Zoom-fwd: Efficient technique for 3D gestual interaction with distant and occulted objects in virtual reality

Assia Messaci, Nadia Zenati, Mahmoud Belhocine, Samir Otmane

Noname manuscript No.
(will be inserted by the editor)

Zoom-fwd: Efficient technique for 3D gestual interaction with distant and occulted objects in virtual reality

Assia Messaci · Nadia Zenati · Mahmoud Belhocine · Samir Otmane

Received: DD Month YEAR / Accepted: DD Month YEAR

Abstract In the last decade, it has been observed an extraordinary acceleration in immersive technologies, including virtual and augmented reality (VR/ AR), as well as innovative experience design. The VR interfaces have widely explored various 3D interaction-based approaches. This technology encourages users to use diverse input devices (e.g., VR headset). However, such devices are still not very reliable, inconvenient and invasive to the person's comfort. Furthermore, users may have an inefficient experience for interacting with objects in a virtual environment. Fortunately, there exists an innovative interaction methods called Natural User Interfaces (NUIs). These interfaces are increasingly introduced in Human Machine Interaction (HMI) systems and they introduce the gestures recognition. They became more useful, improving the user's engagement and sense of presence, providing more stimulating, user-friendly and non-obtrusive interaction methods. In this paper, we present an efficient 3D interaction technique. This technique uses a gesture recognition for 3D interaction tasks (navigation, selection, manipulation and application control).This new apprach called *Zoom-fwd*, it allows a speed and precise interaction with distant and occulded objects.*Zoom-fwd* technique is a software solution for many problems, which can be related to the hardware or even to the software ones. A user study was carried out to investigate the impact of this 3D interaction technique on time completion of the different tasks. To do this study, we compare the *Zoom-fwd* technique with another 3D interaction technique.

Assia Messaci
Centre de Développement des Technologies Avancées CDTA Alger, Algerie
Université Larbi Ben M'hidi Oum-El-Bouaghi. Algerie
E-mail: amessaci@cdta.dz

Nadia Zenati
Centre de Développement des Technologies Avancées CDTA Alger, Algerie
E-mail: nzenati@cdta.dz

Mahmoud Belhocine
Centre de Développement des Technologies Avancées CDTA Alger, Algerie
E-mail: mbelhocine@cdta.dz

Samir Otmane
IBISC, Université Evry, Université Paris-Saclay
E-mail: samir.otmane@ibisc.fr

The results show that the *Zoom-fwd* interaction technique improves the user's task completion performance during selection and manipulation of the virtual objects.

Keywords Virtual reality · 3D interaction · gesture recognition

1 Introduction

Over the past decade, we have observed a growing evolution in the digital world, smart phones, tablets and 3D visualization devices. This evolution is more and more visible with the emergence of virtual and augmented reality (Sherman and Craig, 2003) (Carmigniani et al., 2011) where the user can be completely or partially emerged in a 3D environment. Those environments must provide the possibility to interact with the virtual environment, to navigate around 3D objects and also to select and manipulate them in an interactive way. Many works proposed different classifications for a 3D interaction.The most interesting classification is to divide the 3D interaction into four tasks according to the user's need (Ouramdane et al., 2006b) (Jankowski and Hachet, 2013) (Mendes et al., 2019). The important fundamental tasks of 3D interaction are:

- 1.Navigation, it includes all motions view point of the users in virtual environments;
- 2.Selection, it is the task of target's acquisition or designation;
- 3.Manipulation, it represents the process which allows changing properties of an object or a set of objects (Ouramdane et al., 2006b);
- 4.control application which allows users to run a command for achieving and control a specific goal or objective (Ouramdane et al., 2006a).

These 3D interactions techniques are often performed by tools or devices brought by hand. Recently NUIs are presented to community. NUIs promise to reduce the barriers for computing still further, while simultaneously increasing the user's power, and to enable computing to access still further niches of use (Wigdor and Wixon, 2011) (Shatilov et al., 2019). NUIs-based approach greatly facilitates the interaction experience for user with almost no required learning phase. It is also considered as intuitive and flexible. Since, it enables users to easily customize the interface to better suit their needs and also use it with no interruption (Steinberg, 2012). With the emergence of NUIs, HCI aims to make natural interactions with computers as natural as interactions between humans. Gestures have long been considered as an interaction technique that can potentially deliver more natural, creative and intuitive methods to communicate with computers (Rautaray and Agrawal, 2015). The hand is extensively used for gesturing compared with other body parts because it is a natural medium for communication between humans and thus the most suitable tool for HCI (Hassanpour and Shahbahrami, 2009). Our challenge is to achieve a software solution to overcome the various problems using the different physical interfaces and address some issues related to 3D interaction. In other words, proposing a 3D interaction technique allowing the user to perform his task without sensors problem. Using hand gestures can be the solution while addressing the reccurring problems of 3D interaction. So, in this paper, we address some questions:

- Can software interfaces help where physical interfaces are limited?

- How to offer a user a natural and intuitive way to interact?
- What about the introduction of gestures recognition in 3D interactions?
- How to offer a user a new paradigms and 3D interaction techniques to allow him to address the interaction problems such as distant and occulded objects?

In this paper:

- An efficient 3D interaction technique (selection)is proposed. The design of this 3D interaction technique is inspired from inconvenient existing techniques in literature and limitations of physical interfaces. This technique uses a gestures recognition and allows to accurate a precise, natural, intuitive and easy to use 3D interaction.
- To compare this technique, we have implemented other existing 3D interaction techniques in literature.
- Finally, a user study was conducted to determine if the novel technique is more suitable when performing the interaction tasks.

The rest of this paper is organized as follow: Section 2 describes the related works. Section 3 describes the design and implementation of interaction techniques. Section 4 discusses the user study comparing with other techniques. Sections 5 and 6 presents and discuss the experimental results. Finally, the limitations and future works are presented in Section. 7.

2 Related Works

In this section we present a brief overview of 3D interaction techniques and hand's gestures in immersive environment.

