

A See-through Vision with Handheld Augmented Reality for Sightseeing

Goshiro Yamamoto, Arno in Wolde Lübke,
Takafumi Taketomi, and Hirokazu Kato

Nara Institute of Science and Technology
8916-5 Takayama Ikoma 630-0192, Japan
{goshiro,arno-w,takafumi-t,kato}@is.naist.jp
<http://imd.naist.jp/>

Abstract. We propose a see-through vision with augmented reality (AR) for increasing the quality of the experiences in the activities of sightseeing. In the sightseeing activities, there are certain temporal and spatial limitations when sightseers are gazing on the real objects. In this research, we aim to eliminate the spatial limitations caused by the inability to see farther objects through opaque surfaces by using see-through vision technology with alpha blending, and to make the appearance of the target objects with 3D effect based on parallax effect. We focus on the platform of handheld devices that provides AR experiences with high accurate alignment of virtual objects on pre-captured real scene images based on cube mapping method. This paper describes the engineering methodology that we have used to develop our prototype system and the implementation of the prototype system for confirming our approach. The results about how the prototype system performs on a real site are reported in this paper.

Keywords: Augmented reality, see-through vision, sightseeing, handheld AR.

1 Introduction

Mobile computing technologies have been rapidly developed and these development spreads changes to use augmented reality (AR) for general public because a handheld device provides one of the simplest AR structures with a camera and a display [Möhring et al., 2004],[Henrysson et al., 2005]. The AR technology has the potential to enhance the quality of life (QOL) in daily routines, education, work, and so forth. Sightseeing is one of the activities which is entertaining and intellectual. The essential elements of sightseeing activities include being at the location and facing towards the existing objects. In this paper, we propose a handheld see-through vision with AR technology to give sightseers better experiences. There are certain temporal and spatial limitations with sightseeing when sightseers are gazing on real objects. Temporal limitations are caused by the time constrain that sightseers can only experience the view at current time. Spatial

limitations are caused by the inability to see further objects through opaque surfaces. In this research we are trying to eliminate spatial limitations by using X-ray AR [Sandor et al., 2010][Santos et al., 2013] and Indirect Augmented Reality (Indirect AR) [Wither et al., 2011][Akaguma et al., 2013] technology to give sightseers more enjoyable experiences.

The activity of sightseeing can be carried out in indoor or outdoor environments. For the indoor environment, tracking and motion sensing can be achieved using motion tracking sensors. However, it is difficult to capture tracking points in outdoor environment and it is costly to install motion tracking sensors at every site. Thus, a solution for sightseeing in outdoor environment using motion sensors including accelerometers, gyroscopes, and digital compass on a handheld device is proposed. We implemented cube mapping expression with pre-captured panoramas photos inspired by Indirect AR which is proposed by Wither et al. to reduce perceptual tracking errors with motion sensing. However, 2D information of Indirect AR was not suitable for displaying objects that are close to the user. Therefore, we proposed a novel approach to use 3D effect in conjunction with Indirect AR information. Our approach takes parallax effect into consideration to make the appearance of the target objects feel like 3D with alpha blending based on X-ray AR.

The following sections describes the engineering methodology that we used developing our prototype system, the implementation of the prototype system, and the results of an user experiment using the prototype system in Todaiji temple.

2 Method

In this research, we focus on the situation where the sightseeing spot is a ruin located under the ground. After a ruin was discovered and dug up, it will be filled up within a temporary period. In this case, the sightseers are not able to view the actual scene after the period. To re-present the ruin to the sightseers with an immersive environment, we developed a see-through vision handheld AR application. The application is implemented using the proposed method which includes two phases, cubic environment map construction with alpha blending and a movable viewpoint in the cubic environment map, as shown in Figure 1.

In the first phase, we construct a cubic environment map which contains a 3D model of the ruin and panorama photos of the environment. The 3D model of the ruin is created using the information that were measured when the ruin was yet filled up. The panorama photos of the surrounding environment near the ruin were taken and used as the background of the cubic environment map. The cubic environment map was constructed by these two elements and rendered using alpha blending method.

