applied sciences

MDPI

Article

3D User Interface Design and Usability for Immersive VR

Hind Kharoub, Mohammed Lataifeh and Naveed Ahmed *

Department of Computer Science, University of Sharjah, University of Sharjah, Sharjah P.O. Box 27272, UAE; hkharoub@sharjah.ac.ae (H.K.); mlataifeh@sharjah.ac.ae (M.L.)

* Correspondence: nahmed@sharjah.ac.ae

Received: 3 October 2019; Accepted: 8 November 2019; Published: 13 November 2019

check for updates

Featured Application: Advanced mixed realities 3D interaction, multimodality complex deployments of 3D spatial interactions.

Abstract: This work presents a novel design of a new 3D user interface for an immersive virtual reality desktop and a new empirical analysis of the proposed interface using three interaction modes. The proposed novel dual-layer 3D user interface allows for user interactions with multiple screens portrayed within a curved 360-degree effective field of view available for the user. Downward gaze allows the user to raise the interaction layer that facilitates several traditional desktop tasks. The 3D user interface is analyzed using three different interaction modes, point-and-click, controller-based direct manipulation, and a gesture-based user interface. A comprehensive user study is performed within a mixed-methods approach for the usability and user experience analysis of all three user interaction modes. Each user interaction is quantitatively and qualitatively analyzed for simple and compound tasks in both standing and seated positions. The crafted mixed approach for this study allows to collect, evaluate, and validate the viability of the new 3D user interface. The results are used to draw conclusions about the suitability of the interaction modes for a variety of tasks in an immersive Virtual Reality 3D desktop environment.

Keywords: VR interfaces; mixed realities; adaptive VR interactions

1. Introduction

The field of Virtual Reality (VR) has been an area of interest for both research and development for the past five decades. With the advancement of technology, specifically miniaturization of computing hardware and improved display technology, the field is finally going mainstream [1]. Indeed, the advancement of consumer grade VR devices, e.g., Oculus Rift, HTC Vive, Gear VR etc., has opened new ways for people to interact with technology. In general, the user-interaction paradigm for VR is an extension of how people interact with the computing systems using a graphical user interface (GUI). The concept of GUI revolves around the concept of direct manipulation (DM), where the user uses some input device to directly manipulate the content on a two-dimensional (2D) display, and receives an immediate feedback of its action, e.g., point and click, dragging, etc.

The VR on the other hand is a different medium, where the content is not only displayed in a true three-dimensional (3D) environment, but also because of the VR device, the user loses the ability to directly observe the input devices, e.g., mouse and the keyboard, which renders the current DM user interface paradigm not well suited for 3D immersive VR environments. Therefore, for an immersive VR environment, it is necessary to explore different types of user interface paradigms that are better suited for natural user interaction (NUI) without means of a physical input device. A natural user interface (NUI) allows people to interact with the technology without the need of intermedial devices for the user interaction; these interactions rather take place directly using hand gestures or body movements. Thus, it is very important to understand how an NUI can be used in a 3D user interface as an effective means of Human Computer Interaction (HCI).

Moreover, a number of new input devices, e.g., Kinect, Leap Motion, Holo Lens, Magic Leap, etc. are ushering in the new era of HCI systems. These devices allow the user to control any 3D user interface using hand or body gestures, be it on a desktop system or a VR environment. A 3D User Interface (UI) design is now becoming a critical area for developers, students, and researchers to understand. The use of VR visualization and natural interactive gestures can increase the user's engagement, which results in enhanced usability and user experience [2].

In the past decade, the use of VR has increased tremendously. It has been employed in a number of industries and is powering many different types of interactions. A number of novel and innovative user interfaces have been created for VR applications. These interfaces have changed the way people interact with VR environments and result in truly unique experiences.

A 3D VR application should be useful compared to a 3D non-VR application because VR results in higher immersion, uses natural skills, and provides immediacy of visualization. However, one of the limitations of many current 3D VR applications is that they either support only simple interactions or have serious usability problems [1].

3D UI design is an area ripe for further work, the development of which is often referred to as new exciting areas of research in HCI today, providing the next frontier of innovation in the field [1]. There are already some applications of 3D UIs used by real people in the real world (e.g., walkthroughs, psychiatric treatment, entertainment, and training). Most of these applications, however, contain 3D interaction that is not very complex. More complex 3D interfaces (e.g., immersive design, education, complex scientific visualizations) are difficult to design and evaluate, leading to a lack of usability. Better technology is not the only answer—for example, 30 years of virtual environment (VE) technology research have not ensured that today's VEs are usable [1]. Thus, a more thorough treatment of this subject is needed [1].

Within a 3D human-computer interaction mode, user's tasks are performed in a 3D environment, and the spatial context is extended to the third dimensions. 3D interaction involves either 3D input devices (such as hand gesture or physical walking) or 2D input devices (such as mouse) mapped into a 3D virtual location. 3D UIs as the intermediary between the human and the machine become the medium where different users directly perform the tasks within the spatial context of a 3D domain. In carrying out different tasks, users' commands, questions, intents, and goals are communicated to the system, which in turn must provide feedback with information about the status and flow for the execution of the tasks [1].

As the interaction mode is an important aspect in a 3D environment, more natural interaction modes need to be explored. One type of NUI is the gesture-based interaction. When a person uses hands, arms, face, or body to convey some information, it is deemed as a gesture. In the first place, a gesture is recognized by tracking the motion of hands, arms, face, or body. Afterward, the gesture is interpreted and mapped to a meaningful action within the context of the application. Saffer [3] has identified the following advantages of gestural interfaces against traditional interfaces: (1) more natural interactions; (2) less heavy or visible hardware; (3) more flexibility; and (4) more fun. Gesture recognition can be used in many areas. Some of the application domains that employ gesture interaction are VR, gaming, desktop and mobile applications, telepresence, sign language, and robotics. This work only focuses on hand gestures.

The UI design and usability depend a lot on the type of input device used in a 3D UI. Additionally, the display characteristics can affect 3D interaction performance, for example, when employed for VR, Head Mounted Displays (HMDs) block out the visibility of one's hands, body, and the overall surroundings. If the user cannot see the input device then using it with good hand eye coordination is a challenging task. Therefore, it is imperative to develop a user interface that does not require the actual visibility of the hands or the body, because the user will not be comfortable moving within a physical environment with this limitation.

Among the variety of applications used, the desktop is the most commonly used environment used by billions of people every day. The metaphor is used widely on desktop computers and part of it is employed in the form of the home screen design of mobile devices. As VR becomes even more ubiquitous, a successful translation of the desktop environment from the computers and mobile devices to the 3D environment is imminent. Therefore, it is important to design and analyze the usability and user experience of various user interaction paradigms in an immersive VR desktop environment.

Therefore, the objective of this work is to propose an answer to the following question, what is the most suitable interaction mode for the desktop environment in immersive VR? Consequently, a new dual-layered 3D UI is proposed that allows the user to perform common desktop tasks in a VR environment. The proposed 3D UI allows the user to visualize the content on a 360-degree wide field of view, in the form of multiple curved screens. The gaze mechanism is used to switch between the screen layers and the interaction layer. The interaction layer allows the user to perform simple tasks related to a desktop environment. In order to test the new 3D UI, three user interaction modes are implemented. Finally, a comprehensive user study is performed using a mixed methods approach to validate its usability and user experience both qualitatively and quantitatively.

In order to study the feasibility of the proposed new 3D UI and draw conclusions about the suitability of the three different interaction modes for a variety of tasks in an immersive VR 3D desktop environment, the main contributions of this work are:

1. Designed a novel dual-layer 3D UI that allows the user to interact with multiple screens portrayed within a curved 360-degree effective field of view available for the user.
2. Designed a context-aware UI using the eye gaze mechanism to move between the two layers of the proposed 3D UI and for performing certain tasks.
3. Developed and implemented three different interaction modes: point-and-click interaction, controller-based DM interaction, and a gesture-based interaction to evaluate the proposed 3D UI.
4. Performed a user study using a mixed-methods approach to quantitatively and qualitatively validate the usability and user experience of the new 3D UI using three interaction modes.
5. Employed the current state of art technology, HTC Vive headset and its controllers, and Leap Motion controller for developing three interaction modes for 3D desktop environment in the immersive VR.

In the following sections, first the literature review is presented, followed by the methodology, user study, experiment setup, results, data analysis, and finally the conclusions that discuss the findings, limitations, and future work.

2. Literature Review

In this section, an extensive literature review of 3D UI in the VR environment is provided in terms of the following aspects: 3D UI, natural user interaction, DM interaction, and finally evaluation and usability of 3D UIs.

2.1. 3D User Interfaces

In a 3D UI, the user performs the interactions in a three-dimensional space naturally involving a three-dimensional movement. 3D interactions are universally classified into four main categories defined as navigation, selection and manipulation, and system control [4]. While 3D interaction allows the user to accomplish a task within a 3D UI, immersive VR systems combine interactive 3D graphics, 3D visual display devices, and 3D input devices (especially position trackers) to create the illusion that the user is inside a virtual world. The input is then used by the application to generate a 3D scene rendered from the point of view of the user and sent to the HMD [5]. Figure 1 illustrates this process while Figure 2 shows typical components of a VR system that are employed to implement the process.

3D interaction involves either 3D input devices (such as hand gesture or physical walking) or 2D input devices (such as mouse) mapped into a 3D virtual location. Recently, the arrival of off-the-shelf consumer grade motion capture devices, e.g., Kinect, or Leap Motion allows for a greater accuracy of gesture tracking at a very low cost. This has resulted in a surge in motion-controlled applications for gaming and general use.

Every application needs a special interaction concept, which depends on its working task. Kaushik and Jain [6] studied the trends in virtual interaction for a 3D UI, and issues that can arise when an NUI is employed for the user interaction. Ultimately, 3D user interfaces rely on the interaction modes as discussed in the next section.

