# Acceptable distortion and magnification of images on reflective surfaces in an augmented-reality system

Shoji Yamamoto[1*], Natsumi Hosokawa[2], Mayu Yokoya[2], Norimichi Tsumura[2]

[1] Tokyo Metropolitan College of Industrial Technology, TOKYO, JAPAN,
[2] Graduate School of Advanced Integration Science, Chiba University, CHIBA, JAPAN.

## Abstract

In this paper, we investigated the consistency of visual perception for the change of reflection images in an augmented reality setting. Reflection images with distortion and magnification were generated by changing the capture position of the environment map. Observers evaluated the distortion and magnification in reflection images where the reflected objects were arranged symmetrically or asymmetrically. Our results confirmed that the observers' visual perception was more sensitive to changes in distortion than in magnification in the reflection images. Moreover, the asymmetrical arrangement of reflected objects effectively expands the acceptable range of distortion compared with the symmetrical arrangement.

**Keywords**: Consistency, Magnification, Distortion, Environmental map, Virtual showcase

* E-mail Address : yamasho@metro-cit.ac.jp

## 1. Introduction

Recently, virtual showcases using augmented reality have become widespread in internet-based shopping and virtual catalogs. For example, a dressing-room simulation allows the user to confirm the outward appearance of clothes at home [1]. In addition, virtual furniture systems can verify the size and position of a three-dimensional (3D) model of the furniture using arbitrary points in the user's room for calibration, as shown in Fig. 1 [2]. For such applications, exact 3D reproductions of the products and of the environment are required, since augmented reality should provide high-fidelity representation of commercial products.

Fig. 1

For enhancing the realism of augmented reality systems, the computer-generated reflection model is a very important factor. Usually, this model consists of diffuse reflection and specular reflection [3]. The diffuse reflection mainly represents the color and texture based on the material property, and the specular reflection mainly represents reflections of images and light on the surface. When we reproduce the realistic 3D object in virtual showcase, a rendered object with only diffuse reflections is largely unaffected by changes in illumination and the surround scene. In contract, if specular reflection is included, the result is more realistic because the reflection images depend on the object's position within the surrounding scene. Even a small error in reflections results in a feeling of strangeness and reduces the perceived realism of the virtual showcase.

Ordinarily, cube mapping is used to reproduce the reflection correctly on the 3D object [4-7]. This method is suitable for rendering consistent reflection images using an environment map captured at the appropriate location [8-10]. The location of the environment map should be chosen with care to ensure consistency in the augmented-reality system, because users are sensitive to small differences in reflection images that reveal the positional relationship between the virtual 3D object and its surroundings. To reproduce a faithful reflection image at arbitrary positions, it is necessary to prepare an image of the surrounding scene that corresponds to each possible position. However, preparation of such an environment map is impractical. A feasible workaround to overcome this problem is required. Therefore, we investigate correspondence between reflection images and environmental scenes, and acceptable distortions of reflection images. Based on human visual perception, some degree of inaccuracy in correspondence may be acceptable between the reflection image and environmental scenes [11]. For example, distance is difficult to judge from a reflected scene [12, 13]. This ambiguity may allow a compromise between the accuracy of a virtual or augmented reality system and the computational feasibility for practical use.

## 2. Related work of visual consistency

According to Helmholtz's theory of vision and perception, retinal images are inherently ambiguous recognition regarding the size, shape and distance of objects [14]. These ambiguities are usually resolved using knowledge or assumptions based on real-world experience. These assumptions are used when we observe an object that has some characteristics in common with real objects. The CAPTCHA is a good example to demonstrate this ability [15]. Recognizing the distorted characters of a CAPTCHA is difficult for a machine, but quite simple for a human. Similarly, our visual perception of reflections allows for some degree of distortion.

The distortion of images can be evaluated by visible difference predictors, which are based on the principle of visual perceptual evaluation [16]. Peak signal-to-noise ratio was an early metric used to evaluate differences between images, and was employed in formulas considering the dynamic range characteristics of human vision [17, 18]. Structural similarity was designed to improve on metrics such as peak signal-to-noise ratio [19]. Structural similarity considered image deformation as change of perception in structural information. Although these metrics are useful for the evaluation of image degradation and geometric distortion, it is difficult to apply them to human perception of visual ambiguities.

Ramanarayanan et al. introduced the concept of visual equivalence for use in evaluating image fidelity in computer graphics [20]. They proposed that two 3D models with original and warped illumination map are sometimes visually equivalent, even if the reflection patterns are visibly different. To understand this phenomenon, Lin and Kuo conducted several experiments, rendering various transformations of environment maps onto a modified sphere [21]. The modified sphere included bumps of uniform amplitude and increasing spatial frequency, and the environment map was blurred with Gaussian kernels and Perlin noise before being rendered onto the modified sphere [22]. Their results indicate that visual perception of the reflection image was less sensitive to blurring, and more sensitive to warping of environment maps.

Warping inconsistency is also generated by misaligning the captured environment map and its reflection image in a virtual showcase. In particular, a substantial error is induced by offsetting the reflection image from the surrounding objects in the visual field, as shown in Fig. 2. Figure 2(a) and (b) shows a reflective ball with reflection images that are reproduced from environment maps captured at different positions. An off-center position image of Fig. 2(b) gives an inconsistent visual perception, in contrast to Fig. 2(a). To reduce this error it is necessary to prepare an environmental map corresponding to each possible position. However, preparing numerous environmental maps in this way is impractical. Therefore, in the present paper, we investigate methods to reduce the number of capture positions for the environment map while maintaining a consistent visual perception, taking advantage of visual ambiguity.

Fig. 2

## 3. Preliminary Simulation for Visual Consistency

In our experiment, reflection image recognition is evaluated using an image-based method. Before the actual experiment, we performed a preliminary evaluation using computer graphics to limit the experimental parameters. The variable parameters were the shape of the 3D object, the roughness of the reflective surface, and the viewing angle of elevation, which affect

Table 1

the distortion, sharpness, and composition, respectively, of the reflected image. For this evaluation, we prepared pairs of reflection images. Each reflection image was rendered using an environment map captured at an equidistance position, with one image in each pair rendered at a 10° rotation in the horizontal direction. For the shape parameter, we prepared three objects: a sphere, an amorphous blob, and a Utah teapot. Surface roughness was simulated by rendering using 3×3 or 5×5 Gaussian filters for the environment map; no filter was used for the 'zero surface roughness' case. Three viewing elevation angles were employed: 15°, 30°, and 45° from horizontal.

