

Attention Sharing in a Virtual Environment Attracts Others

Takuji Narumi^(✉), Yuta Sakakibara, Tomohiro Tanikawa, and Michitaka Hirose

Graduate School of Information Science and Technology, The University of Tokyo,
7-3-1 Hongo, Bunkyo-Ku, Tokyo 113-8656, Japan
{narumi, sakakibara, tani, hirose}@cyber.t.u-tokyo.ac.jp

Abstract. Virtual reality offers a highly interactive and flexible experience. It has the advantage of enhancing users' understanding and interest as well as the disadvantage of overlooking the main features in a virtual environment. The excessive amount of information and interactive options in most virtual reality settings may cause users to quit exploring before experiencing the entire content of the virtual environment. In this paper, we propose a new method of inherently encouraging users to continue their experience in virtual environments while permitting free exploration through social interactions. The proposed method generates a joint attention by displaying the movement of the position and gaze direction of other concurrent/previous users. We introduced the proposed method into a virtual museum exploring system and demonstrated it in a real museum to evaluate the effectiveness of our method when used by a large number of people. The results showed that the proposed method enhances users' interest and prolongs the experience time of virtual museum exploring.

Keywords: Attention · Social interaction · Digital museum · Virtual reality

1 Introduction

Virtual reality offers a highly interactive and flexible experience. It is widely recognized that it enhances users' understanding and interest in virtual objects/environments more effectively than learning through passive media. This is because their active selections of interactive objects that attracts their interest makes their experience more subjective and unforgettable [1–3]. For example, when users want to experience a virtual environment, it is more effective for users to explore and understand it through their own navigation than to passively watch instructional videos.

Recently, museums have high expectations for this effect and have introduced interactive technologies, including virtual reality into their exhibition methods to effectively provide supplementary background information regarding their exhibits [1, 4, 5]. Notably, photorealistic virtual content based on spherical image capturing have been a focus of attention to preserve and transmit cultural heritage. Spherical images are suitable for easy construction and use in immersive and realistic virtual environments [6]. A spherical image contains the entire information of the landscape of all angles from a location. The popularization of omnidirectional cameras, which can capture spherical

images instantaneously, has facilitated the archival of a real space. Additionally, the usage of experiential devices such as tablet devices and head-mounted displays has become widespread. The experience of spherical images by using a hand-held device, such as a tablet device, is known to be immersive and effective in the understanding of geometric space [6–9].

However, virtual reality has these merits as well as the disadvantage of overlooking the main features in a virtual environment. In most virtual reality settings, there is so much information and interactive options that users may quit exploring prior to experiencing the entire content in the world.

Particularly, users should be discouraged from quitting a virtual experience without interacting with items the designers of the virtual world consider important. For example, in a photorealistic virtual museum, the most important objects are the exhibits but less important objects or information such as lighting, room arrangements, and spatial orientation should not be ignored because these are usually part of the exhibition design. Furthermore, it is possible to induce the psychological phenomenon that an increase of gazing time by an appropriate amount enhances the preference for the object (the mere exposure effect) [10, 11]. Therefore, it is necessary to make users aware of the important exhibits while they are exploring.

Guidance methods to guide users to pre-defined locations in the virtual environment have been suggested. One of these methods focused on steering users along a pre-defined path while at the same time allowing some extent of free exploring [12–14]. Another method uses an explicit arrow pointing at the target locations [15]. However, these methods are so intrusive that users cannot enjoy free exploration in the environment, and they hardly achieve a sense of accomplishment to find the target, which may decrease the quality of the experiment.

To inherently guide users to pre-defined locations in the virtual environment, while continuing to permit free explorations, Tanaka et al. proposed a method of inducing users to look at a pre-defined point in the spherical image by redirecting the virtual camera [8]. The user's virtual camera direction is shifted to look at a point closer to a target point. Moreover, Tanaka et al. also proposed the guidance method "Guidance field" [16], which slightly alters user's input for locomotion and rotation based on a potential field, which represents the drawing force to a target location. They showed that these modifications successfully guided users to pre-defined locations and made users aware of the target objects in the virtual environment. However, this method did not extend the experience time of users. This result suggested that this method would enable us to draw users' attention, but not enhance their interest to search virtual environments.