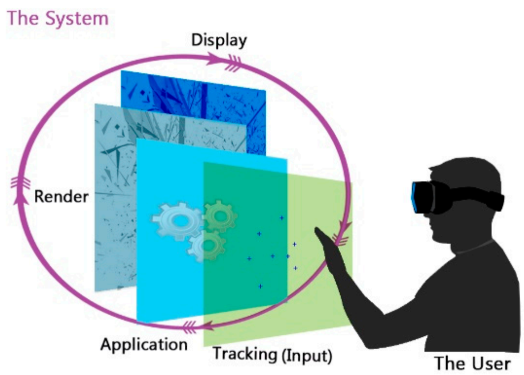


Figure 1. A typical Virtual Reality (VR) system process. The user's movement is tracked and translated into the input for the application. The output of the application is a rendered image sent to the Head Mounted Display (HMD).

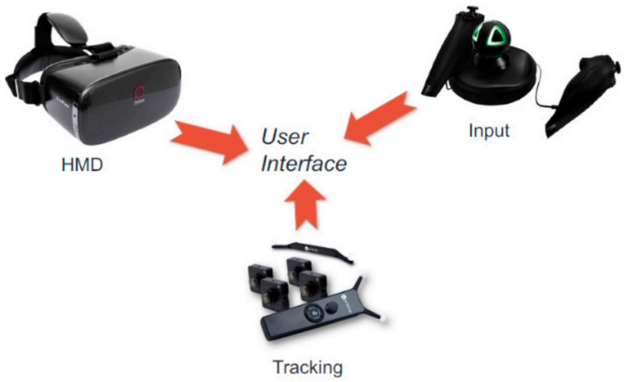


Figure 2. Typical immersive VR system components.

2.2. 3D User Interactions Modes

Gesture-based interaction coupled with VR are the two technologies that truly complement each other. As an HMD-based VR completely blocks out user's hands or body, a gesture-based interaction using hand tracking allows the user to visualize the virtual hands in the VR environment. Additionally, gestures can be easily employed in the VR to perform the tasks that would otherwise require additional interaction elements. Therefore, gestures are an obvious solution to the interaction issues that arise in VR and provides a completely device-free and touch-free solution to many interaction problems.

Ahn et al. [7] explored using slide-based presentations in a virtual environment (VE). They used an infra-red camera for hand gesture recognition, along with a 3D human body tracking system. Additionally, they also employed voice recognition in their system. Pang et al. [8] explored a system for hand gesture recognition and interaction in a real-time vision-based recognition framework. Their work allowed the controlling of the cursor in the desktop environment and they analyzed different tasks that are required to control the cursor.

Recently, the size of the display became so large that the existing methods have more difficultly controlling the application. Kang et al. [9] presented a 3D interaction mechanism using Kinect that allows to control the interface from a distance using the joint tracking information from Kinect to implement the gestures.

A hybrid system comprising of air gestures and touch gestures is presented by Bragdon et al. [10]. These hybrid interactions allow small developer group meetings using Kinect. In addition, Ebert et al. [11] also explored the applications of gestures in various medical systems for interaction tasks.

A VE allows for different types of interactions, the variety of which should be addressed intuitively. Kerefeyn and Maleshkov [12] developed a gesture-based approach using the Leap Motion controller to manipulate virtual objects in a VR environment. In addition, also using the Leap Motion, Fanini [13] presented a 3D user interface for manipulating different contents in a 3D scene. Kry et al. [14] presented a hand gesture-based interaction system that allows the direct manipulation of digital models.

Balaa et al. [15] presented two criteria for categorizing pointing approaches in 3D VEs in the domain of mobile devices. The classifications that are based on problems and factors affecting pointing performance help improve the current techniques and design new ones. Song et al. [16] presented in-air gestures to complement the touch-based input and study the impact of a touch-free user interface in the presence of a touchscreen on mobile devices. Oviatt et al. [17] presented approaches for interpreting speech and pen-based gestural input in an efficient manner. They also explored multimodal input that incorporates both gestures and speech.

Gaze detection, using eye tracking has been employed in a number of interaction mechanisms. The idea of combing gaze with the hand gestures for remote target selection in 2D interaction was presented by [18] and [19]. Zhang et al. [19] proposed a hybrid gaze and hand gesture interaction mechanism for large scale displays using Kinect. They created a photo sorting app that was tested with two mechanisms: gestures, and gaze with gestures. Their study showed that adding gaze with gestures significantly improves the usability of the interaction mechanism. In general, users preferred the gaze mechanism, but it introduced some errors that required further study.

2.3. Evaluation and Study of 3D User Interface

The design of a UI can only be validated through a user evaluation. For VR, being a new medium, a user study is crucial to determine if the UI is really responding to users' needs, and results in a positive user experience. Therefore, the user evaluation is also an active area of research for UI design, both in terms of usability and user experience, especially for VR systems [2].

Villaroman et al. [20] use Kinect for VR to establish the validity of using this setup in learning setups. They show that this user interaction can be employed in HCI courses and achieve favorable results. Kang et al. [9] has also been cited to have employed Kinect for VR and showed that using a gesture-based system in VR has 27% faster interactions compared to using a mouse. As presented in the previous section, Bragdon et al. [10] also showed that a hybrid system using gestures and touch controls can be equally effective.

Bhuiyan and Picking [21] showed the viability of using hand gestures for making telephone calls, controlling the television, and running mathematical calculation. Their work showed that a gesture-based interface can be really beneficial to elderly and disabled users. Ebert et al. [11] showed that a gesture-based image reconstruction method for a CT scan was slower compared to keyboard and mouse. On the other hand, a gesture-based method could mitigate the risk of infection for engaged personnel.

Farhadi-Niaki et al. [22] used Kinect to perform typical desktop tasks through arm gestures. They performed a comprehensive user study, to compare the gesture-based interface to a traditional mouse and keyboard input setup. Their results assured that gestures are pleasurable and closer to natural use, though may cause more fatigue. In a later work, they extended the work [23], and linked figure gestures to a simplified desktop. The gestures performed simple desktop tasks. Their study showed a preference for finger-based gestures that provided more leverage compared to the arm-based gestures in terms of the usability, users also noted to have found them closer to natural interaction.

Barclay et al. [24] presented a new model to evaluate and improve a gesture-based system in terms of its usability to derive an effective gesture-based interface. Manresa-Yee et al. [25] presented a comprehensive literature review and survey of gesture-based interfaces derived from computer vision algorithms. A number of usability factors for these gestures were considered, namely effectiveness, efficiency, and satisfaction. Cabral et al. [26] and Bhuiyan and Picking [21] presented the use of gestures in VR and many other applications. They showed that the gestures can be successfully employed in a VR environment and perform simple day to day tasks, e.g., telephone calls, controlling the television, etc.

Finally, in terms of the devices available for gesture-recognition, recent studies [27–29] have evaluated the accuracy and usability of off-the-shelf devices, e.g., Kinect, or Leap Motion. These studies show that using these devices are a viable alternative to more expensive and specialized gesture-recognition setups.

2.4. Comparison of Related Work

As a number of studies are discussed in the literature that propose 3D user interfaces and their evaluation, most of these works are not directly creating a user interface for a 3D desktop in the immersive VR. Below, six similar studies directly related to this work are discussed.

Thanyadit and Pong [30] presented a mirror metaphor for a desktop VR environment. Users can interact with a virtual object in front and behind a virtual screen using their hands. The study showed that the mirror metaphor can be effectively applied to a VR environment. Zhang [31] used leap motion and HMD-based VR to create a training system for miners. Their study showed the gesture and HMD-based VR training was more useful compared to a controller-based training. Jibin and Gao [32] presented a pen and touch input-based 3D desktop. They showed that performing 3D desktop tasks using touch gestures was more intuitive and efficient.

Farhadi-Niaki et al. [22,23] presented gesture-based system to control a common desktop using arms and fingers. Their study established the viability of using a gesture-based system for common desktop tasks and found that finger-based gestures are better suited compared to the arm-based gestures. They also concluded that gestures cause more fatigue, compared to the mouse. Song et al. [33] presented a handlebar metaphor to manipulate virtual objects. Users pose in a similar way as holding a handlebar and move the bar in a similar manner to manipulate the virtual object. Their study showed that the metaphor results in higher usability and a pleasant user experience. The method maps directly to a skill-based system, where the user improves with practice. They also observed a higher level of fatigue with the use of gestures. Finally, Cournia [34] performed a user study to establish the benefits of gaze-based pointing with respect to the hand-based pointing. They did not find one system being more advantageous over the other. From the fatigue point of view, gaze is definitely better because extending the arms causes tiredness after some time.

3. Methodology

The basic idea of this work is to evaluate a new 3D UI using both device-based and gesture-based user interactions for an immersive VR 3D desktop environment. The proposed UI is a dual-layer 3D UI that allows the user to interact with multiple screens portrayed within a curved 360-degree effective field of view available for the user. Downward gaze allows the user to raise the interaction layer that facilitates several traditional desktop tasks, such as selecting a different type of file to open, deleting a file, duplicating a file, etc. Thus, the three different interaction modes are point-and-click, controller-based DM, and gesture-based interaction. Based on each interaction mode, the three UIs are named accordingly. Then, a mixed-methods approach is used to perform a comprehensive user study for both the usability and user experience analysis of all the three UIs.

Like any system, the proposed 3D desktop system for the immersive VR has inputs (i.e., physical input device or natural gestures as system commands), and outputs that allow the user to control the 3D desktop tasks. The input will be translated through the 3D UI into actions that are used to perform a variety of tasks in an immersive VR 3D desktop. The block diagram of the proposed methodology decomposed into three stages is shown in Figure 3. The first stage is designing the 3D UI. The second stage is implementing and developing the different interaction modes. The third stage is evaluation, to analyze the usability and user experience of the UI using the three different interaction modes. The following sections shed more light on these stages.

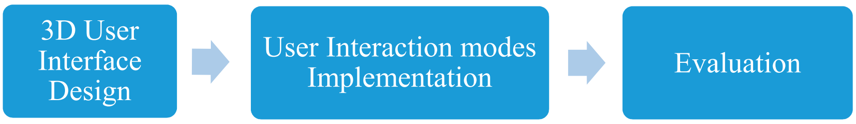


Figure 3. The block diagram of the proposed 3D User Interface (UI) system in the immersive VR.

3.1. 3D User Interface Design

Since immersive VR has a different medium, where the content is displayed in a 3D environment, a new 3D UI had to be designed to suit this environment. The designed 3D interface is different than the traditional 2D desktop, where all the files and tasks are carried out in a 2D space and on the same layer.