Ten observers compared the pairs of images, and chose which was distorted, based on the inconsistency of the reflection images. The results of this evaluation are shown in Table 1. A high score indicates a high sensitivity of that parameter. The Utah teapot shape, zero surface roughness, and 30° viewing elevation resulted in the highest scores for the following reasons. The shape of teapot allows a reflection image to overlook a broad arrangement of objects. A perfect mirror reflection image on the surface of object can make even slight differences in the surroundings noticeable. A simultaneous view of the reference object and floor obtains the best approval when the viewing elevation is 30°. Using these conditions, we next evaluate visual consistency with an actual image-based experiment.

## 4. Experimental system

Our experimental layout is shown in Fig. 3. In this system, we used two cameras; an omnidirectional camera (Ladybug 2; Point Grey Research; 1024×768) to capture 360° environment images, and a digital camera ("view camera"; LU175-SC-IC; Luminera; 1280×1024) and lens (Fujinon, HF9HA-1B; f = 9 mm) to capture two-dimensional images of the composed scene. The scene to be used in the augmented reality evaluation was reproduced from images captured by the view camera. We used an augmented-reality marker, which was set under the omnidirectional camera, to measure its position accurately. The 3D model was created as a virtual object, and the reflection image on the surface of the 3D model was rendered by mapping the 360° environment image to the 3D model. Two pairs of wood blocks were used as the reference objects, which were set in front of the omnidirectional camera to replicate the initial computer-generated 3D model. Fig. 3

The reflection image on the surface of the 3D model was changed by moving the omnidirectional camera, as shown in Fig. 4: left or right displacement distorted the reflection image ("distortion"); forward or backward displacement scaled the reflection image ("magnification"). The observers evaluated how closely the reflection image correlated with the position of the 3D model, based on the degree of distortion or magnification. Fig. 4

Other conditions, such as the position of view camera, focus length of each camera, and illumination were fixed in this experiment. A black curtain was used to envelop the experimental booth, so that the two cameras capture only the wood blocks.

## 5. Procedure

Figure 5 shows the experimental layout set in a symmetrical arrangement. At first, the omnidirectional camera was set to the position of circle 1 (hereinafter called #1), defined as the base position (Fig. 5). Moving the omnidirectional camera to #2, #3, #4 or #5 caused magnification of the reflected image; moving it to #6, #7, #8 or #9 caused distortion of reflection image. Examples of final evaluation images are shown in Fig. 6. Figure 6(a) shows the reproduced reflection image at #1, Fig. 6(b) shows it at #2, and Fig. 6(c) shows it at #6. We performed a subjective evaluation to investigate the sensitivity for the change of magnification and distortion by using these images. Fig. 5,6

Figure 7 shows a layout of the subjective evaluation system, which simulates a virtual showcase. We used a 17-inch display (FlexScan L551, EIZO, 338×270 mm) placed at a distance three-times the height of the display (810 mm) from the observer. The observers subjectively evaluated the consistency of the reflection images with the real objects. Observers assigned a score from 1 (seriously inconsistent) to 5 (excellent consistency) to each image (Table 2).

Fig. 7

Table 2

If observers evaluated the reflection image with a score of ≤ 2, the observer's comments were recorded along with the score. The images for all positions (#1 to #9) were presented in a random order to each of the observers, and this trail was repeated three times. Overall, 10 observers (6 men and 4 women, age 22–25 years) participated in the evaluation, all of whom had normal or corrected-to-normal vision. In total, 30 evaluations were obtained for each image.

# 6. Results of evaluation

## 6.1 Symmetrical arrangement

Figure 8 shows the average scores for all trial. The x-axis indicates the position of the captured environment map, and the y-axis indicates the subjective score for each image. All observers assigned the highest score to images generated at #1, the base position, and also assigned relatively high scores at #2 and #3. By contrast, images at #4 and #5 were assigned low scores by most observers. These results suggest that our perception of magnification and distances has some tolerance, but extreme changes in magnification result in a sense of inconsistency or error in reflection images.

Images generated at #6, #7, #8 and #9 were all assigned a low score by most observers. This result suggests that observers are relatively sensitive to changes in distortion caused by displacement to the left or right. Thus, visual perception is more tolerant to changes in magnification than distortion of the reflection image.

Fig. 8

From these experiments, we assume that the positional relationship between wood blocks and reflection image has an effect on the sensitivity to changes whether the wood blocks are arranged symmetrically or not. This assumption gives us a new interest on whether the visual perception remains more sensitive to changes in distortion than to changes in magnification, even if two blocks lose the symmetry in the reflection image. Therefore, we performed the next experiment by arranging two pairs of wood block asymmetrically in the reflection image.

## 6.2 Asymmetrical arrangement

Figure 9 shows an additional experimental layout where the wood blocks were arranged asymmetrically. The two pairs of wood block were rotated 20° about the base position to make an asymmetrical arrangement. The omnidirectional camera was set in positions #1 through #9 as in the previous experiment, and additional positions #10 through #13 as shown in Fig. 9. The reflection images at #10 through #13 caused only magnification changes. In contrast, the reflection images at #2 through #9 caused changes in both distortion and magnification, since the wood blocks were asymmetrically set. However, the degree of distortion and magnification at each point differed: #2 through #5 mainly generated magnification changes, and #6 through #9 mainly generated distortion changes in the reflection images. It is assumed that relationship between magnification and distortion depends on the angle of rotation of the blocks. The other conditions and procedures were the same as the previous experiment. Figure 10 shows an example of rendered images in this experiment. Fig. 9,10

