ABSTRACT

Quantitative Evaluation of the HTC Vive Virtual Reality System in Controlled Movement

Tyler A. Jost, M.S.B.M.E

Advisor: Jonathan Rylander, Ph.D.

Modern virtual reality systems enable users to be immersed in a virtual world. Validation of the HTC Vive will assure users and developers that games and applications made for the system are accurate representations of the real world. Here, both the translational and rotational capabilities of the HTC Vive are investigated using a robotic arm and an optoelectronic motion capture system. It was found that the average difference between reported translational distances traveled was 0.74 millimeters with an interquartile range (IQR) of 0.66 millimeters for all room-scale calibration trials and 0.63 millimeters with an IQR of 0.29 millimeters for all standing calibration trials. The mean difference in angle rotated was 0.46° with an IQR of 0.81° for all room-scale calibration trials and 0.54° with an IQR of 0.62° for all standing calibration trials. Overall, the HTC Vive shows promise as a tool for clinic, research, and industry.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

First of all, I would like to thank Dr. Jonathan Rylander. A short two years ago, I wanted to leave Baylor for what I thought were greener pastures. Eventually, that did not work out and I ended up staying at Baylor. Jonathan was there encouraging me for all of it, and if I didn't have his support I don't know where I would be now. When I think of the Christian spirit Baylor tries so hard to embody, I think of Jonathan Rylander.

I want to extend special thanks to Grant Drewelow and Jenny Tavares for truly becoming my adopted family. I don't know if I could have spent two years hoofing it to the back of the BRIC if I didn't know at least one of you would be there. I don't think many people go into engineering expecting comradery, but I feel I found a second family because of you two. I will greatly miss planning, researching, and dreaming with you both.

Finally, I want to thank my family and friends. Thank you especially to Alison, Austin, Kat, and Anthony for staying in Waco and making weekends and nights fun and worthwhile. Thank you to Elliott for always pushing me to be a better researcher and person. Thank you to my Dad for making science so special and important, you really jumped on that train before it was cool and I will always admire you as someone who thought deeply about anything and everything. Finally, thanks to my Mom for always being someone I could call and someone I could rely on. I love you all.

CHAPTER ONE

Introduction

Virtues of Virtual Reality

Virtual reality, while not a new field, has experienced rapid growth in the last decade with the release of several mainstream consoles, most notably the HTC Vive and Oculus Rift. The first modern virtual reality system was created by Ivan Sutherland[1], and in many ways reflects the current design of virtual reality headsets. Prior to this, no system existed which dynamically moved and subsequently reflected that movement in the headset (or other display device) of the user. Defining what virtual reality is can be difficult due to how all-encompassing the definitions tend to be. In some literature, anything taking place in a virtual environment (i.e. a video game) is described as virtual reality. However, this is confusing because there is a significant difference between a traditional video game, which has an inherent 3rd person perspective, and a virtual reality video game, which is always in the first person. This initial distinction comes entirely from the nature of the hardware. Systems such as the Xbox Kinect (a game system which tracks the position of the user in order to interact with games) and Nintendo Wii (a game system which tracks the position of a game controller to interact with games) serve to further muddy the waters because they have dynamic input that is received and processed by the system. For example, De Paolis states virtual reality as being, "a computer-generated, immersive, multi-sensory information program which tracks a user in real time"[2]. However, the Xbox Kinect is able to loosely fit this definition. A decade or so before, this might have been a correct definition for existing virtual reality systems, but technology has progressed to a level where it becomes difficult to determine what virtual reality is in a singular definition, even if there is an implicit understanding of what separates an HTC Vive and any other game system.

In light of this, it makes more sense to classify virtual reality and its kind on a continuum ranging from the real world to a complete virtual environment. Modern virtual reality systems can be considered to be the furthest from the real world on the virtuality continuum as defined by Milgram and Kishino[3]. Where a system lies on the virtuality continuum can be decided by three other continuums which determine how much of the world is tracked and understood by the computer (extent of world knowledge), how realistically it can be displayed (reproduction fidelity), and how immersive it is (extent of presence metaphor). Modern virtual reality systems achieve the most "extreme" end of these continuums, which can be described as a fully virtual environment.

The extent of world knowledge continuum describes how the user's entire body is placed in relation to the environment. On the furthest end of the spectrum, the computer has knowledge about each object in the world and can allow the user to interact with these objects. For example, a device such as the Kinect might use a backdrop which has no relation to the user, whereas with the HTC Vive even the background is able to be touched, even if no interaction occurs. Therefore the HTC Vive is further towards the virtual environment end of the virtuality continuum. With modern virtual reality devices, reproduction fidelity is also on the extreme end of the spectrum, with some systems able to generate near photorealistic environments. However, the graphical rendering capabilities are still capable of being upgraded. For example, the newest generation of

HTC Vive technology does not improve its motion tracking ability, but instead focuses primarily on increasing the resolution in the head mounted display (HMD)[4]. While the display in the HTC Vive might still not be as clear as some televisions, it is still far closer to a depiction of reality than prior systems. Finally, modern virtual reality systems can be considered to be the furthest on the extent of presence metaphor continuum due to their ability to provide realtime imaging. This means that the user is not only able to travel around in the world, but able to receive updates to this travel as well as any other updates to the environment as they occur and is common in modern virtual reality systems.

Technical Implementation of Virtual Reality

Currently, there are no published papers by HTC/Valve or Oculus VR (makers of the HTC Vive and Oculus Rift respectively) on how their technology works exactly. However, a YouTube video by Valve engineer Alan Yates[5] outlines the basic concept behind the HTC Vive. The system uses two components for tracking. The first is an inertial measurement unit (IMU), which is a combination of an accelerometer and gyroscope. The IMU is located on the neck of the controller. The second is an optical system which emits infrared (IR) light. This is accomplished with two separate towers, called base stations. The IMU is able to determine location by twice integrating the acceleration of the controller, which gives position. However, this leads to a great deal of error due to the nature of the signal as well as the noise that comes from integration of signals. To correct for this, the base stations correct the position. This is accomplished by emitting IR light which is received by the controller. Position is determined by the time delay between when the base stations first emit the IR and when it is received by the controllers. The base stations are configured in such a way that the controllers are also able to discriminate between them. No such information exists describing the Oculus Rift in such detail, however the Oculus Rift does have towers which emit IR as well as controllers which have an IMU[6].

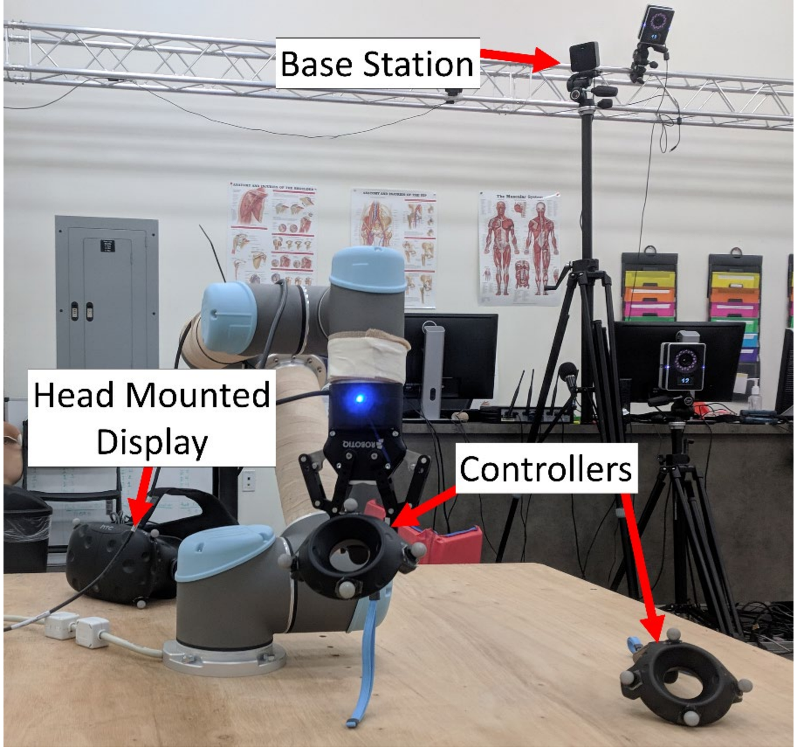


Figure 1.1. HTC Vive components.