The proposed 3D UI is a dual-layer UI consisting of two main layers with curved screens (Figure 4). The upper layer (screen layer) is used for displaying the opened files. It allows the user to interact with multiple screens portrayed within a curved 360-degree effective field of view available for the user. While the lower layer (interaction layer) contains the traditional desktop tasks e.g., delete, duplicate, etc. Downward gaze allows the user to raise the interaction (lower) layer that facilitates several traditional desktop tasks. The effective field of view (FOV) of the 3D UI from the top is shown in Figure 5.

Based on the three different interaction modes, the three UIs are named accordingly. The three interfaces are Point-and-Click UI, Controller-based DM UI, and Gesture-based interface. A visualization of the three UIs design showing the two layers is found in Figures 6–8, respectively.

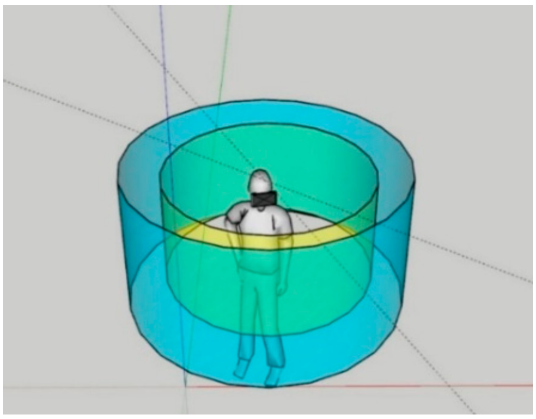


Figure 4. The proposed dual layer UI consisting of two main layers with a curved field of view (FOV).

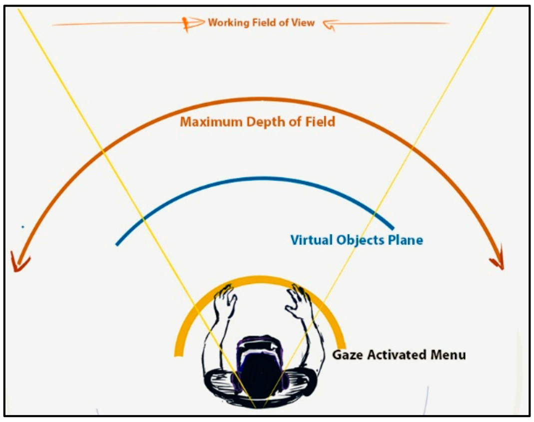


Figure 5. Top view of the 3D UI with FOV.

The gaze-based selection mechanism is used in the 3D UI as a context-aware mechanism to move between the two layers, and to perform certain tasks. Applying the context-aware gaze tracking activates the intended layer/object. For example, looking down makes the lower layer active, while looking up makes the upper layer active.

In addition to that, gaze is used to choose an object within a curved screen at the upper layer to make it active and then the user can perform the task on that object. The term object is used to refer to different types of files available in the 3D UI, e.g., images, videos, and folders. Specifically, for the video files, the gaze mechanism is used to play the video, and when the user moves his gaze away from the video, it will be paused automatically.

The user can display and access the menu that contains the different types of objects by either flipping his/her left hand for the gesture-based interaction, or by flipping the left controller for the controller-based DM interaction. Whereas, in the point-and-click UI, the menu is always presented at the lower layer. A visualization of the lower layer menu in the three UIs is shown in Figure 9, whereas the flipping menus can be seen in Figures 10 and 11.

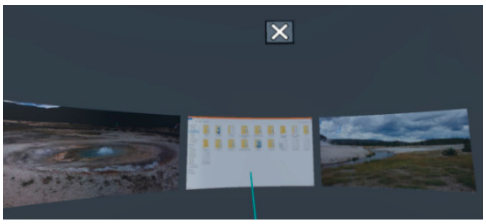


Figure 6. Point-and-Click 3D UI (upper layer).

Figure 7. Controller-based 3D UI (upper layer).

Figure 8. Gesture-based 3D UI (upper layer).

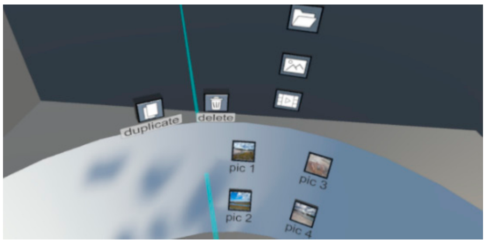


Figure 9. Menu in the lower layer.

Figure 10. Controller-based 3D UI (lower layer), activated when the controller is turned.

Figure 11. Controller-based and Gesture-based 3D UI (lower layer), activated with the lower gaze.

3.2. Interaction Tasks

Several tasks can be performed by the users through the proposed 3D UI for the immersive VR desktop. As mentioned earlier, the context-aware gaze tracking is used to indicate which specific object is active, and then the specific task is performed on this object.

The manipulation and system control tasks available in the 3D UI are as follows:

1. Open/Select: allows to select and open an object from the menu.
2. Close: allows to close an object displayed in the selected space.
3. Scroll: allows the movement (scrolling) of the layer in the x-axis of 3D space.
4. Scale: allows zooming in or zooming out on all objects displayed in the selected space.
5. Duplicate: allows to make a copy of an object on the lower layer by first selecting the duplicate button and then selecting the intended object.
6. Delete: allows to remove an object from the lower layer by first selecting the delete button and then selecting the intended object.

The design of the 3D UI takes into consideration the design principles such as affordance, feedback, consistency, and visibility. For example, to maintain the consistency in the design, similar tasks operations are used, and similar elements are used for the similar tasks. The two layers are used in all the interfaces, where the upper layer is used for displaying objects and the lower layer for interaction to perform the desktop tasks. The rationale being that having consistent interfaces are easier to learn and use. In addition to that, feedback is provided to the user about what has been done, e.g., when the delete/duplicate button is clicked on, a green highlight feedback was provided.

In a traditional user interaction, most of the interaction is typically performed using a keyboard input, mouse movement, and/or button pressing. These combinations can be attached to intuitive gestures defined by the movements of the hands alone or combined with fingers. Not only to make the interaction more natural and easier, but to also consider the three theoretical determinants of anticipated usefulness, ease of use, and behavior intention of the unified framework for the Technology Acceptance Model (TAM) and Reasoned Action Theory (RTA) [35]. These determinants are shared across a range of different models that investigate user's adoption of novel technology, see for instance early VR adoption in therapeutic studies [36]. In the following sections, the three different interaction modes are presented in more detail.

3.3. Point-and-Click Interaction

In this interface, the basic mouse interaction for the traditional desktop is adapted to a 3D laser pointer using one of the Vive controllers. The pointer has a blue laser light that gives an indication about where it is pointing to. Both hands can be independently used to control the pointer. Developing the idea of the laser pointer with some clicks, the user can interact with the 3D desktop in the immersive VR environment. An example of the user interaction in the point-and-click UI is shown in Figure 6. As mentioned above, the Vive controller is used to simulate the 3D laser pointer. The Vive controller has several input methods and only two of them are used, the trigger and trackpad. The trigger button is used for the tasks of opening, deleting, duplicating, and closing, where the users need to point to the intended place and then click the trigger. While the trackpad is used for scrolling and scaling tasks. The users need to move one finger on the touchpad right or left for scrolling, and up or down for scaling (zoom in or zoom out respectively). A summary of the desktop tasks and their corresponding interactions using the point-and click UI is found in Table 1. The menu is always displayed at the lower layer, the user can point and choose any type of objects.

Table 1. Point-and-click interaction.

No.	Task	Point-and-Click Interaction
1	Open	Trigger click
2	Close	Trigger click
3	Scroll	Right and left movement using a finger on the touchpad
4	Scale	Up and down movement using a finger on the touchpad (up: zoom in, down: zoom out)
5	Duplicate	Trigger click
6	Delete	Trigger click

3.4. Controller-Based DM

In this interface, a pair of Vive controllers is used for developing the interaction mode. The user uses both controllers to interact with objects in the immersive VR 3D desktop environment. The Vive controllers have sensors that are tracked by the base stations. A combination of controller's inputs and movements was used to perform the desktop tasks. An example of user interaction in the controller-based DM UI is found in Figure 7. A visualization of both controllers (right and left) is also presented on the 3D UI, where each controller is named by a letter "R" or "L" to distinguish between them. Only two input methods for the Vive controller are used, the trigger button and grip button. The user can open, delete, and duplicate by pushing/taping the intended objects using one of the controllers (right or left one), it does not matter. While to close a certain object, the user needs to press the grip button using one of the controllers (right or left one), it does not matter. To perform the scrolling task, the user should move one controller right or left, while keeping the trigger button clicked on. Whereas to perform the scale task, the user should use both controllers; by moving the controllers away from each other (zoom in) or towards each other (zoom out), while keeping the trigger button on both controllers clicked on. A summary of the tasks and their corresponding interactions using the controller-based UI is found in Table 1. The user can display and access the menu that contains the different types of objects by flipping the left controller only.

3.5. Gesture-Based Interaction

In the gesture-based UI, the user performs a certain set of hand gestures to manipulate the common desktop tasks in the immersive VR environment. The natural UI should always be designed on specific users and with use context in mind [3]. The aim is to focus on the common natural hand gestures that users know and can apply from their lives; taking advantage of their existing knowledge. Introducing completely new gestures is out of the scope of this work as it would enforce a new learning curve that might be hard for the users to learn and remember. The design took into consideration that the same hand gestures could have different meanings in different cultures [11]. Therefore, the aim was to choose the most common natural hand gestures based on the people's level of familiarity with mobile devices and touch screen input. For example, people can directly perform an in-air tap gesture, similar to the tap gesture on the touch screen. Additionally, pinch to zoom and slide to scroll are very familiar gestures. Finally, the only novel gesture is closing a window by closing the fists. It does not directly relate to a touch-based gesture, but it conforms to the task description and consequently should result in a quicker recall.

An example of the user interaction in the gesture-based UI, showing the visualization of the user's hands as well, is found in Figure 8. Each task in the interface is assigned a different hand gesture. The proper script is activated when the user's hand is recognized performing the intended gesture. The Leap Motion device is used for hand gesture recognition. A summary of the tasks and their corresponding hand gestures interactions using the gesture-based UI is found in Table 2. The users can display and access the menu that contains the different types of objects by flipping their left hand only.

