Transformation of Environment Map by Changing the Captured Position in a Cuboidal Room

Toshimitsu Kugimoto[1], Shoji Yamamoto[2*], Ryota Domon[1], Kentaro Hikosaka[3], Norimichi Tsumura[1]

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

Abstract

In this paper, we consider the use of an environment map as a photographic support tool for understanding the condition of surrounding light. It is well known that an environment map should be captured at a position directly above the target object. However, this is not always possible. Therefore, we developed a practical transformation method for an environment map that was captured from a different position. The proposed method is effective in an indoor scene where the direction of a light source differs significantly depending on the captured position. A mapping for a rectangular parallelepiped (cuboid), and perspective information estimated from a corner position of a room is used to calculate the transformed environment map at the target position. The method was applied to the transformation for an actual room, and it was verified that the transformed environment map almost corresponds to the ground truth of the environment map captured at the target position.

Keywords: Environment map, Vanishing point, Rectangular parallelepiped, Image processing

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

1. Introduction

It seems intuitive that the appropriate selection of the light source position is very important to take an impressive picture. The best shot to emphasize the shape and material appearance of an object will be obtained by control of the light settings.[1-3] A professional photographer with a great deal of experience can judge the effect of light settings.[4-6] Instead of relying on the experience of a professional photographer, we researched a virtual support system for amateur photographers using computer graphics.[7-8] This support system is achieved with an environment map captured by an omnidirectional camera, which can acquire all of the necessary information regarding the light position.[9-11]

Our system has the limitation that this environment map must be captured at a position immediately above the target object. This may create complications if the photographer and the target object are separated by some distance. Difficulties arise when it is impossible to approach the target object. To solve this problem, we considered a method for transferring the environment map from the photographer to the target position with a reasonable degree of accuracy.[7-8] This method employs a simple transformation from polar coordinates to 3-dimensional Euclidean space by cube mapping.[12-15] The environment map above the target object is acquired by movement of the center of the cube map from the photographer to the target position and by inverse transformation from 3-dimensional Euclidean space to polar coordinates.

In the case of an outdoor scene with sunshine, our conventional method can change the position of the environment map accurately, because the angle to the sun hardly changes with movement of a few meters. In the case of a cubic room, it is also possible to change the position of light with the wall and ceiling accurately, because the angle in polar coordinates can be measured exactly from the position in 3-dimensional Euclidean space. However, rooms are typically not perfectly cube, and are often cuboid such as rectangular parallelepiped which is defined as a shape that has six faces that are all rectangles. Errors can arise in the transformation of the environment maps for these rooms when conventional cube mapping is used.

In this paper, we propose a practical transformation method from polar coordinates to 3-dimensional Euclidean space in a cuboidal room. Our proposed method uses only a single environment map captured at the photographer's position. The corner positions of the room in the environment map are specified manually, and each pixel of the environment map is replaced for the rectangular parallelepiped according to the perspective information estimated from the corner positions. The movement between the position of the photographer and the target position is computed by the relative distance in 3-dimensional Euclidean space. Finally, the transformed environment map at the target position is reconstructed. These approaches are described in Section 3, and the verification result of our proposed method is described in Section 4. As the verification, we apply our transformation to two types of practical cuboidal rooms. The result of the transformation and the verification of the accuracy are discussed in Section 5.

2. Related work on environment mapping

In computer graphics, environment mapping is an efficient image-based lighting technique for reproducing the appearance of a surface reflection.[16-18] Also, environment mapping is used for rendering a real scene as global illumination, which includes both direct and indirect light sources with infinite distance. Figure 1 shows an environment map captured in a room by using an omnidirectional digital camera [Ricoh, Theta]. Equation 1 denotes the allocation of latitude and longitude in the environment map, which is shown as the equirectangular image in Fig. 1.[19]

$$X = H\left[\frac{90-\theta}{180}\right], \quad Y = W\left[\frac{\phi+180}{360}\right] , \qquad (1)$$

where θ is the latitude from 0 to 180 degree and ϕ is the longitude from -180 to 180 degree, H is the height and W is the width of the image, and X and Y are the coordinates of each pixel starting from the upper left corner. Thus, an original point of θ and ϕ is located at the center of image in Fig.1.

It is well known that cube mapping is a suitable method for rendering specular reflection in computer graphics. This mapping has six square images, which are transformed from an equirectangular image, and these images address the viewing direction from the captured position. Figure 2 shows the sequence of the rendering process when the environment map is used as a reflection image on the surface of a shiny metal object.[20-22] The six faces of the cube map in Fig. 2(b) indicate the front, back, left, right, up and down directions. Furthermore, the environment map is useful as an eccentrically located light source such as global illumination. In this case, the data structure of the cube mapping technique enables fast and accurate lighting computation.[23]

Here, it is generally considered that a captured position of the environment map should be matched with the proximity of the target object, as shown in Fig. 2(c). If surrounding objects such as light sources exist far from the captured position, the difference between the captured position and the target object is insignificant with respect to the accuracy of direction. For example, in the case of the outdoor scene shown in Fig. 3(a), the environment map captured at the photographer's position indicates almost the same direction of the sun, compared with that captured at the target position. Similarly, Figure 3(b) shows the case of an indoor scene with consistency between the cube map and the shape of the room. These observations about the captured position are beneficial for understanding the surrounding light condition by the reason that its position is regardless. However, in the case of an indoor scene without consistency between the cube map and the shape of the room, as shown Fig. 3(c), an obvious difference of the viewing angle of the light position is observed due to the difference of the captured position.

To perform an accurate transformation in a room other than a cubic room, we need to estimate the relative rate of the distance between the position of the photographer and the target position in the room. For estimating the rate without additional measuring equipment, we refer to the reconstruction method of 3D information from only an image, as proposed by Horry et al.[24] Their method “Tour into the Picture” was presented for making simulations of a 2D picture or photograph of a scene from different viewpoints. In this method, a simulation is created from the viewpoint of a camera that can be three-dimensionally “walked or flown through” the 2D image. This method enables us to produce a pseudo three-dimensional space from a two-dimensional image and experience as if walking into the image. In addition, the method roughly specified vanishing points manually and set a spidery mesh to prescribe a few perspective conditions, as shown in Fig. 4. When the spidery mesh is displayed, the vertices of the inner rectangle and vanishing points can be moved by clicking the neighborhood and dragging it. The criteria for the depth of the image are set by the four vertices and the vanishing points.

