Smartphone-based Remote 3D Interaction for Digital Heritage Applications

Alejandro Rodríguez
University of Granada
alejandrora@ugr.es

Alejandro León
University of Granada

***Abstract*—In this paper, we explore the use of smartphones as remote 3D interaction devices for digital heritage applications. We present a set of single-handed interaction techniques by combining the modern 3D sensing capabilities of these devices with touch-screen inputs, covering the basic 3D manipulation tasks in an intuitive and ergonomic way.**

A 3D gallery prototype has been developed, and all the proposed techniques have been integrated in an Android application to remotely interact with the gallery through a normal WiFi connection.

***Index Terms*—3D interaction, smartphone, interactive graphics**

I. Introduction

Heritage is increasingly being transfigured through digitization devices into an exciting new representation of information. This way, the vast information of an environment or heritage object can be restructured and recomposed into a plethora of representations both textual and visual. Once digitized and elaborated, information associated with heritage objects and environments can be easily manipulated and these virtual representations raise enormously their position in accessibility. As much of the generated content include three dimensional data, 3D reconstructions and virtual environments are great tools to expose such content. Some examples of these tools can be found in the work of Webel et al. [1], where low-cost virtual reality setups are explored to experience cultural heritage artifacts, or the work of Manferdini et al. [2] where a web-based visualization tool is proposed in an attempt to reach more users. A very practical view on the creation of multimedia cultural heritage content is presented by Beraldin et al. [3], showing three real cases. As any 3D environment, a set of 3D interaction techniques is required to enable an intuitive manipulation and exploration of the content presented to the users.

Nowadays, smartphones include a set of built-in sensors that provide 3D capabilities to these devices. These 3D sensing capabilities, together with the touchscreen input, a very high compute power, and its widespread use, have made the smartphone an obvious candidate to visualize and interact with 3D content.

In this paper, we propose a set of single-handed 3D interaction techniques for smartphones to provide an ergonomic and intuitive tool to remotely manipulate 3D content. These techniques, together with the possibility of displaying custom application control interfaces, enable the smartphone as a generic remote 3D interaction device, which we believe will increase the ability of the users to interact with digital content in live environments, e.g., the visitors of a museum or an exhibition. Since the proposed techniques do not require external device tracking, an easy set up is possible through the use of a wireless connection between the user's smartphones and the remote application, which can be visualized using any standard screen. We present a remotely controlled 3D gallery of digitized heritage models as a case study.

II. Related work

Over the last years, the 3D capabilities of mobile devices have been widely studied. Different fields related to 3D content have used mobile devices for input and/or output tasks. Mossel et al. [4] proposed different manipulation techniques to handle Augmented Reality applications for mobile devices. Medeiros et al. [5] proposed the tablet as a tracked, 3D interaction tool for immersive engineering environments. Davies et al. [6] proposed an implementation of a Cross Reality system for virtual cultural heritage.

Accounting for the multi-touch input capabilities present in such devices, many existing interaction techniques have been adapted for its use with mobile devices, and also new techniques have been developed for such devices. Steinicke et al. [7] discussed the use of mobile multi-touch devices for 3D interaction tasks. Telkenaroglu and Caping. [8] revisited existing techniques for touch-based 3D interaction and proposed a set of new techniques for mobile devices.

Regarding the 3D built-in sensors, several authors proposed to use them for performing remote interaction. De Souza et al. [9] used the accelerometer data to remotely interact with 3D medical data. Katzakis et al. [10] proposed the smartphone as a game controller, providing a 3D cursor for in-game interaction.

While a combination of an accelerometer and a compass is enough to provide a drift-free tracking of the current orientation of the device, some modern smartphones include a gyroscope sensor which can be used to improve the quality of the tracking through the use of sensor-fusion techniques (the reader is referred to the work of Ayub et al. [11] as an example).

In this work, we track the current device orientation through sensor-fusion, and combine it with single-handed touchscreen input in order to provide a set of intuitive and ergonomic 3D interaction techniques, precise and general enough to perform a natural and fast exploration of digital heritage 3D data without the necessity of using an external tracking infrastructure.