Figure 12 shows the employed gesture recognition algorithm. After acquiring and processing the data from the Leap Motion controller, the system is able to recognize the hand gestures without the need to have predefined gestures or calibration. The gesture classification is subject independent, which means that obtaining subject specific parameters during the training phase is not needed to be used in the testing phase. In addition, given the nature of the gestures, and the data acquired from the Leap Motion, the gestures are recognized using heuristics rather than a classification method.

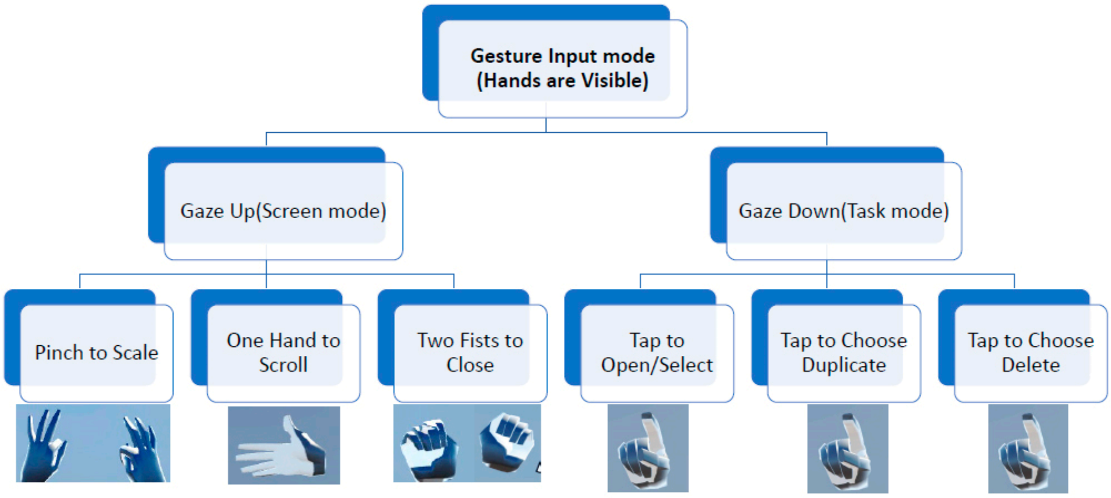
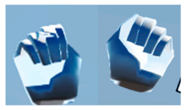


Figure 12. Flowchart of gesture recognition.

Table 2. Gesture-based interaction.

No.	Task	Gesture-Based Interaction	Hand Gesture
1	Open/Duplicate/Delete	Finger push/tap using one hand on the specific icon or button.	
2	Close Window	Two hands fists (closing both hands)	
3	Scroll	Thumb up towards the ceiling and the rest of the four fingers are open, moving/sliding the hand right or left (using one hand)	
4	Scale	Pinch using both hand; the thumb and index fingers close to each other while keeping the rest of the three figures open; moving both hands towards each other (zoom out) and apart from each other (zoom in)	

4. User Study and Experiment Setup

4.1. Experimental Evaluation

To evaluate the viability of the proposed 3D UI, a user study was performed to analyze the three different interaction modes for an immersive VR 3D desktop environment. In addition, other user evaluation methods popular in the HCI field were also performed. The aim was to provide a comprehensive usability study in terms of the following standard usability issues: precision, efficiency, effectiveness, ease-of-use, fatigue, naturalness, and overall satisfaction. In addition to the user experience; the way users feel about the UI using the three different interactions mode for 3D desktop in an immersive VR environment. The work in [37] presented a summary for the common types of AR user evaluation, which can also be applied for VR user evaluation. User evaluation means evaluating the work through user participation. A mix of the user evaluation techniques available in the literature to analyze the usability and user experience for each interface were employed. The mixed of the user evaluation techniques used in this study are:

1. Observational methods;
2. Query techniques.

4.2. Observation Methods

During the testing sessions in the controlled experiment, the following observation methods were used:

Direct observation: the participants were asked to use the think-aloud protocol; where the users were observed performing a task; while they were to describe what they were doing and why, what they think was happening, etc. Most of the users managed to apply the think-aloud techniques during the testing to give direct feedback, others were mostly focused on fulfilling the requested tasks.

Indirect observation: the users' activities were tracked e.g., analyzing the recorded video, eye gaze, footstep movements, and errors occurred.

4.3. Query Technique

Two query techniques were employed:

Semi-structured interview: the participants were interviewed after testing each UI under each condition (the different interactions modes and the VR experience) as well after finishing the whole usability test (exit interview). The interview was guided by a script of one closed question for each interaction. The question was about how they assess their success level for each interaction (easily completed or completed with difficulty or help). The exit interview was about the overall feedback on the whole experience and the UI design. It gives a good balance between feedback and replicability. Some interesting issues were explored in more depth.

Questionnaire: a comprehensive questionnaire including a clear purpose of study was developed, with different styles of question e.g., closed, open-ended, scalar, multi-choice, and ranked. The questionnaire was designed by combing different usability questions based on the standard questionnaires available in the literature to suit the context. The questionnaire was used at the end of the testing session, the participants were requested to rate their satisfaction on a scale of 1 to 5 (5 for strongly agree and 1 for strongly disagree) for features on all the interfaces under each experimental condition, and to complete an overall satisfaction questionnaire.

The questionnaire consisted of five main parts as follows:

Part 1: Personal Information

Part 2: Point-and-click UI

- Standing VR Experience
- Seated VR Experience

Part 3: Controller-based DM UI

- Standing VR Experience
- Seated VR Experience

Part 4: Gesture-based UI

- Standing VR Experience
- Seated VR Experience

Part 5: Overall Experience

In Parts 2, 3, and 4, three sets of questions in each part were designed. One set of questions for user satisfaction on selected usability goals, the second set of questions for user satisfaction on the idea of menu for accessing the different folders/categories of objects, and the last set for selecting the user experience from predefined answers that contain both positive and negative user experiences. Moreover, these three parts were repeated for each UI under both experience conditions, standing and seated positions. The last part in the questionnaire was designed for the overall user experience which contains multi-choice, ranked, closed and open-ended questions.

Based on the data collected from the above evaluation techniques, an appropriate qualitative and quantitative analysis was used to analyze the data and provide a comprehensive usability study. In the following sections, both the qualitative and quantitative analysis are discussed.

4.4. Quantitative Analysis

During the testing session, various forms of quantitative data were measured: the task completion time, the scenario completion time, number of eye gazes, and footstep movements. Task time refers to the individual time taken for each task in the test case scenario. Based on these measures, the average completion time of the task and scenario for each interaction modes under each condition (standing and seated) were calculated. This result was used to determine if there could be any correlation between the usability, the UX and the mean time spent on each interaction mode. Both simple quantitative analysis and statistical technique for further analysis were employed, as described below.

Number of eye gaze and footstep movements were calculated manually from video recordings that were captured during the testing session using a camera placed in front of the participants.

Simple quantitative analysis: numerical methods were used to analyze the completion time as follows:

- Percentages, averages, maximum, minimum.
- Graphical representations of data.

Statistical technique: a two-way repeated analysis of variance (ANOVA) was used to check if the user interfaces were significantly different from each other under various factors. The independent variables factors that were used in the study are:

Factor 1: the input devices (pointer, controller, gesture) or they can be called interaction modes.
Factor 2: the experiences (standing, seated) positions.

4.5. Qualitative Analysis

Both the questionnaire and unstructured interviews provided different elements of the qualitative data. The qualitative data was categorized into questionnaire results, and user feedback from the interviews.

A simple qualitative analysis was used to express the nature of elements and was represented as patterns by categorizing data and looking for critical incidents. Feedback from users' interviews was divided in three categories: positive, negative, and neutral for the overall experience and the UI design and their success level for each interaction (easily completed or completed with difficulty or help). Simple numerical methods were employed to analyze the results of the closed questions from the questionnaire e.g., weighted average and percentage. Following this, a visual representation of the qualitative data was made by putting the frequency of each category into a chart. This way, it was possible to see whether there was a consensus about the UX and usability for each interaction mode.

The last part in the questionnaire was designed for the overall user experience which contains multi-choice, ranked, closed and open-ended questions. User interaction and menu questionnaires can be seen in Appendix A.

4.6. Experiment Setup

The experiment setup included the hardware and software used, the participants that contributed to the experiment, the experimental conditions, and finally the experiment workflow.

This work utilized three types of hardware, HTC Vive, Leap Motion controller, and a laptop with Core i7 and GeForce Ti 1050. HTC Vive was used as the immersive VR display. The HTC Vive controllers were used to implement the pointer and controller-based interfaces. Leap Motion was used to implement the gesture-based interface. For the software part, Unity was used to construct the VR environment as it is a cross-platform game engine with a built-in IDE developed by Unity Technologies. Leap Motion Orion SDK was used for the hand tracking.

4.7. Participants

The participants were selected through random sampling/selection on ease of access and their availability. Since the usability study was conducted in the context of a specific application, the target user population was defined to evaluate the most diverse user population possible in terms of age, and gender, and community group. It was ensured that there were enough participants to get a valid and feasible result. Therefore, a collaboration of 21 volunteer users (11 female and 10 male), aged from 18 to 55 (the majority between 25 and 34), randomly from a pool of 40 was assembled. They were from different community groups, most of them from the university. These community groups included faculty members (4), university students (7), administrative staff (7), and others (3). Some of the participants (5 out of 21) had experience in using VR systems.

4.8. Experimental Condition

In the experiment, a random assignment was applied, where the participants for the sample were assigned at random to experimental control conditions during the testing session to avoid a biased user experience.

The two experimental conditions in the work are as follow:

Condition 1: the interaction modes or input devices (pointer vs. controller vs. gesture).
Condition 2: the experience (standing position vs. seated position).

Each participant tested the interface under each condition. There were three interaction modes/input devices and two experiences; therefore, in total there were six permutations of tests for each user. The participant tested each interaction in both experiences, e.g., the point-and-click interaction was tested twice for standing and seated positions separately. The assignment for the order on how each participant tested each interface is shown in Table 3.

