Inside-out Tracking Controller for VR/AR HMD using Image Recognition with Smartphones

Keisuke Hattori
Komazawa University
3719101k@komazawa-u.ac.jp

Tatsunori Hirai
Komazawa University
thirai@komazawa-u.ac.jp

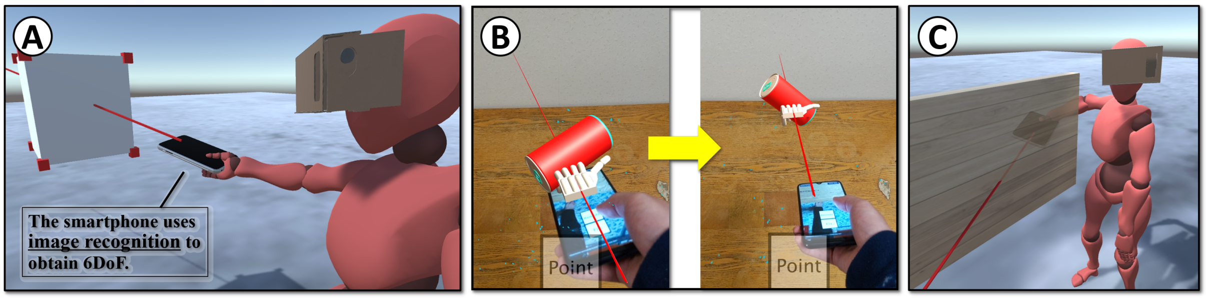

Figure 1: Images of using a smartphone with inside-out tracking as a controller for an HMD. (A)The controller is used to point to an object. (B)An image that uses raycasting to grab a distant object. (C)An image of raycasting with the controller in an occluded spot.

ABSTRACT

To smoothly operate an HMD controller in the 3D space, a variety of operation methods that are linked to the real hand movements have been proposed. Many controller recognition methods are based on outside-in tracking. Therefore, those controllers have the disadvantage that it cannot be operated if it moves into the occluded area from the sensor. In this paper, we propose an inside-out controller for HMD using a smartphone that can estimate its own position and the angle via image recognition. This is a system that transmits the controller's 6DoF(six degrees of freedom) information to the HMD and performs raycasting based on the 6DoF information. Wherever the controller is located, the user can operate it. This method covers the disadvantages of existing HMD controllers.

CCS CONCEPTS

• **Computing Human-centered computing → Mixed / augmented reality**.

KEYWORDS

AR,VR,controllor,inside-out tracking

ACM Reference Format:
Keisuke Hattori and Tatsunori Hirai. 2020. Inside-out Tracking Controller for VR/AR HMD using Image Recognition with Smartphones. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '20 Posters), August 17, 2020.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388770.3407430

ACM ISBN 978-1-4503-7973-1/20/08.
https://doi.org/10.1145/3388770.3407430

1 INTRODUCTION

Many engineers and investment organizations predict that the VR/AR industry could reach a large amount of revenue within a few years, and among them, HMD type devices are attracting attention. The HMD can be attached to the head thus allowing the user's hands to be hands-free. Therefore, the user has to hold the controller in both hands or use hand gestures to operate it. The controller used in many HMDs detects 6DoF by an outside-in method using sensors attached to the HMD or pre-installed sensors, so it can be operated intuitively in conjunction with hand movements. However, if the controller moves out of the sensor's recognition range, it cannot recognize the position. Therefore, existing methods are limited in their scope of operation.

An inside-out tracking controller for HMDs, which can detect 6DoF in order not to limit the operating range, has been studied. GyroWand[Hincapié-Ramos et al. 2016] is a study to detect the 6DoF of the controller and perform raycasting using the IMU. However, because of the accumulation of misalignments in the IMU, it is difficult to operate the system intuitively in conjunction with hand movements. The TrackCap[Mohr-Ziak et al. 2019] is an inside-out tracking controller for HMDs that uses a smartphone's camera to track images. It uses a smartphone's front camera to track the image printed on a user's cap flange and detects the controller's 6DoF. However, if the smartphone is moved to a position where the image cannot be tracked, it cannot be operated.

Therefore, we propose an inside-out tracking controller for HMDs that acquires ambient information and estimates its own position using image recognition from a smartphone. Specifically, the smartphone used as the controller of the HMD estimates its own position and transmits the 6DoF information to the HMD. The user can then raycast based on the 6DoF information sent by the user. This system does not require the installation of other sensors, since the controller itself does the self-position estimation. It also features the ability to operate the controller without limiting its operating range.

