https://doi.org/10.2352/EI.2026.38.2.SDA-339

Aerial-imaging light-field camera with wide-viewing angle for real-time 3D imaging

Kotaro Sakamoto, Shoma Kono, Masanori Morimoto, Masahiro Kawakita
Professional Affiliation: Osaka Institute of Technology, Faculty of Information Science and Technology, Osaka, Japan
Keywords: 3D image, aerial image, light field, lenticular lens, 8K video

Abstract

Conventional light-field (LF) cameras are typically limited to narrow viewing angles for capturing light rays emitted from objects. We proposed an aerial-imaging light field (ALF) camera that forms aerial images of objects and captures them via a lens array with wide viewing angles. Using a prototype lenticular lens array, we achieved a 36.1° viewing angle and captured more horizontal perspectives. The elemental images were then input into an integral 3D display with the same lens array, allowing real-time, glasses-free 3D image display. Based on this functionality, we developed a prototype real-time 3D video communication system combining an ALF camera and an integral 3D display.

1. Introduction

Light field (LF) technology, which acquires, processes, and reproduces light ray information emitted from objects, has attracted considerable attention in recent years [1]. It is expected to provide natural glasses-free 3D images while reducing visual fatigue commonly associated with stereoscopic viewing. Owing to this advantage, LF technology has been applied in various fields of visual media, including head-mounted displays (HMDs) for stereoscopic visualization and omnidirectional image capture and display using panoramic photography.

LF cameras enable the direct acquisition of light-ray information emitted from objects. Conventional LF camera technology typically employs a lens array close to the image sensor within the camera lens [2]. This approach offers advantages such as compact device design and ease of operation. However, it is limited in that it can capture light ray information only within a narrow viewing angle of a few degrees. To address this limitation, we propose an aerial-imaging light-field (ALF) camera that generates aerial images of objects and captures them through a lens array, thereby expanding the range of light ray acquisition [3,4].

By optimizing the lens array in this method, we successfully acquired light ray information over a range exceeding 30°, which represents more than a tenfold improvement compared to conventional LF cameras. Moreover, as a practical application, we generated free-viewpoint images from elemental images and demonstrated a glasses-free 3D communication system. Nevertheless, a limitation of this approach is the relatively low resolution of the reconstructed 3D images.

In this paper we present the development of a 3D image capturing system that combines a triple mirror (TM) screen with a lenticular lens array. In this system, the TM screen is employed to form images in midair, while the lenticular lens array is used to capture them. By integrating these components, the resolution of the captured 3D images has been significantly improved. Furthermore, as an application, we demonstrated real-time 3D capture and display using an ultra-high-definition (UHD) video system, along with the generation of multi-view images.

2. Proposed method

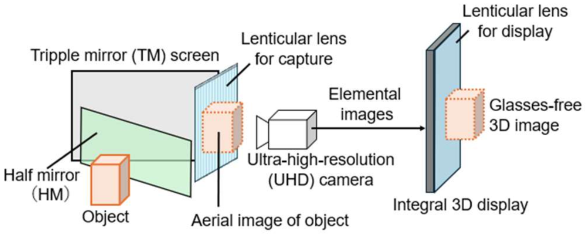

Fig.1 Real-time 3D capture and display using aerial images and lenticular lens

Figure 1 shows the configuration of the proposed system. A retroreflective optical element is employed to form an aerial image of objects, at which a lens array is positioned. Compared with a real image formed by a conventional lens system, an aerial image offers a wider light-gathering angle.

Various methods have been developed for forming aerial images. In our previous studies [3, 4], a dihedral corner reflector array (DCRA) was employed for this purpose. In the DCRA, aerial images are generated by reflecting light rays from an object twice using a microprism array. However, a limitation of the DCRA is that light rays which fail to undergo double reflection produce ghost light.

