

# Realization of precise depth perception with coarse integral volumetric imaging

|                              |   |
|------------------------------|---|
| 著者                           | Kekeya Hideki, Sawada Shimpei   |
| journal or publication title | Proc. SPIE, Three-Dimensional Imaging, Visualization                              |
| volume                       | 8043  |
| page range                   | 80430B  |
| year                         | 2011  |
| URL                          | <a href="http://hdl.handle.net/2241/114081">http://hdl.handle.net/2241/114081</a> |

doi: 10.1117/12.884431

# Realization of Precise Depth Perception with Coarse Integral Volumetric Imaging

Hideki Kakeya\* and Shimpei Sawada

University of Tsukuba  
1-1-1 Tennoudai, Tsukuba 305-8573, JAPAN

## ABSTRACT

In this paper realization of precise depth perception using coarse integral volumetric imaging (CIVI) is discussed. CIVI is a 3D display technology that combines multiview and volumetric solutions by introducing multilayered structure to integral imaging. Since CIVI generates real images optically, optical distortion can cause distortion of 3D space to be presented. To attain presentation of undistorted 3D space with CIVI, the authors simulate the optics of CIVI and propose an algorithm to show undistorted 3D space by compensating the optical distortion on the software basis. The authors also carry out psychophysical experiments to verify that vergence-accommodation conflict is reduced and depth perception of the viewer is improved by combining multiview and volumetric technologies.

### keywords

coarse integral volumetric imaging, optical distortion, depth perception, vergence-accommodation conflict, refractometer, texture mapping

## 1. INTRODUCTION

Though the current 3D technology is mainly focused on entertainment such as movies and games, 3D vision is not always required for these applications and is likely to remain as an optional or additional function as far as consumer electronics is concerned. In some applications, however, 3D vision can be an essential part of the system. For example, 2D displays cannot provide enough visual information to realize real-time interaction with 3D space, such as tele-operation of robot manipulators, remote-control of vehicles, or virtual simulators of these tasks.

To realize direct-manipulation with 3D vision, the display should provide to the viewer not only the sense of depth, but also precise 3D positions of the objects in the scene. Most conventional stereoscopic displays can provide binocular parallax, which is sufficient enough to enhance the sense of depth for the purpose of entertainment. For the viewer to perceive precise depth, however, the display should provide not only binocular parallax, but also motion parallax in both horizontal and vertical directions. Besides parallax, focal depth effect is also important to let the viewer perceive precise depth of the presented object. To grasp not only precise depth but also precise 3D position, geometry of the presented image should be correct and undistorted.

Coarse integral volumetric imaging (CIVI) is a 3D display solution that can provide not only binocular parallax, but also motion parallax and focal depth without geometrical distortion [1-3]. CIVI combines volumetric technique with multiview technique based on integral imaging. Integral imaging, which combines a fly-eye lens sheet (fine convex lens array) with a high resolution display panel, is a prominent 3D display system in the sense that it can show not only horizontal parallax but also vertical parallax. In the conventional integral imaging, the number of pixels that each elemental lens of the fly-eye lens sheet covers is usually the same as the number of views, which means that the viewer perceives each elemental lens as one pixel. Therefore the focus of the viewer's eyes is always fixed on the screen (fly-eye lens sheet), which makes it hard to show realistic images popping up from the screen. CIVI uses a coarse convex lens array whose elemental lens is large enough to cover pixels dozens of times more than the number of views. With this configuration many pixels of each elemental image can be seen at the same time and the view's focal accommodation can be induced off the screen when the real image of each elemental image is generated in the air.

---

\*E-mail : kake@iit.tsukuba.ac.jp Phone&FAX : +81-29-853-5255

When multilayered display panels are used to present each elemental image, multilayered real-image planes are generated in the air. By depicting each pixel on the image plane of the corresponding depth, natural focal depth effect is attained. Smooth expression of 3D pixels between two image planes can be realized by depicting them on two adjacent panels so that the intensity may be in inverse proportion to the distance between the 3D pixels and the image planes.

The real image generated with CIVI lens system includes various optical distortions and aberrations, which become more prominent as the viewing angle becomes wider. CIVI corrects these optical distortions, including field curvature, barrel distortion, and chromatic aberration, on the software basis. Each elemental image on the layered panel is distorted in the inverse direction so that the image observed by the viewer may have correct geometry.

