A User Study on Viewpoint Manipulation Methods for Diorama-Based Interface Utilizing Mobile Device Pose in Outdoor Environment

Masayuki Hayashi* University of Tsukuba
Itaru Kitahara † University of Tsukuba
Yoshinari Kameda ‡ University of Tsukuba
Yuichi Ohta § University of Tsukuba

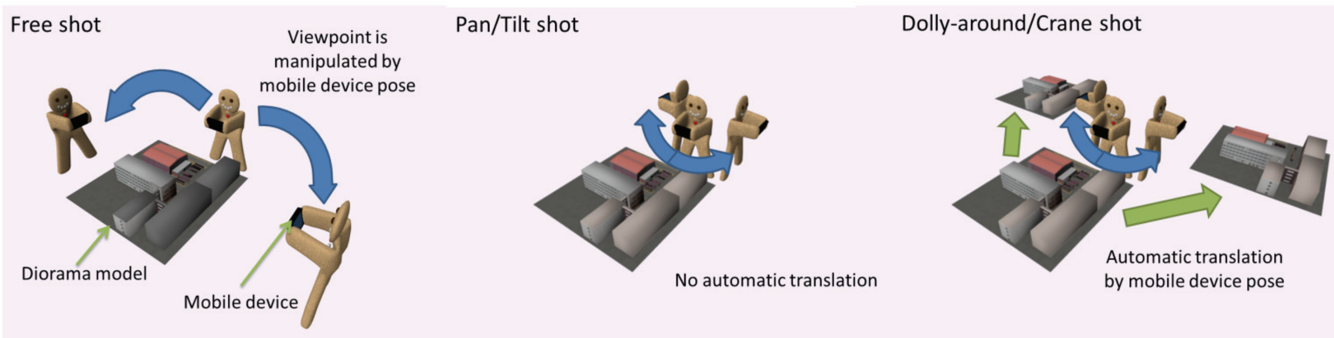

Figure 1. Three methods of viewpoint manipulation were compared in this paper. (Left) Free shot method works as if it was real diorama model placed on a table. (Middle) In pan/tilt shot method, orientation of the viewpoint corresponds to the orientation of mobile device. (Right) In dolly-around/crane (D/C) shot method, in spite of the rotation of the mobile device, the diorama model always appears in front of the mobile device.

Abstract

Diorama-based interface that displays a point of interest (POI) on a miniature of real 3D world is a good approach to share the POI with people working in outdoor environment. The viewpoint to observe the diorama model should be manipulated by users to explore the diorama model and find the POI. However, poorly designed viewpoint manipulation method may cause difficulty to understand the corresponding point of the POI in the real world.

A viewpoint manipulation method should be able to manipulate the viewpoint freely and the viewpoint enables a user to understand the correspondence between the real world and the diorama model easily. In order to realize a viewpoint manipulation method with satisfying the above requirements, this paper compares three viewpoint manipulation methods (free shot, pan/tilt shot and dolly-around/crane shot) that utilize a mobile device pose. We have implemented an AR (Augmented Reality) test bench of the diorama-based interface with photorealistic diorama model to conduct the subjective evaluation experiment. As a result, we found that dolly-around/crane shot is superior to the others in aspect of performance to find POI and subjective impressions.

Keywords: Outdoor mixed reality, User study, Viewpoint manipulation.

* e-mail: mhayashi@image.iit.tsukuba.ac.jp
† e-mail: kitahara@iit.tsukuba.ac.jp
‡ e-mail: kameda@iit.tsukuba.ac.jp
§ e-mail: ohta@acm.org

Index Terms: H.4.3 [**Information Systems Applications**]: Communications applications – *Point of Interest sharing*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *Artificial, augmented, and virtual realities*; I.3.6 [**Computer Graphics**]: Methodology and Techniques – *Interaction Techniques*

1 Introduction

When working in outdoor environments, people may need to find a three-dimensional point of interest (POI) indicated by other people. By putting a POI on a miniature diorama model (or a 3D map) of the environment, people could find the corresponding point to the POI in the real world. Hence, diorama-based interface such as *world in miniature* (WIM) [1] is a good approach to share a POI. In a design of the interface, the viewpoint observing the diorama model should be manipulated by users to explore the diorama model and find the POI. It is important to design the viewpoint manipulation method carefully.