The abovementioned method by Horry et al.[24] is very useful for the simulation of images that include vanishing points and four vertices of a front wall. The vanishing points and vertices can be used for estimating the relative distance in the image. In this paper, we combine the forward and backward relative distance derived from Horry’s algorithm.

3. Reconstruction of transformed environment map

In this section, we explain our proposed method, which can transform the environment map from the position of the photographer to the position of the target in a non-cubic room. First, conventional cube mapping is applied to expand the environment map, as shown in Fig. 5. In the expanded images, we select two images, which have four corners in the front and back walls of a cuboidal room. Next, we estimate the perspective projection from the selected images by applying the method proposed by Horry et al. According to their method, four corners and the vanishing points are specified manually by using a mouse click. The spidery mesh is calculated by using these points automatically. Here, the spidery mesh is useful to recognize the relative positions in these images, because this mesh indicates an equally spaced pitch from the vanishing points to the captured position according to the viewing frustum. Figure 6(a) shows the calculated results for the spidery mesh from the front and back images shown in Fig. 5. Because our environment map is captured by the same camera, the size of the perspective image and the relative position of the vanishing point in the front and back image are the same when the omnidirectional camera is set precisely at the center of the room. However, when the position of the camera has some offset from the center of the room, the size of the quadrangle consisting of four corners between the front and back walls is different, and the perspective distortion also arises from the offset as shown in Fig. 6(b). Under the assumption that the shape of the room is rectangular parallelepiped, we utilize this difference and the distortion. The perspective information is easily estimated from the position of the four corners. Because this perspective information is caused by the offset of the left and right sides from the center of the room, we can estimate the relative position of the captured position. Moreover, we can also estimate the relative position in the forwards and backwards directions using the difference between the sizes of the front and back rectangles. Figure 6(c) shows the results of the 3D reproduction of the cuboidal room using the perspective information. The blue circle indicates the captured position and the green circle indicates the position of the target object. From the reproduction, we can derive the relative position in a 3D cuboidal room between each position and the walls.

Next, we reconstruct an environment map at the position of the target object based on the relative positions of the photographer and the target. Figure 7 shows the variables for transformation of the environment map to assist in explaining our procedure. The black camera indicates the captured position and the Stanford Bunny indicates the target position.[25] All variables in Fig. 7 show the relative distance which is calculated by the spidery mesh of captured environment map. For example, the relative distance from captured position to target position is obtained by counting the number of mesh pitch from captured position and interpolating between pitches linearly. The explanations of variables are given in Table 1, and we input relative distances into variables for the calculation of the reconstruction. Then, we search the pixel positions of the reconstructed image corresponding to the original equirectangular image by calculating the latitude and longitude in each camera and target position in the room. Here, θ'and ϕ' indicate the latitude and longitude of the reconstructed equirectangular image, and θ and ϕ indicate the latitude and longitude of the original equirectangular image. In the first step, we assign θ' and ϕ' to the reconstructed image according to Eq. 1. In the next step, we calculate θ and ϕ corresponding to θ' and ϕ' using the relative distance D_o, as shown in Fig. 7. From this calculation, the pixel value of (θ', ϕ') is obtained by an interpolation of pixel value of (θ, ϕ). Figure 8 shows 10 out of a total of 20 regions the room has been divided into. In this case, our calculation is performed by dividing the room into 20 regions because these calculations need a complicated case classification for trigonometric function and high computational cost for cubic interpolation. This effort for the divided regions achieves a high-speed calculation with distributed computing. As examples, Fig. 9 shows the calculation for the “Region 1” and the “Region 6”. The “Region 1” consists of the positions between the target, ceiling, front wall, and left corner of the front wall. The “Region 6” consists of the positions between the target, ceiling, left wall, and camera. By calculating all regions, we can transform the pixel value from the environment map at the position of the photographer to the environment map at the target position.

4. Results of transformation and discussion

Figure 10(a) and 10(b) show an environment map of the indoor scene captured in a cuboidal lecture room, which is 2.5 meters in height, 5 meters in width, and 7 meters in depth. In this lecture room, we set a target object (white circle) at the center of the room. The environment map in Fig. 10(a) is captured at a position 2 meters behind the target, as the position of the photographer, and the environment map in Fig. 10(b) is captured at the position of the target object as ground truth. Here, it should be noted that in these images, the contrast has been adjusted in order to clarify some landmarks. Figure 10(c) shows the result of the environment map created by our proposed method. This map is generated by transformation from the position of the photographer to the position of the target object in Fig. 10(a). By comparing Fig. 10(c) to 10(b), it is clear that the positions of the light sources attached to the ceiling are in agreement. To evaluate the accuracy of our results, we compare the degrees of latitude and longitude at several points between the ground truth and our results, as shown in Table 2. These selected points are the lighting equipment on the ceiling, the switch and the electric outlet on the wall, and masking tape on the floor, as shown in Fig. 11. From the results, it is obvious that our method can accomplish an accurate transformation in the latitude and longitude angles at all points.

On the other hand, Fig. 10(c) shows some unexpected points at which the transformation failed. The most noticeable are the tables on the left side and the spotlight on the right side. As previously mentioned, our transformation is calculated by accurately measuring the angles of each corner of a cuboidal room. Two objects, such as the tables and spotlight, are located in front of the back wall in this scene. Therefore, our method performs an inadequate transformation when the objects have the different distance from the corners of the cuboidal room, even if an adequate transformation is shown in the fluorescent lamp along the ceiling. These experimental results mean that the proposed method limits objects at which there are on the ceiling, wall, and floor.