In this paper we discuss precision of depth perception that can be attained by CIVI display system. This paper is organized as follows. In section 2 principle of CIVI is reviewed. In section 3 the method to correct distortion of images generated by CIVI display system is explained. In section 4 results of psychophysical experiments on reduction of vergence-accommodation conflict and precision of depth perception are explained. In section 5 summary of the paper is given.

## 2. COARSE INTEGRAL VOLUMETRIC IMAGING

Integral imaging, which combines fly-eye lenses with a high resolution flat display panel, is a prominent 3D display system in the sense that it can show not only horizontal parallax but also vertical parallax [4]. In the conventional integral imaging, the number of pixels each elemental lens of the fly-eye lens sheet covers is usually the same as the number of views, which means that the viewer perceives each elemental lens as one pixel. Therefore the focus of the viewer's eyes is always fixed on the screen (fly-eye lens sheet), which makes it hard to show realistic images far beyond the screen or popping up from the screen.

Besides the orthodox integral imaging described above, we can also think of integral imaging where each elemental lens is large enough to cover pixels dozens of times more than the number of views [5-7]. Kakeya defined this type of integral imaging as coarse integral imaging (CII) [8]. The advantage of CII is that it can induce focal accommodation off the screen, for it generates a real image or a virtual image with the lenses. Thus we can show realistic images far beyond the screen or popping up from the screen. Yet it cannot overcome the problem of vergence-accommodation conflict because the eyes of the viewer are always focusing on the real image or the virtual image generated by the lens array.

One way to solve the problem of vergence-accommodation conflict is to introduce volumetric approach using multilayered panels in addition to the multiview approach [9-12]. In the scheme of coarse integral imaging, Kakeya et al. proposed coarse integral volumetric imaging (CIVI) as shown in Figure 1. In CIVI multiple display panels are inserted to generate volumetric real image (volumetric virtual image can also be generated if the display panels and the lens array are set closer) to keep the parallax between the images from two adjacent lenses small enough. Since artificial parallax is kept small, discontinuity between images from adjacent lenses are also kept small. Vergence-accommodation conflict is also reduced since each 3D pixel is displayed at the real image (virtual image) layer near the right depth. To express pixels between two panels we can use DFD approach [13-15], where 3D pixels are expressed with two adjacent panels, each of which has the pixel intensity in inverse proportion to the distance between the 3D pixels and the panel. Thus natural continuity of depth is realized.

Here it should be noted that the layered real images generated by the lenses are curved and distorted. Not only the generated image plane is distorted, but also the image planes generated by elemental lenses of the lens array are not uniform. The image planes generated by the elemental lenses off the optical axis of the large aperture Fresnel lens do not have line symmetry about the optical axis, but are slanted toward the optical axis. The slant becomes greater as the elemental lens goes farther from the optical axis. Without taking into account this distortion problem, smooth connection of elemental images cannot be realized. To cope with these distortions we can apply DFD for distorted image plane, which can realize natural connections between elemental images. Figure 2 shows the prototype system of the coarse integral volumetric display and an example of the 3D still picture shown with the prototype system.

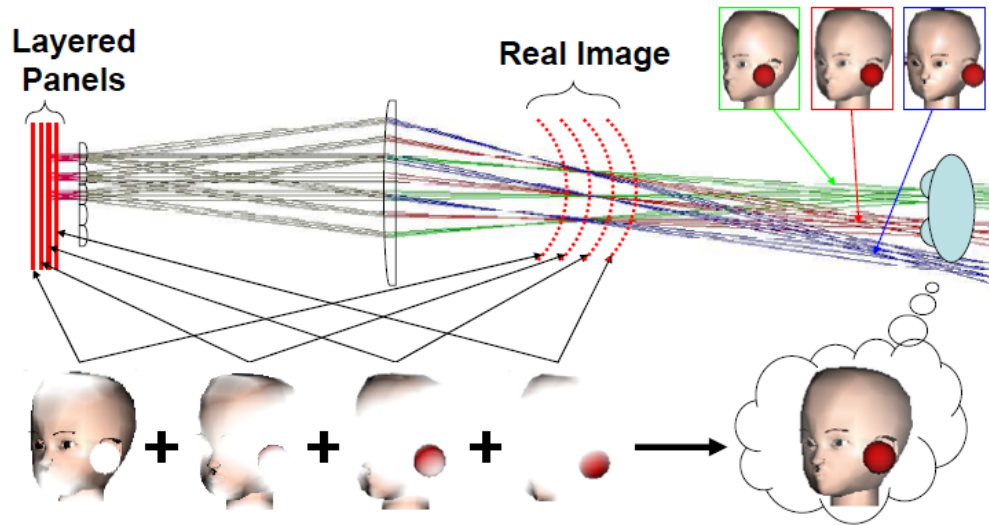


Figure 1: Principle of coarse integral volumetric imaging.