With poorly designed viewpoint, users may not able to find out any cues that match their current view in the real world. As a result, they may feel difficulty to understand the corresponding point of the POI. Wingrave et al [2] proposed scaled and scrolling WIM (SSWIM) which is more suitable than originally WIM for large scale environment. However, it is difficult to apply SSWIM to outdoor augmented reality (AR) since its interaction is designed for immersive environment. Though some papers presented a diorama-based interfaces in outdoor AR context [3-6], there is still no user study on the viewpoint manipulation method for sharing a 3D POI at outdoor MR.

In this paper, we investigate the viewpoint manipulation methods of a diorama-based MR interface by a user study in outdoor environment. In particular, we compare three viewpoint manipulation methods: free shot, pan/tilt shot and dolly-around/crane shot. They are illustrated in the Figure 1. Since the

The 21st International Conference on Artificial Reality and Telexistence
November 28-30, 2011, Osaka, Japan
ISSN: 1345-1278

viewpoint manipulation methods utilize pose of mobile device, the way to handle the mobile device will make a difference.

Free shot is a commonly-used AR method which displays a diorama model on an ARToolKit marker [7] fixed to the real world. Pan/tilt shot is an egocentric viewpoint manipulation utilizing touch-screen instead of real-time position tracking of a mobile device. In dolly-round/crane shot, the viewpoint is moved on the surface of a virtual sphere surrounding a center of rotation by rotation of the mobile device.

These manipulation methods aim to manipulate the viewpoint freely and to make a user easily understand the correspondence between real world and the diorama model. Therefore, we align the viewpoint angles with the mobile device orientation in order to align the orientation of the diorama model with the real world.

We have conducted a user study to evaluate the three viewpoint manipulation methods in outdoor environment. Figure 2 illustrates the experimental situation. Our interface overlays a photorealistic diorama model of surrounding environment of the user on the video image captured by the camera of the mobile device. A POI is shown on the diorama model by an arrow-shaped icon. The premise of the experiment is that the diorama model of the environment, position of the users and position of the POI are given.

We exploited a photo-shooting task to evaluate performance to find out the POI. As a result of the experiment, we found that D/C shot is significantly superior to free shot in aspect of performance to find out POIs and subjective impressions. And D/C shot also tend to superior to pan/tilt shot.

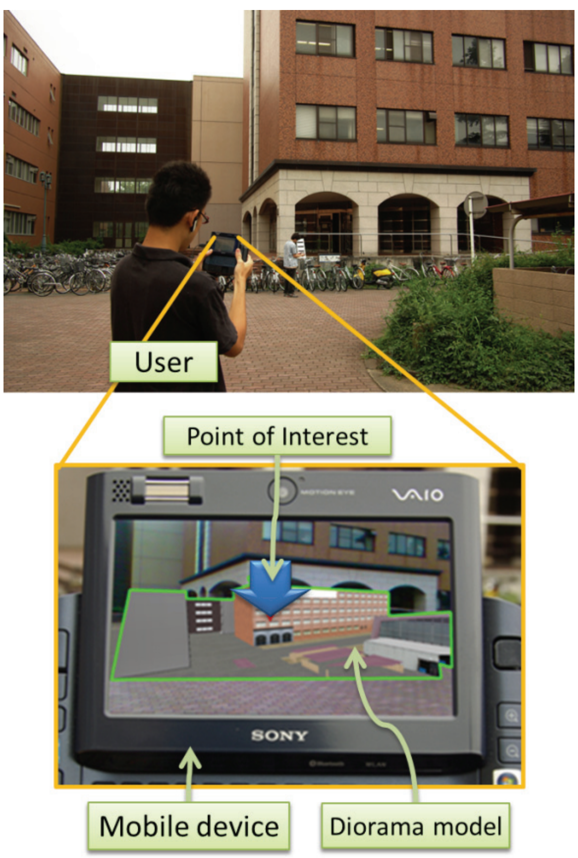

Figure 2. Our proposed diorama-based interface. A diorama model of surrounding environment is displayed on mobile device. An arrow-shaped icon represents the point of interest on the diorama model.

2 Related Work

POI in outdoor scene is sometimes used as geometric annotation. We review approaches of AR interface and user studies on diorama-based interface.