The average scores of all trials are shown in Fig. 11. From the results, we found that almost all of the observers assigned higher scores at #2, #6, and #10. Here, the score of #6 is higher than the score at the case of symmetrical arrangement. In this case, the reflection image at #6 as shown in Fig. 10(f) slightly makes the change of distortion indiscernible compared with the case of Fig. 6(c). The results show that the observers had difficulty recognizing consistency between real objects and the reflection image. Therefore, we assume that an asymmetrical arrangement desensitizes the visual perception of observers. Fig. 11

Considering the results by focusing on each series from #2 to #5, #6 to #9, and #10 to #13, the rapid decrease of scores appears in this result of #6 to #9, even if the scores of #2 to #5 and #10 to #13 decrease gradually. Owing to the asymmetrical arrangement, the amount of distortion of #6 through #9 reduces in comparison with the case of symmetrical arrangement. However, the excessive displacement to the left or right causes inconsistency, resulting in lower scores from observers for #7 through #9. In contrast, the scores assigned gradually decrease for #10 through #13. Since the reflection images generated from #10 to #13 positions are only magnified, this result increases the reliability that visual perception is less sensitive to magnification than to distortion of the reflection image.

## 7. Discussion

We performed a quantitative simulation for the actual size and angle change in the reflection image. Based on the specification of the view camera, the size of the captured image with wood blocks was calculated. Under the condition that this captured image was displayed by full screen mode, we converted the size and position of the reflected objects to the viewing angle of the observer. Although the size and position of reflection image on the virtual object depended on the size and shape of the object, we coordinated the size of the wood blocks in the reflection image using one viewing angle. Thus, we simulated the change in size and position of the reflection image. This change was calculated by ray tracing, as shown in Fig. 12, since the mirror reflection was on the surface of the virtual object between rays from the object and rays from the observer. Fig. 12

Figure 13 shows the results of this simulation. The horizontal line indicates the angle of view where the center position of the wood block is defined as zero. The vertical line indicates the relative width of the wood block based on the width of wood block at #1 base position. Figure 13(a) shows the change of size and position in the symmetrical arrangement, and Fig. 13(b) shows the change of size and position in the asymmetrical arrangement. In these figures, the range which is higher than "normal consistent" score with reference to the results of Fig. 13

evaluation as shown in Fig. 8 and 11. From the results of the symmetrical arrangement as shown in Fig. 13(a), it is clear that visual perception is insensitive to magnification up to 20% decreases in reflection image size. In contrast, visual perception becomes more sensitive for an imbalance in reflection image such as #6 in the symmetrical arrangement. In this case, we can perceive inconsistency for a shift of only 1° in angle of view. In contrast, for the asymmetrical arrangement, the range of tolerance for left and right movement of the omnidirectional camera is about twice that for the symmetrical arrangement, as shown in Fig. 13(b). The change of size ratio at #6 is almost equal between symmetrical and asymmetrical arrangement. This result shows that, for #6, the distortion is mitigated by the asymmetrical arrangement.

Through our experiment, there are two unknown parameters: the size of objects and the curvature of the virtual object. Owing to these unknown parameters, it is difficult to quantify the magnification change caused by movement to #2 through #5 in the symmetrical arrangement or to #10 through #13 in the asymmetrical arrangement. However, new consistency, example for the uniformity of two wood blocks, may be caused by the symmetrical arrangement. Therefore, we instantly notice the imbalance of uniformity generated by distortion. Asymmetrical arrangement does not cause such uniformity and therefore it can obscure the visual perception of distortion. For the construction of an easy-to-use virtual showcase, it is necessary to consider the consistency of visual perception between real and virtual object.

## 8. Conclusion

This paper presented the results of subjective evaluations to investigate visual perception of a reflection image. The reflection image was constructed from environment maps captured by an omnidirectional camera at various positions. Observers' sensitivity to distortion and magnification were evaluated using two sets of experiments. In the first experiment, two wood blocks were arranged symmetrically as reflected objects from the base position of omnidirectional camera. We evaluated the visible difference by using the reflection image captured at different positions relative to the base position. The results of the evaluation showed that observers were less sensitive to changes in magnification than in distortion of the reflection image, although only a 1° change was generated in the angle of view. In the second experiment, the two reflected objects were placed asymmetrically from the same base position. The results of this evaluation showed that almost all observers tolerated the distortion generated by a 1° shift, and the range of tolerance for distortion was expanded to about twice that for symmetrical arrangement. The results show that the differences are mitigated by the asymmetrical arrangement.

In general, we notice the difference when the visual consistency is lost about the size and/or shape in the virtual showcase. Especially, a well-marked consistency such as symmetrical arrangement generates intense sensation as the case of symmetrical arrangement. However, this sensation is obscured by ambiguous condition which has unknown parameters as same as the asymmetrical arrangement. Therefore, consistency of visual perception is important when reproducing virtual objects in real scenes. An asymmetrical arrangement is appropriate to obscure the inconsistency, and can reduce the necessary number of images captured with an omnidirectional camera.

The key finding in this paper is a slight tendency of visual consistency. Additional trials of subjective evaluation by larger groups of participants are necessary to confirm the consistency of visual perception. Moreover, it is important to consider other optical effects, such as shading, shadow, and transparency in the virtual showcase and the relationships between these and visual perception in future work.

## Acknowledgment

This research was partially supported by the Ministry of Education, Science, Sports and Culture, Japan Grant-in-Aid for Scientific Research, Brain and Information Science on SHITSUKAN, 23135530 (2012), 25135707 (2013), and Grant-in-Aid for Scientific Research(C), 15K00415 (2015), respectively.

## References


[1] M. Yuan: IEEE Transactions on Multimedia, **15** (2013) Issue 8, 1958.

[2] JKT Tang: IEEE VSMM, **14** (2014), 345.

[3] S. A. Shafer: Color Research & Application, **10** (1985) Issue 4, 210.

[4] J. F. Blinn, M. E. Newell : Communications of the ACM (1976), 542.

[5] C. R. Hoffman, G. S. Miller: SIGGRAPH 84 Course Notes for Advanced Computer Graphics Animation (1984).

[6] S. E. Chen: Proc. of SIGGRAPH '95 (1995), 29.