Then, as an alternative approach, we focused on the influence of others. Joint attention refers to a social-communicative skill used by humans to share attention directed at interesting objects or events with others via implicit and explicit indications such as gestures and gaze. Because of joint attention, we tend to be attracted to objects or events that others are looking at. Behavioral contagion is a type of social influence and refers to the propensity for certain behaviors exhibited by one person to be copied by others who are in the vicinity of the original actor. Milgram et al. reported that the larger the size of a stimulus crowd standing on a busy city street looking up at a building, the more frequently passersby adopt the behavior of the crowd [17]. These influences on others

can be used for attracting people's attention and enhancing interest to particular objects or events [18].

This phenomenon is already utilized for supporting navigation in virtual environments using virtual agents. For example, virtual humans can give directions or transport users to locations [19]. Other research proposed the use of a flock of virtual animals to indicate interesting places in a virtual environment and confirmed this method's effectiveness [20]. Additionally, to instruct users on how to interact with an exhibit in a museum, a system that records and three-dimensionally superimposes past visitor interactions around the exhibit was proposed [21]. In this system, visitors see the behaviors of previous visitors, and thereby, obtain a better understanding of the exhibit.

In this study, we propose a new method of encouraging users to continue their experience in a virtual environment while permitting free exploration by using the effect of social interactions. As mentioned above, people tend to more frequently direct their attention to the objects/events that someone is directing. Therefore, we employed a method that generates a joint attention by displaying the movement of the position and gaze direction of other users.

2 Attention Sharing in a Virtual Environment by Sharing the Position and Gaze Direction of Others

In this chapter, we first describe the premises of the virtual environment used in the following experiment. Next, we explain the proposed method that displays the movement of the position and gaze direction of concurrent users or previous users.

2.1 Photorealistic Virtual Environment Used in this Study

As described above, the construction of virtual environments with a sequence of spherical images is becoming popular because the virtual spaces constructed with spherical images are more realistic, immersive, and easy to construct compared to virtual spaces constructed with computer graphic models [7, 22–24].

We then developed an application that enables virtual environments to be explored and named it “Window to the Past” [16]. This application depicts a virtual museum space constructed with a large sequence of spherical images. The application uses images of the Modern Transportation Museum in Osaka, Japan, with an area of 10,000 m². The exhibit was closed in April 2014. We archived the museum as 7,096 spherical images with a 360-degree spherical camera Ladybug5 (Point Grey Inc.), and developed a node-edge based walk-through system by using an appropriate arrangement of images (Fig. 1).

Spherical image viewer applications have several user interfaces for rotating the virtual camera such as a joystick, mouse, and on-screen buttons. In this research, the direction of the virtual camera is linked with the orientation of the tablet device. Therefore, when users move the device in a specific direction, the virtual camera moves in the same direction. Interactive interfaces that involve physical motion have been reported as encouraging users to develop a deeper understanding of geometric space and content



Fig. 1. Map of the reconstructed “Modern Transportation Museum (Osaka)” and camera paths (red lines). (Color figure online)

[1, 2]. The orientation of the tablet device in the real world is obtained from the built-in gyro and acceleration sensors in the device.

Moreover, to eliminate the effect of physical and spatial limitations, which would prevent users from rotating the virtual camera arbitrarily, only the angle around the vertical axis can be rotated by swiping the touch screen.

A virtual pad is used as an input interface for locomotion (Fig. 2 (left)). The virtual pad is a touch panel input interface designed to work as a joystick. Sliding the pad is equivalent to tilting a joystick [25]. Moreover, to inform users of the direction in which they can move from their node, arrows are shown around the virtual pad.