Figure 2: Prototype of coarse integral volumetric imaging (left) and the images observed from two different viewpoints with the prototype system (middle and right).

### 3. CORRECTING GEOMETRICAL DISTORTION OF CIVI

As explained in the previous section, the image generated by the lenses is distorted in horizontal, vertical, and depth directions, which should be corrected to show undistorted 3D space to the viewer. As for the correction of distortion in horizontal and vertical directions, we trace light rays from the viewpoint to the display panel and calculate how the images are distorted in those directions. Then texture mapping is used to correct distortion based on the result of optical simulation [16]. As for the correction of distortion in the direction of depth, we calculate refraction of light rays from the points on the display panel and obtain the points where the light rays converge.

The basic idea of the method we propose here is to show the image distorted in the reverse direction on the display panel to present undistorted image to the viewer. Therefore this method can also be applied to cancel out chromatic aberration. If we apply the same calculation described above for red, green, and blue color components independently, chromatic aberration can be corrected.

Now we give a detailed explanation on how the distortion correction proceeds. First the image from each viewpoint is generated by the standard computer graphic technique. Here we assume that orthogonal projection is applied to generate a 2D image with depth values from a 3D space. Then we project the light rays that correspond to the orthogonal projection (parallel light) to the lenses on the optical simulator and calculate how the rays are refracted and which points on the display panel they reach after refraction. Figure 3 illustrates how this calculation proceeds.

Each point of the image generated by orthogonal projection should be displayed on the points of the display panel calculated and plotted by the above optical simulation. To realize it, we can use texture mapping technique as shown in Figure 4. By showing the image distorted in the opposite direction, undistorted image is expected to be observed by the viewer.

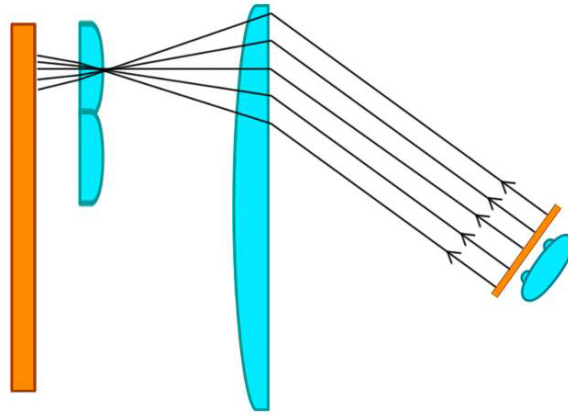


Figure 3: Illustration of how light rays coming from the orthogonal projection are refracted and reach the display panel.

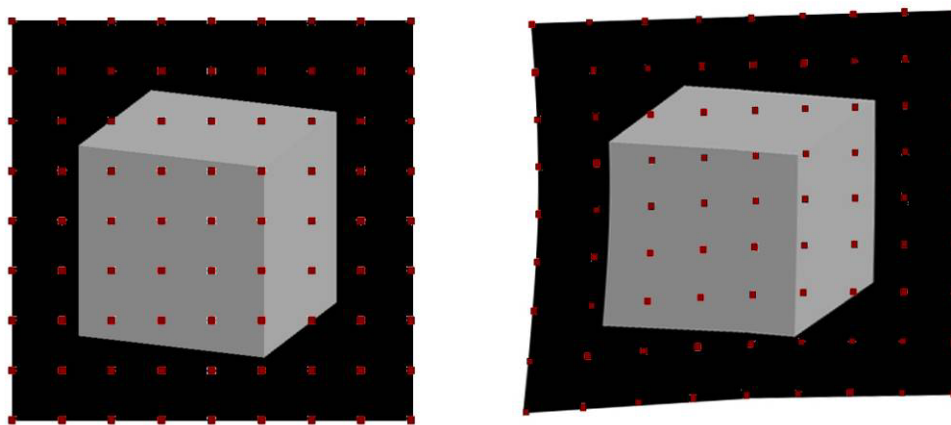


Figure 4: Example of an original image and the image to be display on the display panel.