III. Single-handed 3D interaction techniques

The basics of any 3D user interface were established, and accepted as the *de facto* standard, by Bowman et al. [12]. In order to achieve an ergonomic interaction, our interaction techniques cover all the canonical 3D inspection tasks, and they can be performed single-handed while holding the device in its natural holding position.

A. Calibration of the device orientation

As already mentioned, the sensors present in modern smartphones provide enough information to perform a drift-free, precise tracking of the current device orientation. In our implementation, we obtain a rotation matrix storing the current device orientation from the sensor-fusion algorithm implemented in the Android API. However, the rest orientation of the device, i.e., the pose for which the rotation matrix is equal to the identity matrix, does not necessarily match the orientation of the screen where the 3D content is displayed.

In order to fix this problem without the need to implement external tracking of the device, a calibration step must be performed at the beginning of a remote interaction session. First, we obtain the rotation matrix R_0 of the device while the user holds it parallel to the remote screen. A correction matrix is then calculated as the inverse of R_0, and during the whole interaction session, it will be composed with the current rotation matrix into a calibrated rotation matrix:

$$R_{calib} = R_{current} * R_0^{-1} \tag{1}$$

Thereafter, the rest orientation will match the natural, parallel orientation with respect to the screen, and the touchscreen inputs composed with R_{calib} will induce the expected 3D behavior.

B. Translation

To perform 3D translation of a virtual object, we simply extend the most intuitive 2D translation metaphor to the three dimensions. When the user's finger moves over the screen, a 2D vector is generated using the x and y finger displacements. This vector is then transformed into a 3D vector (with a zero z coordinate) and rotated by composing it with the calibrated rotation matrix:

$$V_{trans} = R_{calib} * (x, y, 0)^t \tag{2}$$

This 3D vector applied as a translation to the remote object will replicate the 3D movement of the finger in the real world onto the virtual object. Fig. 1 shows three interactions where the same touch input produces different 3D translations in the remote application.

C. Scaling

While an intuitive non-uniform 3D scaling metaphor can be challenging to obtain, since the main objective of scaling during inspection is either to zoom in and visualize details on the virtual objects or zoom out to obtain a more general view of them, we limit the scaling to uniform scaling.

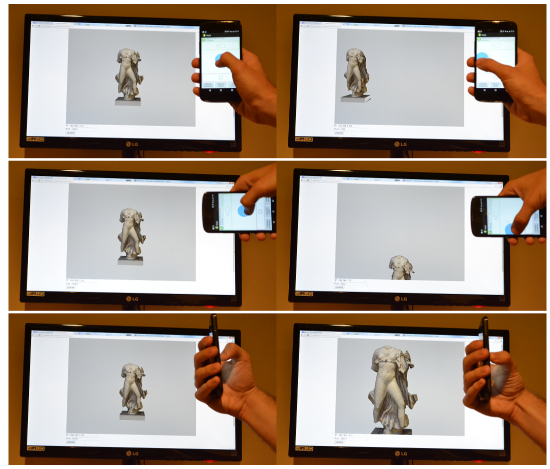

Fig. 1. Each row shows the 3D translation produced by the same touch input when holding the device with different orientations. Our translation operation applies the 3D movement of the finger in the real world to the remote virtual object.

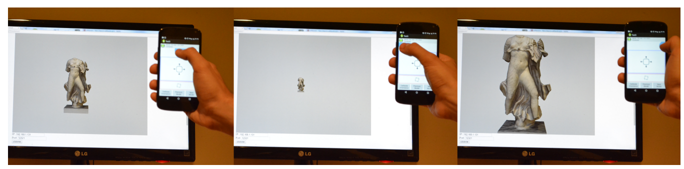

Fig. 2. Remote uniform scaling applied to the virtual object through a slider-type interaction.

This constraint allows us to perform a simple scaling just by converting the variation in the x coordinate of the in-screen position of the finger into a scaling factor. A negative variation of the x coordinate will produce a scaling factor $0 < s < 1$, and a positive variation will produce a scaling factor $s > 1$. Fig. 2 shows an example of this operation.

