# Designing Virtual Equipment Systems for VR

Powen Yao, Tian Zhu, and Michael Zyda

[1] GamePipe Laboratory, University of Southern California, Los Angeles, California, USA
{powenyao, tianzhu, zyda}@usc.edu

**Abstract.** Skeuomorphism in virtual reality can provide more than just visual cues for quick recognition, it can also provide spatial information. A Virtual Equipment System is a system that takes advantage of the spatial information from a user's body by equipping them with virtual counterparts of common wearables. It provides ways to quickly adjust a suite of functions that are conceptually related using spatial knowledge of human anatomy corresponding to the equipment. The system can also be used to provide shortcuts for interacting with properties of other objects related to the user's sensory channels. As a software solution, this Virtual Equipment System can be standardized and implemented across different devices and applied to different extended realities.

**Keywords:** 3DUI, Human-Centered Computing, Virtual Reality.

## 1 Introduction

In the current generation of virtual reality (VR) headsets, the user adjusts various audio-visual settings by using the menu button on the controller to bring up the system menu. Once the system menu appears, the process becomes much like using a mouse on traditional PCs. The controller is used as a pointer to interact with the menu and navigate to the desired option. This is a method that utilizes user knowledge that is easily transferable from traditional PC user interfaces, but it does not take advantage of the spatiality provided by VR or the user's own body awareness.

On certain VR devices such as Oculus Quest, users can adjust sound volume using buttons on the headset. Dedicated physical buttons enable quick access, but the positions and availability of these buttons are not universal across different VR devices. However, with a 6 Degree of Freedom (6DOF) Headset, we have spatial reference points for the headset. We can further associate the space around the headset to different conceptual areas of human anatomy such as eyes, ears, or mouth. We can then use the 6DOF Motion Controller to interact with these spatial reference areas.

Some VR games and experiences have explored using spatial reference points around the headset, but mostly for increased immersion. In games like Job Simulator [1], players can eat a food item by placing it near their mouth. In The Lab [2], players can bring a crystal sphere representing another world to their head to enter that world. An example of using spatial reference points for improved workflow and accessibility can be found in Fantastic Contraption [3]. As an advanced option, players can use the different areas around their head as shortcuts to grab different items (Figure 1).

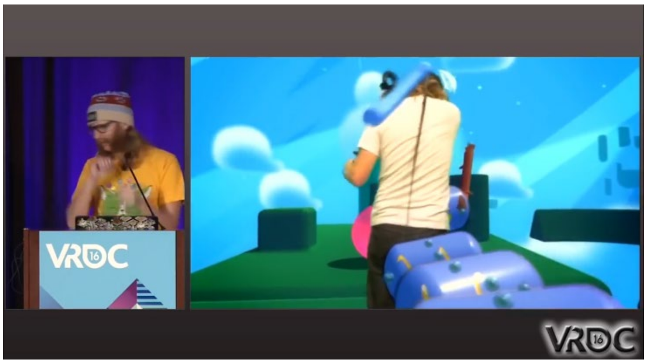


**Fig. 1.** User grabbing a stick from the left ear area of the headset in Fantastic Contraption.

Lastly, we take inspiration from headphones with gesture control such as Sony's WH1000XM3 headphones [4] or the helmet touch interface by Takumi et al [5]. Gesture Control increases the number of options that users have access to, given the same spatial reference point or area.

# 2 PROPOSED SYSTEM

In most VR Setups, the user is equipped with a 6DOF Headset and two 6DOF Motion Controllers. With these physical devices, we have spatial reference points for the head and the hands. Virtual Equipment (VE) is a group of objects that are socketed to these spatial reference points. 6DOF Motion Controllers provide the means to interact with the VE.

VE provides an alternative to the traditional menu interface. Users can grab and put on different equipment. Furthermore, users can interact with their VE through controller gestures or body gestures (e.g. arm gestures or foot gestures). Lastly, the VE can be grabbed and used with other systems for additional interactions.

## 2.1 Location of VE

To determine the locations of VE, we need to have spatial reference points for the user's body. Here, we describe three approaches: Headset and Controller tracking, hardware tracker-based tracking, and camera-based tracking. With additional hardware, a Virtual Equipment System can be bolstered with full body tracking.

The latter two approaches involve full body tracking. Full body tracking provides additional spatial reference points that can then be outfitted with Virtual Equipment. Depending on the tracking, we may gain additional direct spatial reference points, but also indirect locations related to those points. Additionally, full body tracking will provide extra information that can help us improve our evaluation of the ergonomics of the Virtual Equipment System.