2.1 3D Interaction techniques

Many works presented different approaches of interaction techniques in virtual environment. Some of them presented different 3D interaction techniques using metaphors (virtual hand, virtual pointer...) (Argelaguet and Andujar, 2013) and others presented a survey where they reviewed major 3D interaction techniques and their cassification. There was several classifications : in the first one interaction techniques was decomposed into subtasks, the second one proposed an alternative classification based on interaction metaphors into exocentric and egocentric techniques:these terms was used to distinguish between two fundamental frames of reference for user interaction with VEs(Virtual Environments). In the exocentric interaction users interacted with VEs from the outside (the outside-in world reference). Whereas the egocentric interaction presented in Poupyrev et al. (1998), which was the most common for immersive VEs, users interacted inside the environment (the VE embedded the user). In the following, we will list different 3D interaction techniques as well as the advantages and disadvantages of each one. Pietroszek and Lee (2019) described a virtual hand technique that was used to select and manipulate object referring to a hand metaphor. In this technique, the selection and the manipulation was possible when the virtual hand intersects the desired objects. This technique was easy and simple to understand and implement but the limitations of this technique was the selection and manipulation of distant objects. Balaa et al. (2014) presented the GO-GO technique which supported distance selection by using non-linear mapping function by translating measured distance from the user's head to his hand into the controller distance between the real and the virtual hand. This technique was limited in the case of selection and manipulation of small objects and provides an imprecise manipulation, in contrast to moving a virtual hand to the desired object. Bowman and Hodges (1997) and Balaa et al. (2014) presented the Ray-casting technique based on the pointer metaphors, which allowed the user to select and manipulate objects beyond area of normal reach.Selection and manipulation can be performed when the virtual ray hits the desired object, however, it did not provide manipulations along the Z-axis, and hence did not allow occluded objects selection. Argelaguet and Andujar (2013) presented spotlight or flashlight technique by using a cone instead of a ray. Accuracy errors that occur during distant selections were reduced. But more than one object may fall into the light cone. Argelaguet and Andujar (2013) shown also the modification of the spotlight technique, which decreases these ambiguities by providing aperture based and resizable selection cones. Lee et al. (2003) proposed to use Image Plan techniques.The user can interact with the 2D projections of the 3D objects on a virtual image plane located in front of the user. This technique allowed the interaction with distant objects but the interaction was limited only in 2D.Feiner (2003) presented a flexible pointer which was visualized by a curve.This approach was based on two handed control of the curve, whereas the vector formed by the user's hand determine the amount of curvature was determined by the orientation of each hand, this technique solved occultation problems. It was difficult, perhaps impossible, to design a single best 3D selection and manipulation techniques that fit all possible interaction scenarios. Indeed, Bowman and Hodges (1997) proposed a combining manipulation technique HOMER where the user selected an object using a ray-casting technique, and instead of the object being attached to the ray, the user's virtual hand instantly moved to the object and attached to it. The technique then switched to the manipulation mode, allowing the user to locate and rotate the virtual object. This provides true 6DOF manipulation of distant objects. However, HOMER cannot ensure precise interaction with small and occluded virtual objects. The World in Miniature (WIM) technique was discribed by Wingrave et al. (2006)and Hand (1997) which provided the user with a miniature handheld model of the VE, which was an exact copy of the VE at a smaller scale. The user could indirectly manipulate virtual objects by interacting with their representations in the WIM, in this technique precise manipulation of small objects was difficult.The Voodoo Dolls proposed and presented by Pierce and Pausch (2002) was two handed interaction technique that combined and built upon the image-plane and WIM techniques, this technique allowed users to scale the virtual objects by selecting a voodoo doll that had a size relative to the desired environment size, this interaction technique was interesting and powerful. However,it required the use of two 6-DOF devices, which increases the hardware demands on the application. Bacim et al. (2013) presented three novel 3D selection techniques (Squad, discrete zoom and continuous zoom) those 3D interaction techniques provided a progressive refinement to the VE to select objects, those techniques allowed a selection of a small object, these techniques did not provide the tools for precise selection or occulted objects selection. Bellarbi et al. (2017) presented the zoom-in technique, used to manipulate a distant objects in augmented reality, which relied on the idea that if the user zoom the real images, it made it possible to bring the real and virtual objects, closer together. While maintaining the spatial registration between the virtual objects and the real scene.This technique did not provide the precise selection(when the object was in crowded environment and a high precision was nedded to select it) and occulted selection (when the object was hidden by one or more other objects). This variety of techniques, however,it was also a source of difficulty. How did all these techniques relate to each other? Which interaction techniques should be chosen for particular task? NUI could help us to interact efficiency with VE? Using hands gestures could help the user to interact more naturally? Which recognition we should use? These questions persist and need careful reflection. We noticed that the majority of the 3D interaction techniques used explicitly or implicitly users hands, the gestures recognition should be the best modality of interaction. We will present in the next section gestures recognition methods and we will briefly introduce the used devices in detection of the user hands.

2.2 Hand gestures

There were several ways to communicate between people, their hands played an important role, they used the movement of this latters to point a person or an object. To have interaction with each other or with objects, people used their hands to : move and/or transform the state of objects. Consequently the idea of using hands to interact with machines or computers appeared. Hand's gestures or more particularly the gestures recognition facilitated communication with computers in a more natural way.
In the next part we present different gesture recognition methods and some hardware tools used to recognise differents hand's gestures. Several works presented and classified gestures recognition methods. Sagayam and Hemanth (2017) listed the techniques of hand posture and gesture recognition as well as the advantages and disadvantages of each one.Shastry et al. (2010) presented the device-based and vision-based approaches. Device-based technique used various devices or sensors to measure hands gestures movement, it contains mechanical, haptics, ultrasonic, inertial and magnetic devices. Rautaray and Agrawal (2015) presented a mechanical data gloves such as the CyberGlove was used to obtain hand motions informations. In this approach, users had to wear burdensome, uncomfortable data gloves that limited the natural movements of their hands. Haptics devices are very commonly used in our daily. Webel et al. (2008) presented multi touch gestural interactions using HMM. Kaâniche (2009) presented ultra-sonic based motion trackers, this category are composed of three kinds of device:sonic emitters that emits ultrasound, sonic discs that reflect ultrasound and multiple sensors that time the return pulse. The position and the orientation of gestures are computed based on propagation/reflection and speed of time and triangulation. Bourke et al. (2007) proposed recognition systems to detect the normal gestures that are used in our daily activities using accelerometer. Noury et al. (2003) proposed system for multimodal intuitive media browsing in which the user can learn personalized gestures. In vision-based approach, a video stream is used to recognize hand gestures. This approach is based on different gestures appearance under different point of views and conditions. Several hand gesture representations and models have been

proposed in literature. Dardas and Alhaj (2011) presented two major categories of hand gesture representation: 3D model based methods and appearance based methods as shown in Figure.1

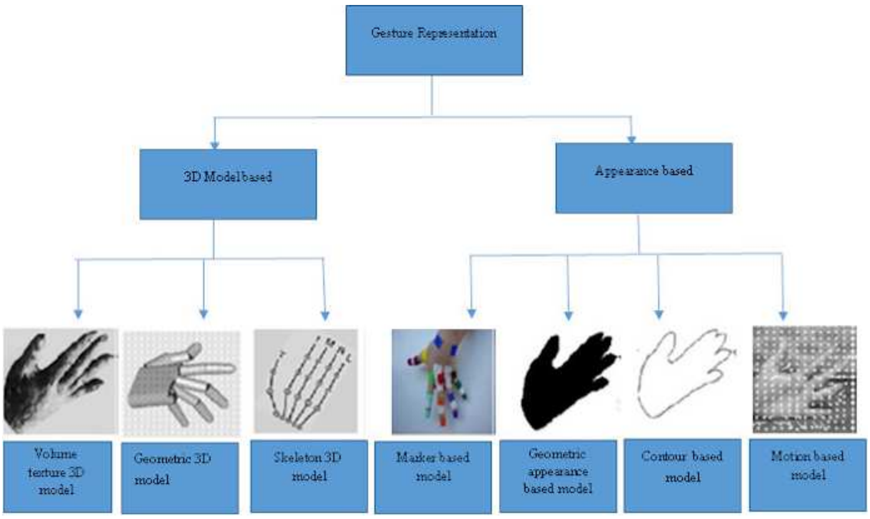

Fig. 1 Hand models, image fromDardas and Alhaj (2011)