[7] R. Mantiuk, K. Myszkowski, S. N. Pattanaik: Proc. of International Conference on Compute Vision and Graphics (2002), 530.

[8] M. F. Cohen, D. P. Greenberg: Proc. of SIGGRAPH '85 (1985), 31.

[9] N. Greene: IEEE Computer Graphics and Applications **6** (1986) Issue 11, 21.

[10] P. Debevec: Proc. of SIGGRAPH '98 (1998), 189.

[11] N. Hosokawa, S. Yamamoto, M. Yokoya, N. Tsumura: The 20th Color Imaging Conference (2012), 100.

[12] A. Murgia, P. M. Sharkey: The International Journal of Virtual Reality **8** (2009) Issue 1, 67.

[13] J. Qian, Y. Petrov: Journal of Vision, **13** (2013) Issue 7, 1.

[14] D. Heyer, R. Mausfeld: *Perception and the Physical World: Psychological and Philosophical Issues in Perception* (Wiley; New York, 2002), Chapter 5.

[15] G. Moy, N. Jones, C. Harkless, R. Potter: Proc. of IEEE computer society conference on Computer vision and pattern recognition (2004), 23.

[16] S. Daly: *The Visible Differences Predictor: An algorithm for the assessment of image fidelity* (In A. Watson, Digital Image and Human Vision, Cambridge, MA: MIT Press, 1993), pp.179-206.

[17] Z. Wang, A. C. Bovik: IEEE Signal Processing Letters **9** (2002), Issue 3, 81.

[18] Q. Huynh-Thu, M. Ghanbari: IEEE Electronics Letters **44** (2008) Issue13, 800.

[19] Z. Wang, A. C. Bovik , H. R. Sheikh , E. P. Simoncelli: IEEE Transactions of Image Processing **13** (2004) Issue 4, 600.

[20] G. Ramanarayanan, J. Ferwerda, B. Walter, K. Bal: ACM Transactions on Graphics **26** (2007) Issue 3, 1.

[21] W. Lin, C,C.J. Kuo: Journal of Visual Communication and Image Representation **22** (2011) Issue 4, 297.

[22] K Perlin: ACM Transactions on Graphics **21** (2002) Issue 3, 681.

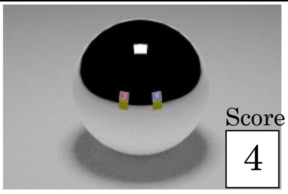
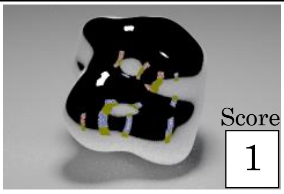
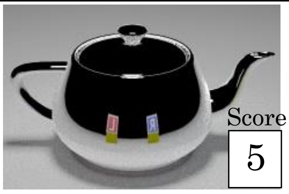
**Table captions**

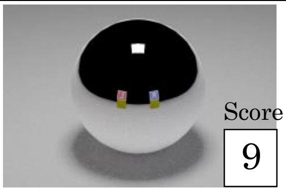
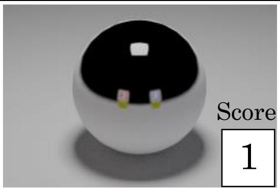
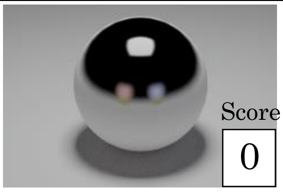
Table 1 Rendered images using each parameter for preliminary evaluation

Table 2 Five grades of evaluation

## Tables

Table 1 Rendered images using each parameter for preliminary evaluation

| Shape | | |
|---|---|---|
| Sphere | Amorphous blob | Utah teapot |
|   |   |   |

| Roughness | | |
|---|---|---|
| None | 3×3 | 5×5 |
|   |   |   |

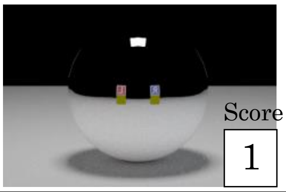
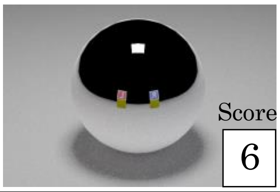
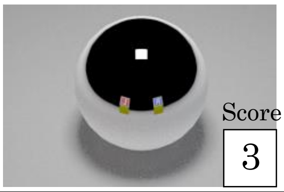
| View angle | | |
|---|---|---|
| 15° | 30° | 45° |
|   |   |   |

Table 2 Five grades of evaluation

| Score | Degree |
|---|---|
| 1 | Serious inconsistent |
| 2 | Slight inconsistent |
| 3 | Normal consistent |
| 4 | Good consistent |
| 5 | Excellent consistent |

**Figure captions**

**Fig. 1** Illustration of an augmented reality system

**Fig. 2** Rendering using environment maps at different positions. (a) center position (b) off-center position

**Fig. 3** Experimental system for evaluation

**Fig. 4** Illustration of reflection images with the change in omnidirectional camera position

**Fig. 5** Experimental layout for the symmetrical arrangement (a) visual layout (b) geometric layout viewed from above

**Fig. 6** Rendering of reflection images at selected positions in the symmetrical arrangement (a) #1 (b) #2 (c) #6

**Fig. 7** Experimental layout for the symmetrical arrangement

**Fig. 8** Results of subjective evaluation in the symmetrical arrangement

**Fig. 9** Experimental layout for the asymmetrical arrangement (a) visual layout (b) geometric layout viewed from above

**Fig. 10** Rendering of reflection images at each position in the asymmetrical arrangement (a) #1 (b) #2 (c) #3 (d) #4 (e) #5(f) #6 (g) #7 (h) #8 (i) #9 (j) #10 (k) #11 (l) #12 (m) #13

**Fig. 11** Results of subjective evaluation in the asymmetrical arrangement

**Fig. 12** Ray-tracing simulation for calculating the actual position of reflection image

**Fig. 13** Result of ray tracing simulation for the actual position of two wood blocks on the mirror surface (a) symmetrical arrangement (b) asymmetrical arrangement

