

Representing Virtual Transparent Objects on OST-HMDs Considering Accommodation and Vergence

Yuto Kimura*

Shinnosuke Manabe†

Asako Kimura‡

Fumihisa Shibata§

Ritsumeikan University

ABSTRACT

We represent virtual transparent objects on OST-HMDs, considering the defocus by accommodation and the binocular disparity by vergence. In augmented reality with conventional stereoscopic displays, it is difficult to reproduce the defocus and the disparity between the images, surface, reflection, and refraction, of a transparent object with different depths. The reason why is that the accommodation focal is always fixed on the display screen, and furthermore, the reflection and refraction images considering the disparity do not exist in real-time, as they require a ray-tracing method for each eye. In this study, we represented transparent objects by reproducing the defocus with blur processing, and the disparity with pseudo parallax refraction. In our experiment, it was confirmed that the transparent object reproduced with the proposed method makes the images more realistic compared to the unprocessed one.

Index Terms: Human-centered computing—Human-computer interaction (HCI)—Interaction paradigms—Mixed / augmented reality; Human-centered computing—Human-computer interaction (HCI)—Interaction devices—Displays and imagers

1 INTRODUCTION

Real transparent objects have some different depth images (*e.g.* object surface, reflection, and refraction). The human eye can freely adjust the vergence and the crystalline lens accommodation for each image. Therefore, to represent transparent objects realistically in computer graphics, it is necessary to reproduce the defocus caused by accommodation, and the binocular disparity caused by vergence. In non-stereoscopic displays, transparent objects are represented by only considering the defocus [3]. In contrast, in stereoscopic displays, the effective representation of the disparity is accounted for. In head-mounted displays (HMDs) and the representation of transparent objects, both the defocus and the disparity should be considered. In conventional HMDs, the crystalline lens accommodation focus is fixed at the display focus. Therefore, the defocus does not occur spontaneously in virtual objects with different depths. In addition, to represent parallax images of reflection and refraction, it is necessary to perform calculations for each eye by using the ray-tracing method. However, this is not done in most cases as it requires high calculation costs. Some HMDs such as light field displays [1] and varifocal displays [4] can provide the defocus and represent transparent objects in conformance with actual visual phenomenon. However, even these are subject to problems such as high computational costs, narrow field of view, and the requirement of large equipment. We use optical see-through HMDs (OST-HMDs), which can observe the real world directly, to make conditions except for virtual objects follow the actual phenomenon.

*e-mail: yuto_kim@rm.is.ritsumei.ac.jp

†e-mail: manabe@rm.is.ritsumei.ac.jp

‡e-mail: asa@rm.is.ritsumei.ac.jp

§e-mail: fshibata@is.ritsumei.ac.jp

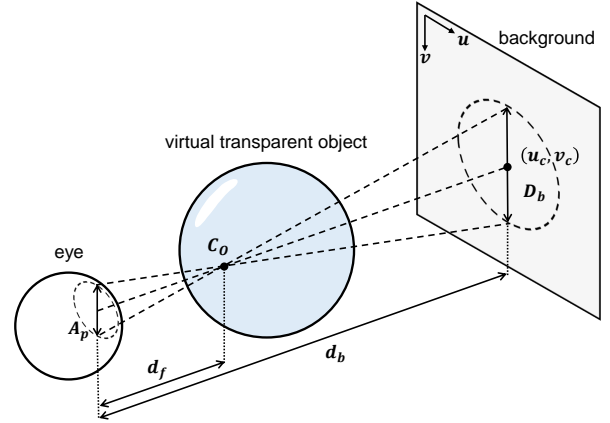


Figure 1. The model for calculating transparent object's color.

In this study, we represented virtual transparent objects realistically on OST-HMDs by reducing the visual mismatch of the accommodation and the vergence between the real and the virtual. Specifically, virtual transparent objects were represented realistically by reproducing the defocus and the disparity using blurring and pseudo parallax refraction, respectively. The defocus was reproduced by appropriately blurring with Gaussian functions, considering the crystalline lens accommodation. The disparity was reproduced by calculating the pseudo parallax refraction image from the refraction vector.