The principle in appearance based hand gesture representation methods is to use markers easily recognizable (gloves, colors, coded targets. . .) to interact with the system. Wang and Popović (2009) proposed a gestures recognition using a colored gloves. These techniques are designed and modeled to make a dynamic gesture recognition easier using RGB camera in virtual reality, although the results shown by the authors are convincing. However, this system is designed to recognize this form of gloves.Takahashi and Kishino (1992) used the hand shape input device Data Glove and investigate the coding of the hand shape/motion information and the recognition method for the hand motion. Bellarbi et al. (2011) used a color marker in some fingers of his hand to track the motion of the hand, the markers are made by a pieces of colored paper, and they used this technique to manipulate digital documents projected on a table. However this technique allows to recognize limited static gestures for 2D interaction. To avoid these problems many approaches have been proposed to interact naturally with the machine using the hand without wearing any markers or having any devices in hand by processing images captured by a camera. Many works used a camera RGB to detect a hands gestures, the immergence of depth cameras in recent years introduced a more robust approaches to detect hands gestures. Wang et al. (2011) proposed a technique called 3 Gear, the idea beyond this technique is that we have a data base containing all possible gestures that are provided by Kinect placed above the user, the principle of this technique is to compare the gestures given by the Kinect with the ones saved in the data base in real time and do the corresponding action. With an almost similar idea,Messaci et al. (2015) used this latter to introduce gestual 3D interaction, this latter used a Kinect to detect a user gestures and they compare the user's hand gestures with the 3Gear database, they assign to each hand gestures a 3D interaction tasks. Khamis et al. (2015) proposed an automatic learning algorithm for efficient search in database of gestures in depth images, this technique has shown better results. However, it is more expensive in terms of execution time. Taylor et al. (2016) and Qian et al. (2014) presented hands modelisation using points articulations, they subsequently applied a cost function to minimize errors between reel hands and 3D models, in order to estimate the position and 3D hands gestures. Wu et al. (2016) and Tompson et al. (2014) introduced a deep learning on 3D gestures recognition by offering a precise and robust techniques, those techniques were efficient but they need a fixe position of user. Holz et al. (2015) presented a new device called "leap-motion", using for recognition and motion track of user hands, consisting of soft part and controller one (hardware). This latter is 13mm x 30mm x 76mm, has two infrared cameras and can detect and track hands in space that can be up to 50 cm. Leap-Motion has become very common in virtual and augmented reality, it is efficiency, mobility and low costing make it accessible to the general public. Many researchs and works used this latter, we cite Bacim et al. (2014) and Jiang et al. (2018) ,Jia et al. (2019), Kim and Lee (2016) and Caggianese et al. (2016). Wen et al. (2020) presented a low-cost glove that tracked the motions of human fingers,they combinied several triboelectric textile sensors and proper machine learning technique, it has great potential to realize complex gesture recognition with the minimalist-designed glove for the comprehensive control in both real and virtual space. In this part we present two important axes of researches: 3D interactions techniques and hands gestures recognition methods, in the second part we present some devices used in detection of hands gestures.In conclusion , we observed that, on one hand all the presented 3D interaction techniques are far from being perfect. These techniques present a major drawbacks, that are the lacks of precision in selection and manipulation of virtual objects (small, large, occulted, far ..) this problem has not been solved in the field of the virtual reality.On the other hand to implement these techniques user is forced to have one or more hardware tools (burden to the user like gloves, unrealiable sensors like Kinect..) that can create a constraints for this latter. To try to solve these two problems (hardware and software)we have decided to work on the software side, by trying to overcame the physical interfaces.Trying to have a 3D interaction technique which allows to react quickly and accurately to the actions of the user despite the problems related to hardware. A novel 3D interaction technique is nedded, the latter will also make it possible to solve problems that have not yet been solved in the field of 3D interaction in virtual reality, such as: interaction with far and occulted objects.

3 Design and implementation

3.1 Setup Design

Before choosing the hardware for the gestures recognition in our experimentation, we have tested several devices which are used for the detection of the user's hands. The first tool is the Kinect, which is placed above the user's hands and classic 3D interaction techniques are implemented(virtual hand and ray casting).We have noticed a fidelity during the reconstruction of the user hand gestures in real time. On the other hand, we have observed some problems related to the field of view of the Kinect. User tends to forget himself during his immersion in the virtual world and can move his hands outside the fields of view of the Kinect. The second tool is Leap Motion(Holz et al., 2015).We noticed that the tool is lightweight, portable and allows the fidelity during reconstruction of the user's hand movements in the virtual world.So, we use it in our experimentation.
Our system setup is composed of Leap Motion which is placed on the virtual reality headset called "Oculus Rift DK2" (Desai et al., 2014). This location allows the user to have the full immersion and easy to move without any constraints. Visual feedback of the gestures reproduction and the result of the control action are displayed into the headset.The setup is shown in Figure. 2

Fig. 2 Setup Design

In order to overcome all the material problems we have encountered (unreliable, field of view, calibration...),we focus on interaction technique design.

3.2 System Workflow

One important step is to identify the workflow of our global system by defining the different modules, their roles as well as the data and information that circulate between them (see Figure.3).

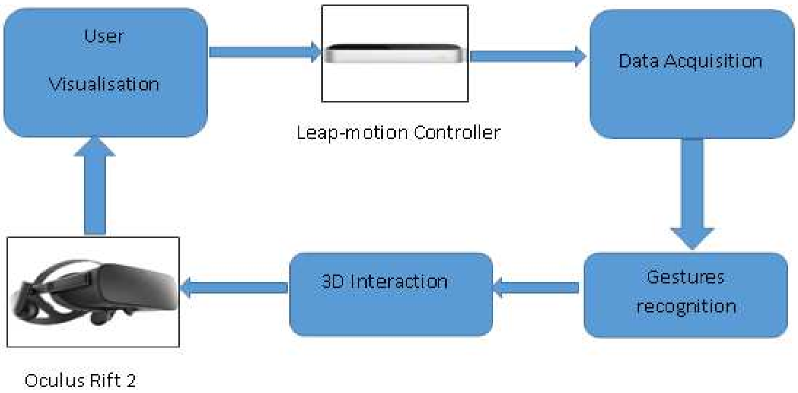

Fig. 3 Software Workflow

Data Acquisition module: This module collects data from the user; which provides from leap-motion (3D spatial information's of joints of user hands) and from oculus rift (orientation). These informations are transmitted to the gestures recognition module.

Gesture recognition module: The model-based method for the gestures recognition is used. This module uses the user hand information provided by the Leap Motion controller. Once the user hands detected, this module determines the 3D hands position and representation. To do this a 3D skeleton model of hands (see Fig. 4)is considered for each hand (left/right). The hand's joints are considered as local landmarks. The 3D position and orientation of each joint's lendmark relative to the wrist's joint lendmark. The wrist's joint is also positioned relatively to the Leap Motion controller, after receiving descriptions of the two hands (each comprising from the wrist's joint and its position) and hands joints (their positions relative to the wrist's joint). So two constraints are used for the definition of the different gestures. The first constraint is the distance between the different joints of the user's hand.The second is the orientation of the different joints. These informations are transmitted to 3D interaction module.

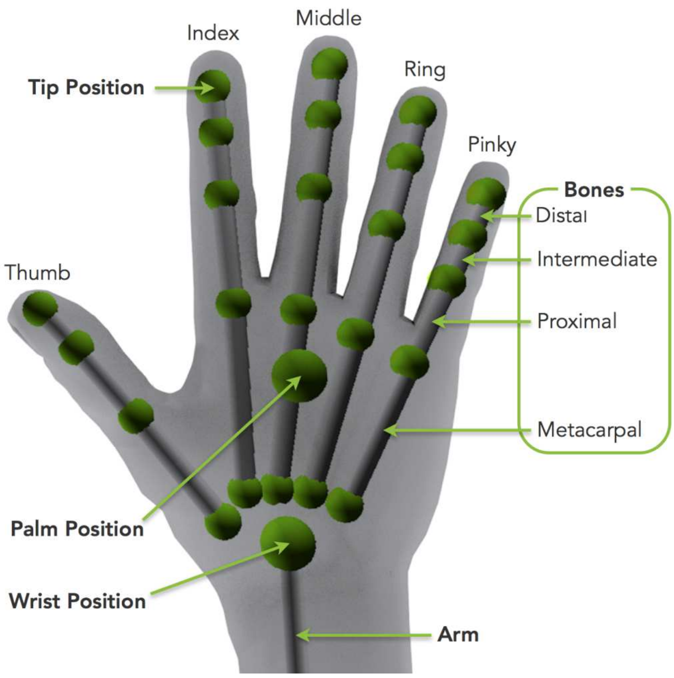

Fig. 4 Joints of hand given by leap-motion controller, image from Davis (2014)

3D interaction: This one manages the 3D graphic render and transmits it to the VEI (oculus rift) the results of the 3D interaction. To design and / or implement this part we need to know which metaphors users should use and which interactions techniques should be used to interact with the virtual environment.