The study employed a total of 21 participants. Based on condition 1 the participants were divided into three equal groups of seven participants each based on the initial interface. Therefore, the point-and-click group (total of seven), controller-based group (total of seven), and the gesture-based group (total of seven). Then each group was divided into two groups. For example, the point-and-click group was dived into group 1 and group 2, where each one started with the point-and click interface but with a different sequence for the remaining interfaces. The same procedure was applied exactly for the controlled-based group and the gesture-based group. For condition 2, the experience, all the groups in condition 1 had the same sequence of experiences.

The sequence for the experience was assigned based on the interface's sequential order in condition 1. For example, in the first interface of all group sequences in condition 1, the user started the testing in standing then seated position. While for the second interface in all group sequences in condition 1, the user started the testing in seated then standing position. Finally, for the third interface in all group sequences in condition 1, the user started the testing in standing then seated position. As mentioned before the random assignment was done to avoid any biased user experiences that may have occurred as much as possible.

Table 3. Implementation of experimental conditions during testing session.

Initial Interface	Participants	Condition 1	Condition 2
Point-and-click	Group 1 [1,4,11,13]	Pointer Controller Gesture	Standing/Seated Seated/Standing Standing/Seated
	Group 2 [5,9,15]	Pointer Gesture Controller	
Controller-based	Group 3 [6,12,20,21]	Pointer Gesture Controller	Standing/Seated Seated/Standing Standing/Seated
	Group 4 [2,10,16]	Pointer Gesture Controller	
Gesture-based	Group 5 [3,8,14,19]	Pointer Gesture Controller	Standing/Seated Seated/Standing Standing/Seated
	Group 6 [7,17,18]	Pointer Gesture Controller	

4.9. Experimental Workflow

In this work, a controlled experiment was performed, using a within-subjects experimental design. In the controlled experiment, the room setup and the hardware setup were kept, particularly the position of the base stations, the same for all participants; to avoid disrupting the tracking process and the required setting up of the testing area again, which may affect the reliability of the experiment. In within subjects design, each participant tested the UI using the three interaction modes under each condition for both standing and seated positions.

The experiment started with a short introduction to the immersive VR 3D desktop environment, its goals, the 3D UI, the hardware (HMD and controllers) used, the three different interaction modes, and a reference to the questionnaire to be completed at the end of the experiment.

After going through the introduction, the observer performed a live demonstration in front of the user on how to interact with the 3D UI using the three interaction modes. All the tasks and how to perform it using the three interaction modes were explained in detail to the user. The demonstration was repeated with the same specific order for all the participants as follows: starting with the point-and-click UI, then the controller-based DM UI, and finally the gesture-based UI.

4.10. Testing Procedures

As shown in Figure 13, during the testing session both the participants and the observer had tasks to perform. The experimental conditions were discussed in detail earlier in this section. The testing session started by asking the participant to perform a sequence of specific tasks by the observer for each interface.

The sequence of specific tasks (scenario) that the participant needed to perform was identified in advance. The use of the scenario was to ensure that users performed every type of possible task on the UI at least once. The same case scenario was applied to all interfaces using the two positions, standing and seated. During the testing session, the observer was monitoring the user performance by recording the completion time for each task and for the scenario in both standing and seated positions using each interaction mode, also by taking down relevant information e.g., errors and success rate. All the testing sessions were videotaped for further observations.

The test case scenario used in the experiment consisted of the following tasks:

- Open video 1 and play it and then open folder #3.
- Zoom in and out the objects on the upper layer.
- Duplicate folder #2.
- Scroll right and left and close the opened video.
- Delete picture #4.

Finally, the participant was asked to complete an online comprehensive questionnaire about each tested UI and the overall experience to gather quantitative data about the UX and usability (see Appendix A for questionnaire). The feedback from the users for each interaction and for overall feedback on the experience and UI design (exit interview) was collected that was used for further analysis. Figures 14 and 15 show two of the participants in a testing session in standing and seated positions.

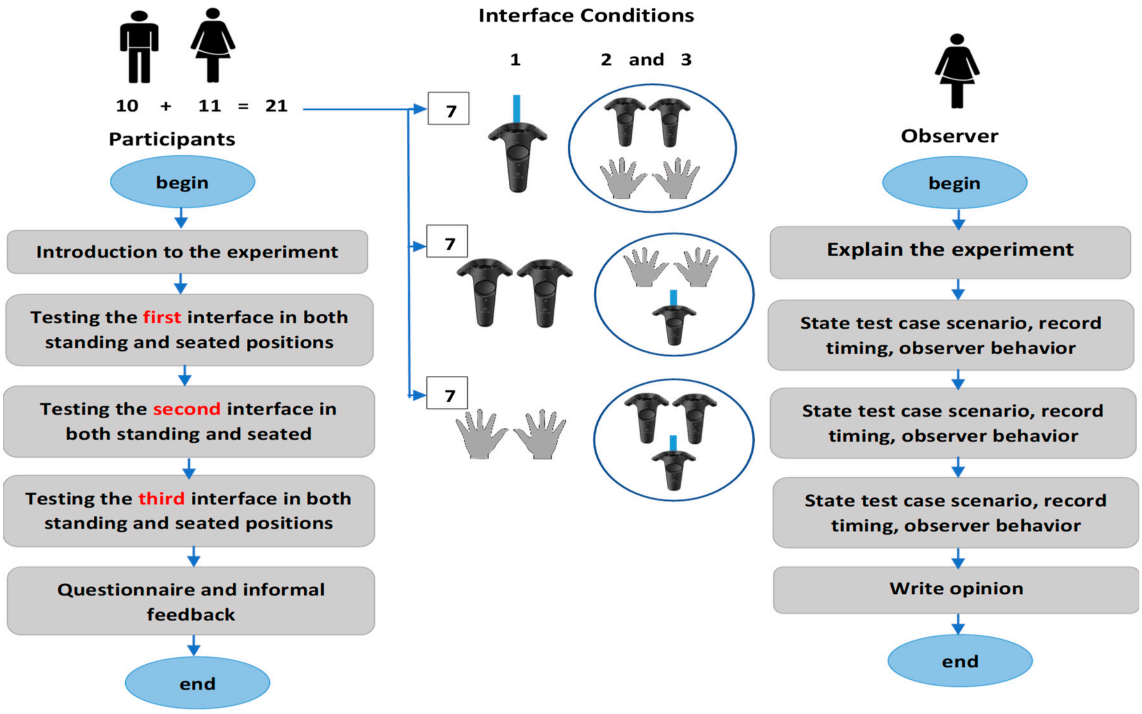


Figure 13. Experiment workflow.

Figure 14. Participant during the testing in standing position.

Figure 15. Participant during the testing in the seating position.

5. Results and Data Analysis

The results from the main study provided a comprehensive dataset to be analyzed through statistical analysis. In the following sections, the results and analysis are presented for scenario completion time, task completion time, two-way ANOVA, observations and interview results, questionnaire results, and finally the summary and findings. All the data analysis was performed using the Data Analysis tool in Excel.

5.1. Scenario Completion Time

The scenario completion time refers to the total time taken to complete the test case scenario that is explained in Section 4. The scenario completion time was measured for each user interaction mode and for each experience.

Completion Time by Interaction Mode: the time taken to complete the scenario by each participant for each user interaction mode can be seen in Table 4, with the minimum, maximum, and mean times at the bottom. In general, the maximum time ranges from 108 s (pointer-based) to 136 s (gesture-based). Similarly. The minimum time ranges from 34 s (pointer-based) to 47 s (gesture-based). This shows that the pointer-based interface is the preferable mode of interaction for the user even in a VR environment because of their familiarity with the interface.

Completion Time by Experience: the time taken to complete the scenario by each participant for each user experience, standing and seated, can also be seen in Table 4.

Table 4. Scenario completion time for each interface.

Experience	**Standing**			**Seated**		
Interface Type	**Pointer**	**Controller**	**NUI**	**Pointer**	**Controller**	**NUI**
Participant 1	59	92	81	57	76	136
Participant 2	86	79	81	50	70	99
Participant 3	42	43	51	35	54	47
Participant 4	44	36	52	50	67	53
Participant 5	50	41	75	41	56	59
Participant 6	42	108	86	48	53	95
Participant 7	52	56	72	42	65	61
Participant 8	50	49	81	45	70	69
Participant 9	65	58	95	64	88	104
Participant 10	40	109	54	39	61	69
Participant 11	108	95	102	108	110	135
Participant 12	57	85	131	61	77	87
Participant 13	76	55	127	44	70	65
Participant 14	89	58	66	34	82	68
Participant 15	107	74	109	85	68	100
Participant 16	52	75	108	50	61	80
Participant 17	44	79	80	92	57	69
Participant 18	43	115	101	56	79	87
Participant 19	51	64	92	57	88	64
Participant 20	49	99	71	69	95	79
Participant 21	42	90	71	53	57	57
Maximum	108	115	131	108	110	136
Minimum	40	36	51	34	53	47
Mean	59.43	74.29	85.05	56.19	71.6	80.4

As can be seen in Table 4, the overall maximum time ranges from 131 s (standing) to 136 s (seated). Similarly, the minimum time ranges from 34 s (seated) to 36 s (standing). The mean time for completing the tasks in standing experience is 72.92 s and in seated experience is 69.32 s. This shows that the difference between their individual means and grand mean (68.12 s) is not significant enough regarding the results for completion time. The grand mean is the mean of sample means or the mean of all observations combined, irrespective of the sample.

Completion Time by Initial Interface: the time taken to complete the scenario by a participant in each initial interface group can be seen in Table 5, with minimum, maximum, and mean times shown at the bottom. The maximum time ranges from 109 s (controller-based) to 101 s (gesture-based). Similarly, the minimum time ranges from 44 s (pointer-based) to 75 s (controller-based). This shows that the controller-based interface is the least preferable mode of interaction for the user.

Table 5. Completion time by initial interface.