In the second phase, we propose the method of movable viewpoint which provides parallax effect of the 3D object to the user. In the conventional methods, the relationship between the cubic environment and the view point is fixed. This caused the effect that the object has 2D appearance even it has 3D information.

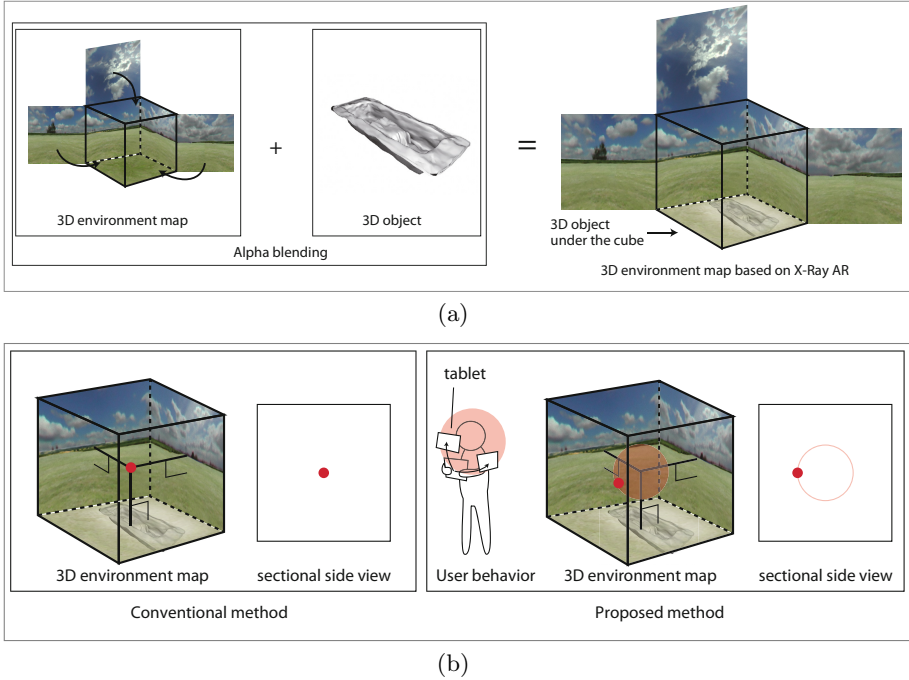


Fig. 1. Proposed method with cubic environment map construction with alpha blending and a movable viewpoint in the cubic environment map. (a) A cubic environment map was created using panorama photos and a 3D model of the ruin. The 3D model is rendered under the cubic environment with alpha blending. (b) In the conventional methods, the viewpoint is fixed at the centre of the cubic environment map. In our proposed method, the viewpoint is moved around the sphere which is located at the centre of the cubic environment according to the orientation of the devices.

Thus we improve the perception by implementing a movable viewpoint according to the orientation of the user’s device. This implementation provides the parallax effect where the user is able to feel the object exists in 3D environment.

The proposed method with two phases is able to provide better perception to the user while viewing the scene. The following chapter details the implementation of the proposed method.

3 Implementation

We implemented a prototype system on iOS devices (Apple iPad/iPad mini) based on the above conceptual idea as the following. In advance, the 3D data of the target ruin is created by measuring the real ruin and using 3D modeling software, and panorama photos is also taken by omnidirectional camera.