**Figures**

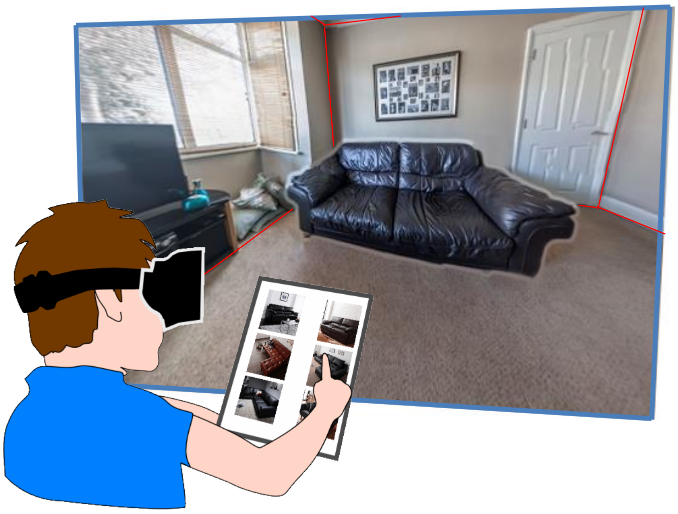
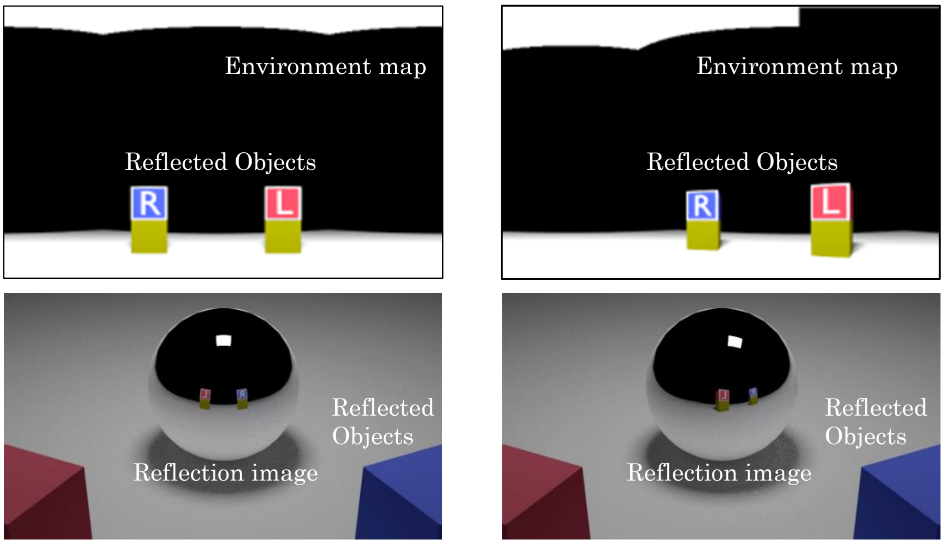


