GESTURAL INTERACTION AND NAVIGATION TECHNIQUES FOR VIRTUAL MUSEUM EXPERIENCES

Fabio Marco Caputo
Dept. of Computer Science
University of Verona, Italy
fabiomarco.caputo@univr.it

Irina Mihaela Ciortan
Dept. of Computer Science
University of Verona, Italy

Davide Corsi
Dept. of Computer Science
University of Verona, Italy

Marco De Stefani
Dept. of Computer Science
University of Verona, Italy

Andrea Giachetti
Dept. of Computer Science
University of Verona, Italy

ABSTRACT

Virtual museums are one of the most interesting applications of Virtual Reality environment, but their success is strongly depending on the development of effective interaction techniques allowing a natural and fast exploration of their contents. In this paper a series of interaction techniques in a full immersive Virtual Reality environment are presented. These techniques are specifically designed to cover basic needs of a virtual museum experience such as navigating in the museum space and accessing meta-information associated to displayed items. Details of the implemented methods and preliminary results collected in user tests performed to compare different navigation choices are presented and discussed.

Keywords

HCI; Virtual Reality; Gestural Input; Virtual Museum;

1. INTRODUCTION

The recent developments on Virtual Reality (VR) and Augmented Reality (AR) technologies made them accessible and usable for general purpose applications [11] ranging from entertainment to professional tools for industrial design or medical healthcare. Each field identifies different interactions tasks between users and system, and requires specifically designed techniques to obtain satisfactory virtual experiences.

In museum applications, virtuality tools, defined as combination of VR and AR, have been proved to be promising for the presentation, as well as documentation, conservation and preservation of cultural heritage (CH) items [1, 2, 5]. One of the main advantages of VR consists in the high number of degrees of freedom for user interactivity [12].

Different kinds of virtual museums require effective core techniques to create satisfactory interactive museum experiences. The development of core mechanics for the expected interactions in virtual museum and the design of effective solution for them is therefore mandatory to build successful applications.

In this paper we focus on two of these core mechanics that are critical for many virtual museum applications: information retrieval from items in the scene such as art pieces, multimedia, etc. and navigation in the virtual museum environment. For both the mechanics, we designed different solutions based on mid-air gesture interaction and tested them with users in simple demo environments to evaluate their advantages and drawbacks. The goal of the study is to derive guidelines and a wide range of solutions for an optimized virtual museum experience. This paper is organized as follows: In Section 2 the proposed techniques are contextualized in the gesture-based interaction literature. In Section 3 each of the techniques designed and implemented is discussed. In Section 4 the experiments performed to analyze and compare the techniques are is presented along with the setup details. In Section 5 preliminary results and the future work directions are discussed.

2. RELATED WORK

The integration of mixed reality technologies as an improvement to the traditional museum experience has been an ongoing phenomenon for the last years [11], [1], [6]. This phenomenon has also branched to examples of complete virtualization of museums that can enhance user experience through the higher freedom in interacting with the artefacts [2] and can provide a series of benefits for socio-educational purposes [6].

As shown in [1], two example criteria for defining a taxonomy of virtual museums can be the content of the exhibition (already existing structures or ex-novo reconstruction of the items in a virtual environment) and the access method to the museum. Based on the latter criterion, there are web-based virtual museums, that can be remotely-accessed by the users, such as ARCO [10] and virtual museums based on VR displays, that can be implemented on-site (interactive kiosks in the museums [6], in archaeologic or excavation sites [2]) or simultaneously online and on-site [11].

In [1], it is pointed out that the interactivity with the work of art is an important pre-requisite of the multimedia representations of CH items and that a stereoscopic view contributes to the immersive experience and moreover to the navigability of the environment [3]. There is a tendency to minimize the contribution of computers in their traditional form, also known as disappearing computer technology [6], where mixed reality engages the user to a high extent, so that the focus of the user is on the interaction with the virtual objects, rather than the media or means of visualization. In order to absorb the users as much as possible in the virtual environment, there are several navigation paradigms [12] and gesture-based navigation is presented as an effective strategy for user interaction [7] with stereoscopic environments. The head-mounted display enables such interaction and has been used for live exhibitions in real museums in a hybrid physical-digital artefact setup [6]. In the following sections, we approach the Oculus Rift and its advantages as an appropriate gesture-based VR navigation technology for virtual museums.