**Headset & Controller Tracking.** Headset & Controller Tracking is the most widely available and easiest method to utilize because it does not need any additional hardware. We utilize spatial reference points for the helmet directly to outfit the user with VE such as headphones, goggles, or microphones. We similarly utilize the controllers to outfit the user with arm bracers. These VE, as shown in Figure 2, are displayed in addition to what the user is wearing and may be different than the physical objects (e.g. having virtual headphones while wearing earphones).

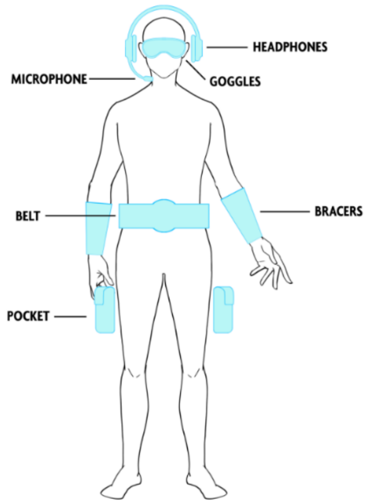


**Fig. 2.** Virtual Equipment available with Headset and Controller tracking that users can interact with using 6DOF Controllers in VR

For VE without a direct spatial reference point, we can estimate the associated location based on statistical data of the human body and the spatial reference points that we do have. For equipment such as a belt or pockets, we can use the height of the VR headset as the user's eye height and determine the user's waist height using anthropometry.

**Advanced Tracking.** One option for better tracking would be to use hardware trackers. Using hardware trackers such as the HTC's VIVE trackers, we can get direct spatial references for other body locations such as feet, knees, and elbows. The accuracy of the full body tracking will depend on the number of trackers we use.

At the time of this writing, the bulkiness and the inconvenience of the trackers makes it inconvenient to use many trackers. However, we can easily place two trackers on the user's left and right feet. Along with the existing headset and the two controllers, this gives us five points of reference. Using inverse kinematic solutions, we can perform inverse kinematics on a human body model to get an estimate of many more key points.

Another solution for better tracking would be camera-based tracking. Using two or more external cameras, we can track the human body using computer vision libraries such as OpenPose. Camera-based tracking has the benefit of being able to track more key points than trackers and does not require the user to wear additional hardware on the body. The downside is that this solution is restricted to a specific physical location and requires additional computing power, which can further tax a system that's already strained by enabling virtual reality.

### 2.2 Types of VE

VE can be grouped into two types: VE related to the human body and VE not related to the human body. The most straightforward example of the first type would be vision and hearing for the sensory systems as well as hands and feet for actuator systems. This category also includes VE that are conceptually associated with the respective body parts due to historic or cultural significance. These associations may not be as universal as the senses. Examples could include wristwatches for the wrist and a sword for the back or waist.

The second type includes anything else that does not fit the first category. This could be because the VE cannot be placed in its logical location due to other VE competing for the same space or because the VE cannot be functionally associated with any human body parts. For example, goggles with a snapshot functionality could be placed where the eyes are, but that location may be occupied by other goggles that allow the adjustment of brightness and enable alternate vision. The system or the user could choose to place the VE in a secondary spot, such as at the feet. However, the equipped VE would not be as intuitive as if they were the first type of VE, leading to more effort on the user's part to associate the VE with the body location.

Other functionalities in VR could take the form of VE and be placed on the user for quick access. For example, network settings could be made into a router that is equipped at the elbow. These VE move away from what is considered equipment in real life and cannot be easily associated with the body parts in which they are equipped, but they can be provided for the user's convenience within an existing VE system.

### 2.3 VE Interactions

VE such as a bow equipped on the back could allow the users to draw the bow and start firing arrows at targets like they could in real life. However, we envision three other types of VE interactions that are not commonly seen in real life: Drag & Drop, Controller Gesture, and Body Gesture

**VE Interaction – Menu Access via Drag & Drop.** The user can use a 6DOF controller to quickly access the menus associated with VE using a special menu node or an "Alt Node". We envision this node to function like the alt key of the keyboard, enabling an object placed in the node to bring up its alternative function. Tentatively, we have attached this node to the controller.