D. Rotation

Rotation is a critical interaction operation for virtual object inspection. While many 3D rotation metaphors using 2D input devices such as a mouse or a touchscreen have been developed and tested, they require some prior training to fully perform precise and natural operations. The reader is referred to the works of Chen et al. [13] and Bade et al. [14] for extensive, comparative studies of 3D rotation techniques for 2D input devices. However, the 3D sensing capabilities of modern smartphones make them a perfectly natural 3-DOF input device to perform 3D rotation, since the rotation that the device is undergoing can be directly applied to the virtual object, imitating the way a real object would be rotated by the user (see Fig. 3).

While this metaphor is very natural, some rotated configurations can be difficult to reach while holding the device with one hand. Either multiple, accumulated 3D rotations are applied to the virtual object or a very uncomfortable wrist motion has to be performed. Also, very subtle rotations can be difficult to apply due to the noise introduced by the sensors or the user's motions. In order to maintain the natural, intuitive benefits of this metaphor and increase the ergonomics

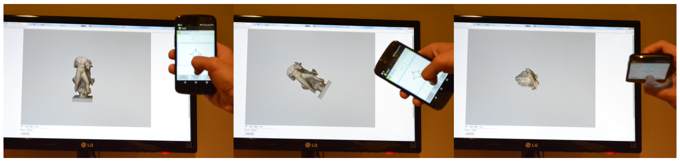

Fig. 3. The direct rotation mapping emulates the virtual object being held by the user, replicating the current rotation of the device.

and precision during the rotation, we propose the Weighted Rotation Mapping.

*1) **Weighted rotation mapping**:* We exploit the fact that any given rotation matrix can be decomposed into an angle of rotation and an axis of rotation. This conversion is known and allows us to modify any of the components of the rotation independently.

Code 1 shows the method to weight the angle of a rotation matrix: the current calibrated rotation is transformed into its axis-angle form (lines 2-3). Through the use of a weighting factor $\alpha > 0$, we modify the amount of rotation applied (line 4). A new, weighted rotation matrix is generated using the weighted rotation amount and the unmodified axis of rotation (lines 5-6).

```
1 matrix3 weight_rotation(matrix3 R_calib, double alpha){
2   double angle = get_angle(R_calib);
3   vector3 axis = get_axis(R_calib);
4   angle = alpha * angle;
5   matrix3 R_weighted = from_axis_angle (axis, angle);
6   return R_weighted;
7 }
```

Code 1. Pseudo-code of the rotation matrix weighting procedure.

When $0 < \alpha < 1$, the amount of rotation is decreased, reducing the impact of the noise (introduced by errors in the sensors or by an unsteady hand) and thus increasing the accuracy in the rotations. When $\alpha > 1$ the amount of rotation is increased, thus allowing to reach a wide range of rotated configurations with small motions of the wrist.

This weighting value can be dynamically modified using a slider-type touch input, therefore allowing the user to increase the ergonomics during large rotations, or achieve precise rotations with comfortable hand motions.

IV. Study Case: Remote 3D Gallery Interaction

The proposed interaction techniques have been implemented for Android devices, and a demo application of a gallery of 3D heritage models has been developed. The device is connected to the remote application via a WiFi connection, communicating through TCP sockets.

The Android interface for the user's device can be seen in Fig. 5. The screen shows three different touch areas, each dedicated to one of the proposed manipulation tasks.

- The top touch area performs uniform scaling. The scaling factor is obtained from the variation in the x coordinate during the swipe motion event.
- The central touch area performs 3D translating operations by composing the 2D touch input with the current device orientation as previously explained.

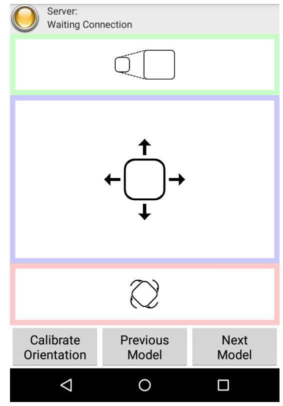

Fig. 5. Interface of the android application for remote interaction with the 3D Gallery.

