Collaborative manipulation of 3D virtual objects in augmented reality scenarios using mobile devices

DEEPAK SRIVATSAV, 2016030
KANAV BHAGAT, 2016046

ACM Reference Format:
Deepak Srivatsav and Kanav Bhagat. 2019. Collaborative manipulation of 3D virtual objects in augmented reality scenarios using mobile devices. 1, 1 (April 2019), 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 ABSTRACT

This project aims to create an application that allows multiple users to connect to a server and experience and manipulate an Augmented Reality Environment at the same time. This is achieved using Unity 3D, Vuforia, and UNet for Networking. We also tackle the issue of the camera going out of focus while using the AR Camera of Unity. By default, the camera does not auto focus, due to which the user might see a blurred image of the marker. To solve this issue, we wrote a small module that would enable the Vuforia AR Camera to Autofocus. Although this slows down detection by some small amount of time, it greatly helps with the detection and tracking accuracy.

2 LITERATURE SURVEY

The usage of mobile devices for the purpose of interaction with video and augmented reality environments has been studied in detail.
Sebastian Boring, in his paper [1] *Touch Projector: Mobile Interaction Through Video*, talks about the usage of a mobile phone to manipulate content on distant screens, such as a wall, or on table. Users would point their devices at these screens and then interact with them. They model the tracking purely based on the targetś display.
We primarily follow the work of Grandi et al. [2] They proposed an environment where multiple users can connect to a server and simultaneously manipulate 3D objects. Users can use touch to select, scale, translate and rotate a 3D object. The users can also group objects together and transform them simultaneously. This enables users to work together and perform tasks such as arranging a living room or scenery. For the purpose of Networking, the Unity Network Environment UNET was used, along with the Vuforia SDK for tracking of the objects. The authors use multiple markers to expand the range of the AR environment.
Object manipulation was done with the help of buttons on the phone's touchscreen. The user could pick one at a time (translation/rotation etc).
In another research publication by the same authors [3], they talk about the usage of mobile phones in Collaborative 3D manipulation. They talk about the ubiquity of mobile phones and the wide range of inertial sensors available on smart phones.

Authors' addresses: Deepak Srivatsav, 2016030; Kanav Bhagat, 2016046.

XXXX-XXXX/2019/4-ART $15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2.1 Algorithm

We use Unity 3D to create the application, and Vuforia to generate the markers. Given a marker which can spawn an object, a user can use touch on the mobile phone to do the following -

(1) Select an item. This is done with raycasting. When the user touches the screen at any position, we do a raycast and try and detect whether the ray hits any collider object. If the ray hits a collider object, we mark the object as currently selected by changing its color. If the user casts a ray towards the object once again, the object will be deselected.
(2) Rotate an object. An object that has been selected can be rotated by choosing the rotate option at the bottom of the screen. Once an object has been selected and the rotate button pressed, the user can use one finger to change the rotation of the object.
(3) Translate an object. An object that has been selected can be translated by choosing the translate option at the bottom of the screen. The user can use one finger to translate along the x axis and two fingers to translate along the z axis.
(4) Scale an object. An object that has been selected can be scaled by choosing the scale option at the bottom of the screen. The user must use two fingers in a pinching motion to scale the object.
(5) Move along with Camera. An object that has been selected can be set to move along with the camera by choosing the hold option present at the bottom of the screen. Once this has been selected, the user must simply move the camera, and the object will move along with the camera, provided the marker is still in view.
(6) The synchronization is achieved using Unity's UNET system. One mobile phone will act as a host while the rest will connect as clients. If the host disconnects, the session will end.
(7) In the paper we refer to, the authors allow users to manipulate the same object. In Unity Networks, there is a concept of *local authority*, that is used to preserve security in the environment. We believe by arbitrarily allowing users to manipulate the same object does not provide an optimal environment. Instead, we allow each user to spawn as many objects as they would like to. A user will have local authority over whatever object they spawn, and will hence be able to select them, group them, and perform transformations on them. We do not allow a user to access an object who he/she does not have local authority over.

3 PROPOSED MILESTONES / COMPLETION

(1) Basic network structure, markers, and translations on individual objects - **midterm goal** - **complete**.
(2) Usage of Camera Gaze to move objects - **midterm goal** - **complete**
(3) Multiple users interact and modify objects in the scene - **final goal** - **complete**. - In this goal, instead of having simultaneous manipulation of the same object, we introduced the idea of a user spawning objects on their own, and manipulating their own set of objects.
(4) Grouping of objects and applying the same transformation on all objects in the group - **final goal** - **complete**
(5) Gesture based translations - **final goal** - **complete** - Completed in mid term.

4 LINKS

Videos of our application
Github
Apk file
Paper Implementation

Fig. 1. Basic Layout of the Application. The Network API is located on the top left corner. The user can either host a session or connect to one from here. On the bottom right, four buttons - for camera gaze based movement, translation, rotation and scaling are present. There is also a button for spawning.