The hypothesis of our experiment is that users tend to perceive the transparent object reproduced by using the proposed method to be more realistic than the unprocessed one. From the experiment, we show that the transparent object reproduced with the proposed method is more realistic than the unprocessed one. We also show the contribution of each effect in enhancing the reproduced image, by including the only parallax and the blur in comparison.

2 PROPOSED METHOD

We propose a method in which virtual transparent objects are rendered by calculating a refraction image suitable for each eye and then blurring them, considered the defocus. Figure 1 shows the relationship between the eye, the virtual transparent object, and the real background. The pseudo parallax refraction image is computed by sampling the intersection (u_c, v_c) between the refraction vector and the background texture. The refraction vector is calculated by Snell's law for each polygon using the polygon's normal vector, the view vector, and the relative refractive index. Then, the defocus is computed by blurring with Gaussian function. The amount of blur is calculated, considering the crystalline lens accommodation according to the gaze point. When the color (R, G, B) of the transparent object surface is C_s and the color of the blurred refraction image is C_b , the output color to the OST-HMDs C_o is given by:

$$C_o = \alpha C_s + (1 - \alpha) C_b. \quad (1)$$

Here, α is the ratio of an alpha blend, and C_b is calculated from:

$$C_b = \frac{1}{2\pi\sigma^2} \sum_i^{K_g} \sum_j^{K_g} \exp\left(-\frac{(u_i^2 + v_j^2)}{2\sigma^2}\right) C(u_i, v_j) \quad (2)$$

Here, σ^2 is the variance of the Gaussian function and is defined as half of the kernel size K_g . The range of u_i, v_j was $(-K_g/2 + u_c < u_i < K_g/2 + u_c, -K_g/2 + v_c < v_j < K_g/2 + v_c)$, as it centered the sampling point (u_c, v_c) obtained by refraction. When the number of pixels of the entire background image was W_p , the actual length of the entire background was W_l , and the diameter of the circular area of the real background was D_b ; the kernel size K_g was defined as:

$$K_g = \frac{W_p}{W_l} D_b. \quad (3)$$

For a pupil diameter of size A_p , D_b could be calculated using similar triangles as:

$$D_b = A_p \frac{|d_b - d_f|}{d_f}. \quad (4)$$

Here, d_b is the distance from the eye to the background, and d_f is the distance from the eye to the focus.

3 EXPERIMENT

We conducted an experiment to confirm that transparent objects reproduced with the proposed method are more realistic. A dense glass sphere was prepared with five representation types (O_1 : real glass O_2 : no effect, O_3 : only blur, O_4 : only parallax refraction, O_5 : blur and parallax refraction shown in Figure 2). In the real scene, a point light source of lump power 575W/G22 was installed in a room with no other light sources. The glass sphere (of refractive index 1.51, and diameter 150 mm) was placed on a desk with a sheet, including a high-frequency component to facilitate detection in a refraction image. A sticker was put on the surface of the glass sphere as a depth cue. In the virtual scene, we reproduced the glass sphere and the light source almost realistically. Virtual glass spheres were rendered with Unity and shaded with Cg/HLSL. The parallax refraction image was calculated only for the first refraction in order to guarantee real-time properties, and owing to the fact that humans do not fully understand the patterns of the refraction image. The reflection of the transparent object was reproduced by using the Phong reflection model. The participant's pupil diameter is required to determine the amount of blur, and it was taken to be 6.0 mm (as it is difficult to measure the participant's pupil in real time, and the range of motion of human pupil diameter is about 2.0~8.0 mm).

The participant wearing Magic Leap One as the OST-HMD was seated, with the neck bent at an angle of 45 degrees to observe the glass sphere on the desk. The participant observed the surface and the refraction image alternately for 5 seconds each, for a total of 20 seconds. During this trial, we changed the amount of blur in each image manually. The participant compared the two pairs (O_i, O_j) of five glass spheres based on Scheffe's paired comparison method. The participant rated O_i against O_j , that implied how realistic the observations were (-3: very low, -2: low, -1: little low, 0:same, +1: little high, +2: high, +3: very high). The participant performed this operation in ${}_5P_2 = 20$ ways, considering the order effects.