HTC Vive Specifics

The HTC Vive is a virtual reality system created by HTC and Valve Corporation. The base system comes with an HMD with a resolution of 1080x1200 pixels per eye[7] and has a refresh rate of 90 Hz. The system also comes with two controllers and the previously mentioned base stations. There also exist separate accessories called Vive Trackers. These are predicated on the same technology, although they do have a different light to digital converter and circuit board[8]. These are capable of being attached to physical objects or other body parts on the user to create a more interactive experience. To use the HTC Vive, a computer must be used to process the data from the HTC Vive itself.

Virtual Reality's Impact on Clinic

Serious games, while having no strict definition[9], can be defined as "games used for purposes other than mere entertainment"[10]. Exergaming is another term that is occasionally used interchangeably with serious games[11], but more specifically refers to games where users exercise to progress in the game[12]. The obvious potential to using exergaming with physical therapy is that a video game is inherently more engaging than traditional modalities such as paper or video. Physical therapy is an inherently uninteresting task due to its repetitive nature, and a home exercise program will only be effective if a patient decides to take personal responsibility in completing their exercises. This presents two problems which can be potentially alleviated by exergaming: creating a game which is engaging to the user in a way that makes them want to complete their tasks, and creating a game which can offer feedback and coach players into performing exercises effectively.

Precision and Accuracy Needed for Rehabilitation Assessment

Before comparing the Vive to prior exergaming systems, it must be understood what level of accuracy and precision would be required to evaluate a patient. Determining the efficacy of varying tools in predicting rehabilitative outcomes can be challenging because tools can range from purely qualitative to purely quantitative[13], [14]. For example, the Berg Balance Scale[15] uses a qualitative assessment of a patient’s ability to ambulate, but is still used to assess recovery of stroke survivors. Another functional assessment, the Shoulder Mobility Screen[14], scores a patient based on the distance between their hands after placing them behind their back. However, the score is based on the distance in reference to the length of the patient’s hands, ranging from 0 to 3. A more precise functional movement assessment is the Functional Reach[13], which uses a tape measure in inches to analyze the forward reaching ability of a patient. Despite being a more quantitative assessment, the Functional Reach showed strong correlation with the more qualitative Berg Balance Scale in analyzing post-acute stroke individuals[13]. Because the Functional Reach test was shown to be sufficient within a precision of inches, it would be reasonable to suggest that the Vive should have a tracking error of less than 1 inch when evaluating reaching tasks.

Predecessors to the HTC Vive

Although exergaming has primarily been focused on the Nintendo Wii, Microsoft Kinect, and the like, the concept of exergaming predates this, with one of the earliest examples being the Atari Joyboard[16]. However, the Wii and Microsoft Kinect have made significant technological strides and represent a new era in exergaming. The efficacy of the Microsoft Kinect in particular has been of interest to researchers[17]–[21]. The Microsoft Kinect (referred to as Kinect, with the first iteration being called Kinect V1) operates via a depth sensor that sends out a grid of infrared light, then uses time of flight analysis on the light that is reflected back[20]. The objective of the Microsoft Kinect is not to track singular points, however. It feeds this time of flight data into a randomized decision forest classifier[22]. This makes the Kinect more similar to a motion tracking system than a singular point tracking system such as the Nintendo Wii, making it easy to repurpose it for a variety of uses. This pose tracking allows the device to calculate joint angles, identify exercises, and even quantify elements of Parkinson's disease[23] and self-esteem[24]. The downside to this method is that it is relatively easy to occlude body parts while performing exercises. If a joint becomes occluded or is not tracked, the Kinect would be forced to guess the position of the joint[18].

Quantitative evaluation of the Kinect's tracking ability yielded a variety of results. Wang et al. reported that joint positions were offset anywhere from 50-100 mm[19] when compared to a research grade motion capture system on humans. In contrast, Mobini et al.[25] created a physical model and found errors ranging from 28-38 mm, although this decrease could have been caused by the controlled nature of their study. The most recent release of the Kinect, the KinectOne (also referred to as the Kinect V2), calculated arm angles with a precision of 3.9 ± 4.0°, and trunk angles of 0.1± 3.8°[26]. Despite these disparities in accuracy, the Kinect represented a vast leap forward in the realm of low-cost motion tracking. It required no markers and was widely available, while presenting a moderately accurate motion tracking system that could be incorporated into video games. The Kinect, for its time, was the best solution available to difficulties faced by home exercise programs.

The cost of healthcare in the United States is significantly higher than other countries[27], and this has led to debate about whether performing exercises at home can be comparable to performing the exercises under supervision[28], [29]. In addition to the Kinect's ability to accurately track motion, it has been shown to increase efficacy of home exercise programs[17], [30]. Exergaming systems are more enjoyable than traditional modalities because they can be gamified[31], meaning game elements are added which can increase the user's likelihood to keep playing the game. For example, a game might reward a user with tokens for completing a number of exercises. These might be used for something trivial, like changing the background of the world, or something more concrete like an upgrade to a tool used in the game. Furthermore, exergaming systems which can track motion enable that tracking to be recorded. This information can then be compiled by the software into a report which can be reviewed by a physical therapist. This way, there is a check that can be made by a human outside of just the computer, and changes can be made to the patient's regimen.

A further step would be to diminish the time physical therapists spend directly coaching patients. It can be argued that, with sufficient tracking technology, a physical therapist should spend less time evaluating exercises because the therapist is unable to know how accurately joint angles and positions are formed. Additionally, an exergaming system is able to make real-time changes and can be used at any time in the comfort of a patient's home. This is not to say that a physical therapist is obsolete or unimportant, as there are intangible factors that are better addressed by a human. In the same vein, physical therapists add an emotional human component which can potentially be important for rehabilitation. However, diminishing time that a physical therapist has to spend in direct contact with a patient saves the patient's money and allows the physical therapist to focus on more pressing matters that might not otherwise receive enough attention.

In many ways, virtual reality is primed to overtake the Kinect and similar exergaming systems. Modern virtual reality systems are immersive, which can bring a greater level of engagement when compared to previous systems. Virtual reality systems do not have tracking occlusions in the same way the Kinect does, and Vive trackers offer the capability to do full body motion tracking for a low price. Furthermore, it is possible to place real world items into the game world with virtual reality, enabling clinicians to incorporate objects like stairs, medicine balls, and other rehabilitative tools into the virtual world. Finally, virtual reality systems offer a large capture volume which allows for more movement and interaction than ever before. Proper implementation of virtual reality can enable better recovery for patients receiving physical therapy.

Virtual Reality's Impact on Research

Virtual reality has a history of being utilized in research application. The distinction must be emphasized between virtual reality that is immersive and fully interactive, such as the HTC Vive, and prior "virtual reality" systems such as the Xbox Kinect and Nintendo Wii because there is a significant difference in the change they are able to affect in subjects. Because modern systems are immersive, they are able to bring a new dimension to existing research and enhance experiments.