This operation functions similarly to a drag-and-drop operation in PC systems. First, the user moves the motion controller to the location of the VE (e.g. headphones), receiving haptic feedback as it enters the equipment. Then, the user pushes and holds the trigger button to pick up the headphones, attaching the VE to the controller. The user can then move the VE headphones freely. If the trigger is released while the VE headphones are inside the Menu Node, the menu associated with the headphones will be opened. If the trigger button is released while the VE is anywhere else, the VE will return to its original position.

**Fig. 3.** From left to right: User moving controller to grab virtual headphones. User releasing a representation of headphones at the menu node. Audio Settings Menu is shown to the user.

Using this interaction with Type 1 VE, users will be able to access general sensory related menus. For example, vision and graphics menus through the VE at their eyes (goggles), the audio menu through the VE at their ears (headphones), and the voice menu through the VE at their mouths (microphones).

Since Type 1 VE also represents the sensory channels the users have, VE can also serve as a shortcut to open the settings menus related to the sensory channel for a specific object. Using the same drag and drop method described earlier, users can drag a Type 1 VE (headphones) from their body location (ear) to an object in the world to bring up that object's sensory menu (audio menu) as illustrated in Figure 4.

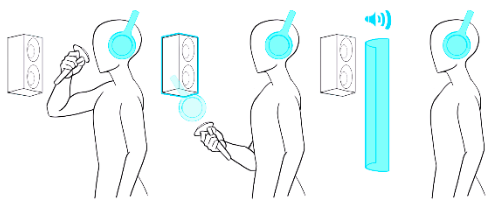

**Fig. 4.** From left to right: User grabbing virtual headphones with the controller. User placing the virtual headphones on an object in the world. Type 1 virtual headphones open the audio menu associated with the object.

**VE Interaction Example - Function Access Via Controller Gestures.** Some VE could include built-in motion gestures to allow quick access to different functions.

Much like grabbing and moving the equipment to the menu node, users can perform gestures by moving the controller to the equipment (e.g. headphones). Instead of grabbing and releasing at the menu node, users would grab and release the headphones slightly away from their original position to perform a controller gesture. Depending on the direction and distance moved, the gesture could trigger different functions as detailed in Figure 5.

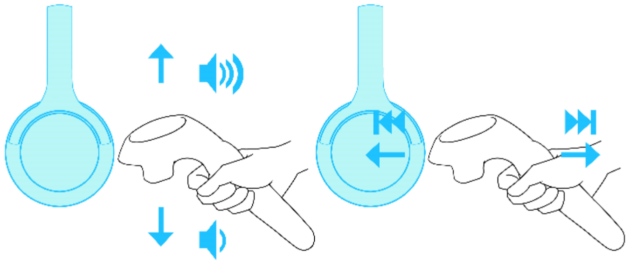

**Fig. 5.** On the left, using the controller to perform up and down gestures on the virtual headphones to increase and decrease volume. On the right, using the controller to perform a forward and backward gesture to go to the next song or previous song.

Like the object-specific sensory menu access, these controller gestures can also be expanded to target a specific object's audio or visual sensory channels. Users could select an object (grabbing with a controller or using a pointer) and perform the controller gesture to adjust that object's audio or visual properties as shown in Figure 6.

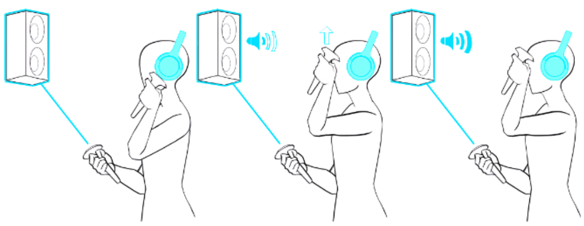

**Fig. 6.** From left to right. User uses the controller's pointer to select an object in the world. User uses the other hand's controller to interact with the headphones via controller gesture. The object's volume increases due to the controller gesture on the audio channel.

**VE Interaction - Function Access Via Body Gestures.** In addition to controller gestures, some VE are situated in places where the users could perform body gestures. For example, the user could have shoes equipped at the feet that can be triggered by tapping the heels together three times, like in the Wizard of Oz. Once the gesture is performed, it will trigger the associated functionality of the VE which could return to Home or bring up movement-related settings menus.

## 2.4 VE Customization

Since the VE is based on the user's body in 3D space, a user's physical characteristics affect our VE System. Improperly placed, the VE will be at best awkward to use, or at worst, inoperable. The ability to customize Virtual Equipment's positions becomes crucial for a good user experience. We outline two methods for user customization.