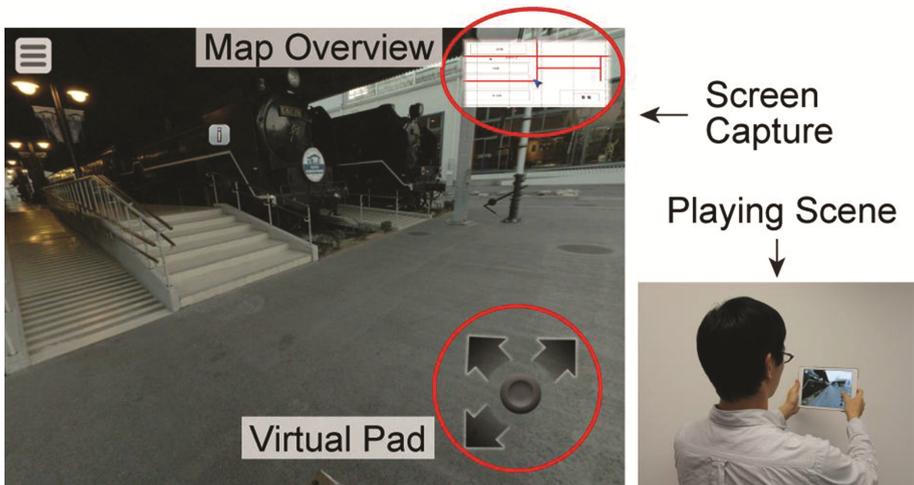


Fig. 2. Screen capture and playing scene of “Window to the Past.”

The distance between nodes was approximately 0.15 m, and the normal walking speed was set at 160 m/min. Thus, approximately 18 spherical images were seamlessly shown in one second. 160 m/min is about twice as fast as the real walking speed. However, it is known that the motion speed in virtual environment is perceived slower than the actual speed [26]. Furthermore, by adjusting the amount of slide of the pad along the radial direction, users can move at arbitrary speed slower than the maximum speed setting (160 m/min).

Part of the overview map, including the camera paths, user's position, and user's orientation, is displayed in the upper-right corner of the screen (Fig. 2 (left)). Figure 2 (right) shows a playing scene of *Window to the Past*. The user is holding an iPad.

This type of setup is ordinary for exploring virtual environments. Therefore, we use this setup as an example.

2.2 Proposed Method for Sharing the Position and Gaze Direction of Other Users

The users perform two operations to explore the virtual environment: translation and rotation. Translation is an operation to move in the virtual environment. Rotation is an operation to look around in the virtual environment. Therefore, it is necessary to present information that shows the state of these two behaviors for sharing other user's experiences in the virtual environment. Then, in order to design an indicator for sharing other user's experiences, we considered the following two requirements:

- can be easily understood as it expresses other user who is interacting with the same system
- can be easily understood the position and gaze direction of other user

Based on these requirements, we designed the indicator as shown in Fig. 3. To fulfill the first requirement, we showed a tablet device and footprint per one user on the main screen. To fulfill the second requirement, we changed the position and direction of the tablet device and footprint according to the position and gaze direction of other user.

The user of the system described in previous chapter is holding an iPad and moves it to search around in the virtual environment. We considered that by displaying the tablet device we could tell the existence of others in at a minimum. By doing so, compared to simply displaying the model of a person, we can clearly express that the indicator is not a person who is simply placed in the virtual environment as a non-player character, but is a person who is interacting with the system as same as the user. In order not to disturb the user's appreciation as much as possible, we did not attach animation effect and made the minimum appearance to understand the position and direction in the virtual environment.

In addition, arrow icons that represent the users are shown in the map. They move and rotate according to the positions and directions of the users. On the map, the icon that represents oneself is displayed in blue, and the icon that represents other users is displayed in green. By changing the color of the icon, the user can distinguish themselves and others from the map.

This system can be used in two ways; real-time sharing and sharing with past experiences.