- The bottom touch area performs 3D rotation operations. When initiated with a single tap, a Weighted Rotation Mapping starts. In this case, the x position of the finger is used to set the α value between 0.2 at the leftmost position and 3 at the rightmost position. When initiated with a double tap, a direct orientation mapping starts, ignoring all the previous rotations applied to the virtual model. In both cases, the rotation ends when the user stops the touch event.

A. System control

As stated by Bowman et al. [12], one of the required tasks for 3D interfaces, and in general any application interface, is *System Control*. This refers to the ability to issue commands to alter the application state. Smartphones allow to display and interact with standard or custom widget-based interfaces together with virtual keyboards and, therefore, system control is trivial since a fully adapted interface can be designed for any remote application.

For our gallery application, just a simple control to iterate through available models is required, and two buttons are shown in the Android interface (*Previous model* and *Next model* buttons in Fig. 5), each issuing a remote command to switch to the previous or next model respectively.

B. Usage test and evaluation

The whole system was set up as a front-end web application communicating with the Android device through a web-socket interface. The raw sensors data and touchscreen input is gathered and handled by the smartphone, so that only concise, light messages are sent to the remote application. The TCP connection provides enough throughput to provide a near-zero delay during an interaction session (the reader is referred to the video provided as additional material). In our test scenario the user must manually connect the device to the remote application, but any service discovery protocol can be used to automatically perform this connection.

An informal preliminary test of the proposed study case was performed with several users. They were asked to evaluate the different 3D tasks while browsing the remote gallery. The three techniques, specially the rotation mapping techniques,

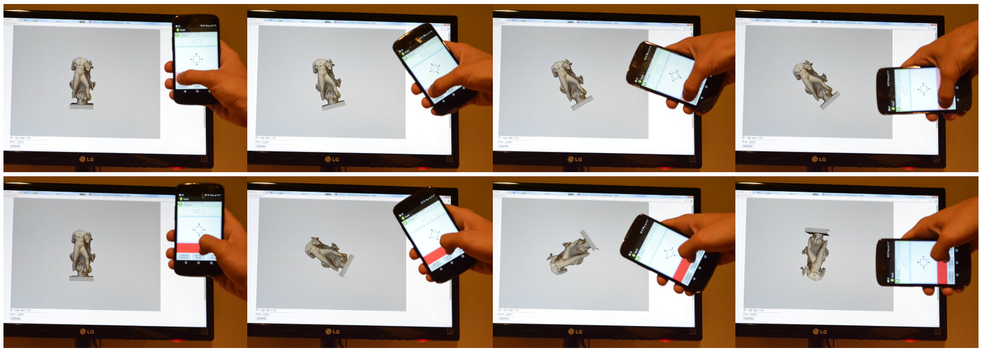

Fig. 4. Weighted rotation mapping examples for the same wrist movement with different weight factors. **Top:** A small weight factor allows precise rotations of the object by lessening the rotation angle. **Bottom:** A big weight factor magnifies the rotation angle, allowing to reach highly rotated configurations with ergonomic wrist movements.

were found intuitive by the participants, many of whom were not familiar with 3D interaction.

V. Conclusions and Future Work

In this work, we have proposed the use of smartphones as remote 3D interaction devices for digital heritage applications. An easy set up of a system using this strategy is possible in live environments without any special hardware requirements, where attenders can use their own smartphones to interact with remote 3D applications shown in standard screens.

We have developed a set of single-handed interaction techniques, general enough to cover all the main 3D manipulation tasks, designed to allow intuitive, ergonomic interactions. A fully functional application example has been developed, demonstrating that a remote 3D application can be effectively controlled by a smartphone, and all the canonical 3D manipulation tasks can be performed likewise. While an informal evaluation of the proposed techniques yielded positive feedback from several users, a proper, formal evaluation and validation must be performed before making formal claims.