**Tailored UI or Tracking-enabled Customization.** We draw upon the clothing analogy to describe user interfaces that are modified by the system to work with a user's physical characteristics such as height, arm length, etc. This is atypical in UI design as most user interfaces are designed for the average human in order to be accessible for the most people

With the basic setup of a 6DOF headset and two controllers, we can utilize the headset to perform a user height calibration as well as utilize the controllers to get a user arm length calibration. With that information, we could then refer to anthropometry data to get estimates for the rest of the body. With full body tracking, we could have UI that's tailored or made-to-fit for the user. Aside from placing each VE in appropriate positions with respect to the user's body, we can adjust interactions based on the user's exact measurements.

For example, when the user performs an up gesture on the virtual headphones, the user's arm length and the rotation of the joints determine how easy it is for the user to move up and down. For people with longer arms, their hands might easily go out of bounds and end up moving the virtual headphones instead of performing an up gesture. It is critical that these interactions in 3d space take into consideration the user's physical attributes when determining the position and orientation in which users must move their controllers or body to interact.

**Virtual Avatar Mannequin.** The system can make its best prediction for what the user would want, but there are times when these predictions are insufficient. We need a way for the users to be able to adjust intuitively and easily. For that, we envision a mannequin that mirrors the user in terms of VE equipped. The mannequin provides a way for users to visualize and adjust the settings for their various VE.

When users interact with the VE that is equipped on themselves, it could either activate a gesture or move it for drag and drop interactions. When the user interacts with the VE that is equipped on the mannequin, it would adjust the position of the VE with respect to the body area they are equipped in. This is useful for situations where the user's VE cannot be interacted with due to it being stuck in a real-world object such as virtual headphones inside actual headphones or a virtual utility belt inside someone's belly.

Aside from adjusting the physical location of the VE, the mannequin can also provide feedback for performing controller gestures as well as adjusting controller gestures, something that's otherwise hard to visualize and see. With a mannequin mirroring the user's every action, the user can see how much they must move their controller to trigger a controller gesture. The user can also spatially adjust the parameters for what would be interpreted as a controller gesture.

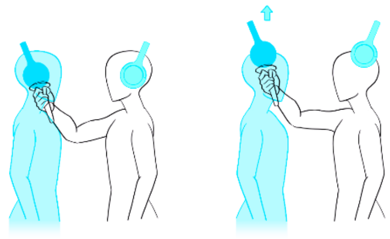

**Fig. 7.** From left to right. User grabbing the virtual headphones on the mannequin using a controller. As the user moves the controller and thus the virtual headphones up, user's equipped virtual headphones also move up

## 3 Evaluation

To evaluate the Virtual Equipment, we need to compare what it takes to perform a task using the different interaction methods in Virtual Equipment versus the traditional methods of using a menu button followed by pointers or dedicated hardware buttons.

Quantitative data that we can capture would be the time it takes, the user's accuracy, and the motion required to perform an action. If full body tracking is implemented, we would have more accurate quantitative data for the action required. We would also like to capture qualitative data to compare the effort and comfort. We will ask users about their experience with VE through user surveys with questions about areas such as difficulty, intuitiveness, comfort, and effort required.

Some basic tasks we want to compare would be adjusting the audio volume, adjusting the brightness of the display, opening the audio settings menu, and opening the visual settings menu. For each task, we will compare a VE system with the currently existing traditional interfaces.

## 4 Conclusion

There are many advantages to a Virtual Equipment System that utilizes the spatial reference points provided by a VR headset and its controllers. A user's spatial reference points can be utilized directly and indirectly, as well as bolstered through tracking. There are a variety of different ways to interact with the Virtual Equipment System, including how the Virtual Equipment can be tailored and customized, and how we can evaluate the performance of these Virtual Equipment.

## References


1. *Job Simulator* (2016). Owlchemy Labs.
2. *The Lab* (2016). Valve Corporation.
3. Fantastic Contraption and why VR Menus Suck, http://www.youtube.com/watch?v=ASXST_iyhl4 , last accessed 2020/03/20
4. Controlling the audio device (Bluetooth connection), https://helpguide.sony.net/mdr/wh1000xm3/v1/en/contents/TP0001703116.html, last accessed 2020/03/20
5. Kitagawa, T., Yamato, Y., Shizuki, B., & Takahashi, S. A Viewpoint Control Method for 360° Media Using Helmet Touch Interface, 1–2. (2019)