

# AR-Technology-Based Locationing System for Interactive Content

Satoshi Saga<sup>(✉)</sup>, Ryota Oki, Shusuke Kawagoe, Wanjia Zheng,  
and Jiacheng Sun

University of Tsukuba, Tsukuba, Tennodai 1-1-1, Japan  
[saga@saga-lab.org](mailto:saga@saga-lab.org)

**Abstract.** In this research, we propose an interactive display system that has a simple implementation. The system is composed of a smartphone, PC, display, and cloud service; it realizes a CAVE-like immersive system by employing simple devices. Relative location information is acquired by the user's smartphone and a marker (or markers). Based on this information, the system renders the appropriate image in real time. In this paper, we describe the implementation and operation of the proposed system.

## 1 Introduction

In recent years, personal head-mounted displays (HMD) such as Oculus Rift have been distributed, and immersive virtual-world display systems, such as gaming environments, have drawn public attention. Not only HMDs but also immersive projection technology (IPT)-based displays such as the Cave Automatic Virtual Environment (CAVE) system have been developed as immersive displays. (The CAVE system is an immersive virtual reality environment where projectors are directed to several walls of a room-sized cube.) IPT display methods require displays or screens located around a user, whereas an HMD covers the user's eyes. Both systems interactively change the displayed image based on the user's movement and produce an immersive sensory experience.

However, both systems have several problems. IPT displaying methods require location sensors, multiple projectors, and multiple screens. These are expensive, lack portability and require a large space. Thus, the IPT display method is unsuited to personal use; its use is limited to research or large events. On the other hand, HMD realizes a personal immersive environment, and is now becoming more widespread. However, the system covers the user's eyes; thus, it cannot be used in everyday life. In addition, the display might induce virtual sickness owing to differences between the user's movement and the displayed image [1].

Here we propose novel but simple immersive display framework that employ smartphone, normal PC, LCD displays, and cloud service. In this study, we propose a simple, novel immersive display framework that employs a smartphone, a basic PC, an LCD display, and a cloud service. Using this framework, we create

simple interactive content and evaluate the functionality of the location information system. As shown in Fig. 1, a smartphone captures the displayed marker and sends the calculated relative location data to a location server in real time. An image server acquires the data and controls the displayed computer graphics. By controlling the displayed image based on location information regarding the user, the framework realizes a motion-parallax-based 3D display. By evaluating the latency of the system, we verify the functionality of the proposed framework and the capability of the framework to display interactive contents.

## 2 Related Research

In recent years, Oculus Rift has drawn attention as a personalized immersive display. The display has gyro, acceleration, and magnetic sensors, and estimates the orientation of the user. The information is used for generating interactive content for the user. Although the device can estimate the orientation of the user, location information cannot be estimated.

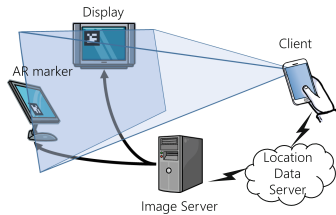
The CAVE system, the basis of IPT, was developed by Cruz, et al. [2]. The system has four screens surrounding the user; it displays immersive content to the screens using projectors. The user wears active-shuttered 3D glasses with a magnetic location sensor. The system estimates the location and orientation of the user from the sensor data and generates location-based binocular vision for the user. The system requires a magnetic location sensor and multiple projectors and screens. This setup is expensive, lacks portability, and requires a large space. Thus, the IPT display method is unsuited to personal use, and its use is limited to research or large events.

Another immersive system, called RoomAlive [3], displays wide immersive vision for a whole room, including the furniture. The system measures all of the shapes in the room and the location of the user and generates an appropriate displayed image. However, the system requires the setup of a camera and a projector system in the room, and measurement of the room beforehand.

Our proposed system estimates user location information and uses simple equipment (only a smartphone, a basic PC, an LCD display, and a cloud service). In the next section, we describe the proposed system.

## 3 Proposed System

In this paper, we propose a novel framework for obtaining more precise information in a simple manner in an urban environment where you can see many displays such as digital signage or large electronic billboards. Conventional augmented reality (AR) technology employs a homography matrix to calculate relative location information between a marker and a camera. This locating technology offers not only translation information but also rotation information. In addition, even if the marker is far away, once the camera captures it, the technology will derive the location information. Thus, we expand the technology to precisely measure the client's location.