In the proposed system (Fig.1), an aerial image formation approach combining a triple mirror (TM) screen with a half mirror (HM) was adopted. A TM screen is a retroreflective optical element that reflects incident light from three mirror surfaces and redirects it back along the same path as the incident light [5]. As shown in Fig. 1, by placing a lens array at the position where the aerial image is formed and capturing it with a camera, elemental images containing light ray information over a wide angular range can be obtained. In this system, a lenticular lens was employed as the lens array. Since a lenticular lens acquires only horizontal light ray information, vertical information is not obtained. This restriction allows the limited image information capacity to be allocated toward improving the resolution of the reconstructed 3D images and expanding the field of view.

3. Experiments

3.1 Aerial-imaging methods

As noted above, various methods for aerial image formation have been reported. Previous studies [3,4] primarily employed the DCRA. In contrast, the proposed approach adopts a combination of a TM screen and a half mirror. In this study, we conducted experiments and comparisons from three perspectives to evaluate which of these two aerial image formation methods is more suitable for the proposed system.

First, we evaluated the resolution characteristics of the reconstructed 3D images. An aerial image with a black-and-white striped pattern was captured through a lenticular lens and displayed on a 3D display. The spatial frequency of the displayed 3D image was then measured to determine its resolution. The measurement results are shown in Fig. 2. The vertical axis represents the spatial frequency (lp/mm), while the horizontal axis indicates the distance between the lenticular lens and the aerial image (mm). The blue dots correspond to the DCRA method, and the orange dots correspond to the TM screen and HM method.

In the case of the TM screen, the resolution was slightly higher than that of the DCRA near the lenticular lens array surface. However, no significant difference was observed between the two systems overall. This is because when the aerial image is located near the lenticular lens, the resolution is constrained by a Nyquist frequency.

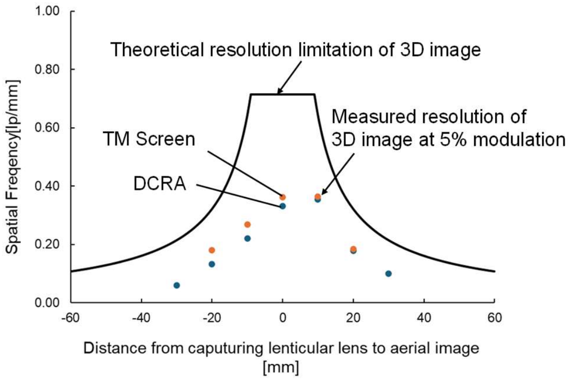

Fig.2 Resolution characteristics of 3D capture and display system

Second, we examined the viewing angle of the reconstructed 3D images when displayed on a 3D display. As shown in Fig. 3, a red character "3" and a green character "D" were used as objects. To introduce parallax, the "3" was placed in the foreground and the "D" in the background. Elemental images were captured using each aerial image formation method and then displayed as 3D images on a 3D display. The viewing angles of the reconstructed images were subsequently measured and compared. The measured viewing angle was 37.1° for the DCRA method and 36.1° for the TM screen method. Compared with our previous system, the TM screen increased the viewing angle by approximately 5°. Although there was a 1° difference between the two methods, we consider this to fall within the measurement error of the field-of-view evaluation. Therefore, no significant difference was observed between the two aerial image formation methods in terms of viewing angle.

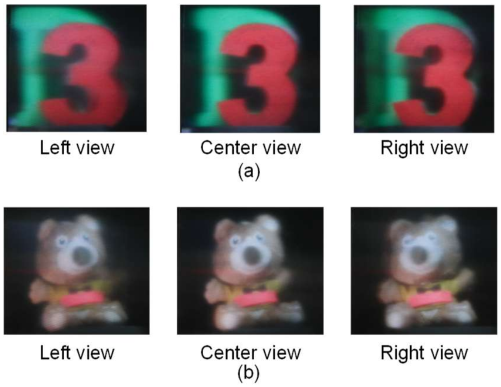

Fig.3 Reconstructed 3D images of objects (a) "3" and "D" characters, (b) "teddy bear".