Moreover, we examine our method in different conditions – the low-ceiling lounge room shown in Fig. 12(a). Figure 12(b) shows the ground truth, and the result is shown in Fig. 12 (c), and Figure 12(c) shows the result of transformation at the case of setting the target location on the table (white circle) as shown in Fig.12(a) . The transformation seems to succeed only when the positions of the light source are observed. However, the position of a beam on the ceiling is different from the ground truth. Detailed observation shows that the transformation tends to move objects to the back, for example, the clock on the wall is located at θ'=30 and ϕ'=266 degree in transformed image as against θ'=36 and ϕ'=257 degree in grand truth image, and the black line on the floor is located at θ'=-23 and ϕ'=330 degree in transformed image as against θ'=-27 and ϕ'=314 degree in grand truth image. At any positions, the transformed result shows the shift to the backward. In the case, it is impossible to calculate a spidery mesh at the back of the image because we fail to find four corners in this image which is captured at very close position to the back wall. Our method needs four corners and vanishing points in both the front and back images for an accurate transformation.

5. Conclusions and future work

We developed a practical method for transforming an environment map that is captured at different positions. In contrast to conventional cube mapping, we manually specify the corner positions of a room to make a rectangular parallelepiped. The location of the corners is used to calculate a spidery mesh with a vanishing point according to the view frustum. By using the shape of the room and the spidery mesh, we can transform the captured environment map from polar coordinates to 3-dimensional Euclidean space with a reasonable degree of accuracy. The movement of the position between that of the photographer and that of the target is performed in 3-dimensional Euclidean space. Then, we calculate updated angles from the target position and reconstruct the environment map. For our calculations, we also propose a region splitting method to consider the transformation of the latitude and longitude when the captured position has some offsets from the center of the room.

In the verification process, we applied our proposed method to an actual environment map that has several check points on the ceiling, wall, and floor. From the results, it is obvious that our method can accomplish an accurate transformation in the latitude and longitude angles at all points. However, the proposed method has some constraints when finding the vanishing points and the corners of the room in the environment map, and the light sources should be set on the walls and ceiling. Therefore, in future work we will consider an adaptive transformation with minimal constraints even if the shape of the room is complex. Furthermore, instead of the manual operations needed for the estimation of corner points in the present method, we will develop an automatic procedure for detecting corners and vanishing points.

Acknowledgment

This research was supported in part by the Ministry of Education, Science, Sports, and Culture, Japan Grant-in-Aid for Scientific Research, 24500267, and Brain and Information Science grant SHITSUKAN, 25135707.

References

1) F. Pellacini, J. A. Ferwerda, D. P. Greenberg: Proc. SIGGRAPH (2000), 55.

2) R. W. Fleming, R. O. Dror, E. H. Adelson: J. Vis. **3** (2003) No. 5, 347.

3) L. Sharan, Y. Li, I. Motoyoshi, S. Nishida, E. H. Adelson: J. Opt. Soc. Am. A Opt. Image Sci. Vis. **25** (2008) Issue 4, 846.

4) C. Grey: *Master Lighting Guide for Portrait Photographers* (Amherst Media; New York, 2004), Chap. 2-4.

5) A. Earnest: *Lighting for Product Photography: The Digital Photographer's Step-By-Step Guide to Sculpting with Light* (Amherst Media; New York, 2012), Chap. 3, pp. 25-35.

6) D. Giannatti: *Lighting Essentials: A Subject-Centric Approach for Digital Photographers* (Amherst Media; New York, 2012), Chap. 3, pp. 30-46.

7) Y. Akaike: ITE Technical reports (2014) Issue 38, 7 (in Japanese).

8) A. Matsufuji: ITE Technical reports (2015) Issue 39, 8 (in Japanese).

9) I. Boyadzhiev, S. Paris, K. Bala: ACM Trans. on Graphics **32** (2013) Issue 4, No. 36.

10) F. Pellacini: ACM Trans. on Graphics **29** (2010) Issue 4, No. 34.

11) A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, M. Cohen: ACM Trans. on Graphics **23** (2004) Issue 3, 294.

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

13) S. E. Chen: Proc. SIGGRAPH (1995), 29.

14) M. F. Cohen, D. P. Greenberg: Proc. SIGGRAPH (1985), 31.

15) T. Y. Ho, L. Wan, C. S. Leung, P. M. Lam, T. T. Wong: IEEE Trans. Vis. Comput. Graph. **17** (2011) Issue 1, 51.

16) B. Cabral, M. Olano, P. Nemec: Proc. SIGGRAPH (1999), 165.

17) P. Debevec: IEEE Computer Graphics and Applications **22** (2002) Issue 2, 26.

18) E. A. Khan, E. Reinhard, R. Fleming, H. Bülthoff: ACM Trans. on Graphics **25** (2006) Issue

3, 654.

19) J. P. Snyder: *Flattening the Earth: Two Thousand Years of Map Projections* (Univ. of Chicago Press, 1997), Chap. 2, p. 62.

20) R. L. Cook, K. E. Torrance: ACM Trans. on Graphics **1** (2002) Issue 1, 7.

21) P. Debevec: Proc. SIGGRAPH (1998), 189.

22) R. Fernando, M. J. Kilgard: *The CG Tutorial: The Definitive Guide to Programmable Real-Time Graphics* (Addison-Wesley Longman Publishing, Boston, 2003), Chap. 7.

23) R. Mantiuk, S. Pattanaik, K. Myszkowski: Proc. ICCVG (2002), 189.

24) Y. Horry, K. Anjyo, K. Arai: Proc. SIGGRAPH (1997), 225.

25) G. Turk, M. Levoy: Proc. SIGGRAPH (1994), 311-318.

Table captions

Table 1. Explanations of variables in Fig. 7.

Table 2. Result of transformation.

Figure captions

Fig. 1. Example of equirectangular image and its notation. H is the height and W is the width of the image, and X and Y are the coordinates of each pixel starting from the upper left corner.

Fig. 2. Environment map, cube map and rendered image.

Fig. 3. Illustration of angle change caused by the difference in the distance from the light source.

Fig. 4. Flow of the transformation method in “Tour into the Picture”. [24] The red circle indicates the corner.

Fig. 5. Environment map and cube map of the image taken in a cuboidal room.

Fig. 6. Explanation of calculation process in our method.

Fig. 7. Illustration of the variables for relative distance. The explanations of variables are given in Table 1.

Fig. 8. The regions of the room as defined for use in the conversion equation (10 of 20 regions).