Fig. 1 Illustration of an augmented reality system




(a) center position (b)off-center position

Fig. 2 Rendering using environment maps at different positions

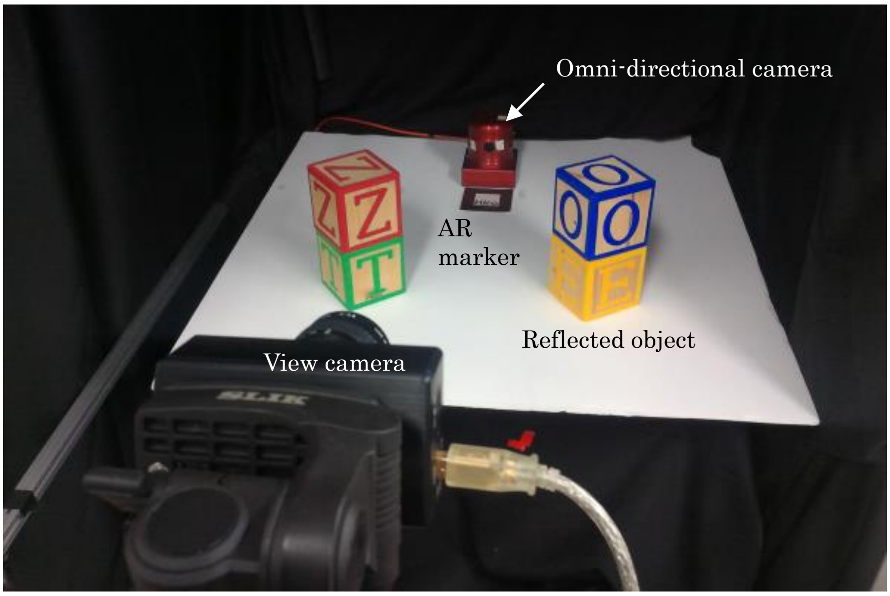



Fig.3 Experimental system for evaluation

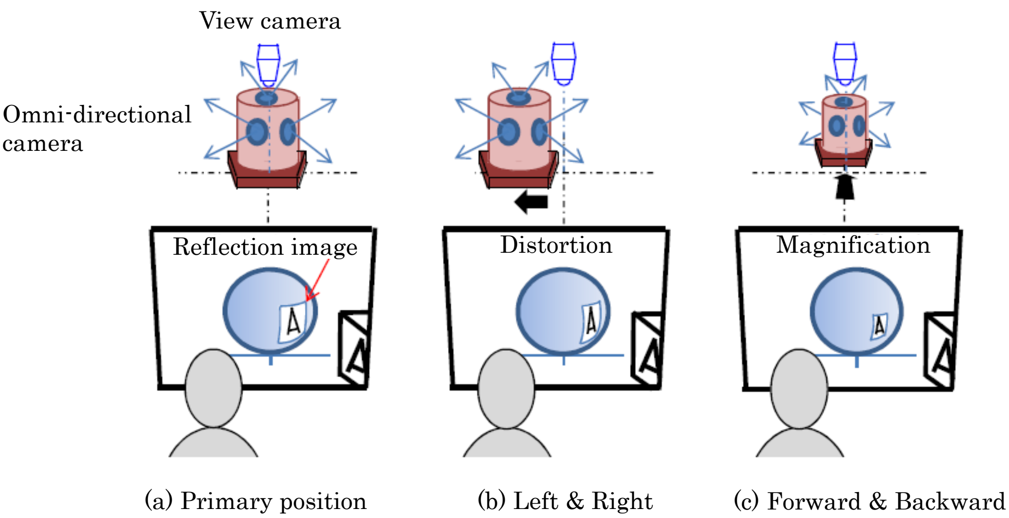
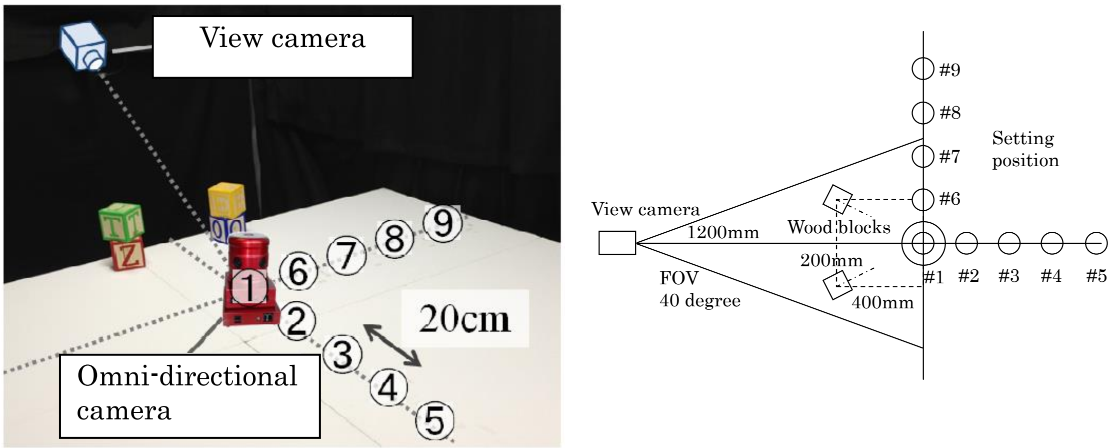
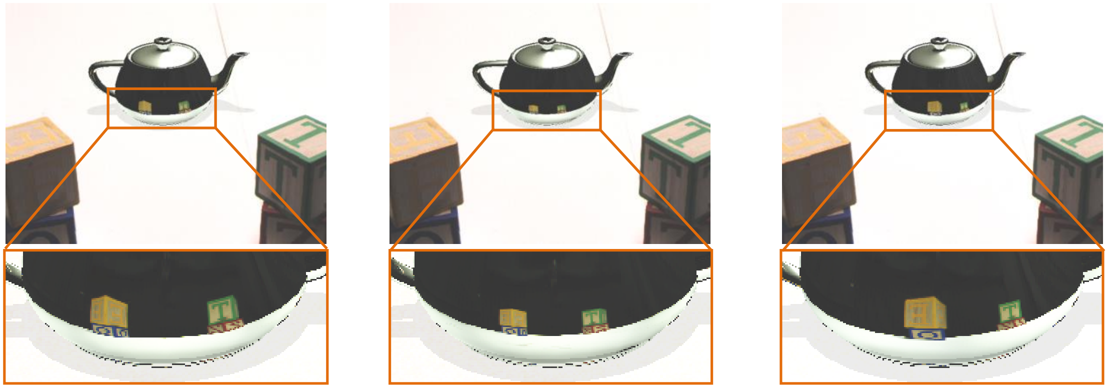



Fig.4 Illustration of reflection images with the change in omnidirectional camera position




(a) visual layout (b) geometrical layout viewed from above

Fig.5 Experimental layout of symmetrical arrangement



(a) #1 (b) #2 (c) #6

Fig.6 Rendering of reflection images at selected positions in the symmetrical arrangement