One field which has received significant benefits from virtual reality is therapy. A number of disorders, including PTSD[32], agoraphobia[33], and arachnophobia[34] have been able to be treated via exposure therapy through virtual reality. In one event recounted by Hoffman[35], patients who received treatment for painful burns reported lower pain levels when distracted by an immersive virtual reality environment than when playing traditional video games. In Hoffman's example the virtual reality environment's original intent was exposure therapy for spiders. Its original purpose was not even tangentially related to the task at hand. This is important because it demonstrates that traditional video games, despite having the sole intent of entertainment, are less effective than virtual reality games which are not necessarily intended to entertain. The reason for this decrease in pain sensation is that so much attention is given towards the environment that not enough is available to fully process signals for pain.

Pan and Hamilton[36] provide further strong arguments for using virtual reality in psychological research beyond exposure therapy alone. The primary reason is the control that virtual reality allows. Experiments can have a single variable changed with everything else held constant with ease. In contrast, sufficiently complex experiments which take place in the real world require can often require great amounts of effort. For example, a study which aims to compare social reactions to race and gender under traditional methods would have to gather many actors and actresses that meet specifically. With virtual reality, it is as simple as changing a character model. Furthermore, virtual reality enables a high level of reproducibility, which has become an increasingly prevalent topic in psychological research[37]. Experiments made in virtual reality will always have their components in the same place relative to the user, and events can be consistently triggered in a repeatable fashion. Another reason virtual reality is valuable in psychological research is its rich environment. Pan and Hamilton explain that social interactions are usually explored by distilling these interactions down into a single component. While these studies have proven useful, they would not be as valid as a study done with virtual reality where a more lifelike representation of reality can be presented. Finally, virtual reality provides the opportunity for subjects to take part in otherwise dangerous scenarios. For example, Zanon et al.[38] used virtual reality to simulate a scenario in which subjects needed to escape from a fire and had the option to save an individual at risk of being injured by the fire.

The paper by Zanon et al. was greatly enhanced by the use of virtual reality. Prior to virtual reality, dangerous scenarios were reduced to only text or cartoons. Additionally, virtual reality allowed for a “flowing experience”, as opposed to one that might be segmented by various scenarios presented through traditional mediums. It might be suggested that this is an overextension of virtual reality, or that the results have been overstated. However, there exists no comparable alternative to virtual reality other than real life for studies such as this. Furthermore, it has been demonstrated that increasing ecological validity (that is, the realistic nature of an experiment) is able to bring decisions made by subjects in experiments closer to real world behavior[39]. Therefore, it could even be argued that psychologists and researchers that do not use virtual reality are not being responsible in terms of experimental design because they will almost always be forced to choose a methodology with lower ecological validity and reproducibility.

Virtual Reality’s Impact on Industry

Virtual reality’s primary impact on industry is its ability to demonstrate objects and tasks without the need for external parts. For example, a company might be able to train a worker without needing a manager there to supervise. Likewise, a company might one day be able to demonstrate its products in every home by making models available online. Complicated tasks, such as building an engine or repairing a computer could be reduced to simple step by step tutorials where a user is free to interact without risk of damage to their own property.

One company making the biggest impact through virtual reality implementation is Osso VR (Palo Alto, CA) [40]. The concept of a virtual reality simulator has been around from before the advent of Osso VR[41] because of the many advantages virtual reality provides. As mentioned before, virtual reality demonstrations save a supervisor from being present. For medical research, this also allows the trainee to practice before operating on a more costly cadaver. Osso VR is even able to present challenges that might be faced by the surgeon, and users who were trained with virtual reality performed twice as well as those who did not.

Entire companies, such as EON Reality (Irvine, CA), have sprung up just to make simulations for training employees. Major companies such as Lowes (Mooresville, NC) [42] have developed virtual reality training software that can be used by customers to train them on their products. While it may not beat the real thing, virtual reality is the closest available tool to simulate and demonstrate products, and it is evident that industry will continue to support it even if home usage declines[43].

Current Research

Due to the recent jump into the commercial realm, there have only been a handful of attempts to quantitatively evaluate the accuracy and precision of commercial virtual reality systems. The first[44] attempts to quantify the efficacy of the Oculus Rift, but only in tracking the cervical portion of the spine. To accomplish this, the Euler angle sequence derived from the rotation matrix describing the Rift's headset was compared to the Euler angle sequence derived from tracking of the head and trunk, which was determined from a commercial active motion capture system. To better understand the overall accuracy and precision of the Rift, a separate inertial sensor was placed against the front cover of the Rift's HMD. The researchers determined that the Rift was, at best, an estimate of the movement of the spine.

The HTC Vive has received more interest, with a paper by Niehorster et al.[45] attempting to generally quantify the translation and rotational accuracy of the Vive. For translation data, the general procedure was to place the headset on a tripod and look at its output data while it was completely static. The authors then reported on the average reported location compared to the "actual" location. To assess the orientation accuracy of the Vive, the researchers would first find the most optimal rotation matrix relating the 3D coordinates of the Vive to the 3D coordinates of its real world location. They used a least squares approximation following the methods of Arun, Huang, and Blostein[46]. After this rotation matrix (called the reference plane by the authors) was determined, the authors would rotate the reference plane by a yaw angle reported by the Vive. They would then decompose this new rotation matrix into pitch and roll values and compare these to the pitch and roll values reported by the Vive. Finally, the authors looked at the RMS values of the waveform reported by the Vive to determine its "jitter" as well as its spread in location. All data was taken with a program called Vizard. The authors came to several conclusions. First that the Vive was permanently slanted, and second that this slant would change every time the Vive lost tracking.

The conclusions drawn from the Niehorster study could potentially be explained by examining their virtual reality setup. A primary example of this is the capture volume used in the study. The base stations were placed 7.4 m apart, which goes way beyond the official recommendation of 5 m [47]. In fact, this would actually trigger a warning from the software informing the researchers that the base stations were too far apart. This is not mentioned by the researchers, but potentially was chosen to test the limits of the system. Unfortunately, this likely negatively affected the results because the Vive was not meant for that distance. Furthermore, the Vive was recorded in a completely static pose. From anecdotal experience, leaving the controllers or headset in a completely static pose will cause them to track inaccurately. In contrast, the controllers will rarely jitter when moving. It can be hypothesized that this is likely due to the IMU located within the controllers and headset, which will experience inertial drift.

The authors of the Niehorster paper were also pigeonholed by their methods. The real world position was determined by a grid drawn out by the researchers. The average positioning error was reported as 1.7 cm. It is unclear as to what tool was used to determine the spacing of the grid, and it is questionable if this grid is accurate enough to validate the Vive. As previously stated, a least-squares approximation was used to determine a rotation matrix relating the Vive coordinates to the real world coordinates. The authors report that this is the best possible calibration between the Vive and the real world. However, the authors had access to a device used to measure orientation (Intersense IntertiaCube4), and should have been able to determine the Vive's orientation using a reported rotation matrix. With these two pieces of information, it is possible to determine the actual rotation matrix relating the Vive and the cube. Admittedly, the method used by Niehorster et al. from Arun, Huang, and Blostein was effective when implemented on data taken from this lab, but it does not present an exact solution. Additionally, orientation was determined using euler angle representation. While euler angles are easier to interpret, they are prone to singularity and are heavily reliant on how the rotation matrix is decomposed.

The first assertion put forth by Niehorster et al., that the Vive's reference plane is internally tilted, has been addressed as a known problem by developers and users[48], and even has a solution discovered by users[49], [50]. However, it is compelling that two separate Vive systems were used in the study and that both were considerably slanted. The second assertion that the Vive's internal reference plane changes each time tracking is lost is less easily explained, but a further dissection can be found in the discussion of this thesis.