		Scenario Completion Time (s)	
Started with pointer	Participant 1	59	
	Participant 4	44	
	Participant 5	50	
	Participant 9	65	
	Participant 11	108	
	Participant 13	76	
	Participant 15	107	
Started with controller	Participant 2	79	
	Participant 6	108	
	Participant 10	109	
	Participant 12	85	
	Participant 16	75	
	Participant 20	99	
	Participant 21	90	
Started with NUI	Participant 3	51	
	Participant 7	72	
	Participant 8	81	
	Participant 14	66	
	Participant 17	80	
	Participant 18	101	
	Participant 19	92	
	Minimum	**Mean**	**Maximum**
Started with pointer	44	72.71	108
Started with controller	75	92.14	109
Started with NUI	51	77.57	101

Completion Time by Gender: the time taken to complete the scenario by male and female participants for each user interaction mode and experience can be seen in Table 6. There were 10 male and 11 female participants. Based on the initial condition of the UI and random assignments used in the experiment, four out of 10 males started with point-and-click UI (group 1), another four out of 10 males started with controller-based UI (group 2), and two out of 10 males started with gesture-based UI (group 3). While three out of 11 females started with point-and-click UI (group 1), another three out of 11 females started with controller-based UI (group 2), and five out of 11 females started with gesture-based UI (group 3). As can be seen in Table 7, the maximum time for completing the tasks by gender for each interaction mode varies according to the users. Furthermore, the summary shows that for male participants, the difference between individual means of each interaction mode and grand mean (67.38 s) is significant enough regarding the results for completion time. Whereas, for female participants, the difference between individual means of each interaction mode and grand mean (77.20 s) is not significant enough regarding the results for completion time.

Table 6. Scenario completion time by gender.

Gender	Groups	Interface	Standing	Seated
Male 1	**Group 1**	**Pointer**	44	50
Male 2			50	41
Male 3			65	64
Male 4			107	85
Male 5	**Group 2**	**Controller**	109	61
Male 6			75	61
Male 7			99	99
Male 8			90	90
Male 9	**Group 3**	**NUI**	51	47
Male 10			66	68
Female 1	**Group 1**	**Pointer**	59	57
Female 2			108	108
Female 3			76	44
Female 4	**Group 2**	**Controller**	79	70
Female 5			108	53
Female 6			85	77
Female 7	**Group 3**	**NUI**	72	61
Female 8			81	69
Female 9			80	69
Female 10			101	87
Female 11			92	64

Table 7. Gender-based time variations.

Gender	Initial Interface	Minimum	Mean	Maximum
Male	Point-and-click UI	41	63.25	107
	Controller-based DM UI	57	80.88	109
	Gesture-based UI	47	58.00	68
Female	Point-and-click UI	44	75.33	108
	Controller-based DM UI	53	78.67	108
	Gesture-based UI	61	77.60	101

5.2. Task Completion Time

The average time taken to complete each individual task by all participants for each combination of user interaction mode and experience (six combinations of tests) is depicted in Table 8. The average task completion time in the point-and-click UI is shorter than the other two interfaces for both experiences. The maximum task completion time is 15.9 s for task 5 (scroll) in gesture-based UI—standing experience, while the minimum is 5.48 s for the task 5 (scroll) also, but in the point-and-click UI—seated experience. For further analysis the average of a specific task which is the open task that occurred twice in the test case scenario was computed (task 1: open a video file, and task 2: open and a folder file). The results can be seen in Table 9. It was found that the average time for the open task (included both experiences) in point-and-click UI is 7.98 s, while in controller-based UI took 11.35 s, and in gesture-based UI took 10.74 s. In general, this shows that the open task can be performed faster in the point-and-click UI than the other two interfaces.

Table 8. Tasks and their completion time.

	Tasks Given	Pointer—Standing	Pointer—Seated	Controller—Standing	Controller—Seated	NUI—Standing	NUI—Seated
Avg of Task 1	Open video 1 and play it	9.1	8.4	15.1	12.6	12.8	10.8
Avg of Task 2	Open folder 3	8.2	6.1	8.1	9.4	10.4	8.9
Avg of Task 3	Zoom in and out all objects	6.6	6.7	7.4	7.0	10.1	9.7
Avg of Task 4	Duplicate folder 2	9.1	7.2	9.1	11.0	11.0	9.8
Avg of Task 5	Scroll right and left	6.1	5.4	6.7	6.7	15.9	11.8
Avg of Task 6	Close the opened video	7.0	8.2	11.0	5.8	9.0	10.5
Avg of Task 7	Delete picture 4	9.5	10.0	13.1	15.4	11.9	14.8
Avg of Overall Scenario		59.4	56.2	74.3	71.7	85.0	80.1

Table 9. Task completion time comparison.

Interface	Avg of Task 1 and Task 2 (Open a File)		
Pointer—Standing Pointer—Seated	8.6 7.2	Average opening—Pointer	7.9
Controller—Standing Controller—Seated	11.6 11.0	Average opening—Controller	11.3
NUI—Standing NUI—Seated	11.6 9.8	Average opening—NUI	10.7

5.3. Analysis of Variance (ANOVA)

Among the different factors under evaluation, a two-way repeated ANOVA was carried out for the two independent variables:

- The interaction modes or input device (pointer vs. controller and gesture).
- The experience (standing vs. seated).

The following analysis shows standard deviation and means for different variables in the forms of Mvariable (e.g., Mpointer is the mean for pointer-based interaction) and SDvariable (e.g., SDstanding is the standard deviation for standing experience). Moreover, F (df, MS) is the test statistic (F-ratio) in which df and MS are the degrees of freedom and mean square, respectively, for the variables. The F-ratio was calculated using MSvariable(s)/MSerror(s) and p is the probability value. All analyses was concluded at alpha of 0.05 significance level and for 21 participants. The results are presented in Table 10.

The null hypothesis was tested for the effect of different interaction modes on the scenario completing time using the two-way repeated ANOVA. The hypothesis here is: "All three interaction modes will give the same interaction experience in terms of time." To reject the null hypothesis or not, the following conditions need to be satisfied:

If the F-value (F) is larger than the f critical value (Fcrit).

If the p-value is smaller than your chosen alpha level.

The analysis illustrates that for variable 1, the interaction modes: F (2,6556.357) = 14.613, p = 0.00000209 (Mpointer = 57.81, SDpointer = 3.08 vs. Mcontroller = 72.95, SDcontroller = 3.03 vs. Mgesture = 82.60, SDgesturer = 3.59). The p-value < alpha (0.05). This illustrates that interaction mode has a significant effect on time.

For variable 2, the experience: F (1,408.960) = 0.911, p = 0.341 (Mstanding = 72.92, SDstanding = 3.09 vs. Mseated = 69.32, SDseated = 2.76), the p-value > alpha (0.05), which illustrates that the experience position does not have a significant effect on time.

Moreover, the analysis shows no significant effect on time for variables 1 and 2 combined F (1,14.198) = 0.0316, p = 0.969, the p-value > alpha. Therefore, it can be concluded that interaction modes and experience did not have any combined effect on the time. Based on the above, the null hypothesis is rejected meaning that interaction modes significantly affect the time taken to complete the scenario in the 3D UI.

Table 10. Two-way variance—ANOVA.

	Source of Variation		
	Columns (Factor 1)	**Sample (Factor 2)**	**Interaction**
SS	13,112.714	408.960	28.396
df	2	1	1
MS	6556.357	408.960	14.198
F	14.613	0.911	0.031
p-value	2.09×10^{-6}	0.341	0.968
Fcrit	3.071	3.920	3.071

5.4. Observation and Interview Results

The results of different observations collected during the testing session with analysis of the video recordings are presented in this section. The observations include the number of eye gazes, the footstep movements, and the errors. In addition, the interview results which include the user feedback for each interaction and for the overall experience of performing the desktop tasks in the immersive VR environment using the 3D UI design are also presented.

Number of Eye Gazes: all the video recordings for the 21 participants were analyzed, specifically the video of initial interface for each participant under the standing experience only. The number of eye gazes were calculated for a specific task, which was task 2 (open a specific folder) in the test case scenario. Ideally, the user needs two gazes only to perform the open task using the 3D UI, one gaze at the lower (interaction) layer to choose the intended file, and the second gaze to look at the upper (screen) layer to view the opened file. The number of eye gazes for each participant for that specific task were calculated manually. It was found that all users were able to perform the task using two gazes only. This shows a good indication that the participants were able to understand the 3D UI design and interact with it effectively.

Number of Footsteps: the users' footstep movements during the testing session by observing the movement outside the defined boundary mentioned in Section 3 were analyzed. Additionally, for certain cases, the video recordings were watched for further confirmation. It is to be noted that the interface does not rely on any type of lower body motion, e.g., walking. The purpose of tracking footsteps is to find if there is any design flaw associated with the interface that results in user making physical movements that are not required. The analysis did not find any significant effect of the footstep movements on the UI or interaction modes. None of the participants went outside the defined boundary, rather a slight movement in terms of rotation was noticed since the 3D UI allows to interact with multiple screens portrayed within a curved 360-degree effective field of view available for the user. It was concluded that there is no significant relationship between the physical movement and the interactions modes.

Error Analysis: during the testing sessions, the observer observed and wrote the comments for any errors that might occur in each UI. The observation did not find any significant errors in any UI. Rather, some participants forget how to perform certain tasks. For example, in the point-and-click UI, one participant forgot how to perform the close task. This suggests that the need to enhance the visibility of the close button, either in terms of size, color, or label. It is intended to be explored in future work. While in the controller-based DM UI, one participant forgot how to perform the close task using the controller twice in standing and seated positions. Another one forgot how to perform the scroll task. An additional participant forgot to switch the controllers; the user was trying to display the menu using the right controller. The reason might be that the training given to the users was short. This will be explored in the future work.

Interview Results: the participants were interviewed after each UI test under each condition of the experience. The interview was guided by a script consisting of one closed question. The question was "how do you assess your success level for completing the scenario using each interaction mode? (Answers: 1. easily completed or 2. completed with difficulty or help)". As seen in Table 11, the results show a higher percentage for completing the scenario easily using the pointer-based interface (90.48% for standing, 85.7% for seated), followed by the gesture-based interface (90.48% for standing, 76.19% for seated), finally the controller-based interface (71.73% for standing, 85.71% for seated).

In addition to that, the participants were asked a general question about the overall experience after finishing the whole usability test (exit interview). Most of the users had a positive experience with the 3D UI and an overall pleasing and productive experience.

Table 11. Participants' success levels for completing the scenario.