3.2.1 Proposed approach

We mentioned in the related works that the problem we want to deal with is the precision of the interaction specially with a distant and occulted objects. To achieve a complete 3D interaction technique, we developed a navigation, selection, manipulation and control application techniques. Our contribution is in selecting objects. We call this technique *Zoom-fwd* because it's based on the zoom principle and we relay on it and go forward. To explain our technique, the following diagram shows the different steps.(Figure. 5)

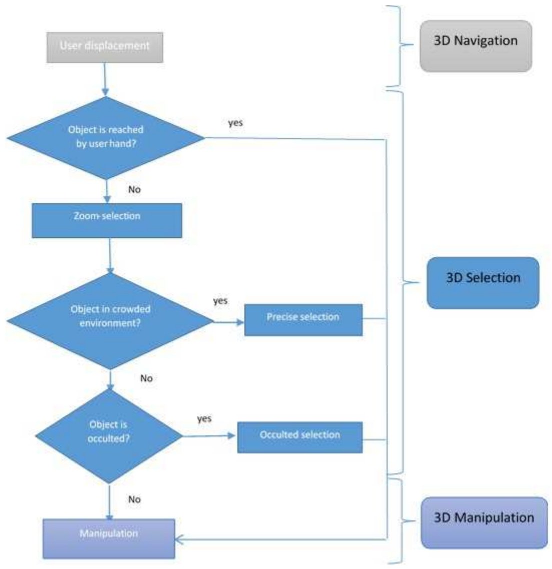

Fig. 5 Proposed approach.

1. **Navigation:** The user moves in the real environment and his movements will be effective in the virtual world as long as it is in the area covered by sensor (see Figure. 6). As this area is limited, we added the possibility to move without moving in the real world, using the gesture hand. The user can move forward, backward, on left or on right in the virtual world.

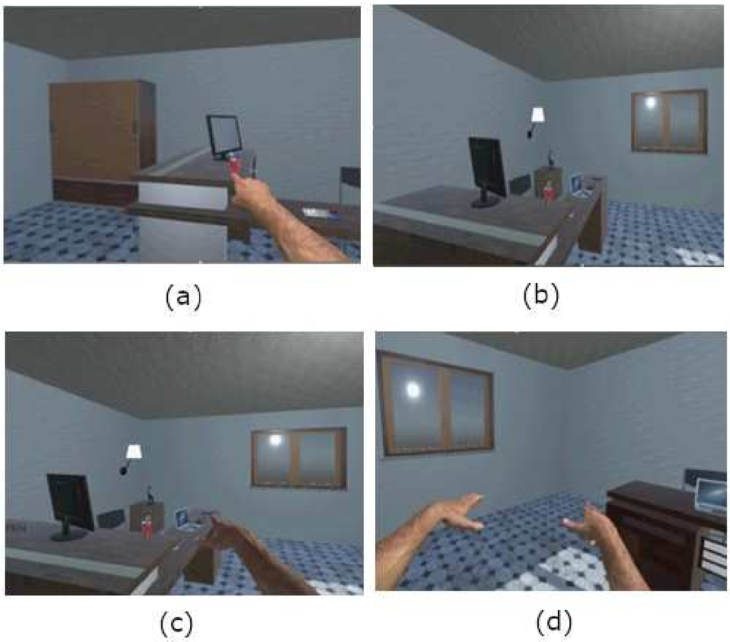

Fig. 6 Navigation in virtual word.

2. **Selection:** to select a desired object we have many steps:

   – Nearby and no occulted objects: the object is reached by the user hand. We choose virtual hand technique,
   – Far objects: the object is not reached by the user hand, then the *Zoom-fwd* technique uses different modules.

   (a) **Zoom-selection**: is a dynamic zoom depending on the distance between a user's hand and the desired object. The dynamic zoom is applied until the desired object is close enough to be within the user's hand reach (Figure. 7).

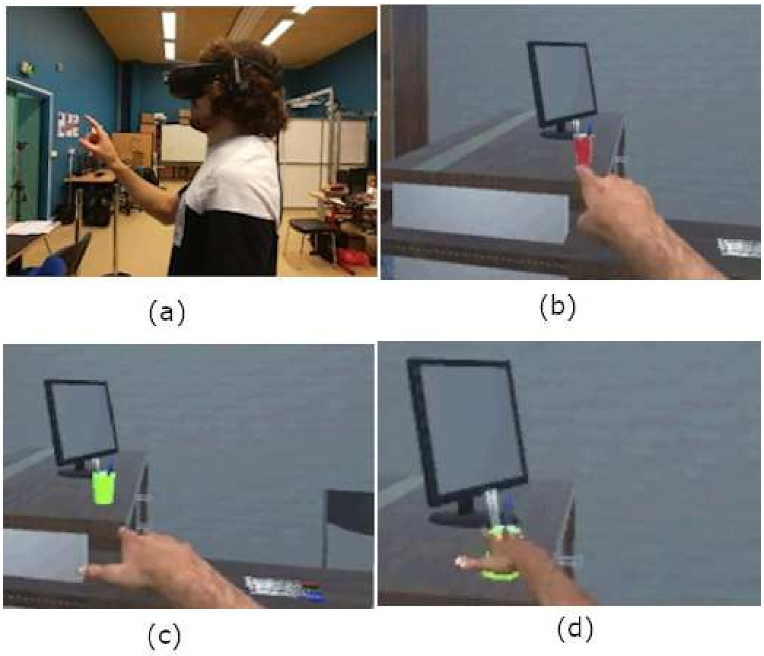

Fig. 7 *Zoom-fwd* technique steps: a) Pointing to the desired object. b) Navigation and reach the right place c) Captured image zooming. d) Grabbing and manipulating the object.

For the purpose of having the distant objet on the user's hand, the idea is to determinate the distances between the selected virtual object and the user's tracked hand. After getting the distances the process calculate dynamically a zoom ratio that has been called zoom factor "F_z", this latter is applied to the virtual camera we can present this problem as following.

- Zoom-selection functioning principle: The previous problem is formulate and shown in Figure(8).
  The C^c and C^o are the coordinate centers of the selected object O and the camera C, respectively.C^h is the coordinate of the tracked user hand H.
  CT_o represents the transformation matrix of the selected object in coordinate of the camera C^c. $Translation.z_o$ represents the Z-axis translation of the selected object obtained from the matrix CT_o.
  CT_h represents the transformation matrix of the virtual tracked hand in coordinate of the camera C^c, $Translation.z_h$ represents the Z-axis translation extracted from the matrix CT_h.
  To have the selected object within the user's tracked hand reach,we need to pass from the $Translation.z_o$ distance to $Translation.z_h$ distance, which is summarized by the factor "F_z" this factor is calculed

by the formula :

$$F_z = Translation.z_o.Translation.z_h^{-1} \quad (1)$$

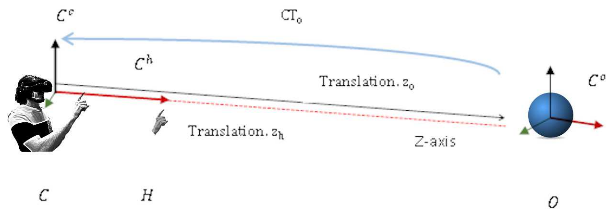

Fig. 8 Zoom-selection technique functioning principle.