Here it should be noted that we have to calculate and obtain the point arrays not only for one display panel, but also for the panels of all layers. The next step is to determine the layer on which each pixel is depicted, which requires information on where the real image of the elemental images on each layered panel is generated. Figure 5 illustrates how we can calculate the position and the shape of the generated real image. We calculate refraction of light rays from each point on the display panel obtained in the previous calculation for correction of distortion in horizontal and vertical directions, and obtain the points where the light rays converge.

Once the positions and the shapes of distorted image planes for each layer are obtained, we can generate the images to be displayed on each layer based on DFD for distorted image plane, where each 3D pixel is drawn on the adjacent two distorted image planes so that the pixel intensity may be in inverse proportion to the distance to each plane as shown in Figure 6. By applying this calculation for all the elemental images, we can realize presentation of undistorted 3D image.

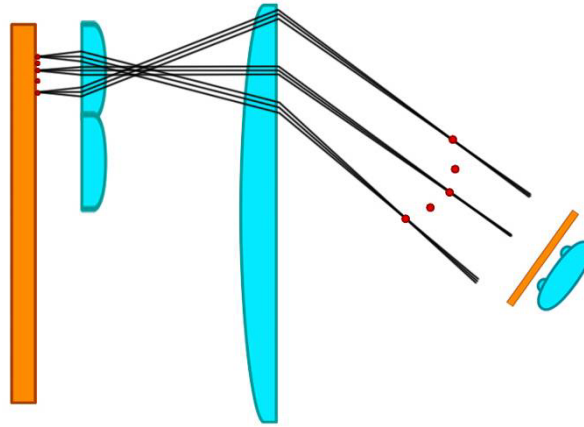


Figure 5: Calculation of position and shape of the real image generated by the lenses.

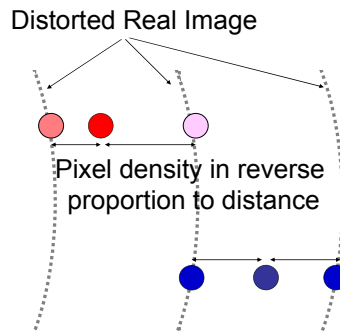


Figure 6: DFD for distorted image planes.

## 4. PRECISION OF DEPTH PERCEPTION

There have been some prior papers related to reduction of vergence-accommodation conflict and precision of viewer's depth perception. Some of the prior works suggest that vergence-accommodation is reduced by multilayered stereoscopy [17,18]. However, the results can vary depending on the detailed conditions of experiments. Here we try an original experimental setup to test whether vergence-accommodation conflict is reduced and precision of viewer's depth perception is improved by combination of multiview and volumetric solutions.

### 4.1 Vergence-Accommodation Conflict

Focal accommodation of the viewer can depend on the kind of image shown to the viewer []. One important point in the experiments is to keep the viewer focus on the presented image. In this paper we set up the experimental system as shown in Figures 7, 8 and 9. Figure 7 shows the design of the experimental system. By using half mirrors we can show images on two virtual planes of different depths. With this configuration we can give three conditions of image presentations: image presentation only on the near plane, only on the far plane, and on dual planes so that images from two planes may overlap. Figure 8 is the picture of the experimental system. Refractometer (NVISION K5001) is used to measure focal accommodation of the viewer. Display 3 is used for calibration so that Displays 1, 2 and the mirrors may be placed at the proper positions and angles. To make the viewer focus on the presented image Landolt rings as shown in Figure 9 are presented. Landolt ring (a) is the standard one. Ring (b) is a blurred Landolt ring, which can obscure the function on focal accommodation. Ring (c) is a white Landolt ring in the black background. When rings (a) and (b) are shown, the background is white and both panels emit light even under single panel conditions. In experiments with ring (c), only one panel emits light under single panel conditions. Diameter of the Landolt ring is 0.3mm on the near plane.

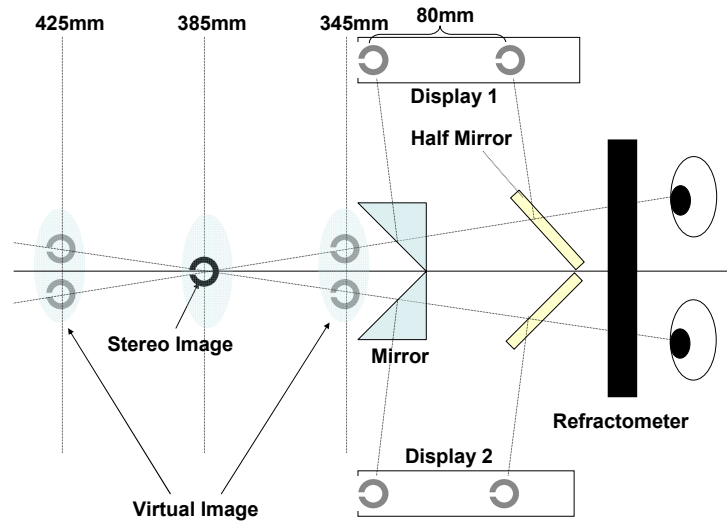


Figure 7: Design of experimental system on focal accommodation of viewers.