**Fig. 1.** Dataflow of the proposed system



**Fig. 2.** A use case of the proposed system; He looks around a virtual work of art.

In this section, we explain the proposed system. A PC will display a marker on an LCD display, and a user will capture the displayed marker with a smartphone. The phone calculates relative location information and sends the data to a cloud-migrated location data server. With this system, we can create the basis of an indoor locating system. In addition, by employing the location information as feedback for displaying the content, we can easily construct interactive content.

### 3.1 System Flow

The system consists of a client terminal (a smartphone), a cloud-migrated location data server, a content server, and a display (Fig. 1). The client terminal acts as the viewpoint for the user; thus, the user holds the terminal in front of their face. AR markers are displayed on the displays. If the camera on the terminal captures a marker, the terminal calculates location information between the marker and the terminal. As the client device, we employ an android smartphone, and as the location server, we use the Heroku service. The location information is sent to the cloud server. To produce computer graphics, the image server acquires relative location information from the Heroku server. In this way, to construct interactive content, we can employ any client device and any image server that has a network connection. The AR marker does not have to be a dot pattern; it can be any natural image that does not interfere with the content itself. The interactive content works as follows: the users can see changes in the appearance of content according to their own movement. If the user moves left, the appearance changes to the image from the left point of view. If the user moves right, it changes to the image from the right point of view. Based on the viewpoint control of the user's location, a motion-parallax-based image is generated. Figure 2 shows a use case of the system.

## 4 Implementation

In this section, we describe the design of the implementation. This system requires fast data transfer to realize real-time responses to the user's movement. Thus, we chose WebSockets and designed a transfer protocol that is fast and compact. WebSockets provides bi-directional communication while keeping

the connection open. The contents of the communication are transfer matrices between client terminal and AR markers, and the identification number of the AR marker.

On the client terminal, the application captures the surrounding environment and searches for AR markers. The application specifies the identification numbers of the markers and calculates transfer matrices between client terminal and AR markers based on the captured markers. Next, the application converts the transfer matrices into binary data using MessagePack [4] and sends the message to the location server, which uses Heroku [5]. For the terminal client, we use a Nexus 7 and an application built with an AR library, Metaio [6].

Next, we prepare a location server. The purpose of the location server is to process and relay the message to the application server. If the server knows the absolute position of the markers, it can also calculate the absolute position of the terminal. By returning the absolute position information to the content server, the content server can provide different positioning services to the user. For the location server, we employ a cloud server, Heroku.

Next, the content server receives the location data for the terminal from the location server. The server parses the relayed message and acquires the transfer matrices. Based on these matrices, the PC renders virtual images. For the prototype system, we employ the gaming engine Unity [7].

## 5 Experiment

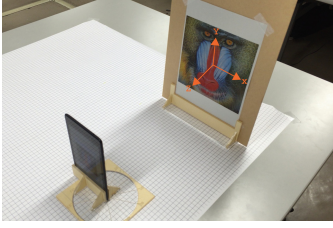
To evaluate the functionality of interactive content with the proposed framework, we performed several experiments. First, we evaluated communication and processing delay in terms of a frame rate. Second, we performed experiments to validate the preciseness of the estimated location.

### 5.1 Interactivity

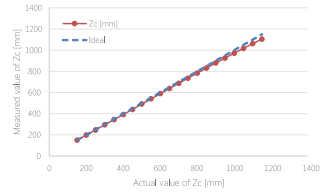
To evaluate the interactivity of the proposed system, we measured communication and processing delay in terms of a frame rate. During 3-second recordings of the transferred messages from the terminal to the content server, we acquired the frame rate of the system. Table 1 shows the measured results. Table 1 shows the measured results.

