Collaborative Manipulation of 3D Virtual Objects in Augmented Reality Scenarios using Mobile Devices

Jerônimo G Grandi*[1] Iago Berndt†[1] Henrique G Debarba‡[2] Luciana Nedel§[1] Anderson Maciel¶[1]

[1] Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[2] Artanim Foundation, Geneva, Switzerland

Abstract

Interaction in augmented reality environments may be very complex, depending on the degrees of freedom (DOFs) required for the task. In this work we present a 3D user interface for collaborative manipulation of virtual objects in augmented reality (AR) environments. It maps position – acquired with a camera and fiducial markers – and touchscreen input of a handheld device into gestures to select, move, rotate and scale virtual objects. As these transformations require the control of multiple DOFs, collaboration is proposed as a solution to coordinate the modification of each and all the available DOFs. Users are free to decide their own manipulation roles. All virtual elements are displayed directly on the mobile device as an overlay of the camera capture, providing an individual point of view of the AR environment to each user.

Index Terms: H.5.2. [Information Interfaces and Presentation]: User Interfaces—Input devices and strategies; H.5.3. [Information Interfaces and Presentation]: Group and Organization Interfaces—Computer-supported cooperative work

1 Introduction

Augmented reality experiences are still depending of the use of see-through personal displays – that should be comfortable and, ideally, wireless – to allow the free movement of the user in the real world. Current mobile phones and tablets seem to be the ideal device to fulfill AR requirements for everyday applications. They have sensors to capture touch, movement, and image and then, it is possible to use computer vision and the integration of acceleration sensors to define the pose of the device in space. The real world captured by the camera embedded on the mobile device is overlaid by virtual elements and rendered to the user's mobile device screen. Interaction can be provided by gestures on the device touchscreen.

The manipulation of three-dimensional objects in AR environments is a complex task that requires the control of multiple DOFs for selecting, translating, rotating, and scaling objects. One alternative toward efficiency and usability is to share the task and solve it in a collaborative manner. Collaborating users may wish to split the different manipulation aspects among them. For instance, while one user translates an object, a second user could rotate it, or control the distance of this object relative to the other user point of view.

The use of the mobile device touchscreen for 3D interaction in mobile AR has been explored to some extent. Notably, Boring et al. used the built-in camera of a mobile phone to directly interact with a distributed computing environment, performing tasks such as selection, positioning and transferring of photos across different displays [1]. In a previous work, our group has proposed an AR interface for the exploration and annotation of medical images in handheld devices [3]. Nonetheless, there is also literature on the mapping of mobile device's sensors and gestures for the selection and manipulation of virtual objects for single users [2], collaborative [4] contexts and in AR environments [5]. However, from our knowledge, none of the previous works approached the problem of collaborative manipulation in mobile AR.

*e-mail: jggrandi@inf.ufrgs.br
†e-mail: iago.berndt@inf.ufrgs.br
‡e-mail: henrique.debarba@artanim.ch
§e-mail: nedel@inf.ufrgs.br
¶e-mail: amaciel@inf.ufrgs.br

2017 IEEE Symposium on 3D User Interfaces (3DUI)
978-1-5090-6716-9/17/$31.00

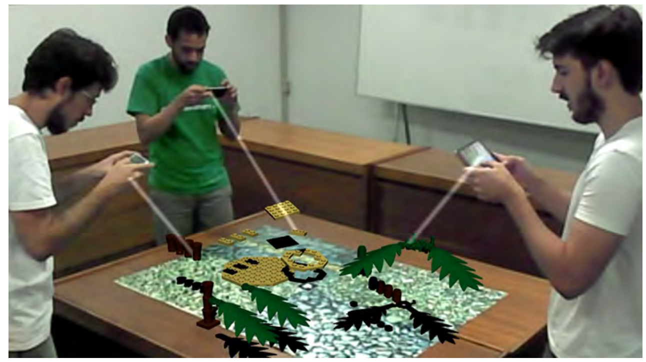

Figure 1: Overview of three persons using our mobile interface to collaboratively manipulate objects in an augmented reality environment.

Our proposal relies on the mobile device position, orientation and touchscreen to select objects and perform the three basic transformations: translation, rotation, and scale. The actions required to perform these actions are already consolidated in the everyday tasks when interacting with mobile apps and games, such as: touch and slide to translate and rotate, pinch and spread to scale the virtual object. Besides the touchscreen gestures, it is possible to translate and rotate the augmented virtual objects by attaching them relative to the pose of the device. In this mode, the user may use the position and orientation of the device to precisely place objects in the scene. Additionally, our approach accepts the connection of multiple mobile devices at the same time, allowing users to collaborate by simultaneously controlling different objects, or different DOFs of the same object. This provides equal participation, individual responsibility and positive interdependence, which are desirable features in a collaborative task.

2 Materials and Methods

Our setup is composed of mobile devices with rear camera, touchscreen and WiFi connection, and a table with image patterns for tracking. The software was developed with the Unity game engine, and the communication between peers is done using Unity Network (UNET). In our implementation, any device can act as a Host (Server + Client), to which the other devices will connect as Clients. This avoids the need for a dedicated server to manage the network. The