The findings of Niehorster et al. were further expanded upon by Peer et al.[51]. Some of the methods used by Niehorster et al. were repeated by Peer et al., including the method of tracking the position of the Vive using a grid. To account for error in the position of the controllers and headset, Peer et al. used a single Vive Tracker and placed it in known spots. Using the coordinates from the Vive Tracker, they were able to form a plane describing its rotation and use this to reduce error. Before accounting for this, Peer et al. achieved errors in the reported height similar to Niehorster et al, but was able to greatly reduce it using this method.

In general, there is much interest in validating commercial virtual reality systems for use in clinic, research, and industry. However, the current research is not highly supportive of the system for use in research. This is especially notable when contrasted with systems such as the Xbox Kinect, which can have tracking errors of up to 38 mm[25] (depending on which version is used). This discrepancy can likely be attributed to the researchers' context, which will greatly influence how they perceive errors. However, due to the problematic design of the previous experiments, there is merit in meticulously designing an experiment which rigidly examines the movement of the HTC Vive and how accurately it tracks.

Purpose

The purpose of this study was to evaluate the HTC Vive controllers for use in clinic, industry, and research. This was done by comparing the relative movement of an HTC Vive controller as reported by the controller to the relative movement reported by a system with a self-reported accuracy of better than 0.5 mm. Movement was done under rigid conditions in which the controller was either translated or rotated in a repeatable fashion under consistent conditions. To test the Vive for use in a variety of environments, the experiments were repeated under both of its internal configurations: room-scale and standing. Analysis of the rigid motion will be able to potentially validate the HTC Vive as a more accurate and precise tool for measurement against existing tools such as the Xbox Kinect. The primary aims of this research are as follows:

Aim 1: Establish a framework for testing six degree of freedom (DOF) virtual reality systems. Prior work on the HTC Vive reflects a low level of scrupulous methodology. Developing a rigid framework for testing virtual reality systems also could prevent redundancy in further papers.

Aim 2: Translational accuracy of the HTC Vive controller. Does the controller track accurately between its starting and ending location when moved in a singular direction? It is hypothesized that, due to the redundancy in tracking from the base stations and internal IMU, the HTC Vive will perform equally well in all directions. This will give users confidence that they can translate the controller in any direction without concern that it will perform consistently worse in any condition.

Aim 3: Rotational accuracy of the HTC Vive controller. Does the controller track the orientation of the controller between its starting and ending location when rotated 90 degrees? Furthermore, does it track equally when it is yawed, pitched, and rolled about each axis? It is again hypothesized that, due to the redundancy in tracking from the base stations and internal IMU, the HTC Vive will perform equally well in all directions. This will give users confidence that they can rotate the controller in any way without concern that it will perform consistently worse in any condition.

Aim 4: Effect of velocity on the HTC Vive controller. Does increasing the velocity from a relatively slow speed (500 mm/s) to a speed more reflective of human movement (1000 mm/s) change the tracking accuracy of the HTC Vive controller? It is hypothesized that the controller will track equally well under each speed. This will give users confidence that the controller will track well under slow and controlled movements as well as more fast-paced gameplay.

Aim 5: Effect of configuration on the HTC Vive controller. The HTC Vive's setup instructions contain two configurations: room-scale and standing. The setup instructions are different for each one, and base stations are set closer to the subject during standing calibrated sessions. Does changing this configuration change the accuracy of tracking of the HTC Vive controller? It is hypothesized that the controller will not track consistently worse in either case because the standing scale is not advertised as a way to increase quality of game experience, but rather as a method to play games in more confined spaces.

Aim 6: Evaluate the HTC Vive's tracking and orientation precision under clinical settings. Will the HTC Vive be able to track equally well when used by a human performing a clinical motion? It is hypothesized that the controller will not track as accurately in a clinical setting.

CHAPTER TWO

Methods

A systematic approach was undertaken to track the position of the HTC Vive controller in a consistent manner. Translation and rotation accuracy and precision of the Vive controllers were tested in different movement directions, movement speeds, and within different calibration collection volumes.

Evolution of Data Collection

As previously stated in the "Current Research" section, there are a number of pitfalls which can prevent the HTC Vive from working properly. Many of these pitfalls were discovered while designing the methodology for translating and rotating the controller. The first major issue encountered was IR interference from the cameras. The HTC Vive and the Vicon motion capture system both rely on IR light to track motion, and both emit light at 850 nm[8]. Initially, the capture volume for the Vicon cameras was extremely small, which caused them to interfere with the controllers' tracking. Noticeable drift occurred, with the controller wildly jumping from spot to spot. This was corrected when the system was in the motion capture laboratory, indicating that the light from the cameras was diffused enough in the larger capture volume. No further aberrant motion was detected when using the system, so it was concluded that all further experiments should take place in the motion capture laboratory.

The second problem which arose while designing the methodology was the orientation of the controller in the robot's gripper attachment. The initial orientation was similar to how a human might hold the controller, with the shaft of the controller orthogonal to the gripper. It was noticed that after about 20 trials, the controller would begin to lose its tracking and “jitter”. This could be reset by shaking the controller, but the problems would soon resurface. Throughout the course of the experiment, the controller would gradually become more jittery when placed in the gripper. It was initially believed that the combination of the base station and robot was preventing the IR light from reaching the controllers. To test this belief, numerous orientations of the base stations and robot were attempted, however no configuration could prevent the controller from jittering. Because of the researchers’ previous experience with the IR signals from the Vicon system, the cameras were shut off to ensure there was no IR interference. However, this did not prevent the jittering either. Finally, the robot was also turned off, in case there was some sort of electrical interference preventing the controllers from properly tracking. However, this did not prevent the jittering. Eventually, it was discovered that placing the controller in parallel with the robotic gripper allowed the controller to track properly. While there was still jittering, the controller could be “reset” by shaking it around, and this reset would last significantly longer than previously. The researchers believe that the drift experienced by the controller is exacerbated by the static pose, eventually causing its position to conflict so severely with the base stations that it begins to move erratically. Another conjecture is that the base stations require movement in order to properly track the controllers, so a static position relies solely on the IMU in the controller.

Data Collection

Data was collected via two systems. The first, a 14-camera Vicon optoelectronic (Vicon Motion Systems LTD, Oxford, UK) motion capture system. The system was calibrated to have an accuracy of better than 0.5 mm. Markers were placed on the controller and subsequently tracked by Vicon. The second was the HTC Vive system itself, which was gathered with a custom Python script.

Collection of World Data

As previously stated, a 14-camera Vicon optoelectronic motion capture system was used to track the controller. Four markers were placed on the head of the controller in a diamond pattern. Markers were strategically placed so as to not block any IR receivers on the controller. Markers were not labeled if they were unable to be properly tracked by the motion capture system (meaning that there were gaps in the data). The system was set to collect at 250 Hz, which was the theorized sampling rate of the Vive output program. Before outputting any marker trajectories, all trajectories were filtered using a Woltring filter, a filter used by Vicon to smooth marker trajectories[52].

Collection of Vive Data

As previously stated, a custom Python script was used to access the Vive controller's position and orientation. Python bindings were used to access Valve's OpenVR virtual reality SDK[53]. On top of this, a wrapper was used to facilitate accessing the Python bindings[54]. The wrapper was developed by an employee of Triad Semiconductor, a company responsible for some hardware contained by the Vive. Because the wrapper converted the rotation matrix of the HTC Vive to quaternion and

Euler angle format, the wrapper was modified so that the rotation matrix was directly outputted instead. This allowed the researchers to access the raw data with no potentially inaccurate conversions. Rotation matrices were chosen because they computational time is not a concern with this level of data, and they do not experience singularities like Euler angle sequences.