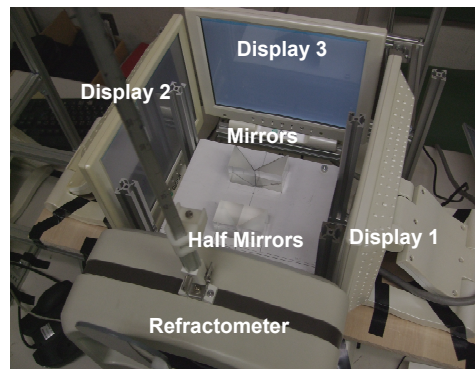


Figure 8: Picutre of experimental system

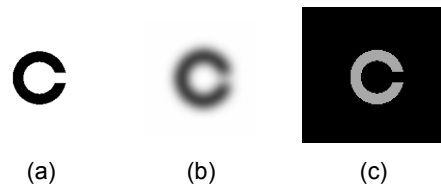


Figure 9: Landolt rings shown in the experiments.

In the experiment 10 subjects are tested 10 times with the interval of 0.4 second and the most frequent diopter value is recorded for each condition. The subjects are asked to keep on answering with the game pad the direction of Landolt ring, which changes after each input of answer. Average diopeters under each experimental condition using Landolt rings (a), (b) and (c) are shown in Figures 10, 11 and 12 respectively. Note that smaller diopter corresponds to the focal accommodation to farther depth. Also note that the intensity of Landolt ring image is in inverse propotion to the distance between the display plane and the stereo image plane under dual plane condition, which means that two image planes are used only when the stereo image is in the middle of two display planes. Control data show the diopeters of the subjects seeing the Landolt ring on display 3 placed 345mm, 385mm and 425mm in front of the viewer.

As Figures 10 and 11 show, the viewer tends to focus on nearer depth even under far plane condition, which can be caused by the fact that background light is also coming from the near plane. This assumption is supported by the

fact that this tendency disappears when the background is black as shown in Figure 12. When the Landolt ring is blurred, focal accommodation is almost static regardless of the difference of experimental conditions. When the Landolt ring is not blurred, dual plane condition can realize accommodation to the middle depth when compared with the near plane condition and far plane condition. This tendency becomes distinct when Landolt ring (c) with black background is used. This result suggests that expression of depth by presentation of stereo disparity on dual image planes, which is used in CIVI display system, can reduce vergence-accommodation conflict.

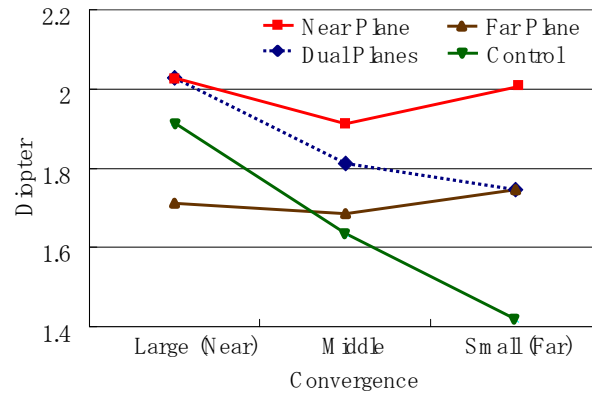


Figure 10: Result of experiments with Landolt ring (a).

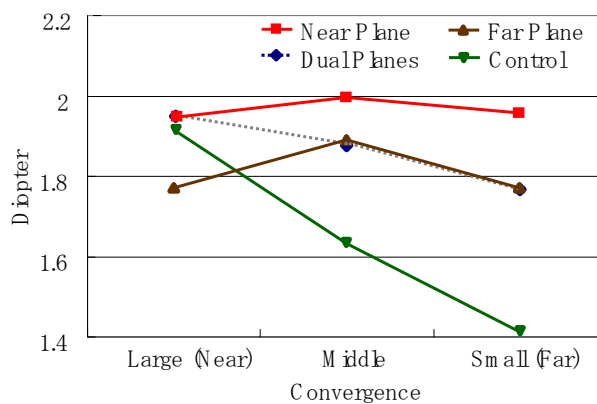


Figure 11: Result of experiments with Landolt ring (b).

