Computers & Graphics 74 (2018) 225–233

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Technical Section

The Smart Pin: An effective tool for object manipulation in immersive virtual reality environments

Fabio M. Caputo*, Marco Emporio, Andrea Giachetti

Department of Computer Science, University of Verona, Italy

ARTICLE INFO

Article history:
Received 30 December 2017
Revised 17 April 2018
Accepted 16 May 2018
Available online 28 May 2018

Keywords:
Mid-air manipulation technique
Virtual reality
Hand tracking
User tests

ABSTRACT

In this paper, we present a novel method for object manipulation in immersive virtual environments. The proposed technique exploits a novel widget, called "Smart Pin", to enable the user to select, translate, rotate and scale objects relying entirely on the positional tracking of a single hand. It is, therefore, suitable for several real-world applications where handheld devices are not available, the surrounding environmental conditions make orientation tracking hard, and the non-dominant hand may be involved in different tasks. We evaluated the method with users on two classical tasks such as detail search and docking, comparing it with a successful two-handed manipulation technique (Handlebar). The measured execution efficiency obtained with the two methods was similar, but most users preferred the Smart Pin for its gestural comfort and ease of use.

1. Introduction

Virtual reality (VR) applications are becoming increasingly popular due to the recent diffusion of low cost effective devices for body tracking and low latency solutions for Head Mounted Displays (HMD), allowing the realization of immersive virtual environments (IVE) designed for a variety of applicative contexts (scientific visualization, virtual museums, simulation, gaming, e-commerce and more). However, a big obstacle to the success of this kind of applications is related to their usability, as the interaction in many practical tasks in VR is not natural as expected, and successful interaction metaphors are still missing for many practical activities in IVE.

A simple example of this problem is the manipulation of virtual objects. Many useful and entertaining VR applications, heavily rely on the ability to grab, move, rotate and scale objects in the field of view. However, as shown in [1] "...development of interaction mechanisms for manipulation remains one of the greatest challenges for VR". In fact, while successful metaphors to control 3D manipulation using 2D desktop interfaces are well established and employed in many 3D editing environments, efficient and flexible solutions well suited for immersive VR are still missing, limiting the potential use of this emerging technology.

Many recent VR setups and applications rely on handheld devices to interact with the scene, e.g. Oculus Touch or HTC Vive. This choice is suitable for personal use and gaming, but is limiting the naturalism of the experience and not appropriate for many use cases (e.g. public displays, work environments, etc). Furthermore, mapping grabbing actions to device buttons and object orientation to the device's one does not solve, for example, issues related to continuous rotation control due to intrinsic limits of hand articulation.

When handheld devices are not an option, it is possible to rely on hand tracking systems and related APIs (Leap Motion, Kinect, RealSense) providing real-time hand and finger position mapping and some gesture recognition capabilities. These trackers, even if not always stable and accurate due to occlusions and noise, can be directly used to build gesture-based interfaces allowing direct object manipulation, similarly to the finger-based manipulation on touchscreens.

A significant hurdle to usability of natural techniques, such as the Simple Virtual Hand [2], is the reliability of hand orientation tracking required to map the rotation of the hand to that of the desired transformation. In order to avoid this problem, a possible solution is to design a manipulation technique that relies only on easy to track positional data, like the 3D coordinates of the hands' centroids. A very well known method of this kind is the so-called "Handlebar", adopted in several practical applications. In the Handlebar metaphor [3], the relative position of the two hands is exploited to derive the object transform mapping, allowing a reasonable accuracy even in case of rough tracking without accurate finger detection. However, this method requires the use of both hands, denying the availability of a free hand for any possible

* Corresponding author.
E-mail address: fabiomarco.caputo@univr.it (F.M. Caputo).

https://doi.org/10.1016/j.cag.2018.05.019

action to be performed in parallel. Furthermore, it can require significant hand displacements that may result in the so-called gorilla arm issue if performed for a long duration [1].

The possibility of manipulating objects with a single hand, as usual on 2D desktop interfaces, would certainly extend the potential applications of immersive VR based interfaces and visualization tools. In [4] a technique based on this concept, MAiOR, is presented. In this technique translations, rotations and scaling actions are controlled using only the information on user hand position tracked using an handheld controller (i.e. handles from the HTC Vive HMD).

In this paper, we propose a novel solution based on the same concept. The idea is to take advantage of a simple widget activated by proximity which allows, providing a visual feedback, a natural selection of separate controls for translational/rotational/scaling degrees of freedom (DOF), introducing DOF separation as suggested in [5] and features, for each manipulation mode, a carefully designed mapping between subsequent hand movements and object transformations. We exploited successful ideas of existing widgets and metaphors, independently for each mode, using direct manipulation only where it works well (translation), and leveraging the ideas of Virtual Trackball/Arcball [6] for rotation.

The widget has been designed to be used for deviceless interaction in VR with low cost hand tracking realized with a Leap Motion device and tested with users with this implementation. Nonetheless, it could be used as well with different tracking hardware and handheld devices to simplify object manipulation.