**Table 1.** Frame rate of the system

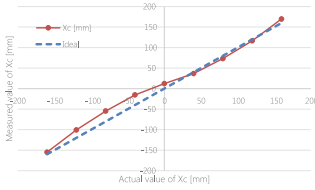
	Send	Receive
Average [ms]	56.6	57.6
Frame rate [fps]	17.7	17.4



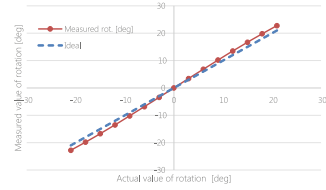
**Fig. 3.** Experimental setup



**Fig. 4.** Measurement result of translation ( $Z$ -axis)



**Fig. 5.** Measurement result of translation ( $X$ -axis)



**Fig. 6.** Measurement result of rotation

## 5.2 Preciseness

To evaluate the preciseness of the location measurement, we measured translation and rotation data and compared them to the actual values. Figure 3 shows the experimental setup.

The center of the AR marker is the origin of the system; the normal direction of the marker is the  $z$  axis. The client terminal is placed in front of the AR marker. We set the client position  $X_c$  to be 0 mm and moved the client position  $Z_c$ , from 150 mm to 1150 mm at intervals of 50 mm and measured the transfer matrices. Figure 4 shows the results. Next, we set the client position  $Z_c$  to be 500 mm and moved  $X_c$  from  $-160$  mm to 160 mm at intervals of 20 mm and measured the transfer matrices. Figure 5 shows the results. Third, we set the client at  $Z_c = 500$  mm and rotated it around the origin by  $\pm 16^\circ$ . Figure 6 shows the results.

## 5.3 Results

From the interactivity experiment results, the frame rate of the proposed system is around 17 frames per second. This is not high enough for a fully interactive system; however, some interactivity can be achieved. From the preciseness experiment result, the translation results show linear behavior that is very close to the actual values. In the  $Z$ -axis results, the maximum error rate is 17%. On the other hand, the  $X$ -axis results show a maximum error rate of approximately 10%. In the results for rotation, the error remains below approximately  $3^\circ$ .

## 6 Conclusion

In this research, we propose an interactive display system that makes a user feel as though they are gazing at a virtual space outside a window. The system is composed of simple devices, including a smartphone, a PC, a display, and cloud services; thus, it realizes an immersive CAVE-like system using only simple devices. Relative location information is acquired by the user's smartphone and a marker (or markers). Based on this information, the system renders the appropriate image in real time. From the experimental results, the system can display roughly 17 frames per second. In addition, the error rate of measurements is kept small. By making the server a cloud service, we can use this location service nearly everywhere. For example, we can use the service in an urban environment where you can see many displays such as digital signage or large electronic billboards indoors (e.g., in underground shopping centers, buildings, or shopping malls). Thus, we can expand the technology and measure client locations precisely, not only for interactions between real and virtual worlds, but also between multiple locations in the real world.

Furthermore, several precise vision-based local positioning technologies such as PTAM [8] or Smart AR [9] have been developed. These technologies do not employ recorded markers, but instead, use general objects to locate the client. With these technologies, our proposed system can be made more precise and immersive. In future work, we plan to incorporate these technologies with our system and propose more powerful applications.

## References

1. Davis, S., Nesbitt, K., Nalivaiko, E.: A systematic review of cybersickness. In: *Proceedings of the 2014 Conference on Interactive Entertainment*, pp. 1–9. ACM (2014)
2. Cruz-Neira, C., Sandin, D.J., DeFanti, T.A.: Surround-screen projection-based virtual reality: the design and implementation of the cave. In: *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 135–142. ACM (1993)
3. Jones, B., Sodhi, R., Murdock, M., Mehra, R., Benko, H., Wilson, A., Ofek, E., MacIntyre, B., Raghuvanshi, N., Shapira, L.: Roomalive: magical experiences enabled by scalable, adaptive projector-camera units. In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, pp. 637–644. ACM (2014)
4. Furuhashi, S.: Messagepack. [msgpack.org](http://msgpack.org), 1, 2013
5. Middleton, N., Schneeman, R. et al.: *Heroku: Up and Running*. O'Reilly Media, Inc., California (2013)
6. Metaio GmbH. Metaio. <http://www.metaio.com/>. Accessed 20 March 2015
7. Unity Technologies. Unity. <http://unity3d.com>. Accessed 20 March 2015
8. Klein, G., Murray, D.: Parallel tracking and mapping for small ar workspaces. In: *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. ISMAR 2007*, pp. 225–234. IEEE (2007)
9. Sony Inc. Smart AR. <http://www.sony.co.jp/SonyInfo/News/Press/201105/11-058/>. Accessed 20 March 2015