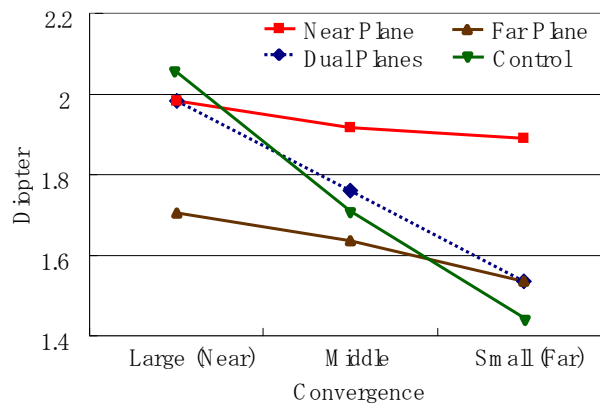


Figure 12: Result of experiments with Landolt ring (c).



## 4.2 Precision of Depth Perception

In this subsection we test whether precision of viewer's depth perception is improved by stereoscopy with multilayered panels. The design of the experimental system is shown in Figure 13. The basic principle of the system design is the same. We combine half mirrors to show stereo disparity with multiple image planes at different depths. The only difference is we use more display panels and mirrors to realize longer distance between two image planes. The images displayed on D1 and D3 are shown to the left eye, while the images displayed on D2 and D4 are shown to the right eye. D5 is used to show the indices of depths.

Figure 14 shows how the experiment proceeds in detail. In the first experiment (experiment A), two displays are set 300mm and 380mm apart from the subject. On the display D5 six index lines are shown with the interval of 10 mm. The first index line is 310mm apart from the subject. With the help of these lines, the subject is asked to answer the depth of the stereo image with the resolution of 5mm. We show stereo images in three different ways: only on the near display (near plane condition), only on the far display (far plane condition), and on both displays (dual plane condition). When we use both displays, we show the dual image whose intensities are in inverse proportion to the distances between the stereo image and the displays. Stereo images with nine different depths between the second line and the fifth line are shown in the experiments. Since three different ways of image presentation are mixed together during the experiments, 27 trials are given for each subject without being notified whether single display or dual displays are used for presentation of the stereo image.

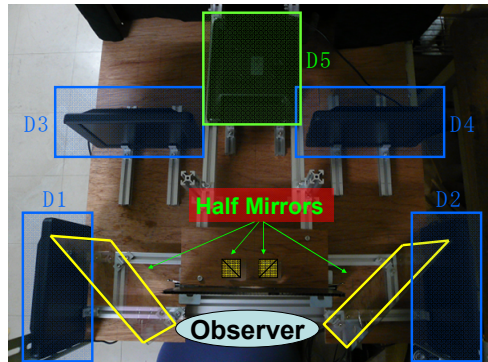


Figure 13: Design of experimental system on precision of depth perception.