Once the zoom facteur(zoom ratio) is calculated, the extrinsic parameters of the camera are updated dynamically, and this results by the following equation.

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (R, t) \begin{pmatrix} 0 \\ 0 \\ F_z \\ 1 \end{pmatrix} \quad (2)$$

with $(x, y, z)'$ is the position of the camera in the real word, (R,t) is the rotation/translation matrix(3,4) representing the extrinsic parameters of the C camera.

(b) Precise selection: one of the major problems of 3D interaction techniques is the precision in the selection of an object, when the latter is in a crowded place and his selection needs a lot of precision. The idea came from the fact that we want to highlight the neighborhood of the object which we need to select, even if it is well surrounded and concatenated to others or have different forms. *Zoom-fwd* technique offers a solution for allowing a precise selection of all the objects, even there are very fine and it is difficult to make an accurate selection. We have therefore highlighted the neighborhood of the object which we want to select by creating a bounding sphere. The latter will make it possible to define the selection space (different objects that have the risk of being selected). All the objects belonging to this sphere can be selected with the help of a hand gesture, the object touched by the hand is selected but if it is not the one that we want to select we move to the next one while moving the hand slightly. If we are far from the first selected neighborhood (we go beyond the space of the sphere) a new neighborhood is displayed and a new sphere is created. Following figure(Figure. 9) shows an example where we have a set of fine pens in a container, the user needs to make a selection of one of them with precision. One sphere is created and it englobe the differents pensils and the container. The first pen touched by the hand is selected(shape is shown, if it isn't the right pen we shift to the right or to the left. Then we select the right one. It has been noticed that this technique allows the user a very precise selection during a very crowded environement and with different forms of objets.

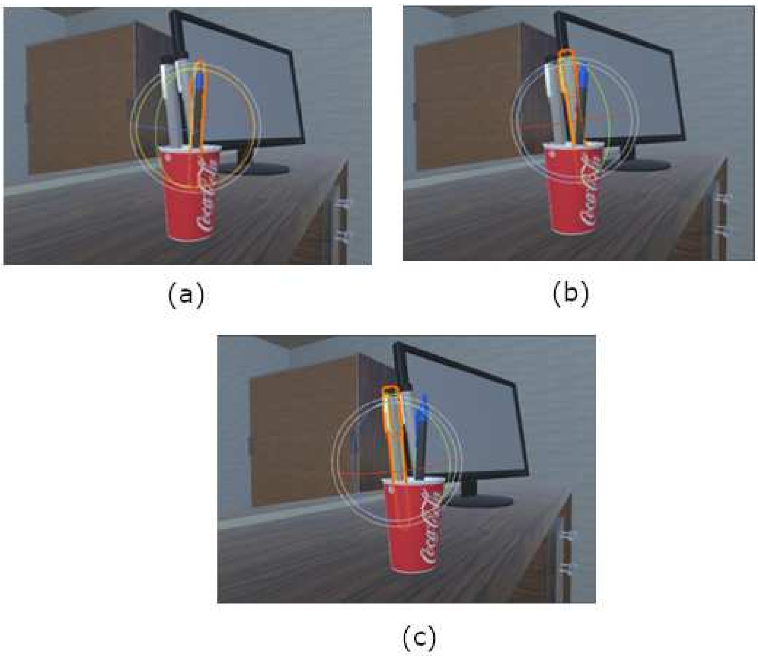

Fig. 9 Precise selection.

(c) Occulted selection: Another major problem of the 3D interaction in VR is the occultation. The hidden object can be partially or completely occulded by another object. This problem makes a selection task very slow or even impossible.
The *Zoom-fwd* technique allows us to display the posture or the shape of the hidden object if they exist(if behind object there is nothing so it displays nothing). Then the user can have the exact position of the hidden one. This

feature is very useful when the object is completely hidden(occulded by a larger object) or cannot be directely seen by the user (the point of view of the user and/or his position hides the object that we want to select). Figure. 10 shows an exemple where we have a computer screen that hiddes completely a container (the object that we want to select). The functionality of the *zoom-fwd* technique allows the slightly displaying the shape of the container when the screen touched by the user's hand(by using selection technique). It has been noticed that this feature allows the user first to see if there are objects completely occulded,and second to have their exact positions what leads us to save a lot of time in terms of achievement the tasks of 3D interaction.

Fig. 10 Occulted selection.

3. **Manipulation:** After using the *Zoom-fwd* technique to select the desired object, the next task is the manipulation. We choose the virtual hand metaphor which is used to move the object from the start position to the desired position.
4. **Control application:** The last task of 3D interaction process is control application. Using hand gesture during this task allowed us to save a lot of time. We apply this latter by validating the passage from step to step and/or task to task.

For our implementation, three different gestures are used: poiting, L gesture and pinch. The pointing gesture uses index finger to do the selection. The L gesture uses two fingers(index and thumb)allowing to stop the selection. The pinch gesture allows grabbing the selected object.

3.3 Research hypothesis

To evaluate the proposed technique, we have implemented a similar selection-manipulation technique, which is ”Homer” technique.We have chosen ”Homer” 3D technique because after having made the state of the art, it is the one which is widely used to interact with distant and hidden 3D objects.We have implemented this latter under the same conditions as those of *Zoom-fwd*. By comparing the two 3D techniques, our main research question is whether the proposed technique improves the 3D interaction of the user in the virtual world. Besides this main question, we want to investigate whether the proposed 3D interaction technique will also influence the user's subjective experience through the usefulness and the satisfaction. Therefore, we have defined the main working hypothesis for our experiment:

- H1) Users will perform the selection-manipulation tasks better using the *Zoom-fwd* technique.
  We think that the *Zoom-fwd* technique will allow users to be faster and more accurate when performing the selection tasks, because it will grant them to perform the selection tasks with more precision and rapidly.

To investigate this hypothesis, we have conducted the following user study.

4 User study

4.1 Participants

Fifteen volunteering participants (8 males, 7 females) students and colleagues were enrolled in this study (N = 15). They are all naive users with limited experience with VR tools (kinect or leap motion). The mean age was 34.5. Twelve of them were right handed. All of them had normal or corrected to normal vision. Nine of them had a previous experience with video games (including smartphone games), with 6 of them playing video games regularly. Twelve of them reported a previous experience with 3D VEs before this experiment.

4.2 Experimental design

A between-subject experimental design was used, with one independent factor (interaction technique). Therefore, all participants performed the experimental task twice, once using each interaction technique. To counter balance for avoiding any learning effects, eight participants started the experiment with the Homer techniques N1, and seven participants started with the *Zoom-fwd* technique N2.

4.3 Apparatus

The natural system using leap motion was developed using Unity3D (version 5.7) with C, because it permits an easy integration of the various used devices(leap interface, etc.). The VE consists of a virtual office (modeled in Blender version

2), the used leap motion SDK 2.0 running on PC with i72GHz Intel® Core™ and 6Go of RAM.

4.4 Experimental task

The carried out task by the participants consisted of doing several tasks using two different 3d interaction techniques. The goal to achieve by the participants was to select or manipulate different objects using two different interaction techniques. We defined different tasks of selection and manipulation: three scenarios of selection and one of the manipulation task.

Scenarios of selection:

- Select a tablet on desk (T1: see Figure. 11)
- Select a red book from the shelf (T2: see Figure. 12)
- Select a hidden statuette (T3: see Figure. 13)

Scenarios of Manipulation:

- Manipulate a book from the shelf to the desk and put it on another one (T4: Move a book from the shelf to the desk and put it on another)

Fig. 11 Tablet selection.