Success of Scenario Completion	Pointer		Controller		Gesture	
	Standing	Seated	Standing	Seated	Standing	Seated
Easily completed	90.48%	85.71%	71.43%	85.71%	90.48%	76.19%
Completed with difficulty/help	9.52%	14.29%	28.57%	14.29%	9.52%	23.81%

5.5. Questionnaire Results

Participants were asked 26 questions to rate their satisfaction for each interaction. The same survey was employed for both standing and seating experiences. The interaction satisfaction questionnaire can be seen in the Appendix A. Furthermore, five questions were asked about their satisfaction with the provided menu system in that particular interaction. This questionnaire can be seen in the Appendix B. For each question, the participants were asked to rate their satisfaction from 1 to 5 (5 for strongly agree, 4 for agree, 3 for neutral, 2 for disagree, 1 for strongly disagree). The rating weighted average for each question was calculated, as seen in Equation (1), where: w = weight of answer choice, and x = response count for answer choice.

$$\text{Rating weighted average} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 \ldots x_n w_n}{\text{Total response count}}. \quad (1)$$

The majority of the weighted average results were 4 and above, which means that the usability goals are met with more than 80% approval.

For the user experience results, the highest percentages in all UIs were for the following answers: "satisfying", "enjoyable", "fun", and "wonderful". This shows that most of the users had a positive experience with the 3D UI and an overall pleasing and productive experience.

5.6. Summary and Findings

Table 12 shows the answers of the interface ranking for the most suitable interface for 3D desktop in VR environment by the participants. The average ranking was calculated as follows, where: w = weight of ranked position, x = response count for answer choice.

$$\text{Average ranking} = \frac{x_1 w_1 + x_2 w_2 + x_3 w_3 \ldots x_n w_n}{\text{Total response count}}. \quad (2)$$

Table 12. Interface ranking by participants.

	Point-and-Click UI	Controller-Based DM UI	Gesture-Based UI
First	11	5	5
Second	8	9	4
Third	2	7	12
Average Ranking	2.43	1.90	1.67

The first choice has a weight of 3, the second choice has a weight of 2, and the third choice has a weight of 1. As seen in the table above, the point-and-click interface has the largest value for average ranking, therefore it is the most preferred UI according to user feedback, followed by the controller-based interface as the second highest choice and finally the gesture-based interface as the third highest choice.

The ranking in Tables 12 and 13 shows that the most suitable interface for the desktop VR environment chosen by participants was the point-and-click UI (57.14%), followed by the gesture-based interface (33.33%), and finally the controller-based UI (9.52%). User's familiarity with the point-and-click interface can be one reason for being the most preferred interface by the participants for the desktop on VR. The percentage for the gesture-based interface shows a promising capability for utilizing the gesture as an interaction mode in the desktop VR. The reason for the lowest percentage for the controller-based interface could be that most of the users are not familiar with controller's device, in addition to the short training given.

Table 13. Participants rating for the most suitable UI for the desktop VR.

Point-and-Click UI	Controller-Based DM UI	Gesture-Based UI
57.14%	9.52%	33.33%

Although the previous results show that the point-and-click UI is the most preferred interface for the desktop in VR from the participant's point of view, a further analysis reveals that participants are still preferring to perform some certain tasks using the other interaction modes. For example, as seen in Table 14 they prefer to perform the scale task using the controller-based interface, but the gesture-based to display the menu.

Table 14. Task-oriented interface preferences.

	Point-and-Click UI	Controller-Based DM UI	Gesture-Based UI
Close	42.86%	14.29%	42.85%
Open	57.14%	9.52%	33.33%
Scroll	47.62%	14.29%	38.10%
Scale	33.33%	38.10%	28.57%
Delete	52.38%	14.29%	33.33%
Duplicate	61.90%	9.52%	28.57%
Display Menu	38.10%	9.52%	52.38%

For each task available in the 3D UI, its corresponding preferred interface from a participants' point of view are as follow:

- Close: Both point-and-click UI and gesture-based UI (42.86%).
- Open: Point-and-click UI (57.14%).
- Scroll: Point-and-click UI (47.62%).
- Scale: Controller-based UI (38.10%).
- Delete: Point-and-click UI (52.38%).
- Duplicate: Point-and-click UI (61.90%).
- Display the menu: Gesture-based UI (52.38%).

Table 15 shows that 52.38 % of the participants prefer to keep the interaction simple in the 3D UI for the desktop in VR using one interaction mode only. While 47.62% prefer to implement a hybrid approach by combining the gesture-based interaction and physical-device based interaction. The results are very close and point towards a hybrid approach as future research.

Table 15. Results of the hybrid approach in implementing the user interaction for desktop VR.

Prefer to Use Only One Interaction Style	Prefer to Use a Hybrid Interaction Style
52.38%	47.62%

6. Conclusions and Future Work

In the following section the final concluding discussion and limitations of the work are presented followed by the future work.

6.1. Discussion

This work presented a new dual-layered 3D UI for a 3D desktop environment in Virtual Reality. The UI allows the user to visualize the desktop screens on a 360-degree wide effective field of view. A downward gaze raises the secondary layer that allows the user to perform common desktop tasks. The UI was evaluated using three different interaction modes. The pointer-based point and click interface along with a controller-based DM interface use the physical device. The UI was also evaluated using a gesture-based controller free user interaction mode. A comprehensive qualitative and quantitative user study was performed to validate both the usability and user experience of the 3D UI.

The user study shows that the pointer-based interface is fastest in terms of task completion time, with more than 80% of the users expressing satisfaction with the interface in terms of both usability and the user experience. The average task completion time in the pointer-based interface is shorter than the other two interfaces for both experiences, which if considered within a TAM point of view is related to the anticipated ease of use and usefulness to complete the task within this mode.

The two-way repeated ANOVA was used with a null hypothesis for the effect of different interaction modes on the scenario completion time. The results show that the interaction modes significantly affect the time taken to complete the scenario in the 3D UI. While the standing or seated experience did not have any effect on time, this can also be read as an indication that participant behavioral intentions and attitudes towards completing the task were not influenced with body positions and the perceived control over actions either being seated or standing. The results of the different observations collected show that the users understand the proposed 3D UI design and interact with it effectively in terms of number of eye gazes needed. The footstep analysis did not find any significant relationship between body movement and the interaction modes. The results of interview show a higher percentage of support for completing the scenario easily using the pointer-based interface, which supported by the questionnaire results for the most suitable interface for desktop VR. Where the results shown in Table 13, confirm that the point-and-click UI (57.14%) is the most suitable interface, then the gesture-based interface (33.33%), finally the controller-based UI (9.52%).

Although the results show that the point-and-click UI is the most preferred interface for the desktop in VR from the participants' point of view, participants still prefer to perform some tasks using other interaction modes. For example, they prefer to perform the scale task using controller-based interface and to display the menu using the gesture-based interface. In general, 52.38% of the participants prefer to keep the interaction simple in the proposed 3D UI for the desktop in VR using one interaction mode only rather than using a hybrid approach.

In conclusion, the mixed methods approach shows that the point-and-click UI is the most suitable user interaction mechanism for the proposed 3D UI for the desktop environment in an immersive VR. However, according to the results the gesture-based interface was the second preferable UI. This indicates a promising capability for utilizing the gesture as an interaction mode in desktop VR. This leads to the following two questions: How can the natural UI (specifically the gesture-based interface) be improved? What factors affect an NUI in the immersive VR environment?

The total immersive nature of VR gives it its uniqueness and sets it apart from other media experiences by giving the audience an extreme close-up and all-encompassing visualization experience.

One suggestion for improving the NUI can be changing the nature of scenario design from the desktop to another scenario, for example, a game. This may lead to different results and improve the usability and user experience of the NUI in an immersive VR.

As mentioned before, the gesture-based interaction is one type of natural user interaction, where a gesture may be defined as a physical movement of the human body or some of its parts (hands, arms, face, etc.) intended to convey information, actions, or meaning. One factor that might affect the NUI is the choice of gestures. Therefore, designing a new set of gestures or letting the users choose their own gestures from a set of predefined gestures to perform the available tasks will provide more flexibility to the UI design. Consequently, it may lead to improving the user interaction.

6.2. Limitations

Although the user interaction modes for desktop VR work with a very good accuracy, there are a few limitations that might be considered.

The hardware limitation of Leap Motion: the sensor is still having trouble accurately tracking a user's hands in situations when one finger covers the other. Fingers right next to each other also pose a problem for the cameras and might not be recognized individually. Additionally, the real-life performance may not translate on to the computer model as expected. In addition, the device is not able to perform well under all lighting conditions.

The setup limitation because of using a wired HMD: Limited freedom to move, required room space, and an advanced cable management system. In addition to that, the participants were asked a general question about the overall experience after finishing the whole usability test (exit interview). Most of the users had a positive experience with the 3D UI and an overall pleasing and productive experience.

6.3. Future Work

As future work, the following tasks might be implemented:

1. Extend the system and analyze it with an extended degree of control over the user interaction modes.
2. Test the system on a larger data sample to demonstrate the robustness of the system.
3. Add the navigation task to the interaction space.
4. Increase the time for training given to the user because there were some time constraints to keep the testing time tractable.
5. Enhance the design of the menu.
6. Change the concept of the lower layer to a temporary layer (contextual layer) as an overlay without the downward gaze. It would require more research in finding the natural gesture to toggle the overlay.
7. Add more feedback to the gesture-based interface to indicate if the user is performing the correct gesture or not, e.g., add some hints in case of difficulty.
8. Finally, implement a hybrid approach by combining the gesture-based interaction and physical-device based interaction.

Author Contributions: Conceptualization: N.A. and M.L.; implementation and testing: H.K. writing—original draft preparation: H.K.; writing—review and editing: N.A. and M.L.; supervision: N.A. and M.L.

Funding: The research was completely funded internally by the University of Sharjah and did not receive any external funding.

Acknowledgments: Special thanks to Anas Einea for his valuable technical support on this work. The authors also thank all who participated in the usability test and in all ways to make this work reach its goal.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Interaction Mode Satisfaction Survey

This questionnaire is generic, and the term this interaction mode is replaced by point-and-click interaction, controller-based interaction, and gesture-based interaction for each user interaction mode.