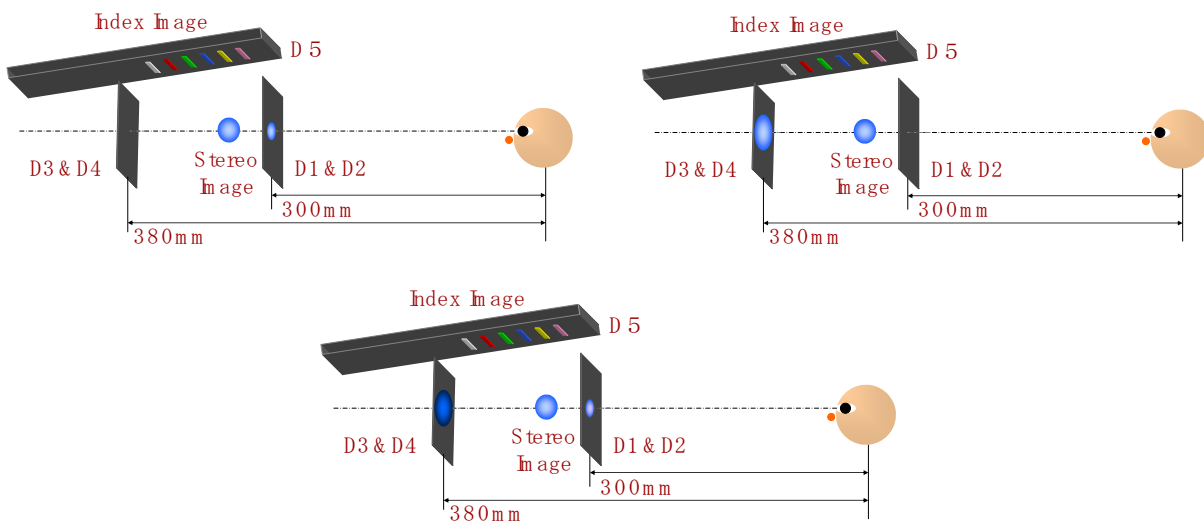


Figure 14: Design of experiment A with two panels 80mm apart.

In the second and the third experiments, two displays are set 300mm and 450mm apart from the subject as shown in Figure 15. In experiment B, the positions of the index lines and the presented stereo images are the same as those presented in experiment A. In experiment C, the index lines are located between 380mm and 440mm from the subject and the stereo images are presented between 390mm and 430mm from the subject.

The results of experiments A, B and C with 8 subjects are shown in Tables 1, 2 and 3 respectively. Here negative error means that the subjects' answer is nearer the subject's position than the right answer. As these tables show, the smallest absolute errors are always obtained under the dual plane condition. Also the absolute error and the standard deviation are smallest in experiments A, which is considered natural because the presented image is nearer and easy to grasp its precise position. The results of *F* tests assure that the differences of standard deviations are statistically significant between the single and dual display conditions except between far plane condition and dual plane condition in experiment C, which means that the dual plane condition can realize stable perception of depth.

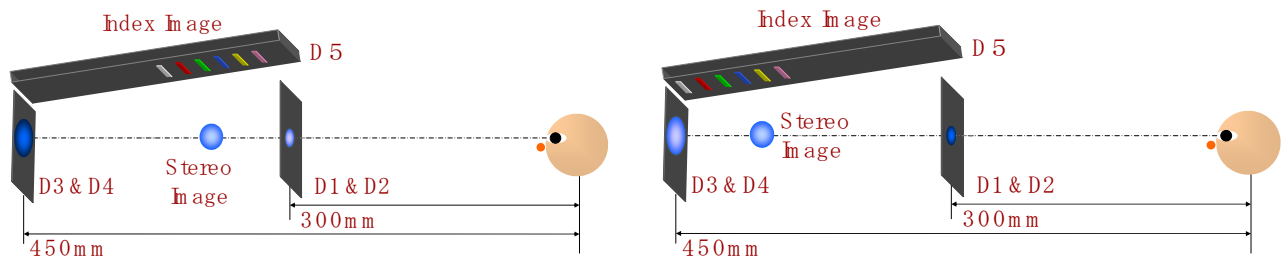


Figure 15: Design of experiments B (left) and C (right) with two panels 150mm apart.

Table 1: Results of experiment A.

|             | Average of absolute error (mm) | Average (mm) | Standard deviation |
|-------------|--------------------------------|--------------|--------------------|
| Near Plane  | 7.4                            | -0.6         | 11.9               |
| Dual Planes | 4.6                            | -1.3         | 6.6                |
| Far Plane   | 8.4                            | -1.0         | 11.0               |

Table 2: Results of experiment B.

|             | Average of absolute error (mm) | Average (mm) | Standard deviation |
|-------------|--------------------------------|--------------|--------------------|
| Near Plane  | 12.7                           | -4.7         | 15.6               |
| Dual Planes | 6.5                            | -1.6         | 9.0                |
| Far Plane   | 11.3                           | -6.5         | 13.6               |

Table 3: Results of experiment C.

|             | Average of absolute error (mm) | Average (mm) | Standard deviation |
|-------------|--------------------------------|--------------|--------------------|
| Near Plane  | 11.4                           | -7.5         | 11.2               |
| Dual Planes | 7.2                            | -2.4         | 8.9                |
| Far Plane   | 13.3                           | -5.3         | 14.7               |

## 5. SUMMARY

In this paper realization of precise depth perception using coarse integral volumetric imaging (CIVI) is discussed. To attain presentation of undistorted 3D space with CIVI, the authors simulate the optics of CIVI and an algorithm to show undistorted 3D space by compensating the optical distortion on the software basis is proposed. The feature of CIVI is combination of multiview and volumetric display technologies. The results of psychophysical experiments show that vergence-accommodation conflict is reduced and depth perception of the viewer is improved by combining multiview and volumetric technologies compared with the stereoscopic system that depends only on presentation of stereo disparity.