2 RAYCASTING SYSTEM

In this paper, we propose a controller for VR/AR HMD that uses the inside-out method to perform raycasting. Specifically, the system uses image recognition with a smartphone to estimate its own position and performs raycasting via network communication. The image recognition method in this paper is based on "Plane Finding", a horizontal detection system in ARCore, which is an SDK provided by Google. Then, the controller sends the 6DoF information obtained from the self-position estimation using image recognition to the HMD. The controller's 6DoF synchronization uses an ARCore feature called "Cloud Anchor".

We have developed a controller using our proposed system that can operate a general VR/AR controller. The general operating functions in this paper are defined as follows.

- Grabbing: can grab an object and move it.
- Rotate: can be rotated to any angle.
- Remote control: objects placed far away can be selected using raycasting.

In this paper, we used two smartphones with a self-location estimation function. It adopted a smartphone as a device with the same functions as a typical inside-out tracking HMD. In other words, one unit was used as an HMD and the other as a controller. The following describes the 6DoF synchronization method of this method. First, a room is created from the communication standby screen provided in the application of the device used as a controller. This is a necessary procedure to create a virtual space for 6DoF synchronization and to let any space be shared. Then, a camera attached to the device is used to detect the surrounding conditions and estimate the device's self-position. Then, the user registers an anchor centered on an arbitrary coordinate and shares the space. Next, the user joins the created room on the communication standby screen on the device used as a display. In the same way as the previous step, user also perform self-positioning of the device used as a display. It is possible to synchronize the 6DoF of both devices by pointing the camera at the coordinates of a pre-registered anchor. The device to be used as a display is raycasting based on the controller's 6DoF information. In addition, a CG imitating a hand is displayed at the location of the controller to make the raycaster visually easier to understand. Then, the 6DoF synchronization of the controller's devices is performed 30 times per second, and the displayed CG appear to follow the controller smoothly.

In addition, a function for 6DoF correction is provided to cover the gap between the actual 6DoF of the synchronized controller and the user's operation position. In the application, it is provided as a "Fix button" to correct the generated error to the user's arbitrary position. As long as the "Fix button" is pressed, the CG objects displayed on the display device are fixed in the AR space and are not followed by the controller. The user can adjust the 6DoF of the controller to the desired location by piling the controller on the fixed position of the CG and releasing the "Fix button".

We introduce an example of VR/AR contents that are difficult to realize without this system and some examples of VR/AR contents that can be used with this system. In the proposed system, the control device is recognized by the inside-out tracking, therefore there is no blockage spot and the operation can be done within an unlimited range, as shown in Fig. 1.(C). In other words, the major advantage of this system is that it can be operated even if there is a wall between the device and the controller. In other words, the ability to operate even if there is a wall between the device and the controller is a major advantage. Then, we introduce a 3D modeling application as an example of VR/AR contents. While general modeling tools are operated by mouse or keyboard shortcut keys, 3D modeling tools in VR/AR space can be operated from the controller position obtained from coordinates, and can be modeled intuitively according to hand movements.

3 USER STUDY

A user study was performed to measure the effectiveness of the controller of the proposed method. The validity of this evaluation is that it can perform the general operating functions of the VR/AR controller as defined in Chapter 2. To evaluate the effectiveness of the HMD as a controller by providing simple tasks that require the defined functions. In this evaluation experiment, we asked four people to give the following evaluations.

Subjects were asked to complete the task of moving a specific object to a specific location. This task involves using a controller to move 15 randomly arranged cylindrical objects to the position of a particular object. This task requires grabbing, rotating, and selecting away objects defined in Chapter 2.

When assessed one by one for each of the four subjects, all subjects were able to achieve the tasks in Chapter 2. The time taken to complete the task for individual subjects was 95 s for subject A, 140 s for subject B, 100 s for subject C, and 140 s for subject D. The results of the evaluation show that the proposed method is effective as a controller for HMDs as defined in this paper. One of the opinions of the subjects was that the position of the controller as seen from the display was out of alignment with the actual position, making it difficult to operate. However, this problem can be solved if the user is able to use the fix button.

4 CONCLUSION

In this paper, we propose a VR/AR raycasting system using image recognition on a smartphone. This system allows the user to perform raycasting even when the HMD controller is occluded. In the future, we would like to develop a system for self-position recognition using LiDAR, the latest image recognition technology, and ARCore's new Depth API. If these techniques are also used, users can perform faster operations because they have higher performance than the image recognition techniques used in this paper.

REFERENCES

J. D. Hincapié-Ramos, K. Özacar, P. P. Irani, and Y. Kitamura. 2016. GyroWand: An Approach to IMU-Based Raycasting for Augmented Reality. *IEEE Computer Graphics and Applications* 36, 2 (2016), 90–96.

Peter Mohr-Ziak, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. 2019. TrackCap: Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*.