Fig. 9. Flow chart for obtaining θ' and ϕ' corresponding to θ and ϕ at “Region 1” and “Region 6” shown in Fig. 8.

Fig. 10. Comparison of transformation results with the original image and the ground truth image.

Fig. 11. Illustration of selected points in Fig. 10(b) corresponding to Table 2.

Fig. 12. Comparison of transformation results in low-ceiling lounge.

Tables

Table 1. Explanations of variables in Fig. 7.

H_a	Relative distance from the object to ceiling	W_l	Relative distance from the object to left wall
H_b	Relative distance from the object to floor	W_r	Relative distance from the object to right wall
W_f	Relative distance from the object to front wall	D_o	Relative distance from the object to camera
W_b	Relative distance from the object to back wall		

Table 2. Result of transformation.

Position	Ground truth		Our result	
	latitude (deg.)	longitude (deg.)	latitude (deg.)	longitude (deg.)
①FL lamp	59.03	-90.15	59.01	-90.85
② FL lamp	55.71	91.47	55.23	88.76
③ FL lamp	25.15	-35.68	25.62	-34.68
④ FL lamp	24.66	35.98	24.96	37.61
⑤ FL lamp	26.56	-145.18	25.55	-144.18
⑥ FL lamp	25.82	141.93	25.52	143.75
⑦ Outret	-17.71	59.51	-18.37	62.63
⑧ Switch	4.5	123.73	3.75	127.22
⑨ Tape corner	-44.55	149.95	-43.8	153.58
⑩ Tape corner	-45.51	-0.78	-47.45	-0.09
⑪ Tape corner	-44.86	-152.45	-44.18	-152.83
Spot light	-9.14	145.99	-12.61	136.57

Figures

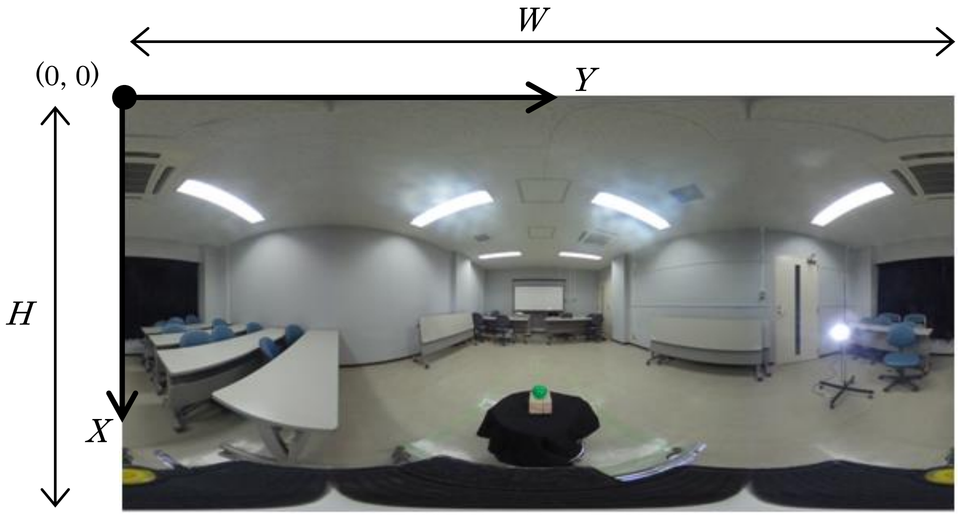
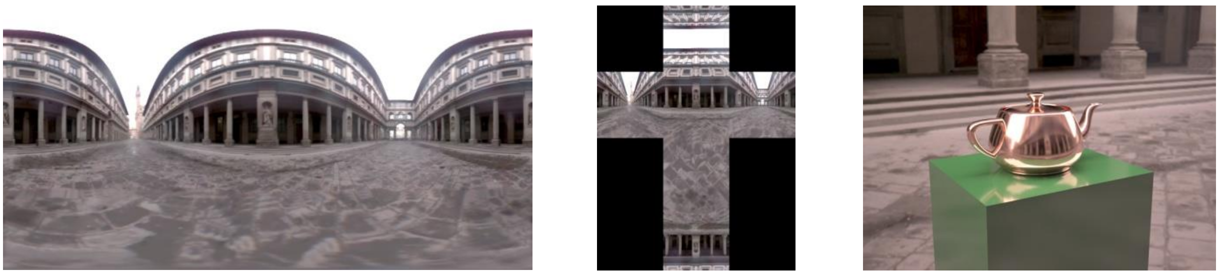


Fig. 1. Example of equirectangular image and its notation. *H* is the height and *W* is the width of the image, and *X* and *Y* are the coordinates of each pixel starting from the upper left corner.



(a) Environment map (equirectangular image) (b) Cube map (vertical cross) (c) Rendered image

Fig. 2. Environment map, cube map and rendered image.

