Poster: Evaluation of Immersive Visualization Techniques for 3D Object Retrieval

Diogo Henriques* Daniel Mendes† Pedro Pascoal ‡ Isabel Trancoso§ Alfredo Ferreira¶

INESC-ID Lisboa / IST / University of Lisbon

Abstract

The number and size of 3D object repositories have been increasing at a considerable rate. Consequently, finding a specific 3D model in such collections is getting more difficult. Current 3D search engines do not take advantage of novel interaction technologies, usually presenting query results in grids of thumbnails. This greatly hinders objects' interpretation and allows little to none manipulation at all. Immersive environments are believed as promising solutions for displaying 3D models, allowing complete representations of the models. These environments can be enhanced with multimodal techniques for a more natural interaction. In this paper we present a prototype that uses immersive visualization and mid-air interactions to explore query results in a dataset with 3D objects, using one of four different visualization modes. We evaluated these modes with 29 users and concluded that our immersive approaches are preferred by users and, albeit novel, perform at par relatively to traditional bi-dimensional grids with thumbnails.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Interaction styles, Graphical User Interfaces, Input devices and strategies

1 Introduction

The vulgarization of 3D modeling applications, such as Blender, 3DS Max or AutoCAD, as well as scanners and low-cost depth cameras that can be used to capture 3D physical objects, has resulted in a considerable increase in the number of available 3D virtual objects. This urges the need for better solutions for searching 3D models in large collections. In recent years, several 3D search engines have been presented. However, most of them still face major drawbacks, mainly concerning the results' presentation. These systems normally present results in a traditional grid of items. Furthermore, items are usually presented through thumbnails, which is clearly an inadequate representation, since it loses relevant 3D information, and might not even provide the best view of the specific model.

Despite the newly emerging technologies for human-computer interaction, most 3D search engines do not take advantage of these new interaction paradigms and visualization.

In this context, we developed a system in which users can search for 3D objects through multimodal interactions, using speech and mid-air gestures. Both exploration and result analysis are performed on a 3D immersive environment, providing a clear representation of the 3D models.

*e-mail: diogo.henriques@tecnico.ulisboa.pt
†e-mail: danielmendes@tecnico.ulisboa.pt
‡e-mail:pmbp@tecnico.ulisboa.pt
§e-mail:isabel.trancoso@inesc-id.pt
¶e-mail:alfredo.ferreira@inesc-id.pt

IEEE Symposium on 3D User Interfaces 2014
29 - 30 March, Minneapolis, Minnesota, USA
978-1-4799-3624-3/14/$31.00 ©2014 IEEE

2 Related work

Besides the traditional approaches for results' exploration of 3D virtual object retrieval systems, such as the ones presented by Funkhouser et al. [2] or Ansary et al. [1], solutions that leverage immersive and interactive technologies have been proposed.

In 2001, Nakazako et al. [4] presented the 3D MARS, which demonstrates the benefits of using immersive environments in the context of multimedia retrieval. Their work focused mainly on query results presentation for a content-base image retrieval system, using a CAVE like setup.

Holz and Wilson [3] explored mid-air gestures to describe the desired shape for a 3D object. Their system system captures gestures with a Kinect camera and then finds the most closely matched object in the collection. Although it tackles the issue of query specification, it does not improve upon visualization of results.

More recently, Pascoal et al. [5] extended 3DMARS approach for retrieving 3D objects, showing that it can overcome some of the issues when presenting 3D object retrieval results. These results are distributed in the virtual space accordingly to their similarity. Users can then explore the results by navigating in the immersive environment, through an head-mounted display. Their system also allows a diversified set of different visualization and interaction devices, used to test multiple interaction paradigms for 3D object retrieval. However, in this solution, some of the results may appear overlapped if they are too similar. Traditional approaches do not overlap results, but their visualization based on thumbnails lack an adequate 3D perception of the objects.

3 Our Prototype

Aiming for an immersive and natural solution for 3D object retrieval, we developed a prototype that takes advantage of recent devices in Human Computer Interaction. Spoken queries are processed through a speech recognition mechanism, generating terms to be used in our retrieval system. To browse and analyse results, our prototype offers a fully immersive 3D space instead of the traditional grid with bi-dimensional thumbnails. A Leap Motion

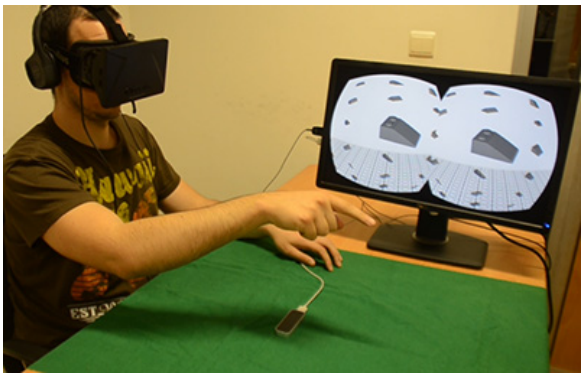

Figure 1: User pointing at an object in our prototype.

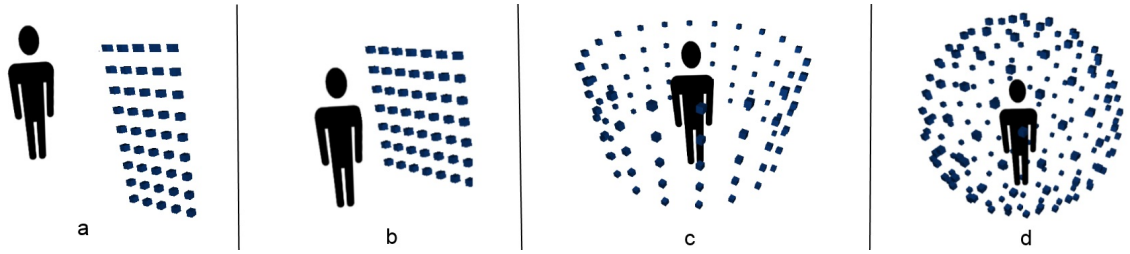

Figure 2: Modes for visualizing query results in an immersive environment: (a) Rectangular Grid, (b) Square Grid, (c) Cylindrical, (d) Spherical.