2.1 AR interface to point a place in outdoor environment

To indicate a POI in outdoor environment, directly putting annotation tag to the real world is common in AR [8], [9]. Although annotation tag makes it possible to intuitively understand the POI, there are two technical issues in practical use; "precise registration" and "good depth cue". Annotation tag could be misaligned according to camera registration error. However, precise and robust viewpoint registration in outdoor environment is still active research topics [10]. Even if registration has been done accurately, it is still difficult to perceive correct distance to annotation tag from user's view in the real world [8].

These problems can be avoided by using 3D map; the position of arrow-shaped icon indicating POI is represented in the diorama model coordinates. Thus, precise registration is not required. And user can perceive the correct distance from their position to the POI by exploring the 3D map. Designing interaction to the 3D map is important to realize effective navigation to POI [11]. Though there are user studies that display only 3D/2D maps [12], [13], we think that, diorama-based approach like WIM [1] is a good approach to share a POI related to the real world. We believe that a good interaction technique with mobile device pose can be more convenient than interaction with buttons, joystick or touch-screen in outdoor environment. Therefore, we utilize mobile device pose for our viewpoint manipulation methods of the diorama model.

2.2 User study on diorama-based mobile interface

A small-scaled CG model of surrounding real environment (i.e., diorama model) is useful to show the geometrical information. It visualizes a place where is not visible from a user's viewpoint. Thus there are many interfaces utilizes a diorama model. In this paper, we call them as diorama-based interface.

The concept of WIM, overlaying a diorama model of surrounding environment on user's view, was originally developed by Stoakley et al [1] to support navigation and interaction with virtual environment. Wingrave et al [2] conducted a detail user study on Scaled and Scrolling WIM. In AR context, Blaine et al [3] presented WIM with head mounted display. They used head-motion to manipulate the viewpoint. Höllerer et al [5] presented wire-frame rendered WIM aligned with the real world for pedestrian navigation using head mounted display. Okuma et al [6] conducted user study about viewpoint manipulation method for museum guide using 3D map.

However, there has not been any investigation about the suitable viewpoint manipulation method of WIM in outdoor environment. We compare three viewpoint manipulation methods utilizing mobile device pose. Our viewpoint manipulation methods can be considered as variations of Okuma's "bird's eye view + automatic rotation" with different interaction.

3 Viewpoint Manipulation Methods

In this section, we explain three viewpoint manipulation methods. For all methods, horizontal orientation of diorama model is aligned with the real world to reduce user's mental rotation [12].

3.1 Free shot

Free shot manipulation method fixes orientation and position of a diorama model to the real world. The viewpoint of the diorama model is fixed to the mobile device. Therefore, a user can manipulate the viewpoint in 6 degree-of-freedom. It is also called as AR view. Figure 1 (left) illustrates a user's motion to look around the diorama model. The user can observe the diorama model displayed on the mobile device monitor as if the model is set on the table in front of the user. Therefore, we initially assumed that free shot is the most intuitive viewpoint manipulation method among the all methods. Pros and cons of this method are follows.

Pros: Intuitive method with highest degree-of-freedom.

Cons: A user needs to walk around the diorama model to translate the viewpoint.

3.2 Pan/tilt shot

Pan/tilt shot (as known as egocentric view) utilizes touch screen instead of the mobile device position to translate the viewpoint in horizontal directions. Hence a user does not need to walk around the diorama model. Only orientation of the viewpoint is manipulated by the mobile device pose. As shown in Figure 1 (middle), position of the diorama model with respect to the real world is fixed unless the user manipulates the horizontal position by dragging manipulation on the touch screen. A vertical dragging from top to down moves the viewpoint to forward. A horizontal dragging from left to right moves the viewpoint to rightward. To looking a place from opposite direction, the user has to turn the mobile device around, and then translate the viewpoint to where the place is seen. The manipulation of vertical position of the viewpoint is omitted in this method. Pros and cons of this method are follows.

Pros: A user does not need to walk around the diorama model to translate the viewpoint.

Cons: Less intuitive and fewer degree-of-freedom than free shot.

3.3 Dolly-around/Crane shot

Dolly-round/crane shot (D/C shot) can be considered as orbital viewing [14] without twist. The viewpoint of the diorama model is moved on the surface of a virtual sphere surrounding a center of rotation. Note that the gravity direction is always aligned with the real world. In D/C shot, the orientation of mobile device is mapped so as to move the viewpoint as shown in Figure 1 (right). It is easy to view the POI on a diorama model from different viewpoint if the center of rotation is placed on the same position of the POI. Quantitative experiment [15] shows that the method is preferable than other head-tracked and non-head-tracked methods.