Set Up of HTC Vive

Prior to any official collection, the HTC Vive was recalibrated according to its internal room setup procedure. There are two internal calibration procedures provided by the vive: room-scale and standing. When collecting data from the room-scale calibration, the base stations were placed 18.5 feet apart. For the standing configuration, the base stations were placed 8.5 feet apart. During both configuration types, the base stations were raised above 6.5 feet, and were angled down between 30-45 degrees. The translation orientation of the controller is shown in Figure 2.1, and the rotation orientation is shown in Figure 2.2.

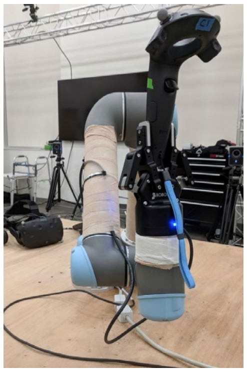


Figure 2.1. Placement of HTC Vive controller during translation.

Movement of the Controller

The controller was moved in a repeatable fashion by using the Universal Robots UR5 (Universal Robots, Odense, Denmark). Robot arms are effective tools for use in motion tracking system validation studies. A key reason is that robot arms can provide consistent, repeatable motion. The UR5 has a precision of ±0.1 mm (4 mils)[55] which allows for study of not only the accuracy of the HTC Vive, but also its error variance in a controlled study. This is because the robot arm provides the ability to repeat the same motion pattern many times over. The robot was programmed with custom code using its internal programming language.

Translation Difference between Vive and World

The robot moved the controller 400 mm in either the X, Y, or Z directions as determined by the robot's internal coordinate system. 400 mm was the chosen distance because it covered approximately the furthest distance the robot could move in all directions while starting from the same origin point. For two of the directions (X and Z), the robot could have moved further, but moving the robot further in the Y direction would have compromised its ability to move properly. Two speeds were chosen to test the Vive's ability to track slow and fast movements: 500 mm/s and 1000 mm/s. 1000 mm/s represents an approximate value of the average movement of the arm during normal gameplay, and 500 mm/s was chosen as a convenient slower speed to test the Vive tracking under ideal conditions. The following list represents the combinations used in the translation trials:

1. Room Scale Calibration translation in the X-Direction at 500 mm/s
2. Room Scale Calibration translation in the Y-Direction at 500 mm/s

3. Room Scale Calibration translation in the Z-Direction at 500 mm/s
4. Room Scale Calibration translation in the X-Direction at 1000 mm/s
5. Room Scale Calibration translation in the Y-Direction at 1000 mm/s
6. Room Scale Calibration translation in the Z-Direction at 1000 mm/s
7. Standing Calibration translation in the X-Direction at 500 mm/s
8. Standing Calibration translation in the Y-Direction at 500 mm/s
9. Standing Calibration translation in the Z-Direction at 500 mm/s
10. Standing Calibration translation in the X-Direction at 1000 mm/s
11. Standing Calibration translation in the Y-Direction at 1000 mm/s
12. Standing Calibration translation in the Z-Direction at 1000 mm/s

Each trial type was collected 30 times, for a total of 360 trials. The Euclidean distance moved by the controller was then calculated. This was done by first averaging the first 50 frames and the last 50 frames for the static beginning and ending position. These positions were chosen because the controller was known to be static at these points, and averaging the locations would reduce noise from both tracking systems. After determining the start and end position, the Euclidean distance between the points was then determined. A ratio, R_{WV} was then calculated relating the two distances:

$$R_{WV} = \frac{D_{World}}{D_{Vive}} \tag{1}$$

After the ratio was established, the absolute difference of the HTC Vive’s travel distance and World travel distance could be calculated by multiplying the ratio by the Vive’s reported travel distance.

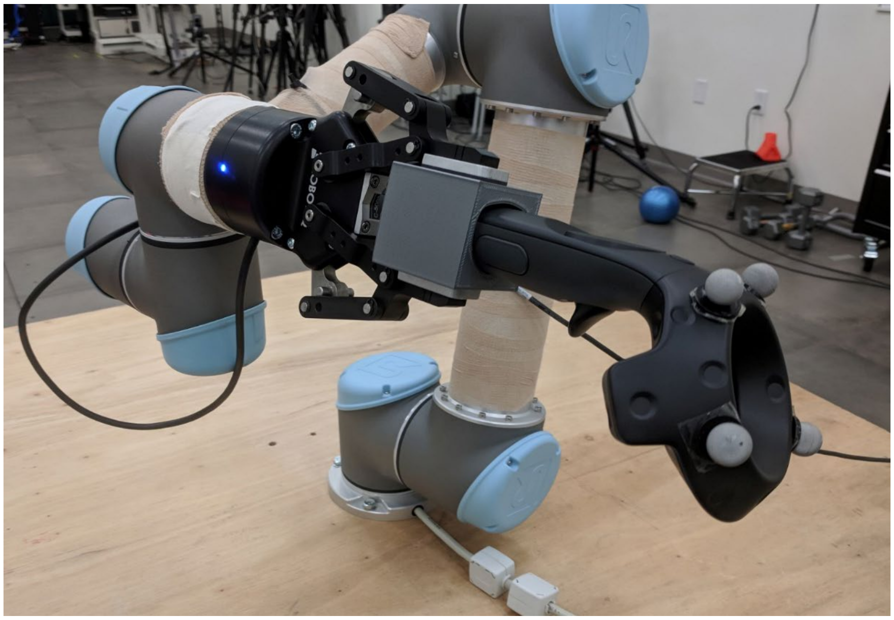


Figure 2.2. The HTC Vive controller in its collar used to give stability.

Orientation Difference of Vive and World

Rotation matrices were obtained from reflective markers placed on the controller and the Vive's internal report and then compared. Four markers were placed on each Vive controller to form a plane and calculate a local coordinate system with a rotation matrix R_W. The plane was defined from three of the points: α, β, and γ. First, two unit vectors ($\overrightarrow{n_{\alpha\beta}}$ and $\overrightarrow{n_{\alpha\gamma}}$) were defined from a central marker α to the two other points.

$$\overrightarrow{n_{\alpha\beta}} = \frac{x_\beta - x_\alpha}{\left\| x_\beta - x_\alpha \right\|} \qquad (2)$$

$$\overrightarrow{n_{\alpha\gamma}} = \frac{x_\gamma - x_\alpha}{\left\| x_\gamma - x_\alpha \right\|} \qquad (3)$$

The first plane vector is arbitrarily set as one of the above ($\overrightarrow{n_1}$) and crossed with the remaining vector to form $\overrightarrow{n_2}$. The two vectors are then crossed to form a final perpendicular plane vector ($\overrightarrow{n_3}$).

$$\overrightarrow{n_1} = \overrightarrow{n_{\alpha\gamma}} \tag{4}$$

$$\overrightarrow{n_2} = \overrightarrow{n_{\alpha\beta}} \times \overrightarrow{n_1} \tag{5}$$

$$\overrightarrow{n_3} = \overrightarrow{n_1} \times \overrightarrow{n_2} \tag{6}$$

Finally, two of the vectors are normalized to become unit vectors.

$$\overrightarrow{n_1} = \frac{\overrightarrow{n_1}}{\|\overrightarrow{n_1}\|} \tag{7}$$

$$\overrightarrow{n_3} = \frac{\overrightarrow{n_3}}{\|\overrightarrow{n_3}\|} \tag{8}$$

A rotation matrix can then be established.

$$R = \begin{bmatrix} \overrightarrow{n_{11}} & \overrightarrow{n_{12}} & \overrightarrow{n_{13}} \\ \overrightarrow{n_{21}} & \overrightarrow{n_{22}} & \overrightarrow{n_{23}} \\ \overrightarrow{n_{31}} & \overrightarrow{n_{32}} & \overrightarrow{n_{33}} \end{bmatrix} \tag{9}$$