(http://www.leapmotion.com/) device enables gesture-based interaction. To create the fully immersive visualization of the environment, a set of Oculus Rift (http://www.oculusvr.com) is used. Unlike Im-ORET [5], which requires users to walk while wearing multiple wire depended devices, affecting their movements, our prototype allows users to remain seated, using Oculus Rift without interfering with interactions.

3.1 Retrieved objects visualization

To view query results of the queries we implemented four modes of visualization, illustrated in Figure 2. We started adapting the traditional grid approach to a 3D immersive environment, to be able to analyse its suitability. The **Rectangular Grid** shows results placed in vertical grid. The top left object in the grid is the higher ranked, while in the down right corner is placed the worst ranked one. The **Square Grid** is similar, but has equal height and width, being the higher ranked objects placed close to the center. Both these approaches appear in front of the user, within the three-dimensional space. We also implemented two approaches to better explore the three-dimensional space. With the **Cylindrical** and the **Spherical** approaches, we surround the user with objects, placing them on the surface of an invisible cylinder or sphere, respectively. In both these approaches, higher ranked objects are placed in front of the user, expanding from there.

Navigation through results is made by freely moving user's head. To see beyond the user's field of view, or focus on a specific section of the results, they can place an open hand in mid-air in front of them and move it to rotate (in Cylindrical and Spherical approaches) or pan (in Rectangular Grid and Square Grid) results. When a detailed view of an object is desired, the user points, with one finger, at the object, which is brought closer to him and rotates, in order to be viewed in different angles, as depicted in Figure 1.

4 Evaluation of visualization modes

We evaluated the proposed visualization modes with 29 participants (22 male, 7 female), whose ages ranged from 17 to 40 years old. 90% were students of computer science. All of them were familiar with search engines for images, experienced in search thumbnails on a screen. The experiment took approximately 20 minutes, which included adaptation to our prototype, perform the tasks and answer a brief questionnaire.The adaptation to our prototype consisted in experience the immersive environment with some objects randomly positioned, to which users could point at and visualize in detail.This was useful to introduce the immersive gesture-based interaction environment, since it was novelty for all users.

In the experiment, a traditional grid on a standard screen was included, which users navigated using a mouse. This acted as a baseline in our evaluation. For each of the five approaches, we asked users to search for an object, which was shown to them using a physical object. This object was randomly chosen, so that users searched for a different object in each visualization mode. The order in which visualization modes were evaluated by each user were also random, to avoid biased results. Since the focus of this evaluation was not search-by-speech, but rather the visualization modes, we did not use it in this experiment.

4.1 Experiment results and observations

During the experiment, we recorded the time taken by each user user to fulfil task requirements and asked them to classify each visualization mode using a five point Likert scale. Analysing times taken by each mode of visualization, statistical differences were not found using the One-Way ANOVA test. Considering that users were experienced using a mouse and traditional screens to search for multimedia content in a daily basis, such as images, our prototype, albeit novel, was able to compete with this traditional approach. Moreover, analysing questionnaires using the Wilcoxon Signed Rank test, we concluded that users strongly agreed that the traditional mode was the less preferred among all techniques (Rectangular Grid: Z=-3.744, p=0.000; Square Grid: Z=-3.430, p=0.001; Cylindrical: Z=-3.542, p=0,000; Spherical: Z=-2,863, p=0.004). Additionally, users agreed that the Cylinder mode was preferred instead of Quadratic Grid (Z=-2.236, p=0.025).

5 Conclusions

Collections of 3D virtual models are increasing considerably and finding a specific object is getting more difficult. Taking advantages of recent advances in human-computer interaction devices, we proposed a fully immersive prototype to retrieve 3D objects. We implemented four approaches for viewing results in the three-dimensional space. A comparative study showed that immersive gesture-based interaction can compete with the traditional 2D grid with thumbnails, which users are acquainted to. Participants agreed that immersive approaches are preferred over the traditional. These results showed the potential of searching for objects in an immersive 3D space, where users can easily explore the results.

Acknowledgements

The work presented in this paper was partially supported by the Portuguese Foundation for Science and Technology (FCT) through projects CEDAR (PTDC/EIA-EIA/116070/2009), TECTON-3D (PTDC/EEI-SII/3154/2012) and Pest-OE/EEI/LA0021/2013, and through doctoral grant SFRH/BD/91372/2012.

References

[1] T. F. Ansary, J.-P. Vandeborre, and M. Daoudi. 3d-model search engine from photos. In *Proc. of CIVR '07*. ACM, 2007.

[2] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs. A search engine for 3d models. *ACM Trans. Graph.*, 22:83–105, 2003.

[3] C. Holz and A. Wilson. Data miming: inferring spatial object descriptions from human gesture. In *Proc. of CHI '11*, pages 811–820, 2011.

[4] M. Nakazato and T. Huang. 3D MARS: Immersive virtual reality for content-based image retrieval. *ICME*, pages 44–47, 2001.

[5] P. Pascoal, A. Ferreira, and J. Jorge. Im-o-ret: Immersive object retrieval. In *Proc. of VR '12*, pages 121–122, 2012.