D/C shot also utilizes dragging manipulation on the touch screen to translate the center of rotation in horizontal directions for exploring the diorama model. Pros and cons of this method are follows.

Pros: A user does not need to walk around the diorama model to translate the viewpoint. Moreover, the user does not need to drag touch screen to look around the center of rotation.

Cons: Less intuitive than pan/tilt shot and free shot.

4 USER STUDY

We have conducted on a user study to compare the three different viewpoint manipulation methods presented above in outdoor environment. To compare the difference between these methods, we fixed the design of interface other than the viewpoint manipulation method by using a test bench interface. The test bench interface has been implemented on tablet PC (Sony VGN-UX92PS) with external inertial sensor (InterSense InertiaCube3). The pose tracking system for the test bench interface is simple. For free shot manipulation method, we used 15[cm] x 15[cm] sized marker set on a tripod stand of 75[cm] as high as common table top to track the mobile device. For pan/tilt shot and D/C shot, we used only inertial sensor since these methods require the orientation but not the position. Besides the viewpoint manipulation method, there are some design factors of our interface, such as the reality of the diorama model or visibility of the mobile device. These factors are not changed through the experiment using our test bench implementation in order to focus on the viewpoint manipulation method.

Figure 3 shows a screenshot of our test bench interface. The screen is divided into two parts; "diorama part" and "GUI part". In diorama part, a user can observe a human-shaped icon to show the user's current location and an arrow-shaped icon to indicate POI on a diorama model. When the icon is occluded from the viewpoint by the diorama model, the occluder object is rendered as a translucent object in order to keep the arrow visible. In GUI part, a slide-bar for controlling the scale of diorama model and mini camera image is displayed. The control for scaling is important to explore a large scale environment. The initial scale of the diorama model is set as 1/150 and could be varied 1/50 to 1/500.

Figure 4 shows the aerial view and a photorealistic textured diorama model of corresponding area used for the experiment. The experiment was conducted as paired comparisons. Participants compared three pairs of viewpoint manipulation methods ("Pan/tilt - Free", "D/C - Free", and "Pan/tilt - D/C"). We divided the participants into six groups in order to counterbalance the presentation order of the three pairs. We had at least two participants in each group so as to counterbalance the presentation order of the viewpoint manipulation methods in each pair. In each comparison, participants repeated user task ten times after few time practices. In the practice, we used same ar

After the repetition of the task, they answered following questions on a 5-point Likert scale.

Q1. Which was easier to understand the POI?

Q2. Which was easier to manipulate the viewpoint?

Q3. Which do you prefer to use?

Participants conducted 10 (trials per method) × 2 (methods in each pair) × 3 (combinations of viewpoint manipulation methods) = 60 trials through the experiment, and it took about 30 minutes.

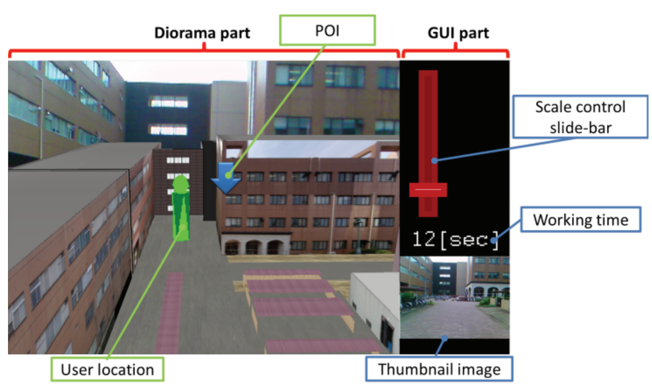

Figure 3. A screenshot of the mobile device display. The screen consists of two parts; one is diorama part where the diorama model, the real world and a user location and POI are displayed. The other is the GUI part with touch controllable slide-bar for controlling the scale, timer for measuring the working time for each task, and the thumbnail video image of the real world.

Figure 4. Experimental environment. (Left) An aerial photo. (Right) the photorealistic diorama model. Size of the modeled area is about 120[m] x 240[m].