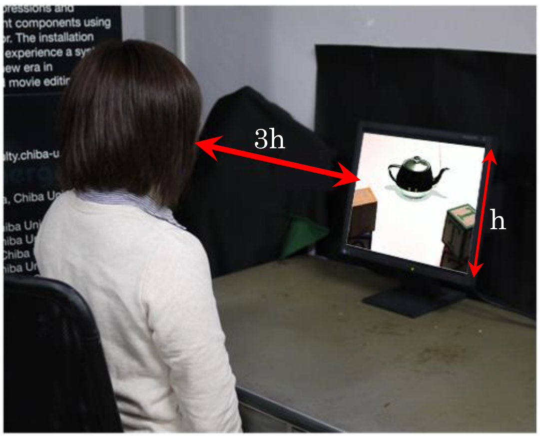



Fig.7 Experimental layout for the symmetrical arrangement

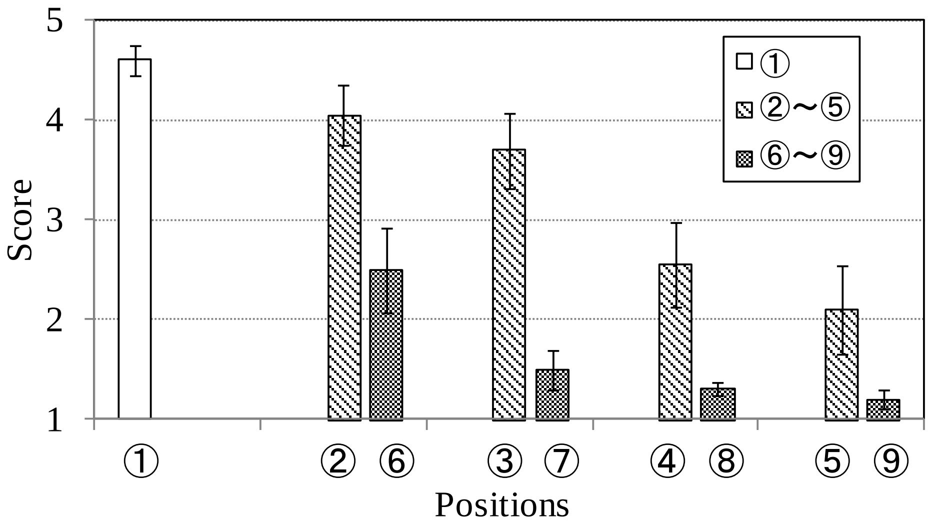
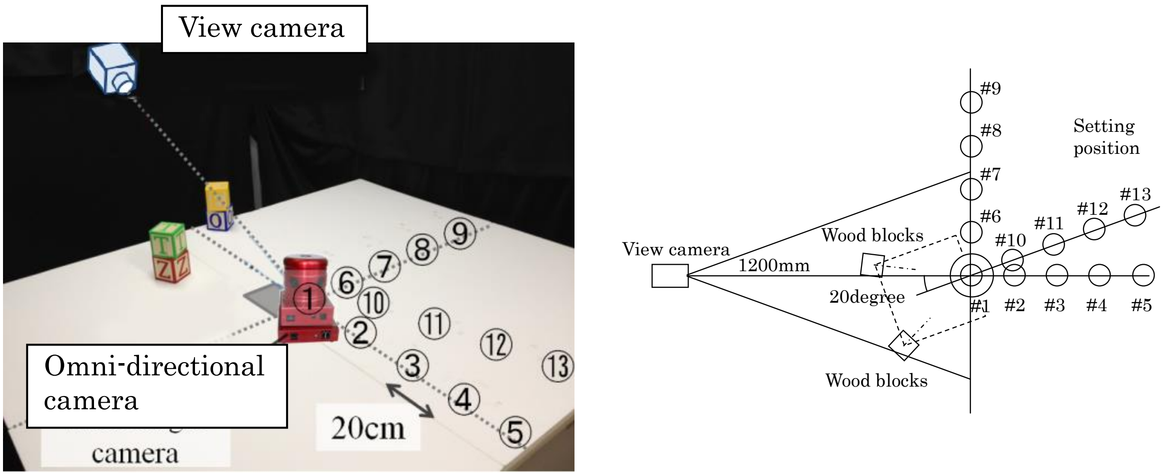



Fig.8 Results of subjective evaluation in the symmetrical arrangement




(a) visual layout (b) geometrical layout viewed from above

Fig.9 Experimental layout for the asymmetrical arrangement

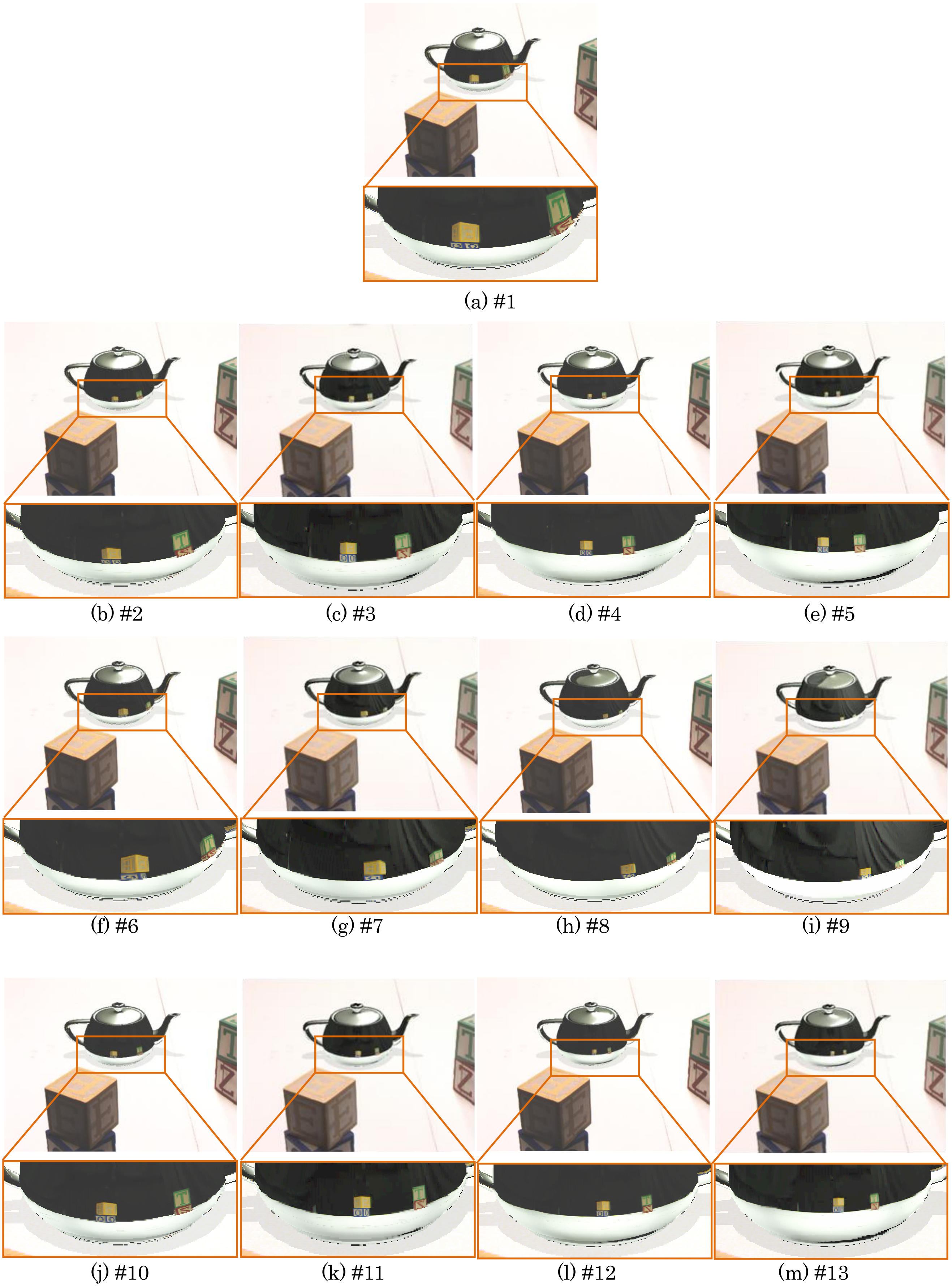



Fig.10 Rendering of reflection images at each position in the asymmetrical arrangement