1. Overall, I am satisfied with the ease of completing the tasks in this scenario using this interaction mode.
2. Overall, I am satisfied with the amount of time it took to complete the tasks in this scenario using this interaction mode.
3. I found the interface is reliable.
4. I found the interface is designed for all levels of users.
5. I am satisfied with the interface speed.
6. Learning to operate the interface is easy for me.
7. I found it easy to get the interface to do what I want it to do.
8. My interaction with the interface is clear and understandable.
9. I found the interface to be flexible to interact with.
10. I found the interface easy to use.
11. I found this is effortless.
12. Using the interface in VR would enable me to accomplish tasks more quickly.
13. Using the interface would improve my performance in VR.
14. Using the interface in VR would increase my productivity.
15. Using the interface would enhance my effectiveness in desktop VR.
16. Using the interface would make it easier to do the common desktop tasks.
17. I can recover from mistakes quickly and easily.
18. I easily remember how to use the interface.
19. The interface is easy to learn.
20. Using the interface in VR saves my time.
21. I feel safe to use the interface.
22. I found the scale (zoom in/out) task natural using the interface.
23. I found the close task natural using the interface.
24. I would prefer to use this interaction mode for desktop VR.
25. Overall, I am satisfied with the support information (menu) when completing the tasks.
26. It is easy for me to become skillful at using the interface.

Appendix B Menu System Satisfaction Survey

1. The idea of the menu for accessing the different folders/categories of objects provides a better accessibility for desktop VR.
2. I am satisfied with the design of the menu.
3. The idea of the menu provides a quick access to the folders of objects.
4. User selection of files for display by looking down is easy.
5. The ordering of menu options is logical.

References

1. Bowman, D.; Kruijff, E.; LaViola, J., Jr.; McMahan, R.P.; Poupyrev, I. *3D User Interfaces: Theory and Practice*, 2nd ed.; Addison-Wesley Professional: Boston, MA, USA, 2017.
2. Re, G.M.; Bordegoni, M. A Natural User Interface for Navigating in Organized 3D Virtual Contents. In *International Conference on Virtual, Augmented and Mixed Reality*; Springer International Publishing: New York, NY, USA, 2014.

3. Saffer, D. *Designing Gestural Interfaces: Touchscreens and Interactive Devices*; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2008; ISBN 978-0-596-51839-4.
4. Murthy, G.R.S.; Jadon, R.S. A review of vision-based hand gestures recognition. *Int. J. Inf. Technol. Knowl. Manag.* **2009**, *2*, 405–410.
5. Jerald, J.J. Scene-Motion-and Latency-Perception Thresholds for Head-Mounted Displays. Ph.D Thesis, University of North Carolina at Chapel Hill, Chapel Hill, CA, USA, 2009.
6. Kaushik, M.; Jain, R. Natural user interfaces: Trend in virtual interaction. *arXiv* **2014**, arXiv:1405.0101.
7. Ahn, S.C.; Lee, T.; Kim, I.; Kwon, Y.; Kim, H. Computer Vision based Interactive Presentation System. In Proceedings of the Asian Conference for Computer Vision, Jeju, Korea, 27–30 January 2004.
8. Pang, Y.Y.; Ismail, N.A.; Gilbert, P.L.S. A real time vision-based hand gesture interaction. In Proceedings of the Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation (AMS), Bornea, Malaysia, 26–28 May 2010.
9. Kang, J.; Seo, D.; Jung, D. A study on the control method of 3-dimensional space application using kinect system. *Int. J. Comput. Sci. Netw. Secur.* **2011**, *11*, 55–59.
10. Bragdon, A.; DeLine, R.; Hinckley, K.; Morris, M.R. Code space: Touch+ air gesture hybrid interactions for supporting developer meetings. In Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Kobe, Japan, 13–16 November 2011.
11. Ebert, L.C.; Hatch, G.; Ampanozi, G.; Thali, M.J.; Ross, S. You can't touch this: Touch-free navigation through radiological images. *Surg. Innov.* **2012**, *19*, 301–307. [CrossRef] [PubMed]
12. Kerefeyn, S.; Maleshkov, S. Manipulation of virtual objects through a LeapMotion optical sensor. *Int. J. Comput. Sci. Issues (IJCSI)* **2015**, *12*, 52.
13. Fanini, B. A 3D interface to explore and manipulate multi-scale virtual scenes using the leap motion controller. In Proceedings of the Seventh International Conference on Advances in Computer-Human Interactions (ACHI 2014), Barcelona, Spain, 23–27 March 2014.
14. Kry, P.; Pihuit, A.; Bernhardt, A.; Cani, M. Handnavigator: Hands-on interaction for desktop virtual reality. In Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, Bordeaux, France, 27–29 October 2008.
15. Balaa, E.; Raynal, M.; Bou Issa, Y.; Dubois, E. Classification of Interaction Techniques in the 3D Virtual Environment on Mobile Devices. In *International Conference on Virtual, Augmented and Mixed Reality*; Springer International Publishing: New York, NY, USA, 2014.
16. Song, J.; Soros, G.; Pece, F.; Fanello, S.; Izadi, S.; Keskin, C.; Hilliges, O. In-air gestures around unmodified mobile devices. In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, Bordeaux, France, 27–29 October 2009.
17. Oviatt, S.; Cohen, P.; Wu, L.; Vergo, J.; Duncan, L.; Suhm, B.; Bers, J.; Holzman, T.; Winograd, T.; Landay, J.; et al. Designing the user interface for multimodal speech and pen-based gesture applications: State-of-the-art systems and future research directions. *Hum. Comput. Interact.* **2000**, *15*, 263–322. [CrossRef]
18. Chatterjee, I.; Xiao, R.; Harrison, C. Gaze+ gesture: Expressive, precise and targeted free-space interactions. In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, Seattle, WA, USA, 9–13 November 2015; pp. 131–138.
19. Zhang, Y.; Stellmach, S.; Sellen, A.; Blake, A. The costs and benefits of combining gaze and hand gestures for remote interaction. In *Human-Computer Interaction*; Springer: Cham, Swithzerland, 2015; pp. 570–577.
20. Villaroman, N.; Rowe, D.; Swan, B. Teaching natural user interaction using OpenNI and the Microsoft Kinect sensor. In Proceedings of the 2011 Conference on Information Technology Education, Calgary, AB, Canada, 11–13 October 2012.
21. Bhuiyan, M.; Picking, R. A gesture-controlled user interface for inclusive design and evaluative study of its usability. *J. Softw. Eng. Appl.* **2011**, *4*, 513. [CrossRef]
22. Farhadi-Niaki, F.; GhasemAghaei, R.; Arya, A. Empirical study of a vision-based depth-sensitive human-computer interaction system. In Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction, Matsue-city, Japan, 28–31 August 2012.
23. Farhadi-Niaki, F.; Etemad, S.A.; Arya, A. Design and usability analysis of gesture-based control for common desktop tasks. In Proceedings of the International Conference on Human-Computer Interaction, Las Vegas, NV, USA, 21–26 July 2013; Springer: Berlin/Heidelberg, Germany, 2013.

24. Barclay, K.; Wei, D.; Lutteroth, C.; Sheehan, R. A quantitative quality model for gesture-based user interfaces. In Proceedings of the 23rd Australian Computer-Human Interaction Conference, Canberra, Australia, 28 November–2 December 2011.
25. Manresa-Yee, C.; Amengual, E.; Ponsa, P. Usability of vision-based interfaces. In Proceedings of the Actas del XIV Congreso Internacional Interacción Persona Ordenador, Madrid, Spain, 17–20 September 2013.
26. Cabral, M.C.; Morimoto, C.H.; Zuffo, M.K. On the usability of gesture interfaces in virtual reality environments. In Proceedings of the 2005 Latin American Conference on Human-Computer Interaction, Cuernavaca, Mexico, 23–26 October 2005.
27. Marin, G.; Dominio, F.; Zanuttigh, P. Hand gesture recognition with leap motion and kinect devices. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014.
28. Potter, L.E.; Araullo, J.; Carter, L. The leap motion controller: A view on sign language. In Proceedings of the 25th Australian Computer-Human Interaction Conference: Augmentation, Application, Innovation, Collaboration, Adelaide, Australia, 25–29 November 2013.
29. Weichert, F.; Bachmann, D.; Rudak, B.; Fisseler, D. Analysis of the accuracy and robustness of the leap motion controller. *Sensors* **2013**, *13*, 6380–6393. [CrossRef] [PubMed]
30. Thanyadit, S.; Pong, T. Desktop VR using a Mirror Metaphor for Natural User. Brighton, United Kingdom: ISS '17. In Proceedings of the Interactive Surfaces and Spaces, Brighton, UK, 17–20 October 2017.
31. Zhang, H. Head-mounted display-based intuitive virtual reality training system for the mining industry. *Int. J. Min. Sci. Technol.* **2017**, *27*, 717–722. [CrossRef]
32. Jibin, Y.; Gao, Y. The design and research of 3D desktop interface based on the Pen+ Touch. In Proceedings of the ICRA 2016-IEEE International Conference on Robotics and Automation, Stockholm, Sweden, 16–20 May 2016; pp. 129–133.
33. Song, P.; Goh, W.B.; Hutama, W.; Fu, C.W.; Liu, X. A handlebar metaphor for virtual object manipulation with mid-air interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 1297–1306.
34. Cournia, N.; Smith, J.D.; Duchowski, A.T. Gaze-vs. hand-based pointing in virtual environments. In Proceedings of the CHI'03 Extended Abstracts on Human Factors in Computing Systems, Hague, The Netherlands, 1–6 April 2003; pp. 772–773.
35. Davis, F.D.; Bogozzi, R.P.; Warshaw, P.R. User acceptance of computer technology: A comparison of two theoretical models. *Manag. Sci.* **1989**, *35*, 982–1003. [CrossRef]
36. Bertrand, M.; Bouchard, S. Applying the technology acceptance model to vr with people who are favorable to its use. *J. Cyber. Ther. Rehabil.* **2008**, *1*, 200–220.
37. Dünser, A.; Grasset, R.; Billinghurst, M. *A Survey of Evaluation Techniques Used in Augmented Reality Studies*; University of Canterbury: Christchurch, New Zealand, 2008; pp. 1–27.