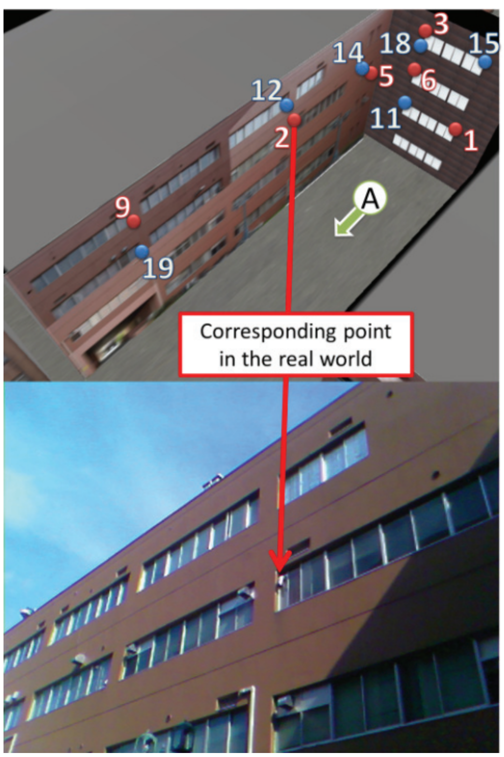

Figure 5. In our experiment, participants had to locate the corresponding point in the real world (top) to the POI in the diorama model (bottom), and took photo of the location to complete the task.

4.1 User task

The most important feature of the diorama-based interface is ability to indicate 3D POI. When a POI is indicated by an arrow on the diorama model, the user can easily understand the position of the indicated POI even if it is invisible from him/her. In order to do correct "photo-shooting task", the user has to understand the POI location accurately. Thus, we choose the task for evaluating the three types of viewpoint manipulation methods. In this task, a POI is indicated on the diorama model to participants. Then participants report the position by taking a photo of the POI. Participants were allowed to walk around the AR marker. They were requested to finish the task as soon as possible. Before starting the task, participants stood at initial position (point (A) in Figure 4) and turned the mobile device to its initial orientation (direction of the arrow). This initialization was done in each repetition, and hence the drift error of inertial sensor was very small during the task.

To complete the task, participants had to locate the corresponding point in the real world (top of Figure 5) to the POI in the diorama model (bottom of Figure 5), and took photo of the location. To evaluate performance, we measured the working time from appearing the diorama model and POI on the display to shooting the photo by a participant. We prepared two groups of ten POIs that are carefully selected in order not to be uneven distribution. In each viewpoint manipulation method, we measured the working time repeatedly ten times with POIs selected from one of the groups in random order. To evaluate subjective impressions, we used paired comparison between all the methods as explained in the section 3.

4.2 Results

Figure 6 shows box-plot of the working time of overall trials of photo-shooting task as a result of the performance evaluation with 15 participants (13 males and 2 females). We ran one-way ANOVA of the statistical software package SPSS for the results. We found significant difference between the means of working time for the viewpoint manipulation methods $F(2,925)=4.797$, $p=0.008$ with 5% significance level. A Tamhane post-hoc test revealed that D/C shot (12.2[sec]) was significantly faster than free shot (13.6[sec], $p=0.016$) and pan/tilt shot (13.7[sec], $p=0.030$), and there are no significant difference between pan/tilt shot and free shot ($p=1.000$).

Figure 7 shows the result of subjective questions for each pairs with 17 participants (15 males and 2 females). We analyzed the result of each question using Scheffe's method of paired comparison with Nakaya's variation. As a result, we found that D/C shot was significantly better than free shot in aspect of all questions at 1% significance level. We also found that participants preferred to use pan/tilt shot than free shot at 5% significance level about question Q3. Although we found no significant differences between pan/tilt shot and D/C shot, D/C shot was tend to be preferred than pan/tilt shot.