Fig. 12 Book selection.

Fig. 13 Statuette selection.

As we said in section 4.2, two groups of participants who do the selection tasks in different orders. The first group (seven participants) was instructed to select a tablet on the desk, to select the red book from the shelf, and to select the hidden statuette and move the book from the shelf to the desk and put it on another. This first group used T1, T2, T3 and T4 starting by the Homer techniques then using *Zoom-fwd* technique. The second group (eight participants) was instructed to select first a red book from the shelf. Select the tablet from the desk, select hidden statuette and move the book from the shelf and put it on another. This second group used T2, T1, T3 and T4 starting by the *Zoom-fwd* technique then using Homer technique to avoid eventual learning of different tasks.
Based on our task analysis, the system was designed to compare two 3d interactions techniques, one existing technique and the proposed one. To compare these techniques, we defined different tasks explained in section 4.3, the tasks were chosen to put in evidence different challenges related to the 3D interaction techniques. The first task of selection is rather simple, the instructed users select tablet on the desk in front of them this task allows the users to be familiar with the different gestures.
The second was more complex, the instructed users select the red book from the shelf, this latter are middle of a lot of books and its selection requires precision which is one of the challenge of 3D interaction techniques.
The third task was the more complex one, the users was instructed to select a hidden statuette. This latter is hidden by different books and its selection allows to deal with another challenge of the 3D interaction in VR which is occultation.The last task, is a manipulation one, the users are instructed to select a red book from the shelf and put it on the desk on another one. A priori, the participants did not know the difficulties of tasks, the participants did not know where the different objects to select or manipulate were located.
So they started each trial by visiting the room, be familiar with the different gestures and choose the best point of view and the best way to perform the desired tasks. This experimental trial was then repeated two times for each participant. The participants were instructed to perform the task as quickly as possible (to evaluate the completion time).

4.5 Experimental procedure

The experimental procedure started with filling the demographics questionnaire. After that, the participant was presented with the set-up, a detailed tutorial was presented. It consisted of a series of steps in the form of short instructions and explications on how to perform the different tasks, how to use the leap-Motion device, and how to use the gestures to navigate, select and manipulate different objects. This training phase was alternated with short manipulation phases for familiarization with the set-up. All the steps could be repeated if necessary, and permitted to progressively understand each step of the task to perform. The last step of the training consisted of performing a whole selection/manipulation procedure. To guarantee that the participant had well understood the experimental task. Once the tutorial was finished, the experimental session started for the first group with the first technique.
It consisted of performing two trials of the tasks. Finally, the participant was asked

to fill in a questionnaire to subjectively evaluate the selection/manipulation technique. This procedure was repeated for the second selection/manipulation technique, and finished with a comparison questionnaire, and a global system usability questionnaire.

4.6 Measurements and data analyses

To compare the two techniques, both objective and subjective measurements were recorded. The user performance was evaluated through the completion time of selection tasks (three tasks because it's the most important part of the interaction process) and the time spent in the manipulation task. The time calculation for the task started once the user began to use the navigation gesture to once the task is finished. In addition to the objective measures , we have collected subjective data. It consisted of responses to a questionnaire for each technique, using a 5-point Likertscale (see2). The questions included different criteria: utility, satisfaction and learning (difficult to use the technique).
The questions (Q1-Q9) were extracted from the USE (Usefulness, Satisfaction and Ease) questionnaire (Lund, 2001), we have also asked the participants to indicate which technique was easier to use, had a better performance with, which one was specifically preferred for performing all the tasks? Finally, the System Usability Questionnaire (SUS) (Brooke, 1996) was proposed to obtain a general usability score of the techniques.

All data analyses were performed using R version 3.3.3 (R Core Team, 2017) using RStudio (RStudio Team, 2016, Boston, MA) with the appropriate statistical tests. We have used a confidence level of 95% for all our statistical analyses. Therefore, a result is considered significant when $\alpha < 0.05$.

5 Results

First, the collected data was analyzed to determine whether parametric or non-parametric tests can be used. We have checked the normality assumption of the data through the Shapiro-Wilks test on the data completion time. The results indicate that all data don't follow a normal distribution. Therefore, the paired non-parametric Wilcoxon Signed Rank test was used to compare the means values. The non-parametric Friedman rank sum test was used to compare the mean scores of the subjective questionnaire data (ordinal data). Results are summarized in table 1.

Table 1 Descriptive and statistical analyses for the objective data.

	Mean time		**Wilcoxon-test**	
Tasks	**Homer**	***Zoom-fwd***	**V**	**p-value**
Tablet selection	7.17	2.23	120	**.0007247**
Book selection	9.25	2.66	120	**.0007229**
Statuette selection	3.07	0.75	120	**.0007229**
Book manipulation	3.64	1.21	104	**.001367**

5.1 Time

The paired-Wilcoxon test shows a significant effect of the technique on the mean completion times for the first task [V = 120, p-value = 0.0007247]. Participants performed the task faster with the *Zoom-fwd* (69% less time). The same test shows a significant effect of the technique on the mean completion times for the second task [V = 120, p-value = 0.0007229]. Participants spent less time in the *Zoom-fwd* technique (71% less time). The test shows a significant effect of the technique on the mean completion times for the third task [V = 120, p-value = 0.0007229]. Participants spent less time in the *Zoom-fwd* technique (76% less time). The test shows a significant effect of the technique on the mean completion times for the fourth task [V = 104, p-value = 0.001367]. Participants spent less time in the *Zoom-fwd* technique (67% less time). The graph which is a boxplot Figure. 14 clearly shows us the superiority of the *Zoom-fwd* technique compared to the Homer technique for all the tasks.

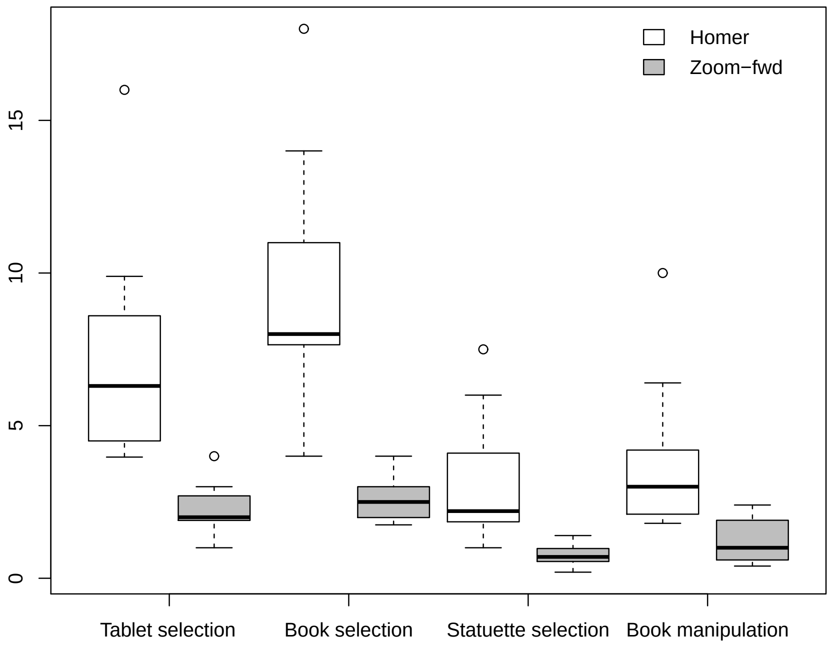

Fig. 14 Time completion of interaction tasks.