3. INTERACTION TECHNIQUES

3.1 Information Display

One of the enhancements a virtual museum can benefit from, compared to regular museums, is the easiness of interaction with objects and ways to retrieve information and multimedia content from them. A typical scenario is the one in which the user has the choice to inspect different objects in the scene. In a virtual museum a user can interact directly with the objects, look at them from any point of view or more in general have access to a different number of options not often possible in real museums. A common issue is however the object selection and the displaying of information, whether they should be either directly linked to the item, or a display of further possible actions to perform on it. Displaying everything at once for each object is unpractical for different reasons including, but not limited to, visual cluttering, and it is better to allow the user to select a specific object and enable the display of linked information through an interaction trigger. In this work four solutions are presented for a specific scenario with multiple objects in the scene. Each of the presented solutions features a gesture-based technique to select the wanted object and to trigger the display of its related information. Their different strengths and weak points will not only offer a different range of solutions for more specific tasks but also provide good insights on viability based on the user performance and other data collected in the test stage.

Display Buttons: The user sees a button for each object. By pressing one of the buttons, the associated object is selected and becomes the focus of the view. On the right, all the text information is displayed. The button is supposed to give the best affordance among the different methods as it naturally suggests a trigger interaction, while also hiding all the information related to an object, without causing any excessive cluttering.

Swipe: In this mode the user sees part of the view space reserved to text information display. There is an initial selected object by default picked from the set of objects available in the scene and all its information are also already displayed in the reserved space. By using a "swipe" gesture the user is able to cycle the currently selected object among the rest in the set and automatically display the information in the reserved space. In this method, the trivial trade-off, is between the available view space and the easiness of selecting a series of different objects.

Object Selection: The user is able to select the wanted object by touching it with a finger. This action automatically makes the object move in front of the user for a better display and also triggers the view of all the text information in a similar way the Display Button method does. This method is best to achieve natural interaction compared to the Display Buttons method as the interaction with the object is direct and can also be useful for applications in which a user can also perform other actions like touching parts of the object, change the view angle, modify its aspect and so on.

Object Picking: The technique is similar to Object Selection but in this case the user has to drag and drop each object on an anchor point in the scene (which is not anchored to the user camera) in order to trigger the display all the information (**Figure 1**). This is convenient in cases where the user has to select the object while not being forced to directly look at it. This happens when the user wants to still be able to look around freely or in specific applications to have a second object selected for close comparison.

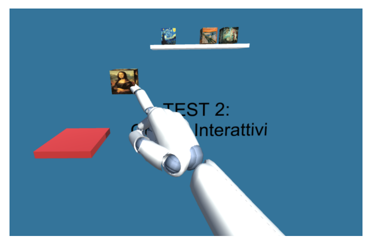

Figure 1: Object Picking technique. The selected object is being dragged towards the anchor point. The rest of the items in the scene remain visibile and potentially available for further actions.

3.2 Environment Navigation

As suggested in [8] virtual museums are not necessarily bound to real space simulation. Keeping that in mind another critical issue in case of virtual museums featuring explorable space is the control of user movement. The considered scenario is again one in which multiple objects are accessible in the museum. In **3.1** a single set of items all visible in a single scene was considered while now a wider scenario is taken into account with multiple set of objects displaced around the virtual environment. This kind of scenario require the user to be able to move and cover bigger distances so different solutions are presented to enable free movement around a virtual space. In all these solutions an hand interaction controls the subject's forward speed while the steering is controlled with the HMD by gazing at the direction where the user wishes to go to.

Palm Rotation: The user can increase the movement speed by placing the palm of the hand in front of the hand tracker with the palm perpendicular to the ground and the fingers pointing forward (**Figure 2**). Any other different hand positioning interrupts the movement. This gesture was chosen to be intuitive and easy to learn and remember.

Palm Position: Similar to Palm Rotation but in this case the user just has to put hand in front of the hand tracker and over a fixed distance threshold from the body. Again any other different hand positioning stops the movement. This variation was implemented to test the effectiveness, in terms of performance and accuracy, of a pose trigger (Palm rotation) against a position in space kind of trigger (Palm Position).

Forward Button: A widget-like button is showed in a Heads-up Display (HUD) style and by pressing it the user is able to move forward (**Figure 3**). Releasing the button stops the movement. This is less natural, takes away part of the user view on the scene and it might be hard to combine with the selection and inspection methods presented in 3.1. Unlike the previous two solutions in which the triggering point where the gesture is recognized by the system is invisible and hard to detect, the advantage of this method is the ability for the user to clearly see the the trigger and have a more precise control on the movement speed.