The paper is organized as follows: Section 2 presents an overview of related work motivating our research and Section 3 introduces the novel method. Section 4 describes some implementation issues and the experimental setup used in our tests. Section 5 describes the user evaluation design and Section 6 analyzes the results.

2. Related work

Mid-air manipulation has been recently investigated in several papers proposing different kinds of solutions and trying to evaluate their usability in specific tasks for both standard or stereoscopic displays and HMD based immersive environments.

This kind of interaction can be based on different sorts of devices, resulting in differences in tracking accuracy, access to buttons, force feedback, etc., and presenting diverse specific problems. For this reason some interesting papers compare manipulation solutions based on different hardware setups, trying to derive useful hints for application design.

Mendes et al. [7] compared different manipulation methods using a 2D touch based approach and mid-air bi-manual methods with two variants: 6DOF Hand, which uses the dominant hand to grab, move and rotate objects, and the distance between both hands for scale and 3DOF Hand, where the dominant hand only moves the object, while scale and rotation are given by the non-dominant hand.

Vuibert et al. [8] also made a comparison of different methods including physically constrained devices, a wand-like device, tangible object replica and finger-based interaction. Another wide range study on direct mid-air manipulation is presented in [9]. Here manipulation with 6 degrees of freedom flystick, and finger-tracking systems are analyzed and many aspects of the problems are addressed, like grasp/release heuristics, a finite state machine for object manipulation, handling of complex hand/finger interactions with objects. Results of these multi-modal comparisons show that users often like mid-air interactions more than other methods, finding the accuracy in manipulation sufficiently good for many proposed tasks.

This fact, together with application-driven constraints, suggested that we should focus our interest on mid-air deviceless interaction, using low-cost trackers to track hands and fingers with no need of wearing markers or using handheld devices and on co-located hands and objects. This kind of interaction is extremely important for emerging, partially or fully, immersive applications like those based on Head Mounted Displays (HMD), autostereoscopic monitors, or even light field monitors [10] where physical input systems or other input channels are not available.

One of the main problems tackled by mid-air interaction research is the search for more “natural” interfaces that can greatly enhance the performance of the system as well as the user experience [11]. This is a very challenging task, as it requires the study of optimal metaphors and implementations that should consider: the lack of buttons, object selection, the detection of the start and end of gestures as well as their decomposition, the difficulties and fatigue of mid-air gesturing, and the accurate mapping of hand motion to object motion with optimal constraints reducing the effect of sensor errors. The simple attempt to directly map 5 or 6 DOF motion on virtual objects may, in fact, be “natural”, but it is not able to create an effective interaction [12]. Furthermore, using low-cost trackers like the Leap Motion[13], but also due to the intrinsic difficulty in controlling hands in mid-air movements, it is necessary to cope with the relevant limits in selection accuracy in a setup with a single tracking sensor and the impossibility of directly tracking large rotations [14].

Another relevant problem for the usability of effective mid-air virtual object manipulation is the lack of haptic feedback, that should ideally be replaced by different feedback mechanisms [15]. The solutions proposed in the literature for mid-air manipulation can be classified into two categories, bi-manual and uni-manual techniques. In the bi-manual category, the Handlebar metaphor [3], is probably the most intuitive and popular solution for manipulation in VR. The typical implementation of this method uses hand gestures for grabbing and estimates the object transform based on the hands' position, exploiting resemblance with real-world actions like rotating and stretching a bar. This particular mechanic allows the user to handle simultaneously 7 DOF. Several research works [16,17] and also some commercial applications exploited it with slight variations. Another way to exploit the bi-manual interaction is to split the DOF control between the two hands. Caputo and Giachetti [18] propose the 2HR technique, mapping the object rotation onto the user's non-dominant hand rotation. User tests demonstrated, however, the poor usability of the method. In the same work, the Finger-Point technique features a rotation action, also performed with the non-dominant hand, by mapping the index movement on a 2D plane into 3D rotations applied on the object. This indirect rotation showed promising results suggesting that indirect approaches for the rotation action are viable and even preferred by users. Following this kind of indirect approach, there are also bi-manual techniques relying on gestures to indirectly manipulate a selected object through the use of widgets or tools. Examples can be found in [19,20]. Manipulation modes, transitions and accurate control are not, however, addressed. In the uni-manual category, it is possible to find a variety of solutions in the literature. The Simple Virtual Hand [2] is the most direct one. This technique enables object selection through a grasping gesture. After that, position and orientation changes of the object are directly mapped onto the hand's ones. This technique is quite intuitive, but it is not suitable whenever a robust hand tracking is not available, as in the case of a single low-cost device, where hand orientation tracking can be unreliable [18]. An alternative kind of approach is to design indirect techniques by either the use of metaphors or widgets. Kim and Park [21] proposed a Virtual Handle with a grabbing metaphor, based on a bounding sphere around the selected object and the creation of a local reference frame for the following