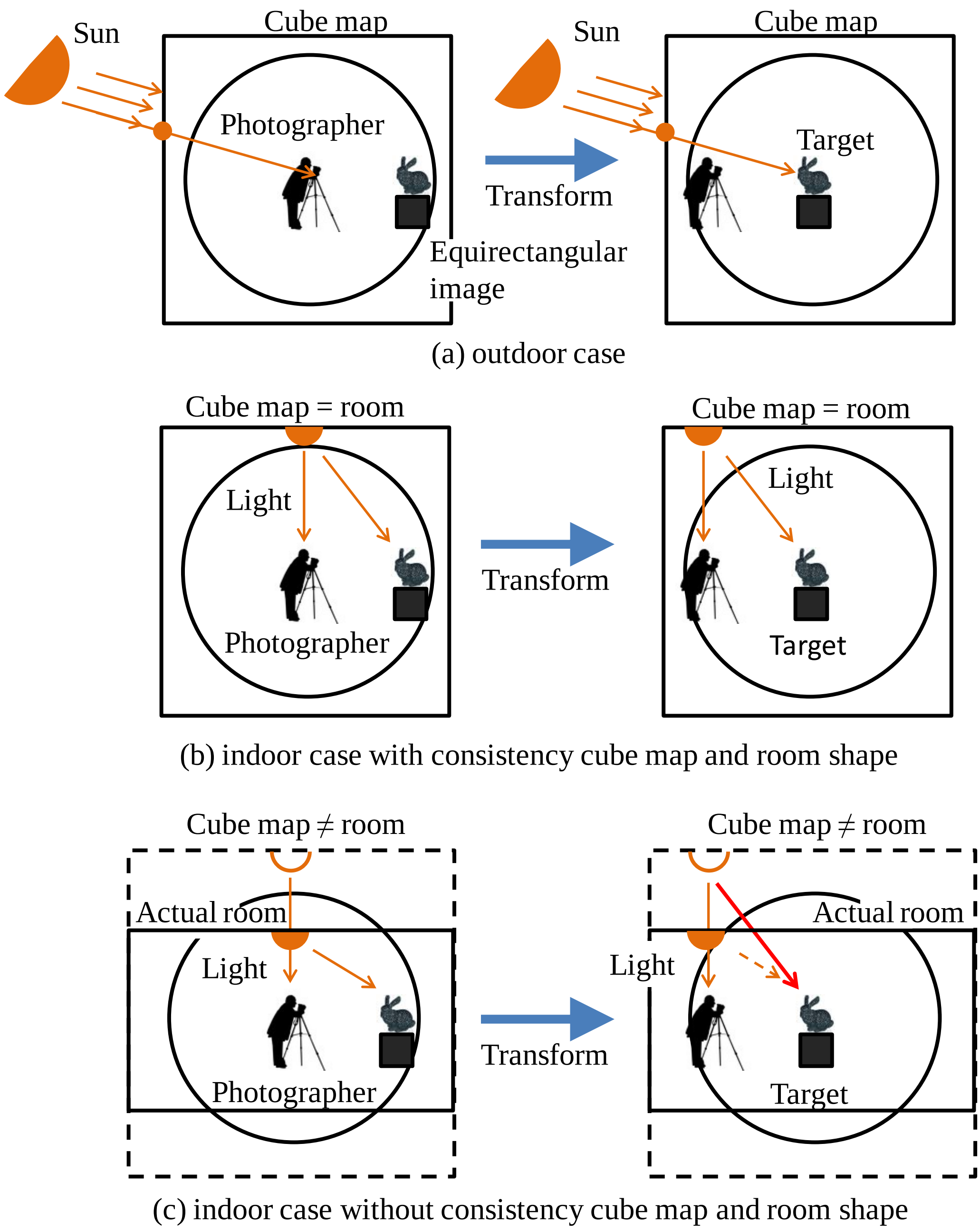
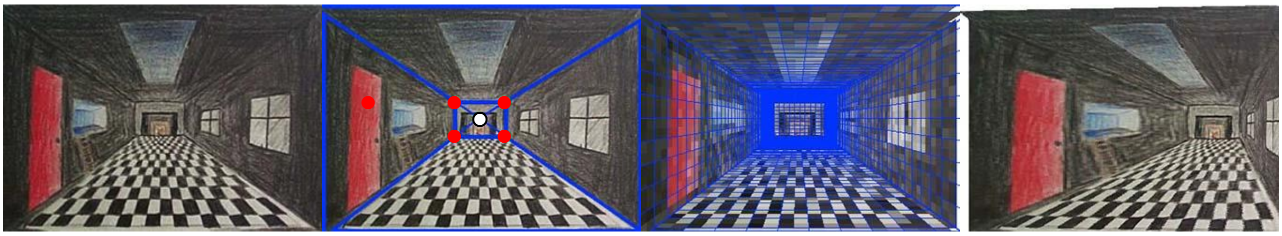


Fig. 3. Illustration of angle change caused by the difference in the distance from the light source.



(a) Input image (b) Generation of spidery mesh (c) Result of perspective calculation (d) Result of viewpoint change

Fig. 4. Flow of the transformation method in "Tour into the Picture". [24] The red circle indicates the corner.

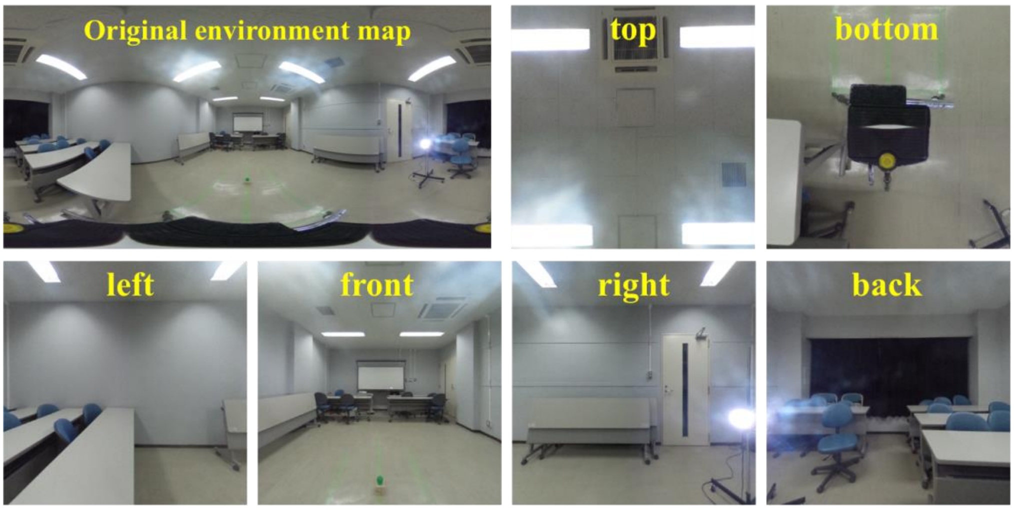
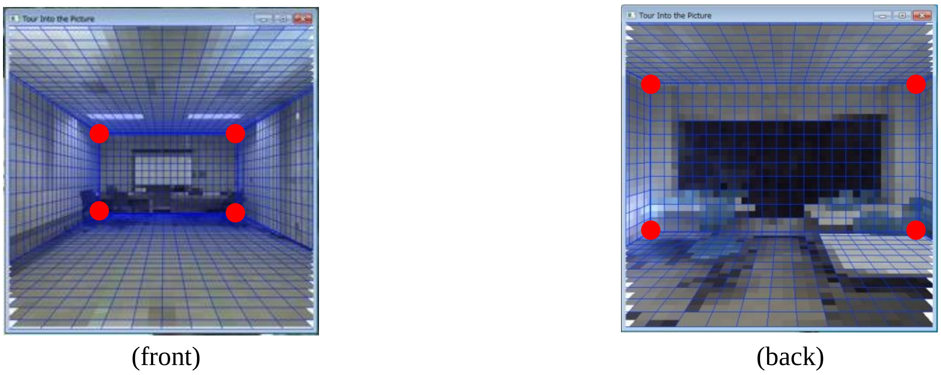
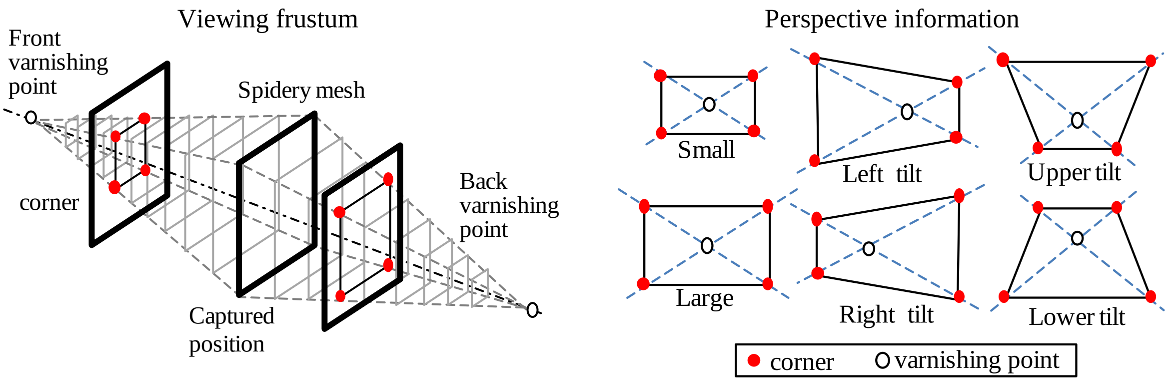


