cameras are generated through view interpolation processing.

The newly developed imaging system consists of an array of 24 color cameras, one color-depth camera, and additional lenses, as shown in Fig. 1(a) and Fig. 2. Additional lenses, including a lens array, are placed in front of the color-depth camera to acquire color videos and depth information from four viewpoints. The displayed multi-view video is generated using the videos captured by the imaging system and view interpolation processing (Fig. 3). The depth information of the videos of the camera array must be estimated to generate a video of a viewpoint between the cameras. The depth information is estimated with high accuracy using the color-depth camera. The video corresponding to the viewpoint between the cameras is generated using the color videos and depth information obtained from the cameras adjacent to the viewpoint. This procedure is repeated to generate the displayed multi-view video (Fig. 1(b)).

3 Display System

The display system reconstructs a 3D video by displaying the multi-view video generated by the imaging system. The principle of this display system is to reconstruct a large number of rays at a high density by projecting and displaying multiple multi-view videos and to appropriately spread the rays according to the ray spacing on a screen with top-hat-type diffusion characteristics [1]. The integral imaging display [5, 6], which is a similar light field display system, has the advantage that a thin display system can be easily configured. However, its disadvantage is that it is difficult to increase the resolution of a 3D video because the resolution is limited by the number of lenses in a lens array. However, the new display system can easily increase resolution because the resolution of a 3D video is equal to that of a single-view video in a multi-view video.

The new display system consists of two 8K projectors, four 4K projectors, and a display optical system for superimposing multi-view videos to reproduce a 3D video, as shown in Fig. 1(c) and Fig. 4. The system uses three features to improve the quality of a 3D video. First, the projector has a pixel shift function to increase the resolution of the 3D video. Although the original video resolution per projected viewpoint is 960 × 540 pixels, the function doubles the resolution to be equivalent to 1920 × 1080 pixels (HD equivalent). Second, the depth range is doubled because the ray density is increased by displaying multiple high-resolution multi-view videos through an optical element that shifts the direction of rays. Finally, the viewing angle is increased by the multiple projectors. The six projectors are arranged in parallel, and a display optical system is developed according to the arrangement. As a result, a horizontal viewing angle of 32.9° is achieved by increasing the number of rays and projecting them over a wide area.

4 Results

We performed experiments on the imaging and display of subjects using the system. There were two types of subjects: (a) street performer and (b) human face. The videos of subjects were captured by the imaging system, and a displayed multi-view video was generated. Table 1 shows the specifications of the display system. The number of display viewpoints was 384 because it was doubled by time division. The resolution was also doubled to HD equivalent resolution by time division.

Figure 5 shows the results of the 3D videos displayed on the display system. It was visually confirmed that a higher-definition 3D video was displayed compared to the conventional method [1-3]. Figure 6 shows the 3D videos of subject (a) from various viewpoints. It was confirmed that the videos corresponding to the