The HTC Vive controller directly reports a rotation matrix. Two rotation matrices from the beginning of the trial to the end of the trial can then be defined as R_1 and R_2. The total rotation angle θ can then be found between the two rotation matrices[56]. The rotation matrix Q, which relates R_1 and R_2, was found using the following steps:

Let:

$$Q \triangleq R_1 R_2^T \tag{10}$$

Where ≜ represents a definition and T represents the transpose of the matrix. The trace (tr) of the matrix is then defined as:

$$tr(Q) = 1 + 2\cos(\theta) \tag{11}$$

Finally θ, the total angle rotated, can be extracted as:

$$\theta = \cos^{-1}\left(\frac{tr(Q) - 1}{2}\right) \tag{12}$$

The controller was rotated by the robot so that it would yaw, pitch, and roll about three orthogonal axes (X, Y, and Z). This was chosen to ensure that the Vive did not favor any type of rotation so that clinicians could utilize any type of rotation equally. A 3D-printed holder was printed so that the controller could be re-oriented 90° in each direction. This is what enabled the robot to yaw, pitch, and roll about a single axis. Due to the number of trials, preliminary data was taken comparing rotational accuracy at different speeds. Because there was no significant difference found, all rotations were taken at 1000 mm/s. The following list represents the combinations used in the translation trials:

1. Room Scale Calibration yaw about the X-axis at 1000 mm/s
2. Room Scale Calibration pitch about the X-axis at 1000 mm/s
3. Room Scale Calibration roll about the X-axis at 1000 mm/s
4. Room Scale Calibration yaw about the Y-axis at 1000 mm/s
5. Room Scale Calibration pitch about the Y-axis at 1000 mm/s
6. Room Scale Calibration roll about the Y-axis at 1000 mm/s
7. Room Scale Calibration yaw about the Z-axis at 1000 mm/s
8. Room Scale Calibration pitch about the Z-axis at 1000 mm/s
9. Room Scale Calibration roll about the Z-axis at 1000 mm/s
10. Standing Calibration yaw about the X-axis at 1000 mm/s
11. Standing Calibration pitch about the X-axis at 1000 mm/s
12. Standing Calibration roll about the X-axis at 1000 mm/s

13. Standing Calibration yaw about the Y-axis at 1000 mm/s
14. Standing Calibration pitch about the Y-axis at 1000 mm/s
15. Standing Calibration roll about the Y-axis at 1000 mm/s
16. Standing Calibration yaw about the Z-axis at 1000 mm/s
17. Standing Calibration pitch about the Z-axis at 1000 mm/s
18. Standing Calibration roll about the Z-axis at 1000 mm/s

Each trial type was collected 30 times, for a total of 540 trials. The absolute difference in angle rotated was then compared.

Testing the HTC Vive in Human Movement

In order to test the HTC Vive in human movement, several adjustments had to be made to the signal in order to make a final comparison. Lack of unified start/stop times, inconsistent sampling rate, and a rotated and translated coordinate system keep the signals from being quantitatively compared.

Spatial Alignment of the HTC Vive and Vicon World Coordinates

One roadblock to interpreting the Vive output is that it exists in its own coordinate system. Assuming that the virtual reality system has a consistent coordinate system (meaning the coordinate system is not defined by the position of the base stations), there exists a singular rotation matrix that could align the coordinates such that a movement in the Vive's X-direction would correspond to a movement in the world's X-direction.

As previously stated, prior research involving the HTC Vive attempted to find this rotation matrix using a least-squares approach. This approach is best used when only coordinates are available, and the orientation of this system is not known. However, both the Vive and Vicon orientations are known. A method for calculating the rotation matrix relating two coordinate frames is described by Craig[57] in the following equation which relates the Vive's coordinate system (V) to the world coordinate system (W).

$$
{}^{V}_{W}R = \begin{bmatrix} \widehat{X_W} \cdot \widehat{X_V} & \widehat{Y_W} \cdot \widehat{X_V} & \widehat{Z_W} \cdot \widehat{X_V} \\ \widehat{X_W} \cdot \widehat{Y_V} & \widehat{Y_W} \cdot \widehat{Y_V} & \widehat{Z_W} \cdot \widehat{Y_V} \\ \widehat{X_W} \cdot \widehat{Z_V} & \widehat{Y_W} \cdot \widehat{Z_V} & \widehat{Z_W} \cdot \widehat{Z_V} \end{bmatrix} \tag{13}
$$

On first pass, the initial inclination might be to apply this directly to the signals, either at a specific point where they best align or averaging each point. However, this will not yield a correct matrix because the signals typically do not align well. The signals do not start at the same time, and the Vive is less consistent in its sampling rate than the Vicon system. This means that it will occasionally lag and occasionally surpass the Vicon system. Even if the systems are perfectly aligned and interpolated to the same sampling frequency, the inconsistent recording from the Vive prevents the systems from temporally aligning. Figure 2.3 shows the conflicting coordinate frames.

Another major hurdle is that the rotation matrix used to represent the orientation of the controllers is different based on which markers are set as α, β, and γ. While the markers were generally consistent, some trials had consistently occluded markers which meant that the redundant marker had to be used. This was not problematic for calculating the amount rotated, but it will yield different orientations between trials.

One thing that can be noted about the Vive system is that it consistently corresponds with the world coordinates, but in a way which is inconsistent with the global frame established by the Vicon system. Every motion in the world's X-direction is recorded as a motion in the –Z-direction by the Vive, every motion in the world's Y-direction is recorded as a motion in the –X-direction by the Vive, and every motion in the world's Z-direction is recorded as a motion in the Y-direction by the Vive. This has been consistent across each trial. Due to this consistency in mismatched coordinates, a sample dataset could be created which corresponds to these rules instead of using real data.

To generate the sample data, a plane was created using arbitrary points. A rotation matrix was then developed for the plane using the method described earlier. The coordinates for the plane were then altered using the pattern relating the Vive and world coordinates (i.e. X became –Z, etc.). The rotation matrix for this new plane was then calculated. Finally, equation (13) was used to find the rotation matrix relating these two coordinate frames. Because these two planes corresponded to the same pattern as the Vive and world coordinates, this new rotation matrix could be used to rotate the Vive coordinates to correspond to the world coordinate frame. After rotating the Vive coordinate frame, both systems can then be made to start at an origin point of (0,0,0) (by subtracting the starting point from the rest of the signal), allowing the two systems to be aligned spatially.

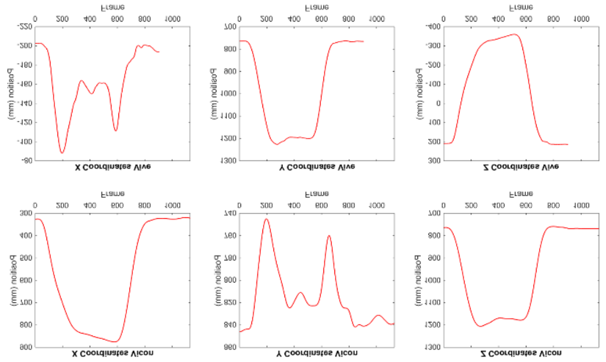


Figure 2.3. The conflicting coordinate systems of the Vicon system (top) and Vive (bottom).

Temporal Alignment of the HTC Vive and Vicon World Signals

Currently, the only way to align the start and stop times simultaneously pressing the start and stop buttons on Vicon Nexus (the program for world data collection) and the corresponding Python script (the method for HTC Vive data collection). While this could potentially call the validity of alignment into question, the subject was holding the controller steady at the beginning and end of the trial. Future work will explore alternative methods of alignment. Since the controllers were static in the beginning and ending of the controlled trials, the average location and orientation was found over this time. To achieve optimal start and stop times in the human movement trials, the signals were visually inspected and manually cropped.