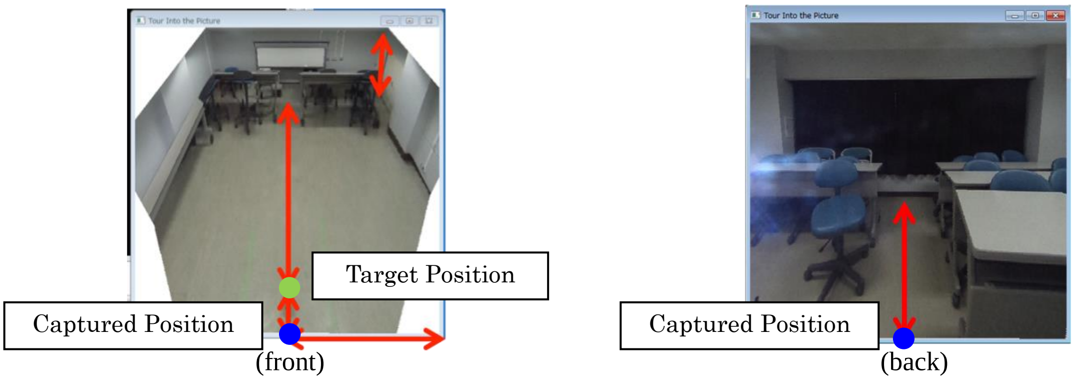
Fig. 5. Environment map and cube map of the image taken in a cuboidal room.



(a) Spidery mesh calculated by Horry's method. The red circle indicates the corner.



(b) Illustration of the viewing frustum and perspective information



(c) Reproducing results for 3D room structure

Fig. 6. Explanation of calculation process in our method.

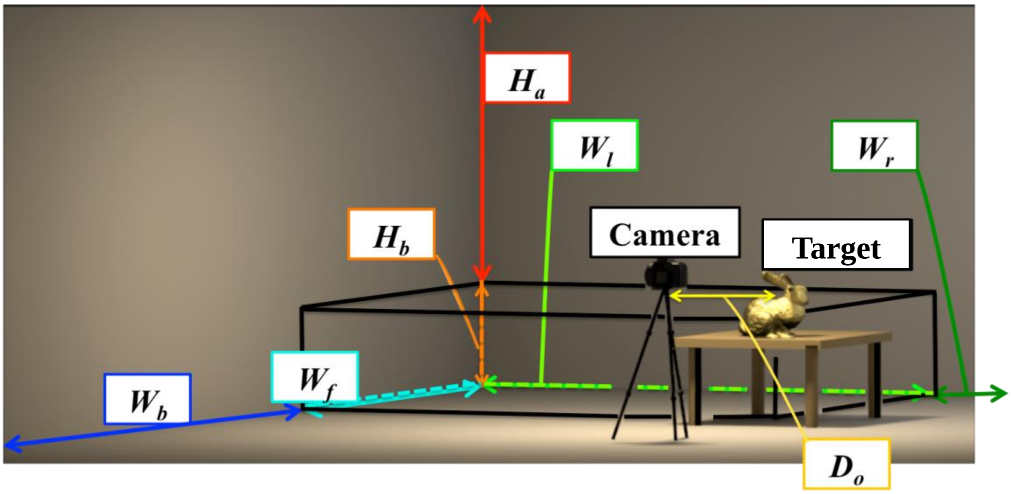


Fig. 7. Illustration of the variables for relative distance. The explanations of variables are given in Table 1.