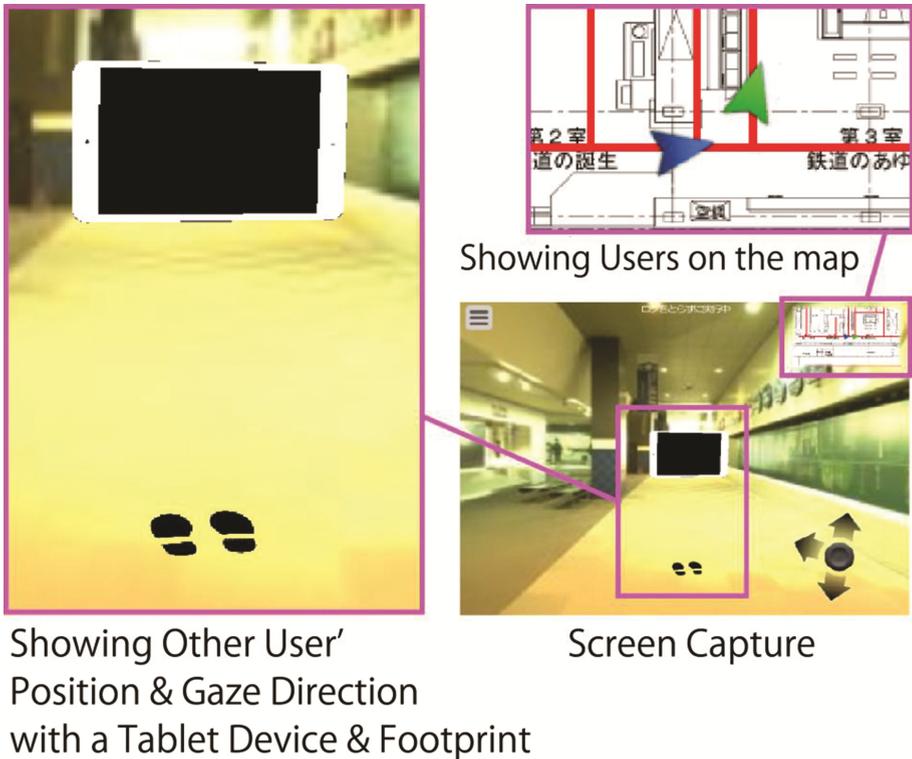


Fig. 3. Showing other user's position and gaze direction with an transparent avatar with a tablet device and footprint on the main screen and arrows on the map view (Color figure online)

In real-time sharing, when this system is played at the same time on several tablet devices, the position and gaze direction of other users are immediately shared through the network. Based on the shared information, the indicators are drawn on the screen. Then the users can see each other's position and orientation in real time. The system can express the attention of people in the virtual environment by displaying position and orientation of others who are exploring the virtual environment together. By doing so, the system aimed to spark user's interest in the experiences in the virtual environment, extend the experience time of the system, and increase the number of interactions such as button operations.

Moreover, it is anticipated that synergistic effects due to communication in the real world between users will occur if the users who experience at the same time are nearby in the real exhibition space. For example, when one of the users has an impressive experience, s/he sometimes tells it to others around her/him. It will lead others to make the same interaction/experience. Then this kind of synergistic effect will make the interactive experience more effective.

In addition to sharing information on other users who are experiencing at the same time, by displaying the history of other users in the past, it is considered that the user can

feel more bustling in the virtual environment. This will enhance the interest arousal effect. This is also useful for guidance in the virtual environment. It is considered that the user can be guided to appreciate the virtual environment as intended by a museum curator by displaying the history of other user who has appreciated in an appropriate way.

3 Experiment in a Real Exhibition

3.1 Overview of the Experiment

We evaluated our method, which shows other user's position and gaze direction, by conducting a large-scale experiment in a real exhibition at Kyoto Railway Museum (Kyoto, Japan). Figure 4 shows an overview of this exhibit. Three iPad Aird act as the tablet devices, and a simple explanation of the Modern Transportation Museum and instructions as to how to use the application are placed as shown in the figure. The participants are people who visited the exhibition and experienced the system during the 56 days of the experiment (From April 29 to July 3 2016, excluded closed period).



Fig. 4. Overview of the exhibit and playing scene

The visitors varied from children to seniors. They did not know the purpose of this experiment and regarded it as a normal exhibit. To analyze the play log of the application, we set up the application to automatically terminate the experience when the user taps the finish button displayed at the upper left of the screen or returning the tablet to the display stand. During the user experience, the orientation and position in the virtual environment was logged. We only analyzed the data of first-time users because our focus was to capture user behavior when exploring the virtual environment without any previous knowledge. We identified first-time users by asking the subjects whether they had already experienced “Window to the Past” before starting their experience. Since the experiment was conducted in real museum in order to collect a large number of subjects, all conditions were not perfectly controlled. For example, although we show the explanations of the usage of the application and the input interfaces before subjects start their experience, it is not clear whether they really read and understand the explanations. The situation that users start to use applications without perfect understand often

occurs in practical use, and it is meaningful to analyze a large number of data obtained from such practical situations.