Traditional techniques in comparing signals, such as root mean squared error, require context in order to inform about the similarity of two signals. Because of this, it is not as useful when discussing results with clinicians and laypeople. One alternative to this explored in this thesis is the absolute distance traveled from the starting point to the furthest point. The primary advantage of this is that it gives a number which is easy to interpret. This distance metric is similar to range of motion (ROM), a common clinical measure of amount moved.

Finally, direct human movement was able to be taken recording basic functional movements. A subject was asked to abduct their arm 30 times, then asked perform an arm extension 30 times. These movements were recorded with both the Vive and Vicon systems.

CHAPTER THREE

Results

Once the ratio between distance reported by the HTC Vive and Vicon system was established, the distance moved by the Vive was converted into real units. The average ratio between the systems was 999.7 with a standard deviation of 1.8, indicating that the Vive is reporting its units in meters as compared to the World which reports in millimeters. Using a conversion factor of meters to millimeters, the average difference between reported translational distances traveled was 0.74 mm with an interquartile range (IQR) of 0.66 mm for all room-scale calibration trials and 0.63 mm with an IQR of 0.29 mm for all standing calibration trials. The mean difference in angle rotated was 0.46° with an IQR of 0.81° for all room-scale calibration trials and 0.54° with an IQR of 0.62° for all standing calibration trials.

Translation Results

As previously stated, the average ratio between the systems for the translation trials was 999.7 with a standard deviation of 1.8. Figure 3.1 demonstrates the ratios for each trial type.

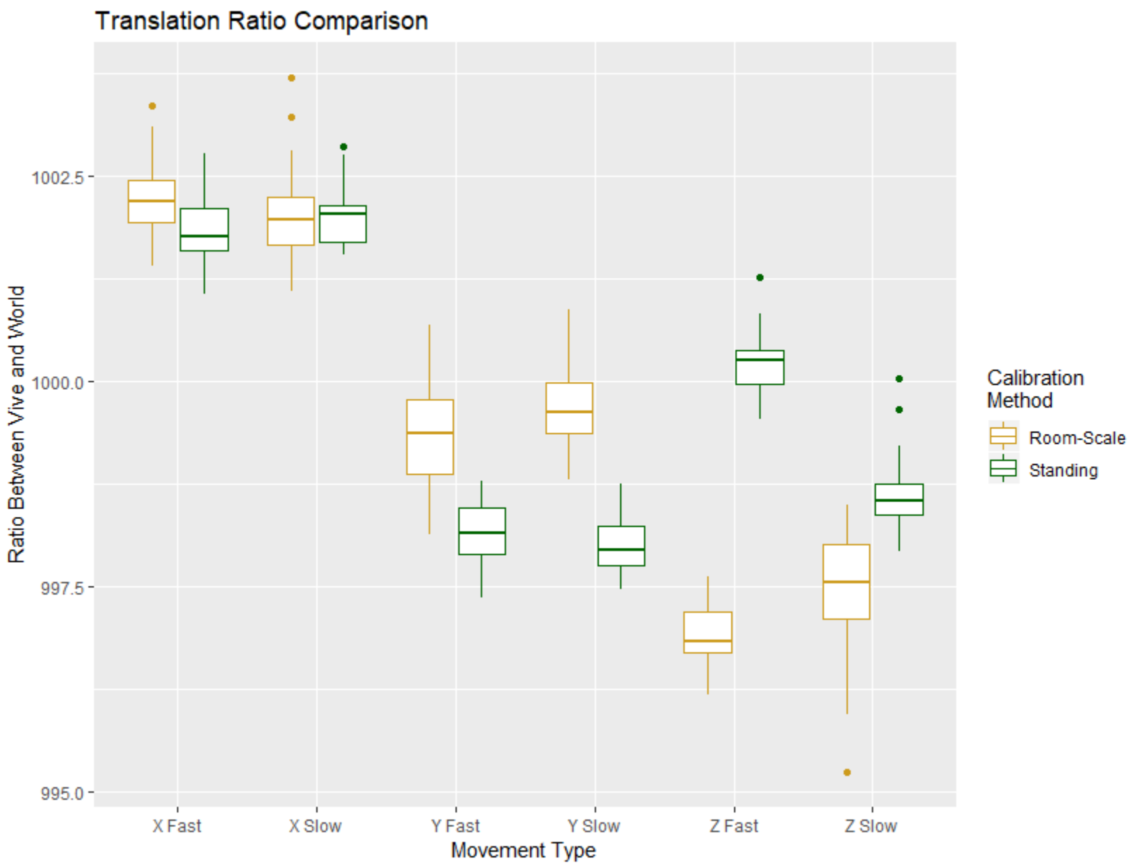


Figure 3.1 Comparison in distance moved across trial types.

It can be seen that each direction has a slightly different ratio, but they hover around 1000 (the conversion factor for meters to millimeters). After this ratio was determined, all of the Vive data can be converted to millimeters in order to match the Vicon system's unit of measurement. The Euclidean distance traveled can be converted using these units and then compared, as shown in Figure 3.2 and Figure 3.3.

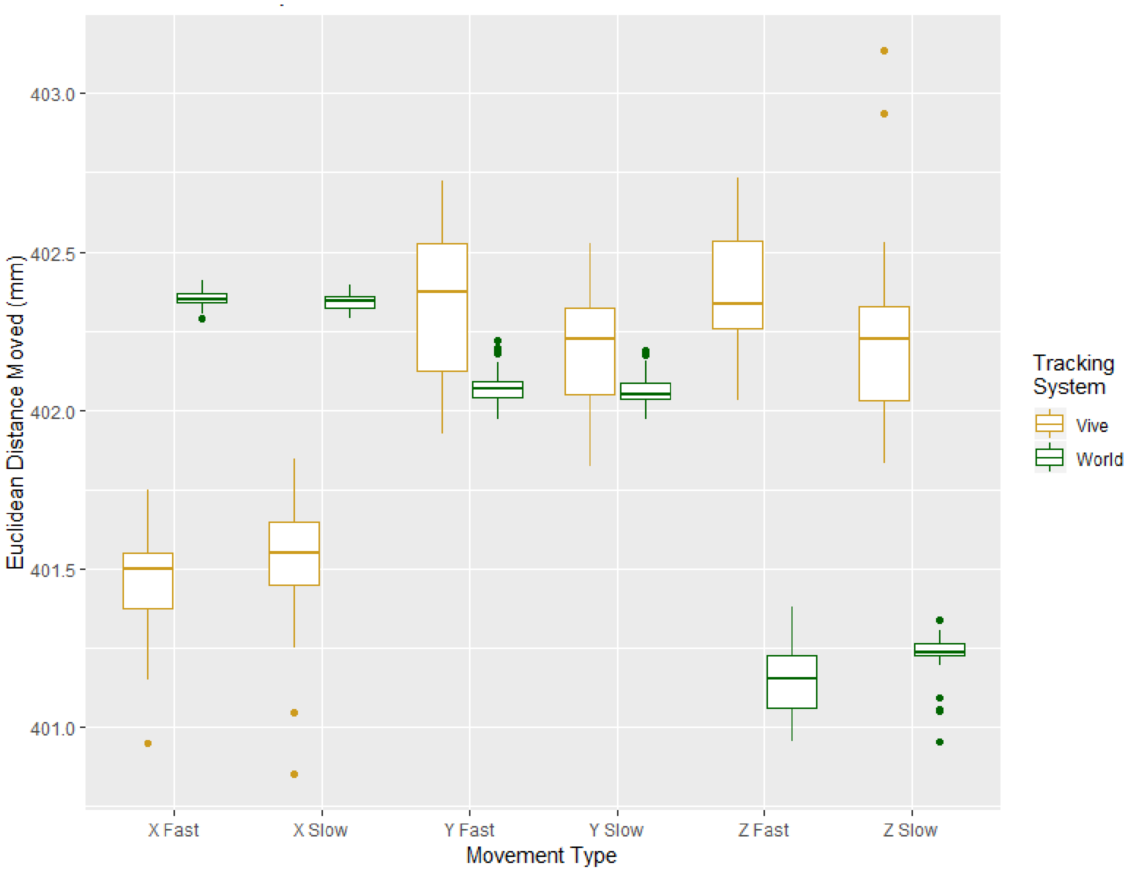


Figure 3.2 Amount translated from a room-scale calibration across conditions.