Fig. 2. An example with Multiple Objects (Single User)

REFERENCES

[1] Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA,

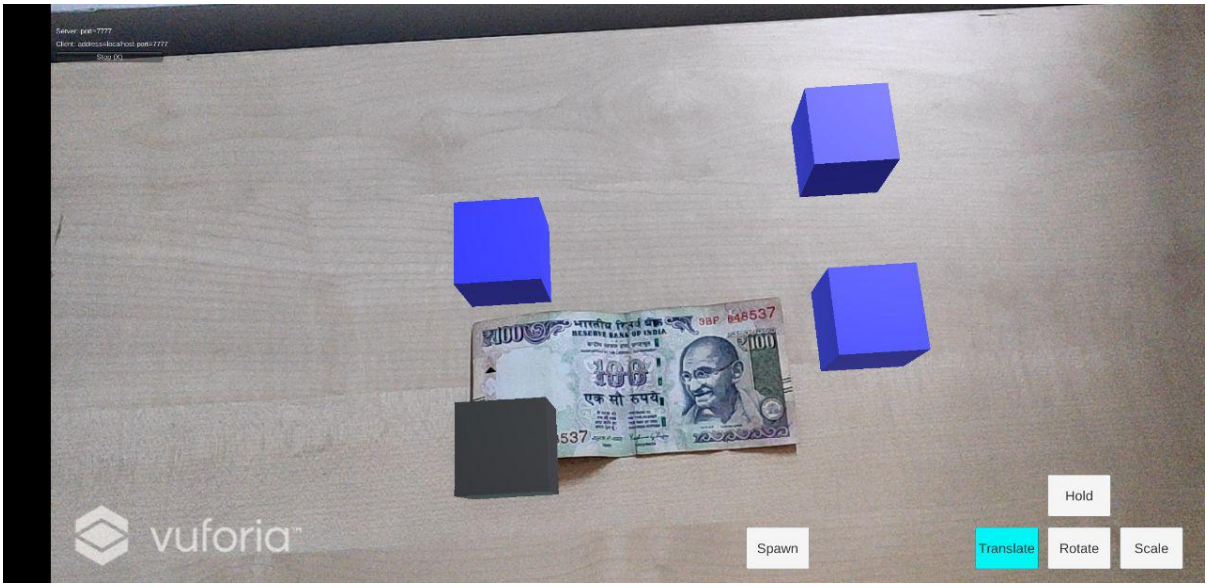

Fig. 3. Selecting objects and Grouping them - Color changes to blue

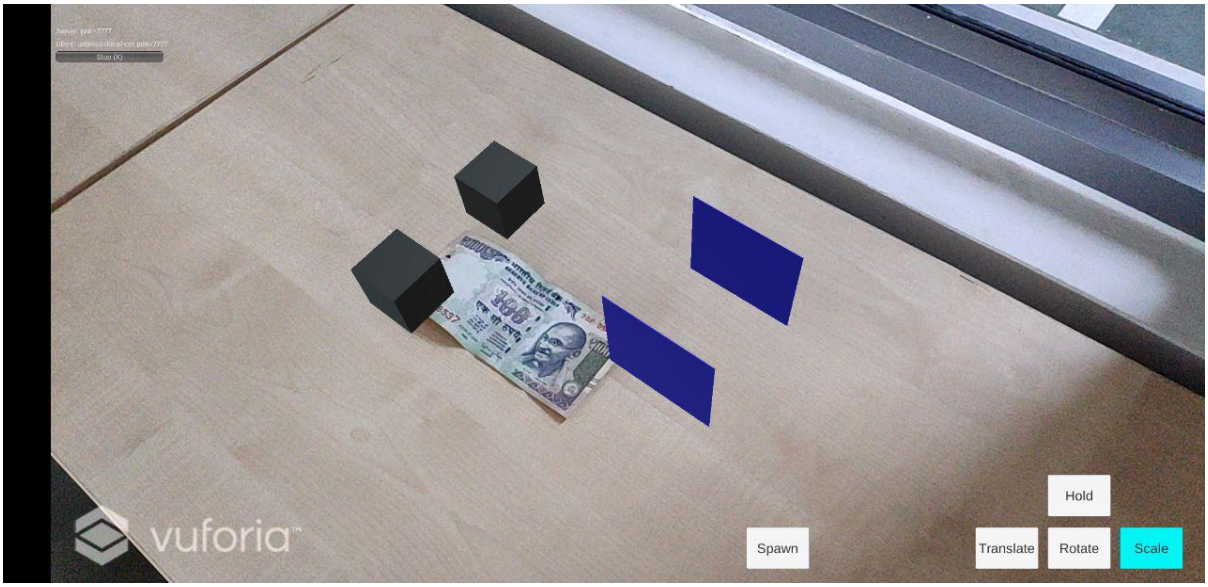

Fig. 4. Viewing the environment from a different point of view, while the marker is present at the same location. A scaled version of grouped objects is also shown

2010. ACM.

[2] JerÃťnimo Grandi, Iago Berndt, Henrique Debarba, Luciana Nedel, and Anderson Maciel. Collaborative manipulation of 3d virtual objects in augmented reality scenarios using mobile devices. pages 264–265, 01 2017.

[3] JerÃťnimo Grandi, Iago Berndt, Henrique Debarba, Luciana Nedel, and Anderson Maciel. Collaborative 3d manipulation using mobile phones. pages 279–280, 03 2016.