5.2 Subjective data:

The analysis of grouped questions through the non-parametric Friedman tests show a significant effect of the techniques on the criteria: Utility and satisfaction (see Figure. 15),on the other hand, the same test show a non significant effect of the techniques on the ease of learning criteria. We have precisely observed a significant effect on the participant's mean scores for the most questions (Q1, Q2, Q3, Q7, Q8, Q9) and no significant effects were found for the other questions.Those results are shown in Figure. 16.

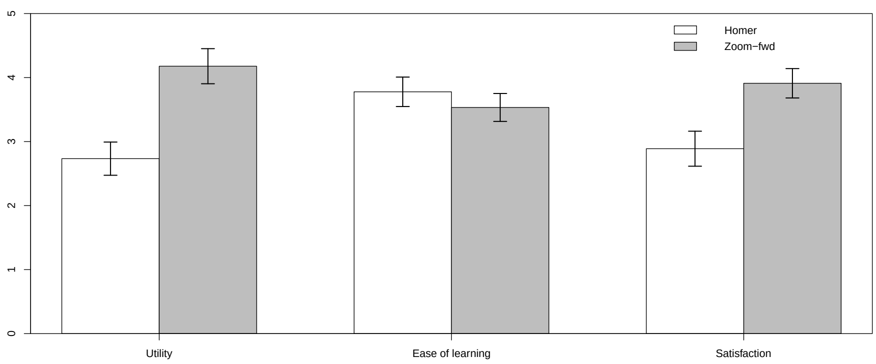

Fig. 15 Grouped subjective results by criteria

In addition, the results show that 62% of the participants found the *Zoom-fwd* technique easier to use, there was 55% of participants prefer to use the *Zoom-fwd* technique and had more performance with it and prefer to end all the tasks with this technique(see Figure. 15). Finally, the SUS score reports a mean value of 70.22%, which stands for a grade C ("Acceptable" we are very close to the "Good").

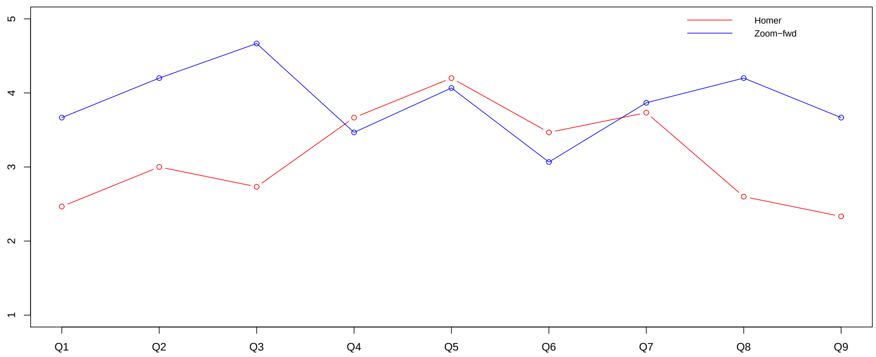

Fig. 16 Scores for each question

Those results are summerize in table2.

Table 2 Items of the post experimental questionnaire and means of scores given for each technique.

Category	ID	Question	Homer	*Zoom-fwd*
Utility	Q1	It is useful?	2.46	3.66
	Q2	It saves me time when I use it?	3	4.2
	Q3	It does everything I would expect it to do?	2.73	4.66
Ease of learning	Q4	It is easy to learn to use it?	3.66	3.46
	Q5	I learned to use it quickly?	4.2	4.06
	Q6	I easily remember how to use it?	3.46	3.06
Satisfaction	Q7	I feel I need to have it?	3.73	3.86
	Q8	I am satisfied with it?	2.6	4.2
	Q9	It is pleasant to use?	2.33	3.66

6 Discussion

In this section, we discuss the results of our experimental study and their implications for the design of the *Zoom-fwd* technique. The results show that using the *Zoom-fwd* technique facilitated the completion of the selection tasks as compared to Homer technique. This is observed through the decrease of the total time spent to perform each task, which was significantly lower in this technique. In addition, the participants spent nearly the same time for each tasks using *Zoom-fwd* technique as compared to Homer technique. This reflects that the first technique is more stable during the selection tasks. For the second selection task (book selection) the time completion using the *Zoom-fwd* technique is practically reduced then with the Homer technique for all the users. This shows us that *Zoom-fwd* technique is more recommended to do the precision tasks which is one of the challenge of 3D interaction techniques. In the third task which is the more complex one (select a hidden statuette), once again the time completion with the *Zoom-fwd* technique is lower than that with Homer technique for all the users, this shows us that the *Zoom-fwd* technique is better for situations with hidden or occulded objects. For the manipulation task, the time spent by the participants using *Zoom-fwd* technique is less than the time spent using Homer technique. This is also supported by the subjective comparative questionnaire on the preferred technique to perform the different tasks (55% of the participants preferred the *Zoom-fwd* technique). In addition the results show that the completion time of the different tasks is very close for all users, which lead us to conclude that the *Zoom-fwd* technique is more stable than Homer technique. Therefore, H1 is validated. Finally, in terms of usability, the SUS score reports a mean value of 70.22%, which encourages the made choices for the design of this 3D interaction technique.

7 Conclusion

We presented in this paper a novel 3D interaction technique called *Zoom-fwd*.This technique uses a gesture recognition for 3D interaction tasks (navigation, selection, manipulation and application control).The *Zoom-fwd* technique allows a faster and more accurate interaction with distant and occulted objects. Our main objective was to propose 3D interaction technique that will try to overcome the found problems using the different physical interfaces. In addition, the proposed technique uses the gestures recognition which makes it very natural and easy to use. The results of our user study show that the proposed technique offers a better task completion performance. This increase in performance was more particularly observed during the precision and occultation tasks (book selection and manipulation, statuette selection).It was concluded that this work has added value in the field of 3D interaction, by offering a new 3D interaction technique which saves more time and accuracy.

8 Limitations and future work

According to the subjective results we noticed that some users find the proposed 3D interaction technique *Zoom-fwd* is difficult to use, it can be one of the major problem. It was also noticed that despite the effort to perform the evaluation, we will always be able to improve it: the sample remains insufficient, even the tasks remain basic, so far we can evaluate more parameters.
As we are interested in another part of our research to collaborative work or more precisely to the collaborative 3D interaction in a virtual environment we will be able to test this technique in order to improve the comanipulation between two collocated users.
Finally, for this study, a 3D interaction technique is proposed, a 3D interaction technique remains an open research field. A future step would be to study the added value of using this technique for two or more users.

9 Ethics statement

The lack of institutional ethics committee caused that the experimental protocol was not approved by ethics committee. An informed written consent was obtained from all the subjects participating in this study prior to their participation.