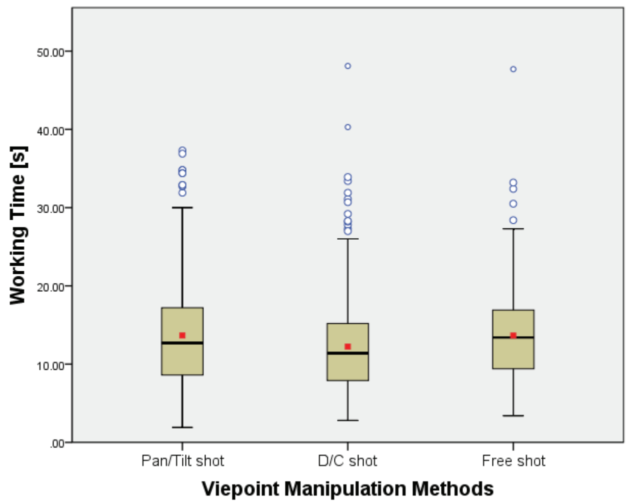

Figure 6. Result of "working time" which is measured as a performance evaluation. In the box plot, the red squares indicate the means and blue circles indicate outliers. Dolly-around/crane (D/C) shot was significant faster than free shot and pan/tilt shot (significance level is 5%).

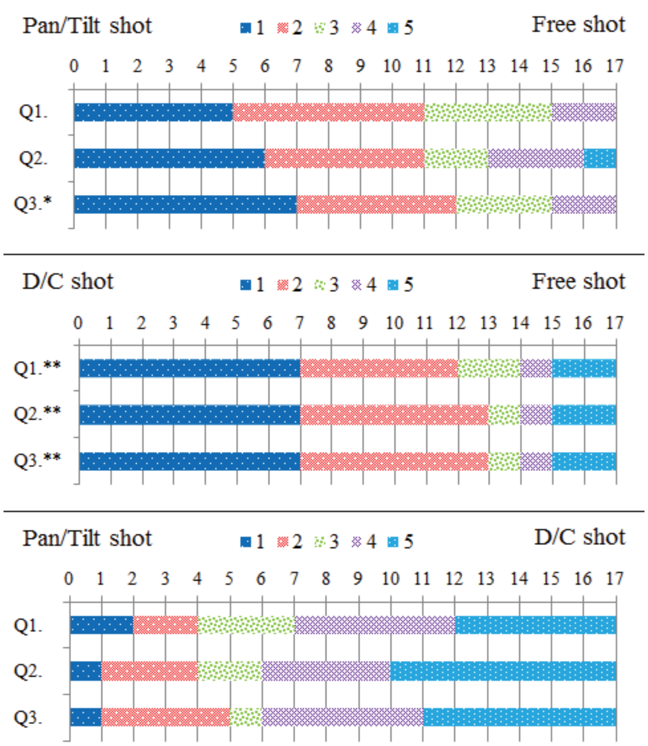

Figure 7. Result of subjective questions of paired comparison in the viewpoint manipulation methods. The graph shows the number of participants of each score. * and ** denotes significant at 5% and 1% level in correspond.

4.3 Discussion

In this experiment, the result shows that D/C shot was obviously superior to the other methods in aspect of performance and subjective impressions. By contrast, free shot could not obtain a good evaluation.

We initially expected that more participants prefer the free shot method in subjective impressions since this is the most intuitive method that has a real world metaphor of a diorama model on a table. One possible reason of the unexpected negative impression for this method is the limitation of implementation about the registration of a mobile device. Tracking of ARToolKit marker [7] was sometimes unstable in the user study, so that some participants reported that they feel discomfort to find the POI. Six participants reported that free shot was easy to use in early trials. However, the other two methods were better when they got used to the methods. Moreover, three participants were not preferred to walk around the AR marker. Therefore, there is also a possibility that the free shot was not suitable for finding a POI.

Pan/tilt shot have been preferred than free shot, though there is no difference in the working time. It seems that the time to translate the viewpoint by touch screen is as long as the time to walk around the AR marker (sometimes extended by tracking error). Therefore, using touch screen instead of walking to translate the viewpoint did not improve the performance evaluation in this experiment.

In D/C shot, over 70% (12) participants preferred this method than free shot, and almost 60% (11) participants preferred than pan/tilt shot. It seems that D/C shot makes it easy to compare POI and its corresponding point in the real world. The possible reason of the result is that diorama model was always visible in the center of the view. Hence it could be easy to compare POI and its corresponding point in the real world. This feature of D/C shot also has a bad effect. The diorama model often occluded a point of real world when they aimed to take a photo. Though it was possible to see thumbnail of camera image without diorama model, some participants demanded to control the visibility of the diorama model.

Though POI on the diorama model often was occluded by the diorama model during the experiment, it does not seem to be a big problem. Simple translucent rendering of occlude object seems effective in the experimental environment. We think we need further investigation to address the self-occlusion problem.