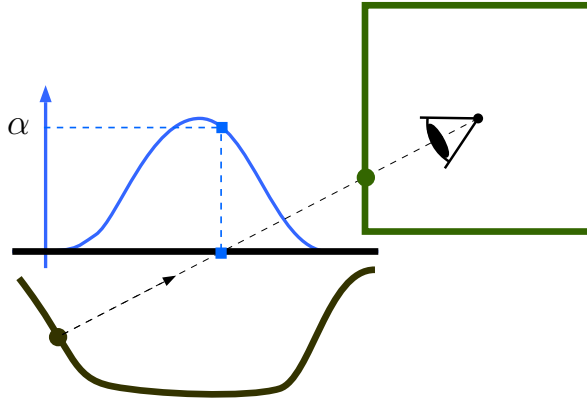


Fig. 2. Alpha lookup: To compute the alpha value that is used to combine the color of a ruin surface point (brown dot) and its occluding point of the environment (green dot) a ray can be cast and its intersection with the alpha plane (black line) can be used to look up the actual value.

The system we implemented is based on Indirect AR idea. As conventional tracking and mapping methods often prove unreliable in many outdoor scenarios, a main benefit of this approach is to avoid registration errors by substituting the commonly used video feed with pre-captured panorama photos of the surrounding environment. Although being purely virtual, the use of panoramic photos which are displayed partially depending on the users location and orientation often creates the illusion of actually seeing the actual environment.

For our prototype system, we rely on a cubic environment map which is mapped onto a cube that is centered around the virtual camera. In order to register the cubic map with the real world environment, we use the build-in sensors (namely accelerometer, gyroscope and compass) of the test device to compute the users orientation and adjust the virtual camera accordingly.

Additionally, a 3D reconstruction of an ancient ruin is part of the scene. The ruin is located below the bottom cube plane. As sketch of the virtual setup is given in Figure 2.

To create the impression that the ruin is located below the ground of the cube we follow work of Sandor et. al. [Sandor et al., 2010] and emulate a x-ray vision that renders the parts of the environment that occlude the ruin transparent. For this purpose, we blend the fragment color of a projected ruin surface point in screen space with the corresponding color of the occluding environment map pixel. The blending coefficient is chosen according to an user-generated alpha texture that is mapped onto a plane located above and oriented parallel to the ruin. Geometrically, the alpha value used for a certain surface point of a ruin can be obtained by shooting a ray from the surface point in direction of the camera origin, computing its intersection with the alpha plane, and looking up the alpha value at that point (see Figure 2).

For the implementation, we chose a two render pass approach: In the first render pass the alpha plane is rendered to an offscreen render target to determine alpha value for each pixel of the screen. In the second render pass the environment is rendered before the ruin with alpha blending enabled, where the final fragment color $\mathbf{c}_{i,j}$ is obtained by

$$\mathbf{c}_{i,j} = \alpha_{i,j}\mathbf{s}_{i,j} + (1 - \alpha_{i,j})\mathbf{d}_{i,j} \quad (1)$$

where $\alpha_{i,j} \in [0, 1]$ denotes the alpha value at pixel coordinate (i, j) of the offscreen render target and $\mathbf{s}_{i,j}$ the fragment color due to the ruin and $\mathbf{d}_{i,j}$ is the fragment color due to the environment.

Following the approach of , to further increase the illusion of seeing through the environment the used alpha map was modified by combining it with edge information of the ground texture.

To amplify the user's feeling of being part of a 3D scene and to model the users potential motion, we decided to place the virtual camera on a ellipsoid that is centered around the origin. We compute the camera's position as the intersection of a ray that is given by the direction

$$\mathbf{r}(t) = t\mathbf{d} \quad (2)$$

of the camera and the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \quad (3)$$

where a, b and c define the form of the ellipse and $\|\mathbf{d}\| = 1$. The position on the ellipse is then $\mathbf{p} = t_i\mathbf{d}$, where t_i is the result of substituting Equation 2 into Equation 3 and solving for t :

$$t = \left(\left(\frac{d_x}{a} \right)^2 + \left(\frac{d_y}{b} \right)^2 + \left(\frac{d_z}{c} \right)^2 \right)^{-\frac{1}{2}} \quad (4)$$

where d_x, d_y and d_z denote the components of \mathbf{d} .