Mobile Control: The user has to use a mobile phone's gyroscope as control device. By tilting it forward the in-world camera begins to move. In order to stop the movement the phone must be brought back to the original pose. With this solution a new kind of feedback is offered to the user, compared to all the previous ones. The mobile phone serves as a smart-object to control the speed by using the inclination angle and offers the possibility to use haptic feedback of vibration as alert for obstacles or notify the user about particular events taking place in the museum.

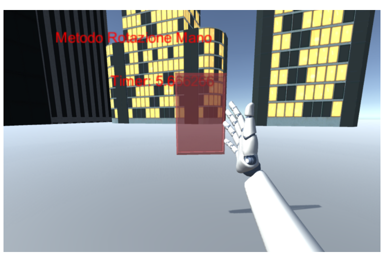

Figure 2: Palm Rotation navigation technique. The user is approaching the first waypoint on the path (red door).

4. EXPERIMENT SETUP

The total immersive experience is achieved through the use of a head mounted display (HMD) and an IR hand tracker. Specifically low-cost technologies were chosen: Oculus Rift DK2 [4] as HMD and Leap Motion Controller [9] for hand tracking. This was also to have a test setup closer to a real case scenario in which high-performance devices are not available. **Figure 4** shows the configuration used with the hand tracker mounted directly on the HMD to ensure the hands are always in the hand tracker field of view.

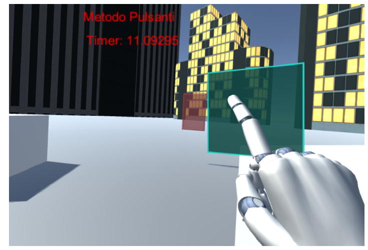

Figure 3: Forward Button navigation technique. The user is approaching a waypoint by pressing the button to control the movement speed (azure square).

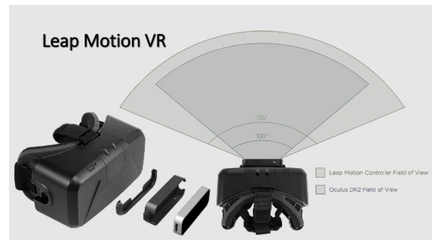

Figure 4: Oculus Rift and Leap Motion configuration

4.1 Data collection

Validation of the implemented methods for both information display and navigation was performed through experiments with subjects in demo environments. For information display testing, users received a hint about one of the available artworks (four in total for each execution) and had to figure out which one the hint was talking about by retrieving the available information with the different interaction techniques. This was repeated for all the different methods in a randomized order for each user. For navigation testing, users had follow with the different methods a path marked with checkpoints. The checkpoints became visible one at the time and always within the user's field of view until the goal was reached. Again this was be repeated for all the methods in a randomized order.
Execution time was used as main measure of performance to compare the efficiency of the techniques against each other. The randomized order prevented the presence of obvious bias in the collected data. Other time splits are measured to identify critical point and possible bottlenecks in each method. In particular: the time spent reading the information against the time spent to access each object in the first task and the number of times the user stops and resumes the navigation to identify possible problems with maintaining the control gesture. At the end of each session every user compiled a questionnaire to rate different aspect of the experience including, but not limited to learnability of the methods, easiness of execution, perceived stress and overall preferred technique.

5. PRELIMINARY RESULTS

At this stage of the study, data was collected, for thirty sessions for both tasks. The most interesting results derive from the total time of task completion. These results already show, performance-wise, relevant differences between the proposed solutions. In **Figure 5** the times for the first task, selection and information retrieval, are shown. The worst performance is associated to the Object Picking technique. This was expected due to the extra step of drag and drop on the anchor point. The best performances come from the Swipe and Object Selection techniques. More data are needed to prove actual significant improvement compared to the other techniques. However, this again agrees with the design for this methods, in fact Swipe relies on the reserved space on screen to offer a quick selection of the objects by just using a single gesture that doesn't require strictly specific directions or angles to be recognize by the system. Object Selection only requires the user to touch an object which is more natural than the Display Buttons. This last technique, that lies in somewhere in the middle on the performance scale, doesn't seem to bring any real advantage, with its focus on affordance, compared to its natural counterpart (Object Selection).