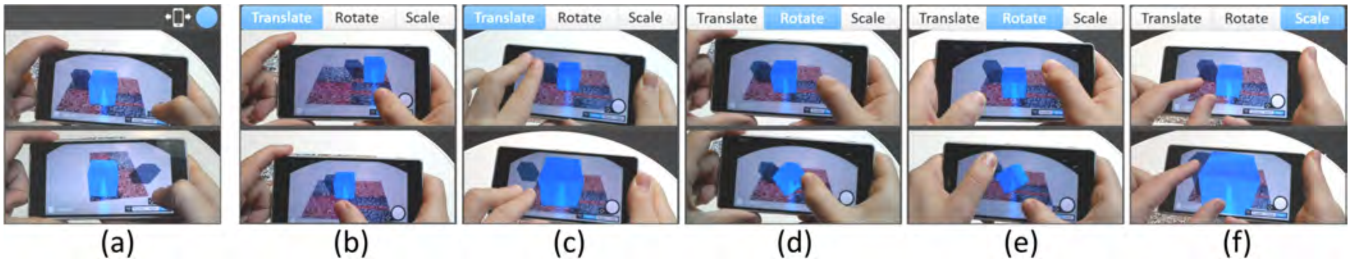

Figure 2: Walkthrough of input to output mapping of the mobile device to the selected object: by holding the circular button, the object gets attached to the phone, and the movement of the phone is transferred to the object (a); in the translation mode, by sliding one finger over the touchscreen the object translates toward that direction (b), by sliding two fingers along the vertical axis of the screen it moves toward/against the phone (c); in the rotation mode, by sliding one finger the object rotates around the axis perpendicular to the slide vector (d), by twisting two fingers it rotates around the axis defined by the screen plane (e); uniform scale is applied by touching the screen with two fingers and producing a pinch/spread gesture while in the scaling mode (f).

rear camera and the Vuforia SDK were used to track the mobile device physical pose relative to the fiducial markers. We used multiple markers to extend the interaction range. These markers were placed on top of a table, as seen in Figure 1.

The following subsections present the techniques proposed for interaction.

2.1 Selection and Grouping

Selection is performed by touching the desired virtual object on the touchscreen. A single tap selects only one object. It is also possible to select more objects at the same time. Having selected an object, the user can touch and hold on a second object to start the multiple selection mode. Once in this mode, single touches select or deselect objects. A touch on an empty area of the touchscreen deselect all objects.

Grouping is used to save a multiple selection. The user can make a group by pressing the "group" button in the lower left corner of the screen. This button only appears if more than one object is currently selected. A selected group can be disassembled by pressing the same button, which now contains the text "ungroup". A group obeys the same interaction rules as a single object. So, it is possible to select multiple groups and merge them into a new group and edit a group by touching and holding an object to add or remove them from the group.

2.2 Manipulation

Manipulation can be performed when an object or a group is selected. We implemented two modes to transform the virtual objects. These modes can be combined to enhance precision. The first mode attaches the object to the physical pose of the device. Thus, the object translates and rotates with the device while keeping an invariant rigid transformation relative to it (Figure 2a). This mode is activated by pressing and holding the circular button in the lower-right corner of the interface, and is halted once the button is released.

The second mode uses touch gestures on the touchscreen to translate, rotate and scale objects. The desired transformation is selected through buttons on the interface, and a user can only have one of them active at a time. All transformations are applied relative to the touchscreen plane orientation (i.e. the device orientation). The transformations can be applied using one and two fingers gestures. The touch and slide of one finger is used for translation and rotation. Translation moves the object along the plane defined by the screen, toward the same direction as the touch. Rotation rotates the object around the axis perpendicular to the touch movement (Figure 2bd). In translation mode, touching and sliding two fingers along the vertical axis of the touchscreen moves the object closer or farther away from the camera (Figure 2c). In rotation mode, touching and rotating with two fingers around each other rotates the object around the axis perpendicular to the screen plane (Figure 2e). Finally, in Figure 2f, uniform scale is performed using pinch and spread gestures with two fingers.

Non-geometrical manipulation is also possible. If a particle system is selected, for instance, it is possible to change its properties of life time, rate and size using the aforementioned gestures respectively. We display level bars for prompt feedback on the parameter changes.

2.3 Simultaneous Manipulation

Our technique supports simultaneous manipulation of virtual objects by multiple users. Users can work collaboratively. They can either manipulate and group separate objects to combine later, or interact with the same object at the same time. Contributions are added.

Moreover, we draw rays from the device location to the currently selected objects. This is useful to inform about the current focus of interest of other users interacting in the same AR environment.

3 Final Comments

We presented a smartphone-based collaborative user interface for 3D object manipulation in AR environments. Smartphones are omnipresent and people are familiar with touch and orientation gestures. This is expected to provide users with a very high affordance. Furthermore, all transformations can be performed by an unlimited number of participants from different viewpoints at the same time. This potentializes the accuracy of collaborative manipulation.

Regarding tracking and visualization, we use fiducial markers to provide a common visualization frame in the real world for the virtual objects and the mobile devices. The setup does not require additional hardware. Any person equipped with an ordinary smartphone is able to connect to the application and interact with it.

References

[1] S. Boring, D. Baur, A. Butz, S. Gustafson, and P. Baudisch. Touch projector: mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2287–2296. ACM, 2010.

[2] H. Debarba, L. Nedel, and A. Maciel. Lop-cursor: Fast and precise interaction with tiled displays using one hand and levels of precision. In *2012 IEEE Symposium on 3D User Interfaces*, pages 125–132. IEEE, 2012.

[3] H. G. Debarba, J. Grandi, A. Maciel, and D. Zanchet. Anatomic hepatectomy planning through mobile display visualization and interaction. In *Medicine Meets Virtual Reality*, pages 111–115, 2012.

[4] J. G. Grandi, I. Berndt, H. G. Debarba, L. Nedel, and A. Maciel. Collaborative 3d manipulation using mobile phones. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 279–280, March 2016.

[5] T. Tanikawa, H. Uzuka, T. Narumi, and M. Hirose. Integrated view-input interaction method for mobile AR. *2015 IEEE Symposium on 3D User Interfaces, 3DUI 2015 - Proceedings*, pages 187–188, 2015.