4 Results

The prototype system is used in our experiment conducted at Todaiji temple which is the famous world cultural heritage site in Japan. In the precincts of Todaiji, the Sobo ruin, which is Monks lodging house, was discovered and dug up for a limited time only. At that time, we measured the ruin and use the measured information to acquire 3D model as shown in Figure 3 (a). After the ruin was filled up, we captured photos of surrounding environment of the ruin using an omnidirectional camera at the several fixed points (Figure 3 (b)). The pre-collected data is the element to create cube environment map.

As shown in Figure 4, the results of views in the conventional environment and proposed environment have good see-through vision. Both environments



Fig. 3. Pre-collected data for creating cube environment map. (a) A 3D model of the ruin based on the measurement in multiple viewpoints. (b) A pre-captured panorama photo of the real environment.

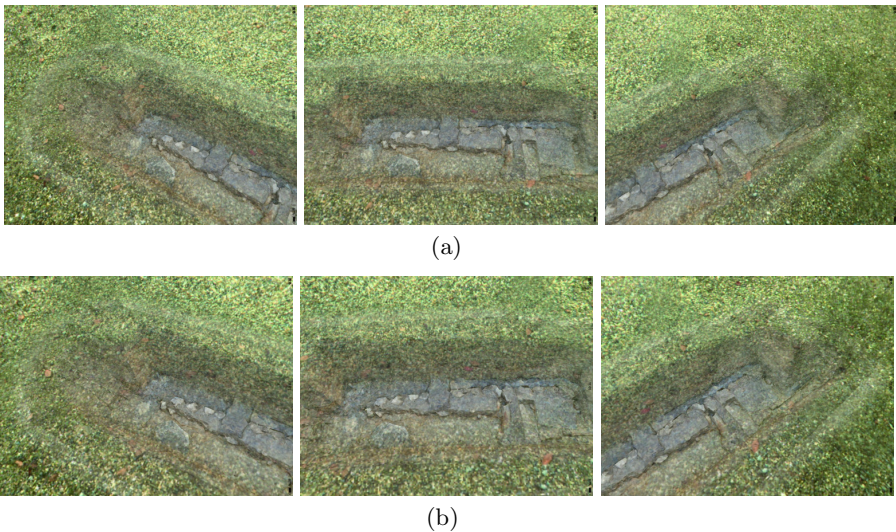


Fig. 4. The results of views of the ruins. (a) Viewing the ruin in conventional cube environment. (b) Viewing the ruin in the environment created by proposed method.

provide realistic vision to the user and create the immersive conception of gazing the ruin. Furthermore, when we display the views of the ruin in 2D pictures, the conventional methods and the proposed method yielded very similar results. However, the proposed method provides the parallax effect to the user where the distance of two viewing directions changes dynamically (Figure 5 (b)). The user is able to feel more existing in the 3D environment compare to the conventional method. The conventional method fixed the distance between two viewing