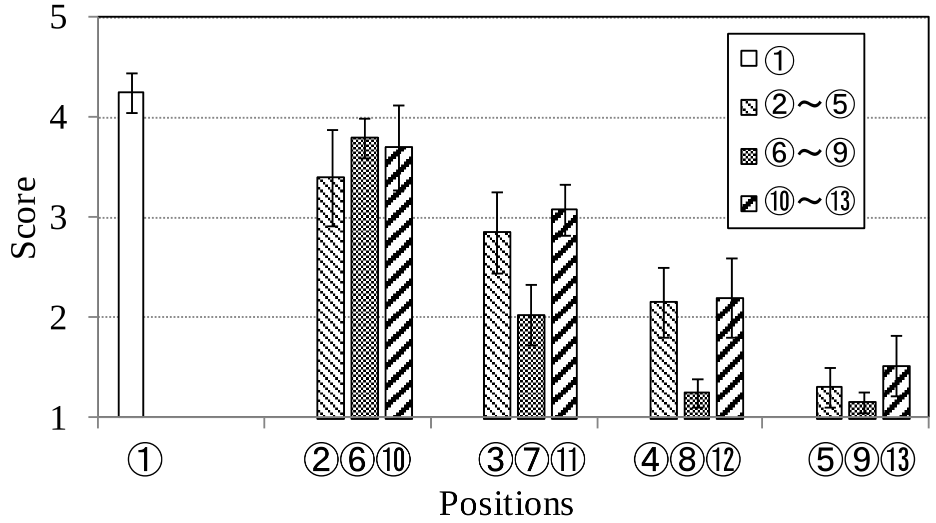



Fig.11 Results of subjective evaluation in the asymmetrical arrangement

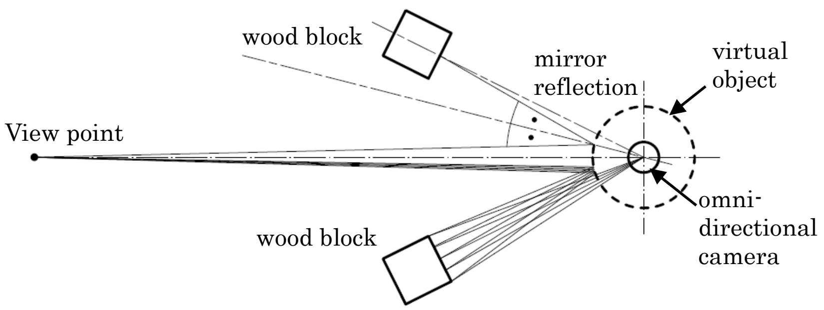
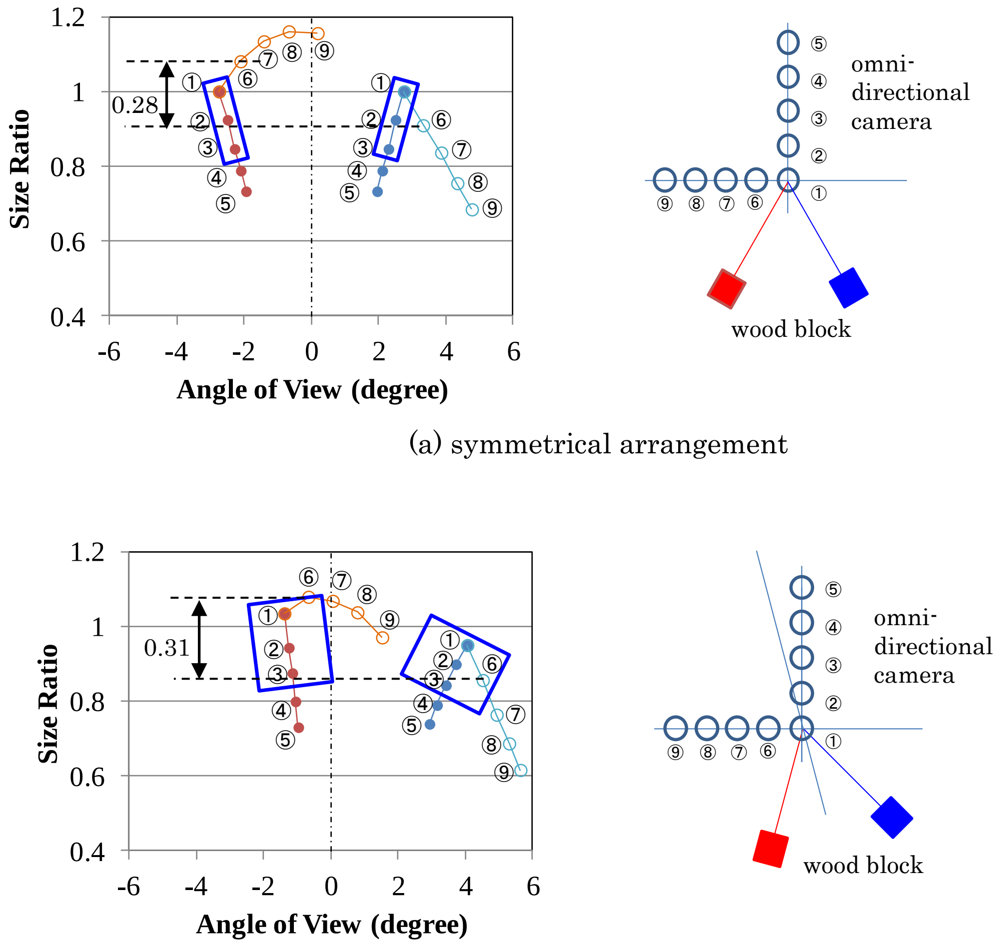



Fig.12 Ray tracing simulation for calculating the actual position of reflection image




(a) symmetrical arrangement



(b) asymmetrical arrangement

Fig.13 Result of ray tracing simulation for the actual position of two wood blocks on the mirror surface.