We compared two conditions; with sharing and without sharing conditions. Under a with sharing condition, the position and gaze direction of other users are immediately shared through the network with the proposed method, as shown in Sect. 2.2 when the system was played at the same time on more than two tablet devices. Under a without sharing condition, the position and gaze direction of other users are not shared. Then all users played the system solely.

3.2 Results and Discussion

Under the with sharing condition, three users played at the same time in 49% of total experience time. Two users played at the same time in 34% of total experience time. Meanwhile, only one user played in 17% of total experience time. Maximum number of users who played at the same time was three in 64% of trials, two in 25% trials, and one in 11% of trials. Resulting in 11% of the participants under the with sharing condition were played as same as under the without sharing condition. However, assuming that the number of participants entering the exhibition room per unit time is constant, the probability of experiencing by only one person throughout the entire experience is considered to be higher for people with shorter experience time. If we analyze except these 11% of participants, it is considered that the experience time under the with sharing condition will be longer than actual. The sharing function was working in 83% of total experience time. It is considered that this rate is high enough to investigate the effect of the proposed system. We did not exclude these data from the sharing condition.

We excluded outliers for following analysis. Outliers are values that are 1.5 times greater than the third quartile of the interquartile range or 1.5 times smaller than the first quartile range of the quartile range. Moreover, we excluded data of users whose experience time is less than 10 s because the users do not appreciate the virtual environment properly in such cases. The total number of participants under the with sharing condition were 7,396, and 4,968 were the number of participants under the without sharing condition.

Figure 5 shows the box plot of the experience time per user by experimental conditions. The median of the experience time was 89.4 s for those under the with sharing condition and 70.6 s for those under the without sharing condition. Mann-Whitney's U test revealed that the experience time was significantly longer for those under the with sharing condition than for those under the without sharing condition ($p < 0.01$). Participants with experience time of less than 60 s were 33.3% of the total number under the with sharing condition and 44.0% under the without sharing condition. Because the number of users who finished with experiences of less than 60 s decreased, the proposed system has the effect of preventing loss of user's interest. These results suggested that the proposed method can increase the user's interest in the virtual environment, and lengthen the experience time.

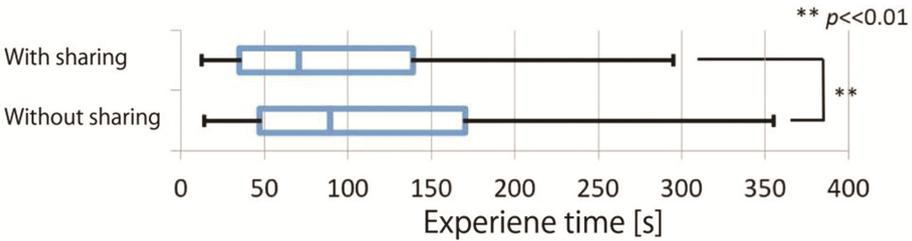


Fig. 5. Box plot of the experience time by the experimental conditions

Figure 6 (left) shows the average moving speed of the user for each elapsed time from the start of the trial. In the case of the system used in this experiment, it is rare that the user is constantly moving in the virtual environment. The users normally repeat moving and stopping. Therefore, it can be thought that the average moving speed is an indicator of activity level in the virtual environment (Higher moving speed indicates that the user moves more actively in the virtual environment). The result showed that average moving speed is higher for those under the with sharing condition than for those under the without sharing condition. In addition, the gap between average moving speeds under each condition increased with time. These suggest that the user moved more actively in the virtual environment under the with sharing condition, and the proposed method increased the moving distance in the virtual environment.

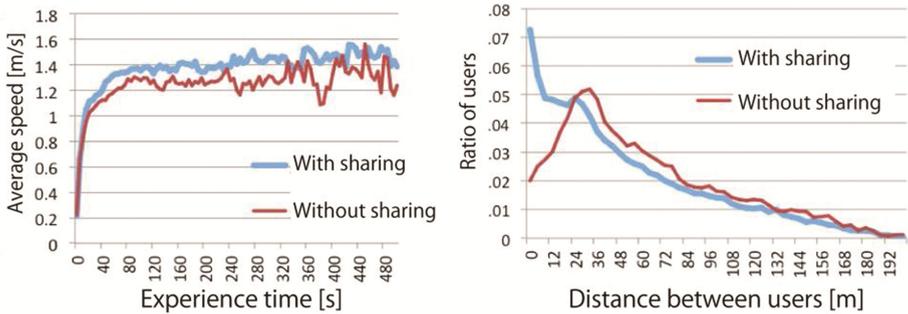


Fig. 6. Average moving speed (left) and Distance between users (right) in the virtual environment