In this work we have exposed the smartphone as a 3D input device, however, a proper combination of its input and output capabilities can enrich remote applications, for example, showing additional interactive 2D and 3D information in the smartphone screen. We intend to explore and evaluate such capabilities. Also, more application-specific interaction metaphors are needed to extend the usage of smartphones to more complex and rich virtual environments. Another interesting extension of this work is the design and evaluation of collaborative scenarios with smartphone-based interaction.

Acknowledgment

This work is supported by the "Formación de Profesorado Universitario, Plan Propio de Investigación" program of the University of Granada. This work is also supported by the project TIN2014-60956-R of the Spanish Ministry of Economy and Competitiveness with FEDER funds. The models used in this work were obtained from the Sketchfab website: *sketchfab.com*. The source code and working demos are available at *wdb.ugr.es/local/alejandrora/s3di*.

References

[1] S. Webel, M. Olbrich, T. Franke, and J. Keil, "Immersive experience of current and ancient reconstructed cultural attractions," in *Digital Heritage International Congress (DigitalHeritage), 2013*. IEEE, 2013.

[2] A. M. Manferdini and F. Remondino, "Reality-based 3d modeling, segmentation and web-based visualization," in *Digital Heritage*. Springer, 2010, pp. 110–124.

[3] J.-A. Beraldin, M. Picard, S. F. El-Hakim, G. Godin, V. Valzano, A. Bandiera *et al.*, "Combining 3 d technologies for cultural heritage interpretation and entertainment," in *Proc. SPIE*, vol. 5665, 2005, pp. 108–118.

[4] A. Mossel, B. Venditti, and H. Kaufmann, "3dtouch and homer-s: intuitive manipulation techniques for one-handed handheld augmented reality," in *Proceedings of the Virtual Reality International Conference: Laval Virtual*. ACM, 2013, p. 12.

[5] D. Medeiros, L. Teixeira, F. Carvalho, I. Santos, and A. Raposo, "A tablet-based 3d interaction tool for virtual engineering environments," in *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. ACM, 2013, pp. 211–218.

[6] C. J. Davies, A. H. D. Miller, and C. Allison, "Mobile cross reality for cultural heritage," in *Digital Heritage International Congress (DigitalHeritage), 2013*. IEEE, 2013.

[7] F. Steinicke, K. H. Hinrichs, J. Schöning, and A. Krüger, "Multi-touching 3d data: Towards direct interaction in stereoscopic display environments coupled with mobile devices," in *Proc. AVI Workshop on Designing Multi-Touch Interaction Techniques for Coupled Public and Private Displays*. Citeseer, 2008, pp. 46–49.

[8] C. Telkenaroglu and T. Capin, "Dual-finger 3d interaction techniques for mobile devices," *Personal and ubiquitous computing*, vol. 17, no. 7, pp. 1551–1572, 2013.

[9] M. de Souza, D. D. B. Carvalho, P. Barth, J. V. Ramos, E. Comunello, and A. von Wangenheim, "Using acceleration data from smartphones to interact with 3d medical data," in *Graphics, Patterns and Images (SIBGRAPI), 2010 23rd SIBGRAPI Conference on*. IEEE, 2010, pp. 339–345.

[10] N. Katzakis, M. Hori, K. Kiyokawa, and H. Takemura, "Smartphone game controller," in *Proceedings of the 74th HIS SigVR Workshop*, 2011.

[11] S. Ayub, A. Bahraminisaab, and B. Honary, "A sensor fusion method for smart phone orientation estimation," in *Proceedings of the 13th Annual Post Graduate Symposium on the Convergence of Telecommunications, Networking and Broadcasting, Liverpool*, 2012.

[12] D. A. Bowman, E. Kruijff, J. J. LaViola Jr, and I. Poupyrev, *3D user interfaces: theory and practice*. Addison-Wesley, 2004.

[13] M. Chen, S. J. Mountford, and A. Sellen, "A study in interactive 3-d rotation using 2-d control devices," in *ACM SIGGRAPH Computer Graphics*, vol. 22, no. 4. ACM, 1988, pp. 121–129.

[14] R. Bade, F. Ritter, and B. Preim, "Usability comparison of mouse-based interaction techniques for predictable 3d rotation," in *Smart graphics*. Springer, 2005, pp. 138–150.