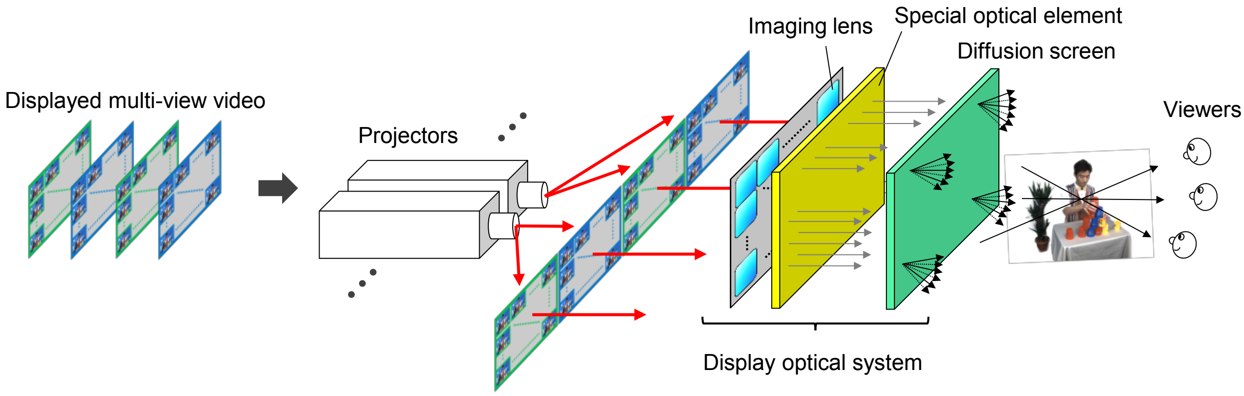
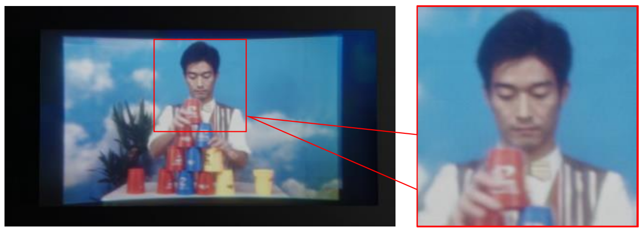


Fig. 4 Display system configuration

viewpoints were observed within the viewing zone.

Table 1 Specifications of the display system

Number of views	24 (H) × 8 (V) × 2
Resolution (1 view)	HD (1920 × 1080 pixels) equivalent
Frame rate	30 fps
Viewing angle	32.9° (H) × 6.0° (V)



(a) Street performer

(b) Human face

Fig. 5 Display result of 3D video (re-shooting)

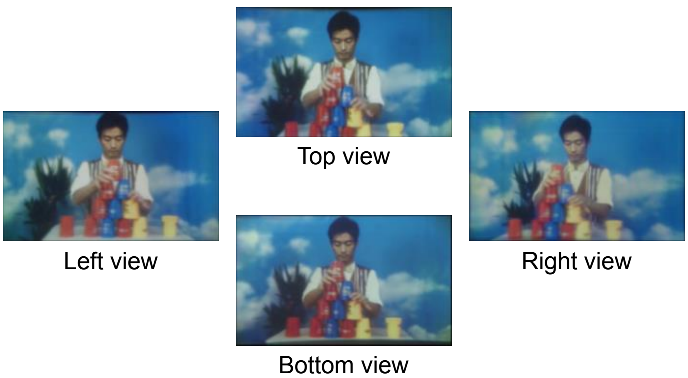


Fig. 6 3D videos from various viewpoints (re-shooting)

5 Conclusions

We developed a 3D video system with HD equivalent resolution and conducted experiments on the imaging and display of subjects. We realized 3D videos with HD equivalent resolution with a horizontal viewing angle of approximately 33° from imaging to display.

References

[1] H. Watanabe, N. Okaichi, T. Omura, M. Kano, H. Sasaki, and M. Kawakita, “Aktina Vision: Full-parallax three-dimensional display with 100 million light rays,” Sci. Rep., Vol. 9, 17688 (2019).

[2] M. Kano, H. Watanabe, M. Kawakita, and T. Mishina, “3D Imaging with 330,000 Pixels Using Camera Array (in Japanese),” ITE Winter Annual Convention, 23D-1 (2018).

[3] https://www.nhk.or.jp/strl/open2018/tenji/23_e.html

[4] https://www.nhk.or.jp/strl/english/open2021/tenji/4/index.html

[5] G. Lippmann, “Épreuves, réversibles donnant la sensation du relief,” J. Phys. Theor. Appl., Vol. 7, No. 1, pp. 821-825 (1908).

[6] F. Okano, H. Hoshino, J. Arai, and I. Yuyama, “Real-time pickup method for a three-dimensional image based on integral photography,” Appl. Opt., Vol. 36, No. 7, pp. 1598-1603 (1997).