In **Figure 6** are the times for the second task, involving the environment navigation techniques. With the current data it's only possible to say that the Forward Button technique outperforms the other techniques in terms of execution time. This is due to the high control level on the movement granted by the displayed button. In our solutions the navigation is achieved with the use of hands interaction methods. Because of this, real natural interaction style can't be achieved, as the hands aren't the part of the human body used to directly move in the real surrounding space. This is one of the possible reasons that prevent the other implemented navigation techniques to bring any advantage in terms of time performance.

In the next stage of the work more data will be collected to be analyzed in depth. This will give a more general vision of all pros and cons of each solution, in order to offer a wider range of methods that can be chosen for specific applications. The experiment setup will also serve as a platform for new or more refined methods and their implementations, to be compared with the existing ones. In a new work, a simulation of a full immersive virtual museum will be developed to evaluate the behavior of various solutions on a real application prototype. The goal will be to offer better insights about convenient and effective combination of techniques for both the core mechanics, necessary for this kind of experience, that were identified and discussed in this paper.

Acknowlegements: This work was partially supported by the Scan4Reco project funded by EU Horizon 2020 Framework Programme under grant agreement no 665091.

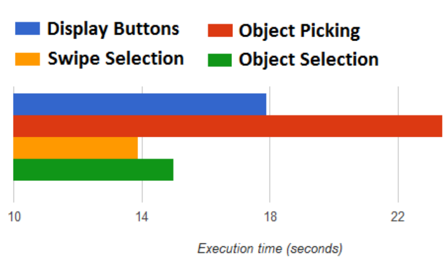

Figure 5: First task completion times

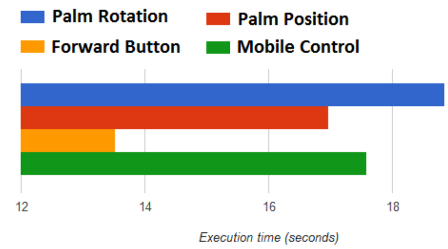

Figure 6: Second task completion times

6. REFERENCES

[1] F. Bruno, S. Bruno, G. De Sensi, M.-L. Luchi, S. Mancuso, and M. Muzzupappa. From 3d reconstruction to virtual reality: A complete methodology for digital archaeological exhibition. *Journal of Cultural Heritage*, 11(1):42–49, 2010.

[2] F. Cameron and S. Kenderdine. *Theorizing Digital Cultural Heritage: A Critical Discourse (Media in Transition)*. The MIT Press, 2007.

[3] S. E. Chen. Quicktime vr: An image-based approach to virtual environment navigation. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 29–38. ACM, 1995.

[4] P. R. Desai, P. N. Desai, K. D. Ajmera, and K. Mehta. A review paper on oculus rift-a virtual reality headset. *arXiv preprint arXiv:1408.1173*, 2014.

[5] A. Gaitatzes, D. Christopoulos, and M. Roussou. Reviving the past: cultural heritage meets virtual reality. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 103–110. ACM, 2001.

[6] T. Hall, L. Ciolfi, L. Bannon, et al. The visitor as virtual archaeologist: explorations in mixed reality technology to enhance educational and social interaction in the museum. In *Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 91–96. ACM, 2001.

[7] G. A. Satalich. *Navigation and wayfinding in virtual reality: Finding the proper tools and cues to enhance navigational awareness*. PhD thesis, University of Washington, 1995.

[8] W. Schweibenz. Virtual museums. *The Development of Virtual Museums. ICOM News Magazine*, (3), 2004.

[9] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler. Analysis of the accuracy and robustness of the leap motion controller. *Sensors*, 13(5):6380–6393, 2013.

[10] M. White, N. Mourkoussis, J. Darcy, et al. Arco-an architecture for digitization, management and presentation of virtual exhibitions. In *Computer Graphics International, 2004. Proceedings*, pages 622–625. IEEE, 2004.

[11] R. Wojciechowski, K. Walczak, M. White, and W. Cellary. Building virtual and augmented reality museum exhibitions. In *Proceedings of the ninth international conference on 3D Web technology*, pages 135–144. ACM, 2004.

[12] J. Zara. Virtual reality and cultural heritage on the web. In *Proceedings of the 7th International Conference on Computer Graphics and Artificial Intelligence*, pages 101–112, 2004.