Figure 6 (right) shows the distribution of distances between users who played at the same time. All subjects start moving from the same starting point. Therefore, in order to eliminate the effect of overlapping of positions immediately after the start, data of participants who are apart for more than 20 m from the start position were analyzed. These results suggest that the distance between the users tends to decrease for those under the with sharing condition. This effect might be utilized for guidance in the virtual environment because the users tend to move to other users shown in the environment.

4 Conclusion

In this paper, we proposed a new method of inherently encouraging users to continue their experience in a virtual environment while continuing to permit free exploration through social interactions. The proposed method generates a joint attention by displaying the movement of the position and gaze direction of other concurrent or previous user experiences. We introduced the proposed method into a virtual museum exploring system and demonstrated it in a real museum to evaluate the effectiveness of our method when used by a large number of people. The results showed that the proposed method prolonged the experience time of virtual museum exploring, and in particular, decreased the number of users who finished with experiences of less than 60 s. The results also suggested that the proposed method enhances users' interest and makes them move more actively in the virtual environment. In addition, the distance between the users in the virtual environment tends to be short when they use the proposed system in real-time sharing mode. These results showed the effectiveness of the proposed system.

In this study, we only investigated the proposed system in real-time sharing mode. Then, we need to investigate whether the proposed system in sharing with past mode also encourages users to continue their experience in a virtual environment. Moreover, the effect of shortening the distance between the users can be utilized for guidance in the virtual environment. We will seek a novel guidance method based on the attention sharing method and investigate its effectiveness in future work.

Acknowledgement. This work was partially supported by the MEXT, Grant-in-Aid for Scientific Research (A), 25240057.

References

1. Narumi, T., Hayashi, O., Kasada, K., Yamazaki, M., Tanikawa, T., Hirose, M.: Digital diorama: AR exhibition system to convey background information for museums. In: Shumaker, R. (ed.) VMR 2011. LNCS, vol. 6773, pp. 76–86. Springer, Heidelberg (2011). doi:[10.1007/978-3-642-22021-0_10](https://doi.org/10.1007/978-3-642-22021-0_10)
2. Arakawa, T., Kasada, K., Narumi, T., Tanikawa, T., Hirose, M.: Augmented reality system for overlaying a scene in a video onto real world and reliving the camera operator's experience. *Proc. IEEE Virtual Reality* **2013**, 139–140 (2013)
3. Cairncross, S., Mannion, M.: Interactive multimedia and learning: realizing the benefits. *Innovations Educ. Teach. Int.* **38**(2), 156–164 (2001). Taylor & Francis
4. Kajinami, T., Hayashi, O., Narumi, T., Tanikawa, T., Hirose, M.: Digital display case: museum exhibition system to convey background information about exhibits. In: 2010 16th International Conference on Virtual Systems and Multimedia (VSMM), pp. 230–233 (2010)
5. Imura, J., Kasada, K., Narumi, T., Tanikawa, T., Hirose, M.: Reliving past scene experience system by inducing a video-camera operator's motion with overlaying a video-sequence onto real environment. *ITE Trans. Media Technol. Appl.* **2**(3), 225–235 (2014)
6. Hwang, J., Jung, J., Kim, G.J.: Hand-held virtual reality: a feasibility study. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 356–363 (2006)