Finally, we compared the effects of ghost light between the two aerial image formation methods. For this comparison, aerial images generated by each method were examined together with elemental images captured through a lenticular lens. Figure 4 shows the images obtained with each method. In the case of the DCRA, two ghost lights appeared at the top and bottom of the aerial image (Fig. 4(a)), and their influence was clearly visible in the elemental images captured through the lenticular lens (Fig. 4(b)). This is attributed to the fact that the ghost lights in Fig. 4(a) occur near the lens, thereby exerting a significant effect on the elemental images.

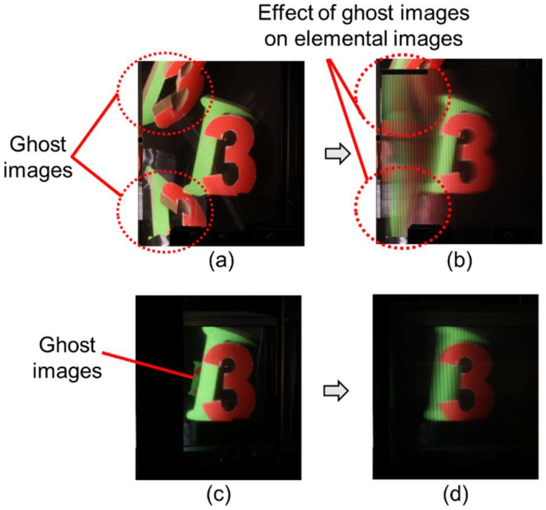

Fig.4 The appearance of ghost light. (a) Ghost and (b) Elemental images in case of DCRA, (c) Ghost and (d) Elemental images in case of TM screen.

In contrast, for the TM screen method, ghost light was observed at the center of the screen (Fig. 4(c)). However, its impact on the elemental images was minimal (Fig. 4(d)). This is likely because the ghost light is formed at a depth position farther from the lens array surface, thus reducing its influence on the elemental images.

From these results, it can be concluded that the TM screen is more suitable for the proposed method, as the effect of ghost light on elemental images is significantly smaller compared to the DCRA.

3.2 Real-time 3D capture and display using 8K video

As one application of the proposed method, we developed a prototype 3D real-time capture and display system using 8K video (Fig. 5). For aerial image formation, a TM screen combined with an HM was employed. Elemental images were captured with an 8K camera (BOSMA, DC0201) and projected onto a lenticular lens using an 8K projector (JVCKENWOOD, DLA-V80R). The objects used in the experiment were the characters "3" and "D," along with three dices. These objects were arranged in a circular pattern and rotated on a turntable to capture the elemental images. The reconstructed 3D image is shown in Fig. 6. The results confirm that the system is capable of real-time 3D capture and display.

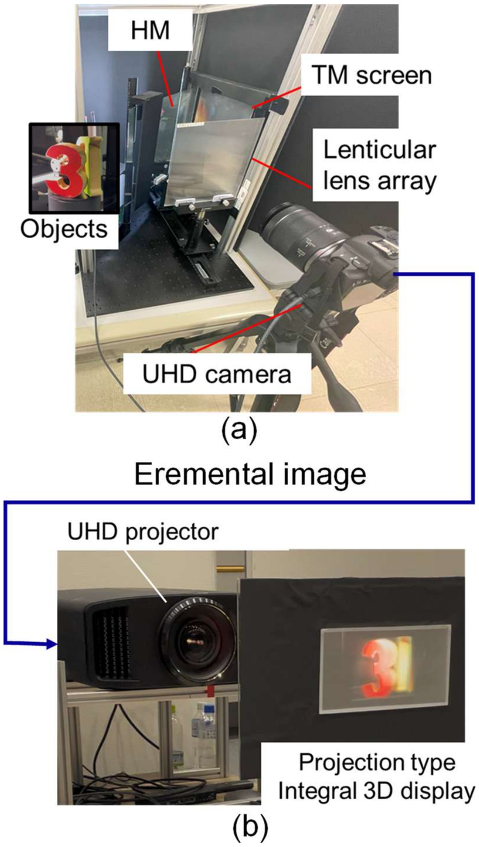

Fig.5 Real-time (a) 3D capturing system and (b) 3D display using 8Kvideo system

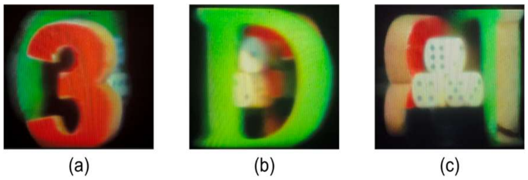

Fig.6 Realtime displayed 3D images of rotating objects. (a) Character "3", (b) "D", and (c)dices.