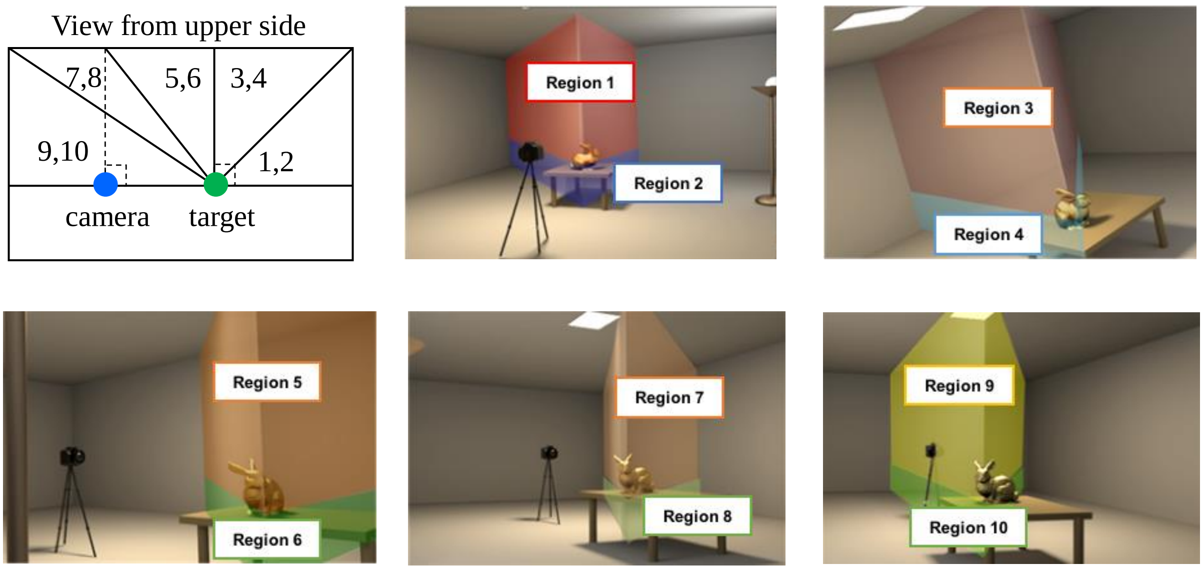


Fig. 8. The regions of the room as defined for use in the conversion equation (10 of 20 regions).

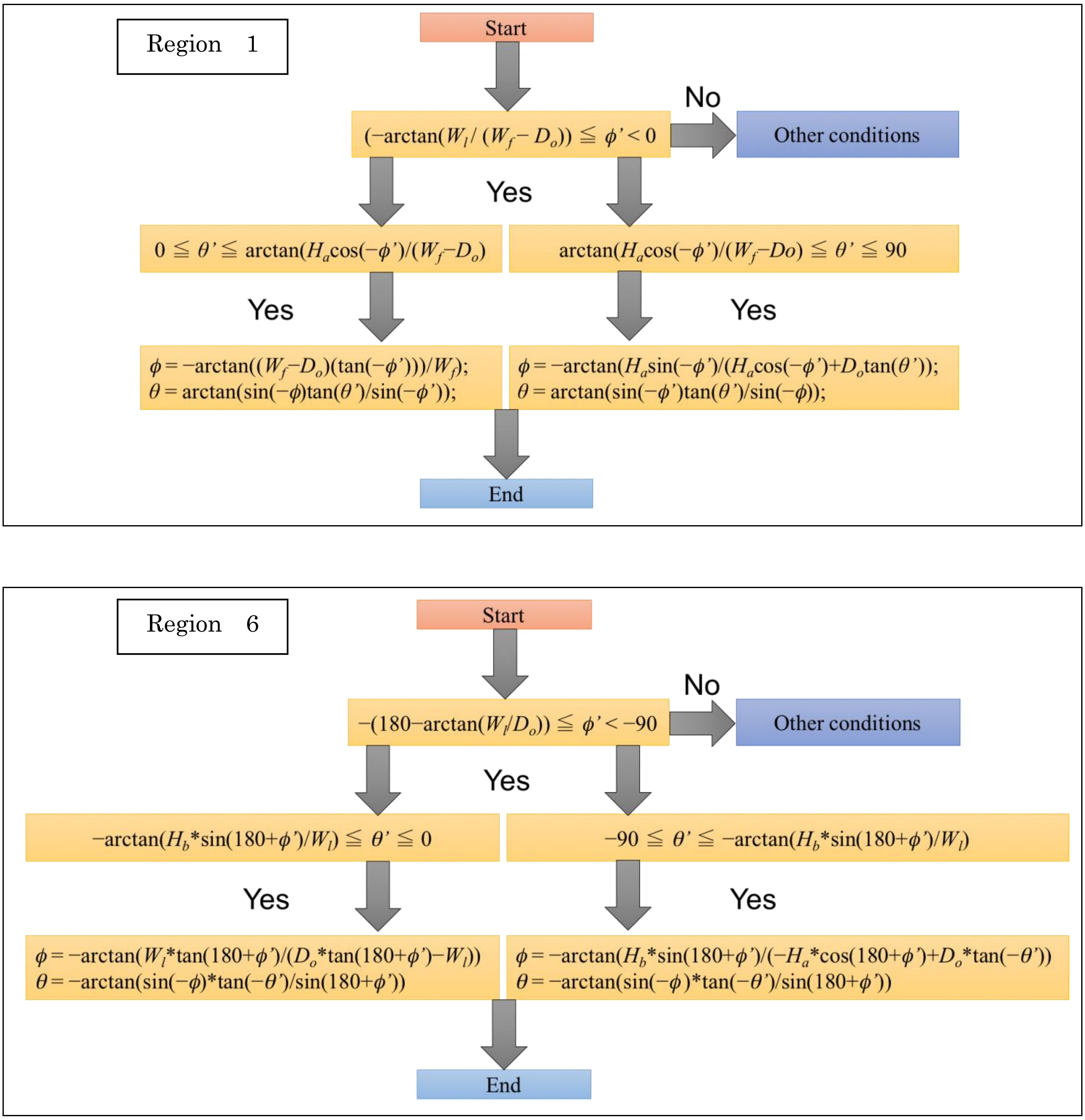
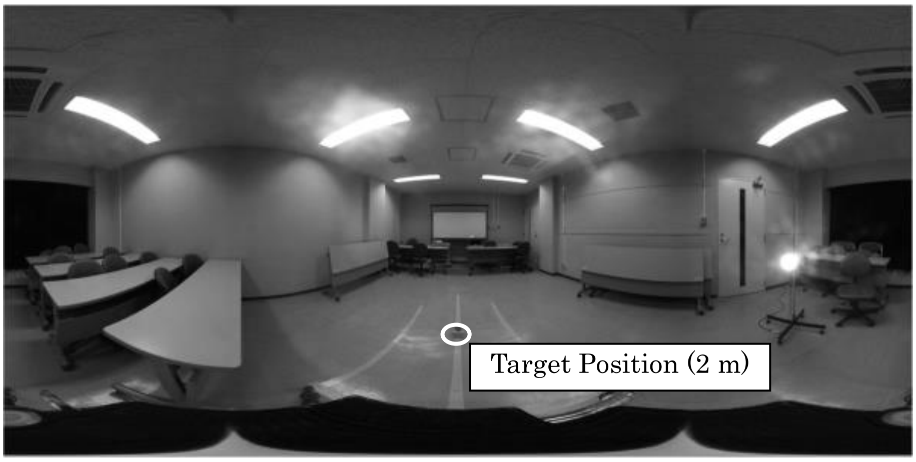