7. Chen, S.E.: Quicktime VR: an image-based approach to virtual environment navigation. In: Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, pp. 29–38 (1995)
8. Tanaka, R., Narumi, T., Tanikawa, T., Hirose, M.: Attracting user's attention in spherical image by angular shift of virtual camera direction. In: Proceedings of the 3rd ACM Symposium on Spatial User Interaction, pp. 61–64 (2015)
9. Okada, N., Imura, J., Narumi, T., Tanikawa, T., Hirose, M.: Manseibashi reminiscent window: on-site AR exhibition system using mobile devices. In: Streitz, N., Markopoulos, P. (eds.) DAPI 2015. LNCS, vol. 9189, pp. 349–361. Springer, Heidelberg (2015)
10. Shimojo, C., Simion, C., Shimojo, E., Scheier, C.: Gaze bias both reflects and influences preference. *Nat. Neurosci.* **6**(12), 1317–1322 (2003). Nature Publishing Group
11. Zajonc, R.B.: Attitudinal effects of mere exposure. *J. Pers. Soc. Psychol.* **9**(2p2), 1–27 (1968). American Psychological Association
12. Galyean, T.A.: Guided navigation of virtual environments. In: Proceedings of the 1995 Symposium on Interactive 3D Graphics, pp. 103–ff. ACM (1995)
13. Ropinski, T., Steinicke, F., Hinrichs, K.: A constrained road-based VR navigation technique for travelling in 3D city models. In: Proceedings of the 2005 International Conference on Augmented Tele-existence, pp. 228–235 (2005)
14. Abasolo, M.J., Della, J.M.: Magallanes: 3D navigation for everybody. In: Proceedings of the 5th International Conference on Computer Graphics and Interactive Techniques in Australia and Southeast Asia, pp. 135–142 (2007)
15. Chittaro, L., Burigat, S.: 3D location-pointing as a navigation aid in Virtual Environments. In: Proceedings of the Working Conference on Advanced Visual Interfaces, pp. 267–274 (2004)
16. Tanaka, R., Narumi, T., Tanikawa, T., Hirose, M.: Guidance field: potential field to guide users to target locations in virtual environments. In: Proceedings of 2016 IEEE Symposium on 3D User Interfaces (3DUI), pp. 39–48 (2016)
17. Milgram, S., Bickman, L., Berkowitz, L.: Note on the drawing power of crowds of different size. *J. Pers. Soc. Psychol.* **13**(2), 79 (1969)
18. Narumi, T., Yabe, H., Yoshida, S., Tanikawa, T., Hirose, M.: Encouraging people to interact with interactive systems in public spaces by managing lines of participants. In: Yamamoto, S. (ed.) HIMI 2016. LNCS, vol. 9735, pp. 290–299. Springer, Cham (2016). doi:[10.1007/978-3-319-40397-7_28](https://doi.org/10.1007/978-3-319-40397-7_28)
19. Van Dijk, B., Zwiers, J., op den Akker, R., Nijholt, A.: Navigation assistance in virtual worlds. In: Boyd, E., Cohen, E., Zaliwski, A.J. (eds.) Proceedings 2001 Informing Science Conference, pp. 1–9 (2001)
20. Ibanez, J., Delgado-Mata, C.: Flocking techniques to naturally support navigation in large and open virtual worlds. *Eng. Appl. Artif. Intell.* **25**(1), 119–129 (2012)
21. Narumi, T., Kasai, T., Honda, T., Aoki, K., Tanikawa, T., Hirose, M.: Digital railway museum: an approach to introduction of digital exhibition systems at the railway museum. In: Yamamoto, S. (ed.) HIMI 2013. LNCS, vol. 8018, pp. 238–247. Springer, Heidelberg (2013). doi:[10.1007/978-3-642-39226-9_27](https://doi.org/10.1007/978-3-642-39226-9_27)
22. Endo, T., Katayama, A., Tamura, H., Hirose, M., Tanikawa, T., Saito, M.: Image-based walk-through system for large-scale scenes. In: Proceedings of 4th International Conference on Virtual Systems and Multimedia (VSMM 1998), vol. 1, pp. 269–274 (1998)
23. Angelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J.: Google street view: capturing the world at street level. *Computer* **6**, 32–38 (2010)

24. Uyttendaele, M., Criminisi, A., Kang, S.B., Winder, S., Szeliski, R., Hartley, R.: Image-based interactive exploration of real-world environments. *IEEE Comput. Graph. Appl.* **24**(3), 52–63 (2004)
25. Jankowski, J., Hulin, T., Hachet, M.: A study of street-level navigation techniques in 3d digital cities on mobile touch devices. In: *Proceedings of 2014 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 35–38 (2014)
26. Banton, T., Stefanucci, J., Durgin, F., Fass, A., Proffitt, D.R.: The perception of walking speed in a virtual environment. *Presence* **14**(4), 394–406 (2005). MIT Press