10 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Argelaguet F, Andujar C (2013) A survey of 3d object selection techniques for virtual environments. Computers & Graphics 37(3):121–136

Bacim F, Kopper R, Bowman DA (2013) Design and evaluation of 3d selection techniques based on progressive refinement. International Journal of Human-Computer Studies 71(7-8):785–802

Bacim F, Nabiyouni M, Bowman DA (2014) Slice-n-swipe: A free-hand gesture user interface for 3d point cloud annotation. In: 2014 IEEE Symposium on 3D User Interfaces (3DUI), IEEE, pp 185–186

Balaa E, Raynal M, Issa YB, Dubois E (2014) Classification of interaction techniques in the 3d virtual environment on mobile devices. In: International Conference on Virtual, Augmented and Mixed Reality, Springer, pp 3–13

Bellarbi A, Benbelkacem S, Zenati-Henda N, Belhocine M (2011) Hand gesture interaction using color-based method for tabletop interfaces. In: 2011 IEEE 7th International Symposium on Intelligent Signal Processing, IEEE, pp 1–6

Bellarbi A, Zenati N, Otmane S, Belghit H, Benbelkacem S, Messaci A, Hamidia M (2017) A 3d interaction technique for selection and manipulation distant objects in augmented reality. 2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B) pp 1–5

Bourke A, Obrien J, Lyons G (2007) Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. Gait & posture 26(2):194–199

Bowman DA, Hodges LF (1997) An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In: Proceedings of the 1997 symposium on Interactive 3D graphics, pp 35–ff

Brooke J (1996) SUS: a "quick and dirty" usability scale. Usability evaluation in industry 189(194):4–7

Caggianese G, Gallo L, Neroni P (2016) An investigation of leap motion based 3d manipulation techniques for use in egocentric viewpoint. In: International Conference on Augmented Reality, Virtual Reality and Computer Graphics, Springer, pp 318–330

Carmigniani J, Furht B, Anisetti M, Ceravolo P, Damiani E, Ivkovic M (2011) Augmented reality technologies, systems and applications. Multimedia tools and applications 51(1):341–377

Dardas NH, Alhaj M (2011) Hand gesture interaction with a 3d virtual environment. The Research Bulletin of Jordan ACM 2(3):86–94
Davis (2014) Getting started with the leap motion sdk
Desai PR, Desai PN, Ajmera KD, Mehta K (2014) A review paper on oculus rift-a virtual reality headset. arXiv preprint arXiv:14081173
Feiner AOS (2003) The flexible pointer: An interaction technique for selection in augmented and virtual reality. In: Proc. UIST'03, pp 81–82
Hand C (1997) A survey of 3d interaction techniques. In: Computer graphics forum, Wiley Online Library, vol 16, pp 269–281
Hassanpour R, Shahbahrami A (2009) Human computer interaction using vision-based hand gesture recognition. Journal of Computer Engineering 1:3–11
Holz D, Hay K, Buckwald M (2015) Electronic sensor
Jankowski J, Hachet M (2013) A survey of interaction techniques for interactive 3d environments
Jia J, Tu G, Deng X, Zhao C, Yi W (2019) Real-time hand gestures system based on leap motion. Concurrency and Computation: Practice and Experience 31(10):e4898
Jiang X, Xiao ZG, Menon C (2018) Virtual grasps recognition using fusion of leap motion and force myography. Virtual Reality 22(4):297–308
Kaâniche M (2009) Gesture recognition from video sequences. Theses, Université Nice Sophia Antipolis
Khamis S, Taylor J, Shotton J, Keskin C, Izadi S, Fitzgibbon A (2015) Learning an efficient model of hand shape variation from depth images. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2540–2548
Kim M, Lee JY (2016) Touch and hand gesture-based interactions for directly manipulating 3d virtual objects in mobile augmented reality. Multimedia Tools and Applications 75(23):16,529–16,550
Lee S, Seo J, Kim GJ, Park CM (2003) Evaluation of pointing techniques for ray casting selection in virtual environments. In: Third international conference on virtual reality and its application in industry, International Society for Optics and Photonics, vol 4756, pp 38–44
Lund AM (2001) Measuring usability with the use questionnaire12. Usability interface 8(2):3–6
Mendes D, Caputo FM, Giachetti A, Ferreira A, Jorge J (2019) A survey on 3d virtual object manipulation: From the desktop to immersive virtual environments. In: Computer graphics forum, Wiley Online Library, vol 38, pp 21–45
Messaci, Zenati, Bellarbi, Belhocine (2015) 3d interaction techniques using gestures recognition in virtual environment. In: 2015 th International Conference on Electrical Engineering (ICEE), pp 1–5, DOI 10.1109/INTEE.2015.7416813
Noury N, Barralon P, Virone G, Boissy P, Hamel M, Rumeau P (2003) A smart sensor based on rules and its evaluation in daily routines. In: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No. 03CH37439), IEEE, vol 4, pp 3286–3289
Ouramdane N, Davesne F, Otmane S, Mallem M (2006a) Evaluation of the follow-me technique for grabbing virtual objects in semi-immersive virtual environment
Ouramdane N, Otmane S, Davesne F, Mallem M (2006b) Follow-me: a new 3d interaction technique based on virtual guides and granularity of interaction. In: Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications, ACM, pp 137–144

Pierce JS, Pausch R (2002) Comparing voodoo dolls and homer: exploring the importance of feedback in virtual environments. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 105–112

Pietroszek K, Lee N (2019) Virtual hand metaphor in virtual reality.

Poupyrev I, Ichikawa T, Weghorst S, Billinghurst M (1998) Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. In: Computer graphics forum, Wiley Online Library, vol 17, pp 41–52

Qian C, Sun X, Wei Y, Tang X, Sun J (2014) Realtime and robust hand tracking from depth. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1106–1113

Rautaray SS, Agrawal A (2015) Vision based hand gesture recognition for human computer interaction: a survey. Artificial intelligence review 43(1):1–54

Sagayam KM, Hemanth DJ (2017) Hand posture and gesture recognition techniques for virtual reality applications: a survey. Virtual Reality 21(2):91–107

Shastry KR, Ravindran M, Srikanth M, Lakshmikhanth N, et al. (2010) Survey on various gesture recognition techniques for interfacing machines based on ambient intelligence. arXiv preprint arXiv:10120084

Shatilov KA, Chatzopoulos D, Lee LH, Hui P (2019) Emerging natural user interfaces in mobile computing: A bottoms-up survey. arXiv preprint arXiv:191104794

Sherman WR, Craig AB (2003) Understanding virtual reality: interface, application and design

Steinberg G (2012) Natural user interfaces. In: ACM SIGCHI conference on human factors in computing systems

Takahashi T, Kishino F (1992) A hand gesture recognition method and its application. Systems and Computers in Japan 23(3):38–48

Taylor J, Bordeaux L, Cashman T, Corish B, Keskin C, Sharp T, Soto E, Sweeney D, Valentin J, Luff B, et al. (2016) Efficient and precise interactive hand tracking through joint, continuous optimization of pose and correspondences. ACM Transactions on Graphics (TOG) 35(4):143

Tompson J, Stein M, Lecun Y, Perlin K (2014) Real-time continuous pose recovery of human hands using convolutional networks. ACM Transactions on Graphics (ToG) 33(5):169

Wang R, Paris S, Popović J (2011) 6d hands: markerless hand-tracking for computer aided design. In: Proceedings of the 24th annual ACM symposium on User interface software and technology, ACM, pp 549–558

Wang RY, Popović J (2009) Real-time hand-tracking with a color glove. ACM transactions on graphics (TOG) 28(3):63

Webel S, Keil J, Zoellner M (2008) Multi-touch gestural interaction in x3d using hidden markov models. In: Proceedings of the 2008 ACM symposium on Virtual reality software and technology, ACM, pp 263–264

Wen F, Sun Z, He T, Shi Q, Zhu M, Zhang Z, Li L, Zhang T, Lee C (2020) Machine learning glove using self-powered conductive superhydrophobic triboelectric textile for gesture recognition in vr/ar applications. Advanced Science p 2000261

Wigdor D, Wixon D (2011) Brave NUI world: designing natural user interfaces for touch and gesture. Elsevier

Wingrave CA, Haciahmetoglu Y, Bowman DA (2006) Overcoming world in miniature limitations by a scaled and scrolling wim. In: 3D User Interfaces (3DUI'06),

IEEE, pp 11–16
Wu D, Pigou L, Kindermans PJ, Le NDH, Shao L, Dambre J, Odobez JM (2016) Deep dynamic neural networks for multimodal gesture segmentation and recognition. IEEE transactions on pattern analysis and machine intelligence 38(8):1583–1597