3.3 Multi-view images

To estimate the resolution of the displayed 3D image, multi-view images were generated from the elemental images (Fig. 7). Camera calibration was performed using a black-and-white checkerboard to correct lens distortion, after which the images were trimmed to match the lens size. Viewpoint images were then extracted from the elemental images to generate the multi-view images.

Using a lenticular lens array, 28 horizontal views were obtained. By restricting the viewpoints to the horizontal direction only, the number of pixels per viewpoint increased by approximately 60 times compared with our previous system [4].

Although 28 viewpoints were generated and used to construct each multi-view image, 13 viewpoints (the first 8 and the last 5) contained severe artifacts due to lens boundary effects and were excluded only in the evaluation stage. As a result, 15 valid viewpoints were used for visual evaluation.

It should be noted that the spatial resolution of each viewpoint is determined during the multi-view image generation process based on all 28 viewpoints. Therefore, excluding invalid views does not affect the resolution per viewpoint. The resolution improvement factor (approximately 60 times compared to the previous system) is thus based on the full set of 28 viewpoints.

The overlap is attributed to the fact that image correction after camera calibration was performed manually. In future work, the number of usable views can be further improved by automating the correction process.

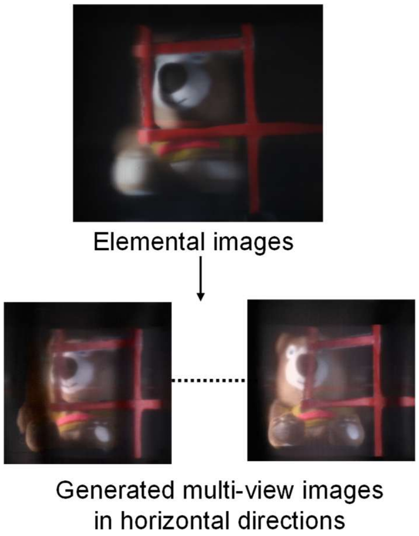

Fig.7 Multi-view images generated from elemental images.

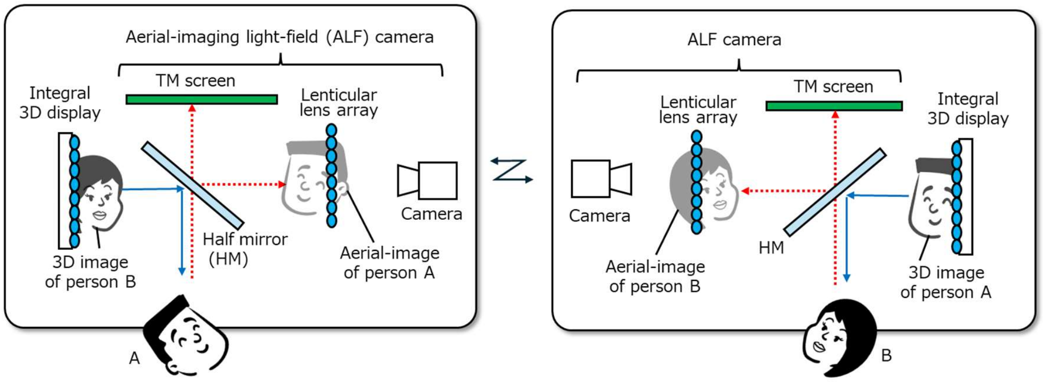

Fig.8 Real-time glasses-free 3D video communication using ALF camera and integral 3D display