Fig. 9. Flow chart for obtaining θ and ϕ corresponding to θ' and ϕ' at “Region 1” and “Region 6” shown in Fig. 8.



(a) Original image (lecture room)

(b) Ground truth image

(c) Result of transformation

Fig. 10. Comparison of transformation results with the original image and the ground truth image.

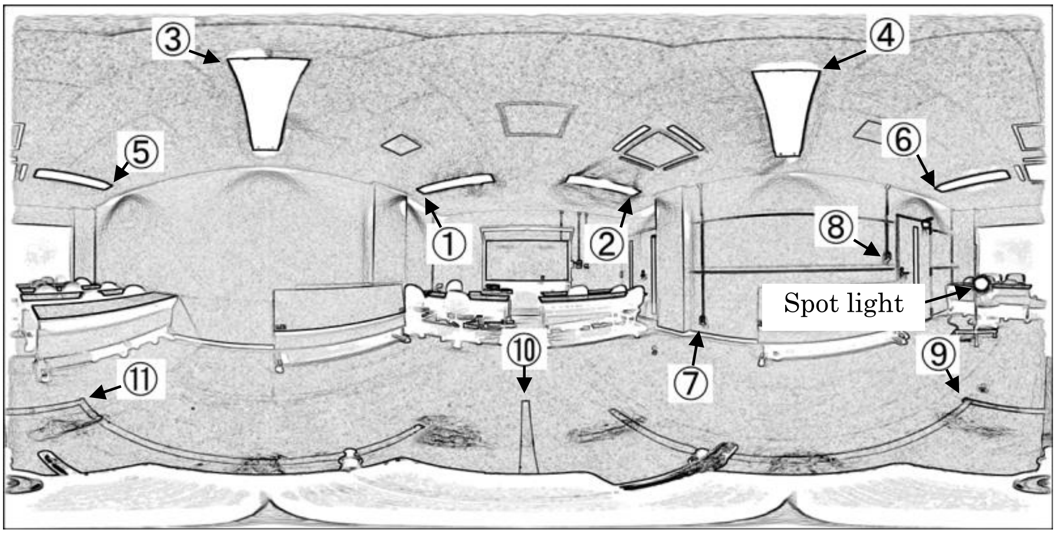
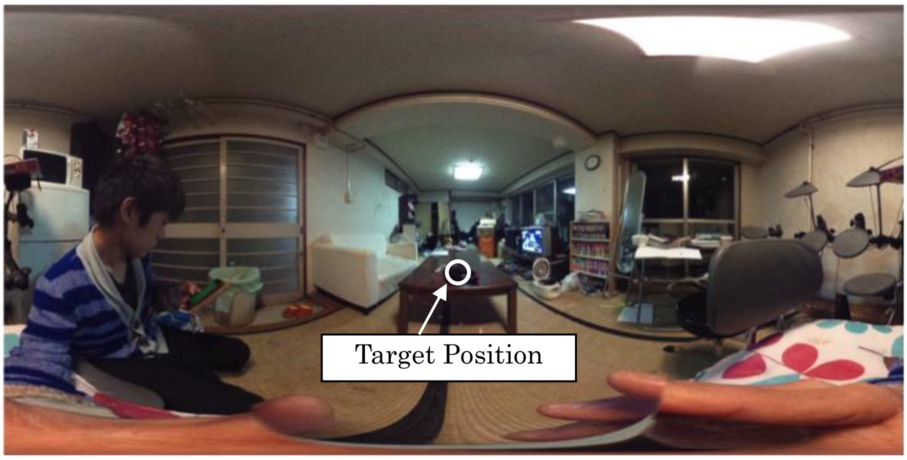
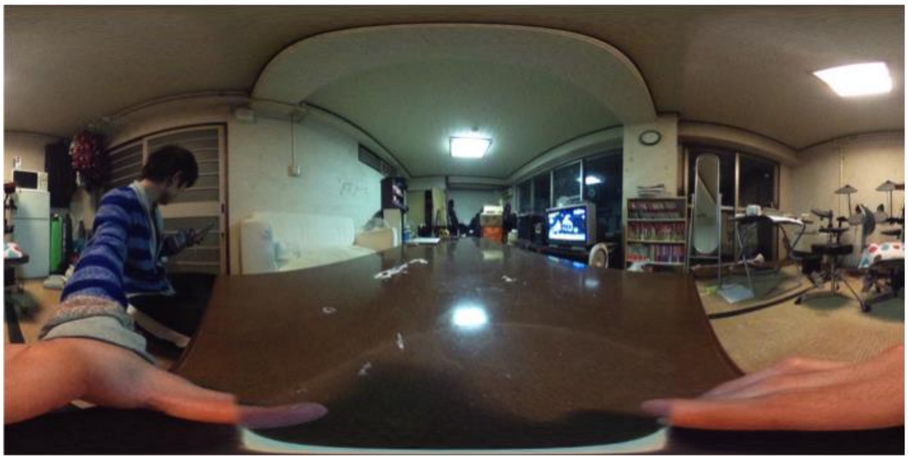


Fig. 11. Illustration of selected points in Fig. 10(b) corresponding to Table 2.



(a) Original image



(b) Ground truth image

(c) Result of transformation

Fig. 12. Comparison of transformation results in low-ceiling lounge room.