## ACKNOWLEDGEMENT

This research is partially supported by the Grant-in-Aid for Scientific Research, MEXT, Japan, Grant number: 22680008. The experiments are carried out with the help of Takeru Matsuda and Wang Qi.

## REFERENCES

- [1] H. Kakeya, "Improving Image Quality of Coarse Integral Volumetric Display", Proceeding of SPIE Vol. 7237, 723726 (2009).
- [2] H. Kakeya, T. Kurokawa, Y. Mano, "Electronic Realization of Coarse Integral Volumetric Imaging with Wide Viewing Angle", Proceedings of SPIE Vol. 7524, 752411 (2010).
- [3] T. Kurokawa and H. Kakeya, "Coarse integral imaging without pseudo image," Proceedings of SPIE Vol. 7863, 78631M (2011).
- [4] G. Lippmann, "La photographie integrale," Comptes Rendus Acad. Sci. 146, 446–451 (1908).
- [5] B. Lee, S. Jung, S.-W. Min, and J.-H. Park, "Three-dimensional display by use of integral photography with dynamically variable image planes," Opt. Lett. 26, 1481–1482 (2001).
- [6] J.-H. Park, S. Jung, H. Choi, and B. Lee, "Integral imaging with multiple image planes using a uniaxial crystal plate," Opt. Express 11, 1862–1875 (2003).
- [7] S.-W. Min, B. Javidi, and B. Lee, "Enhanced three-dimensional integral imaging system by use of double display devices," Appl. Opt. 42, 4186–4195 (2003).
- [8] H. Kakeya, "Coarse integral imaging and its applications", Proceedings of SPIE Vol. 6803, 680317 (2008).
- [9] R. Yasui, I. Matsuda, H. Kakeya, "Combining volumetric edge display and multiview display for expression of natural 3D images," Proceedings of SPIE Vol. 6055, 60550Y (2006).
- [10] H. Ebisu, T. Kimura, H. Kakeya, "Realization of electronic 3D display combining multiview and volumetric solutions," Proceedings of SPIE Vol. 6490, 64900Y (2007).
- [11] Y. Kim, J.-H. Park, H. Choi, J. Kim, S.-W. Cho, and B. Lee, "Depth-enhanced three-dimensional integral imaging by use of multilayered display devices," Appl. Opt. 45, 4334–4343 (2006).
- [12] Y. Kim, H. Choi, J. Kim, S.-W. Cho, Y. Kim, J. Park, and B. Lee, "Depth-enhanced integral imaging display system with electrically variable image planes using polymer-dispersed liquid-crystal layers," Appl. Opt. 46, 3766–3773 (2007).
- [13] S. Suyama, H. Takada, K. Uehira, S. Sakai and S. Ohtsuka, "A Novel Direct-Vision 3-D Display using Luminance-Modulated Two 2-D Images Displayed at Different Depths," SID'00 Digest of Technical Papers, 54.1, pp. 1208-1211, (2000).
- [14] S. Suyama, H. Takada and S. Ohtsuka, "A Direct-Vision 3-D Display Using a New Depth-fusing Perceptual Phenomenon in 2-D Displays with Different Depths," IEICE Trans. on Electron., Vol. E85-C, No. 11, pp.1911-1915 (2002).
- [15] S. Suyama, S. Ohtsuka, H. Takada, K. Uehira and S. Sakai, "Apparent 3-D image perceived from luminance-modulated two 2-D images displayed at different depths," Vision Research, 44, pp. 785-793 (2004).
- [16] S. Sawada and H. Kakeya, "Coarse integral volumetric imaging with flat screen and wide viewing angle," Proceedings of SPIE Vol. 7863, 78631L (2011).
- [17] K. J. MacKenzie, S. J. Watt, "Eliminating accommodation-convergence conflicts in stereoscopic displays: multiple-focal-plane displays can elicit continuous and consistent vergence and accommodation responses," Proceedings of SPIE Vol. 7524, 752417(2010).
- [18] K. J. MacKenzie, R. Dickson, S. J. Watt, "Vergence and accommodation to multiple-image-plane stereoscopic displays: 'Real world' responses with practical image-plane separations?" Proceedings of SPIE Vol. 7863, 786315 (2011).