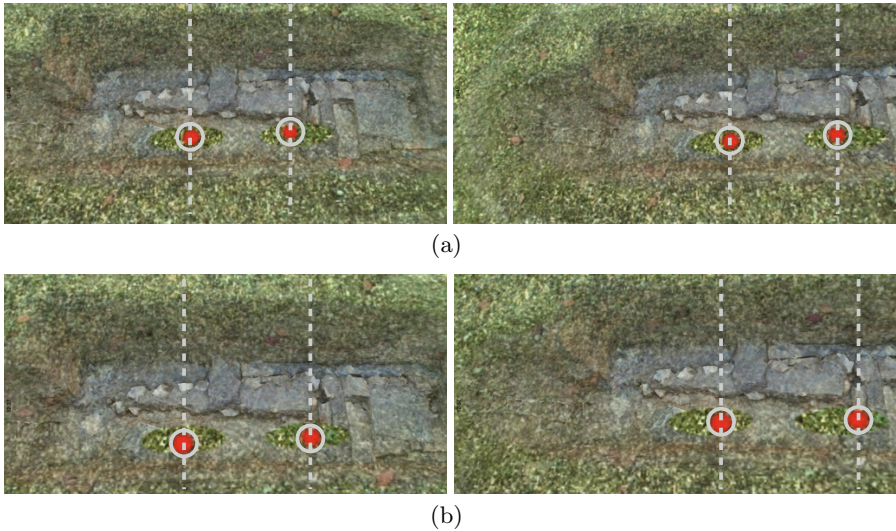


Fig. 5. The effectiveness of parallax effect between conventional cube mapping method and proposed method. The circles are located on the surface of the cube environment. (a) Conventional cube mapping method provides only 2D information even the 3D model was used. (b) Proposed method provides parallax effect between the surface of the cube environment and the 3D model.

directions where the user has less 3D concept even the 3D model if the ruin was used in the implementation (Figure 5 (a)).

The limitation of the proposed method is the reliability of the scene. This limitation is caused by the moving viewpoint where distortion happens at two plane images of the cube were merged. Instead of cube-based mapping, implementing sphere mapping method might mitigate the distortion effect.

Furthermore, front camera of the device can be used to collect additional information, including user's facing direction, facial expression, and ambient lighting. These informations could be useful to create more attractive and interactive AR experiences to the user.

5 Conclusion

This paper focuses on increasing the quality of the experiences in the activity of sightseeing with AR technology. We choose the ruin which is in the ground as the target environment, and propose see-through vision technology that can provide the sightseers additional experiences than traditional touring way. To realize 3D appearance of the object, we implemented two phases, cubic environment map construction and movable viewpoint in the cubic environment map, based on Indirect AR and X-ray AR.

In the first phase, the alpha blending method is used for rendering 3D model of the object and panorama photos are used for constructing cubic environment map. The second phase, movable viewpoint in the cubic environment map, can provide a better 3D perception to the user than fixed viewpoint. However, if the viewpoint was moved farther from the center of the cubic map the corners of the cubic environment map will cause collation of the 3D effect. This drawback can be overcome by using sphere mapping instead of cubic mapping.

We also conduct an experiment at Todaiji temple using our prototype system. Future work includes adding the effect of shadow and improve stereoscopic vision with movable viewpoint.

Acknowledgements. This research was partially supported by Grant-in-Aids for Scientific Research (A), No. 23240024, the “Ambient Intelligence” project sponsored by MEXT, and NAIST Advanced Research Partnership Project. We are very grateful to Todaiji for giving us the opportunity to conduct this experiment.

References

- Möhring, M., Lessig, C., Bimber, O.: Video See-Through AR on Consumer Cell Phones. In: Proceedings of International Symposium on Augmented and Mixed Reality (ISMAR 2004), pp. 252–253 (2004)
- Henrysson, A., Billinghurst, M., Ollila, M.: Face to Face Collaborative AR on Mobile Phones. In: Proceedings International Symposium on Augmented and Mixed Reality (ISMARf 2005), pp. 80–89 (2005)
- Witther, J., Tsai, Y.T., Azuma, R.: Indirect Augmented Reality. *Computers & Graphics* 35(4), 810–822 (2011)
- Sandor, C., Cunningham, A., Dey, A., Mattila, V.V.: An Augmented Reality X-Ray System based on Visual Saliency. In: Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 27–36 (2010)
- Akaguma, T., Okura, F., Sato, T., Yokoya, N.: Mobile AR using pre-captured omnidirectional images. In: Proceedings ACM SIGGRAPH Asia 2013 Symp. on Mobile Graphics and Interactive Applications, Article No. 26 (2013)
- Santos, M.E.C., Chen, A., Terawaki, M., Yamamoto, G., Taketomi, T., Miyazaki, J., Kato, H.: Augmented Reality X-ray Interaction in K-12 Education. In: Proceedings of the 13th IEEE International Conference on Advanced Learning Technologies (ICALT 2013), pp. 141–145 (2013)