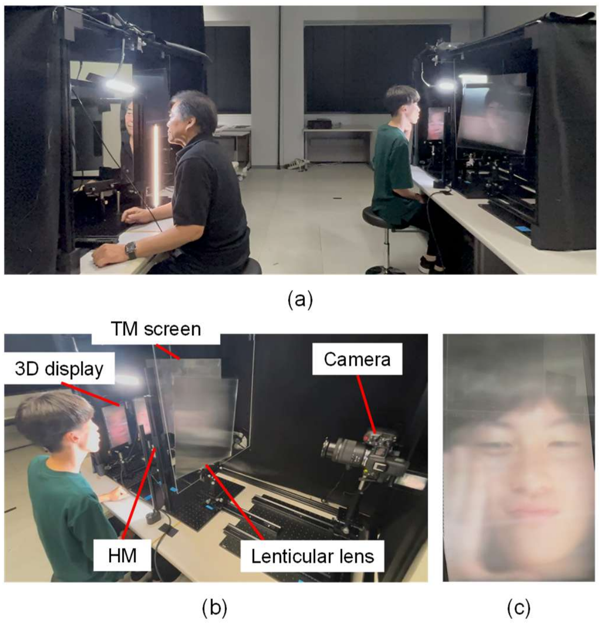

Fig.9 Prototyped 3D communication system.
(a) Experimental setup, (b) configuration of optics, and (c) real-time displayed 3D image.

3.4 Integral 3D video communication system

As a further application, we developed a two-way communication system based on the proposed shooting and display technology (Fig. 8, Fig. 9(a)). Elemental images were captured using a UHD video camera (Canon EOS R5C) and transmitted via HDMI to two 3D displays, where the images captured by both parties were rendered as 3D images. By reflecting the displayed 3D images with a half mirror (HM), the system enables mutual communication through the viewing of each other's 3D images (Fig. 9(b)). Nevertheless, the elemental videos were affected by light emitted from the 3D displays, since the displays and cameras were aligned in a straight line. To address this issue, a polarizing filter was installed on the camera lens, which effectively suppressed the interference. Figure 9(c) shows the 3D video displayed using this method. These results indicate that real-time glasses-free 3D video communication systems are achievable with the proposed approach.

4. Conclusion

To enhance the resolution of the LF camera for acquiring light ray information, we developed a prototype system incorporating a lenticular lens array. Experimental results demonstrated that aerial images generated using a TM screen are suitable for this approach. As applied technologies, we further implemented real-time 3D capture and display with 8K video, generated multi-view images, and developed two-way 3D communication systems.

References

[1] C. Jia, F. Shi, M. Zhao and S. Chen, "Light field imaging for computer vision: a survey," Front. Inf. Technol. Electron. Eng., vol. 23, no. 7, pp. 1077–1097, 2022.
[2] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Stanford Tech Report CSTR 2005-02, pp.1-11, 2005.
[3] M. Kameyama and M. Kawakita, "Light field capturing method using aerial image and lens array for integral 3D display," IDW'22, 3Dp1-7L, 2022.
[4] S. Kono, M. Kameyama, and M. Kawakita, "Aerial imaging light-field camera and integral 3D video communication system," Optica Imaging Congress 2024 DTh5F.5, 2024.
[5] T. Okoshi, "Three-Dimensional Imaging Techniques," Academic Press (2012) J. Doe, Sample Book, Springfield VA: IS&T, 1999.

Acknowledgments

This work was supported by JST SCORE JPMJST2051, JSPS Grant-in-Aid JP 25H01154, and the Hoso Bunka Foundation. We also express our gratitude to Ruan Jiaying for her meticulous proofreading and enthusiastic assistance in organizing the manuscript.

Author Biography

Kotaro Sakamoto received his B.E. degree from Osaka Institute of Technology in 2025. He is currently pursuing an M.S. degree at the Graduate School of Information Science, Osaka Institute of Technology. His research interests include 3D imaging and related technologies.


electronic
IMAGING


Image Engineering


STUDENT/YP SHOWCASE
PURDUE