5 Conclusion

We have compared three viewpoint manipulation methods utilizing mobile device pose, free shot, pan/tilt shot and dolly-around/crane (D/C) shot for diorama-based interface. We have conducted a user study to evaluate the performance and participant's subjective impressions to find a 3D point of interest (POI). The experimental environment was a part of the campus of our university, and the photorealistic diorama model of the area. As a result, D/C shot was relatively superior to the other methods in aspect of performance to find a POI and subjective impressions. Our experiment has some limitations; the result of free shot was influenced by the instability of camera tracking. The effect of self-occlusion of the diorama model was limited since the experimental environment is not so complicated. We need further experiment with more robust tracking and more complex environment.

Acknowledgements

This work is partially supported by Grant-in-Aid for JSPS Fellows (23310).

References

[1] R. Stoakley, M. J. Conway, and R. Pausch, "Virtual reality on a WIM: interactive worlds in miniature," *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 265–272, 1995.

[2] C. A. Wingrave, Y. Haciahmetoglu, and D. A. Bowman, "Overcoming world in miniature limitations by a scaled and

scrolling WIM," *IEEE Symposium on 3D User Interfaces*, pp. 11-16, 2006.
[3] B. Bell, T. Höllerer, and S. Feiner, "An annotated situation-awareness aid for augmented reality," *Proceedings of the 15th annual ACM symposium on User interface software and technology (UIST)*, vol. 4, no. 2, pp. 213-216, 2002.
[4] T. Höllerer, S. Feiner, T. Terauchi, G. Rashid, and D. Hallaway, "Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system," *Computers & Graphics*, vol. 23, no. 6, pp. 779-785, Dec. 1999.
[5] T. Höllerer et al., "User interface management techniques for collaborative mobile augmented reality," *Computers & Graphics*, vol. 25, no. 5, pp. 799-810, Oct. 2001.
[6] T. Okuma, M. Kourogi, K. Shichida, and T. Kurata, "User Study on a Position-and Direction-aware Museum Guide using 3-D Maps and Animated Instructions," *CD Proceedings The first Korea-Japan workshop on Mixed Reality (KJMR08)*, p. 8, 2008.
[7] H. Kato and M. Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR)*, pp. 85-94, 1999.
[8] J. Wither, S. DiVerdi, and T. Höllerer, "Annotation in outdoor augmented reality," *Computers & Graphics*, vol. 33, no. 6, pp. 679-689, Dec. 2009.
[9] C. Sandor et al., "Egocentric space-distorting visualizations for rapid environment exploration in mobile mixed reality," *Virtual Reality Conference (VR), 2010 IEEE*, pp. 47–50, Oct. 2010.
[10] J. Karlekar, S. Zhou, W. Lu, Z. C. Loh, Y. Nakayama, and D. Hii, "Positioning, tracking and mapping for outdoor augmentation," *Proceedings of IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 175–184, 2010.
[11] A. Nurminen and A. Oulasvirta, "Designing Interactions for Navigation in 3D Mobile Maps," *In L. Meng, A. Zipf, S. Winter (Eds.), Map-based Mobile Ser- vices: Design, Interaction and Usability, Springer, Lecture Notes in Geoinfor- mation and Cartography, London*, pp. 198-224, 2008.
[12] A. Oulasvirta, S. Estlander, and A. Nurminen, "Embodied interaction with a 3D versus 2D mobile map," *Personal and Ubiquitous Computing*, vol. 13, no. 4, pp. 303-320, Jul. 2008.
[13] K. Laakso, O. Gjesdal, and J. R. Sulebak, "Tourist information and navigation support by using 3D maps displayed on mobile devices," *Workshop on Mobile Guides, Mobile HCI 2003 Symposium, Udine*, no. September, pp. 3-8, 2003.
[14] D. Koller and M. Mine, "Head-tracked orbital viewing: an interaction technique for immersive virtual environments," *Proceedings of the 9th annual ACM Symposium on User Interface Software and Technology*, pp. 81-82, 1996.
[15] J. C. Chung, "A comparison of head-tracked and non-head-tracked steering modes in the targeting of radiotherapy treatment beams," *Proceedings of the 1992 symposium on Interactive 3D graphics (SI3D)*, no. 919, pp. 193-196, 1992.