The participants were 21 adults (19 men, 2 women) in their 20s with no eye abnormalities. The experimental results are shown in Figure 3. The results show that the transparent object is more realistic in the order of O_1, O_5, O_4, O_3 and O_2 . Since there is a statistically significant difference between O_2 and O_5 , it can be inferred that the proposed method makes the transparent objects more realistic than the unprocessed ones. Furthermore, disparity is a more effective method for representing transparent objects than blur, and the proposed method combines them both, which enhances the image significantly compared to either one.

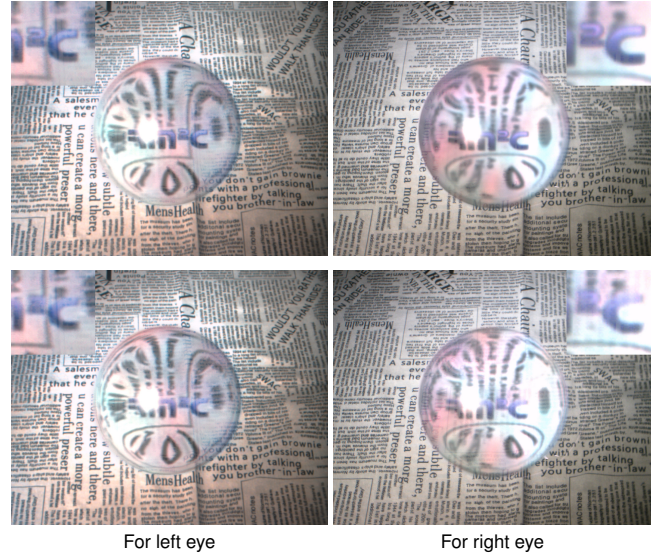


Figure 2. Transparent objects with the proposed method presented to the real world. *Top*: Rendering result during surface gaze. *Bottom*: Rendering result during refraction image gaze. Images were taken with cameras (Point Gray Flea3) on the left and right displays of Magic Leap One. Each image is obtained by Exposure Fusion [2] with multi-exposure sequence.

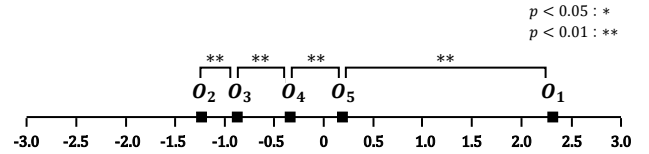


Figure 3. Result of each transparent object by scheffe's pair comparison.

4 CONCLUSION AND FUTURE WORK

In this study, we represented transparent objects by reproducing the defocus caused by crystalline lens accommodation, and the binocular disparity caused by a pseudo refraction image. Specifically, the defocus was reproduced by blur processing with the Gaussian function using the kernel obtained from the pupil diameter and the distance from the eye to the focus and to the background. The disparity was reproduced by the parallax refraction image computed by sampling from the intersection between the refraction vector and the background texture. In an experiment comparing glass spheres with five representation types including the real glass sphere, it was confirmed that the proposed method is the most realistic of the four types, excluding the real one; and that disparity is more effective than blur. In future, by dynamically determining the amount of blur by measuring the pupil movement in real-time.

ACKNOWLEDGMENTS

This work is partly supported by JSPS Kakenhi #19K22882

REFERENCES

- [1] D. Lanman and D. Luebke. Near-eye light field displays. *ACM Trans. Graph.*, 32(6):220:1–220:10, Nov. 2013.
- [2] T. Mertens, J. Kautz, and F. V. Reeth. Exposure fusion. In *Proc. 15th Pacific Conference on Computer Graphics and Applications (PG 2007)*, pp. 382–390, 2007.
- [3] P. Rokita. Generating depth of-field effects in virtual reality applications. *IEEE Computer Graphics and Applications*, 16(2):18–21, 1996.
- [4] S. Liu, D. Cheng, and H. Hua. An optical see-through head mounted display with addressable focal planes. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 33–42, 2008.