The primary takeaway from these figures is that, while the HTC Vive demonstrates a high level of accuracy, its precision is not as good when compared to the Vicon motion capture system.

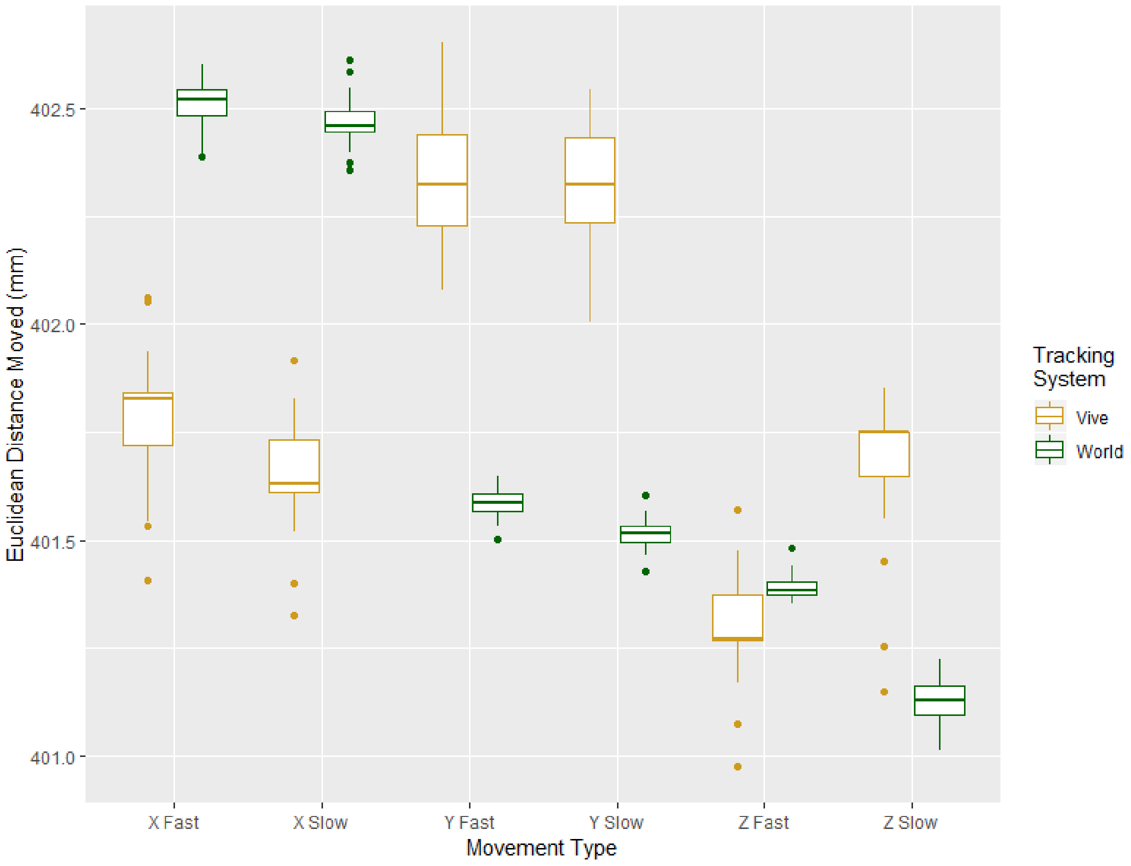


Figure 3.3 Amount translated from a standing calibration across conditions.

This can be seen in both the room-scale and standing calibrations.

Finally, Figure 3.4 shows the difference in Euclidean distance between the converted Vive coordinates and the reported world coordinates.

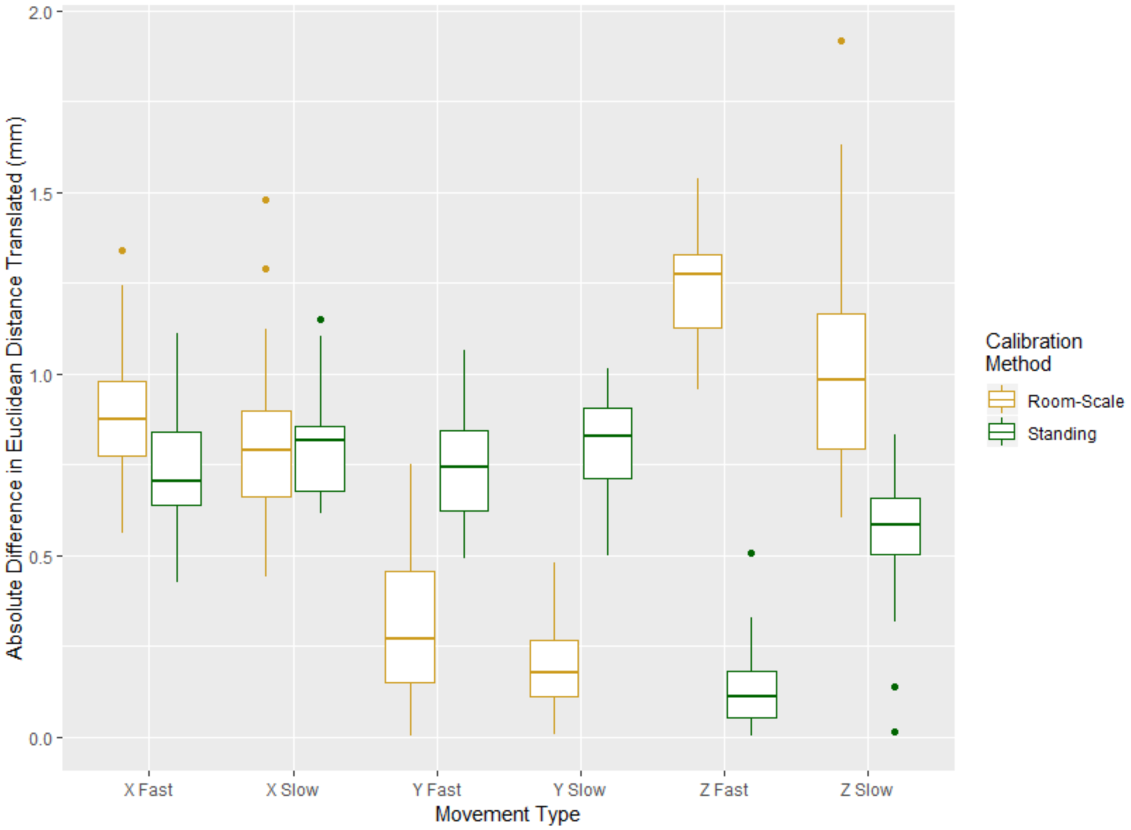


Figure 3.4 Difference in amount translated across conditions.

This figure demonstrates every condition and calibration method at once. It is clear that no condition or calibration method consistently yields superior results.

Orientation Results

Figures 3.5 and 3.6 demonstrate the difference in amount rotated between the Vive and Vicon systems.

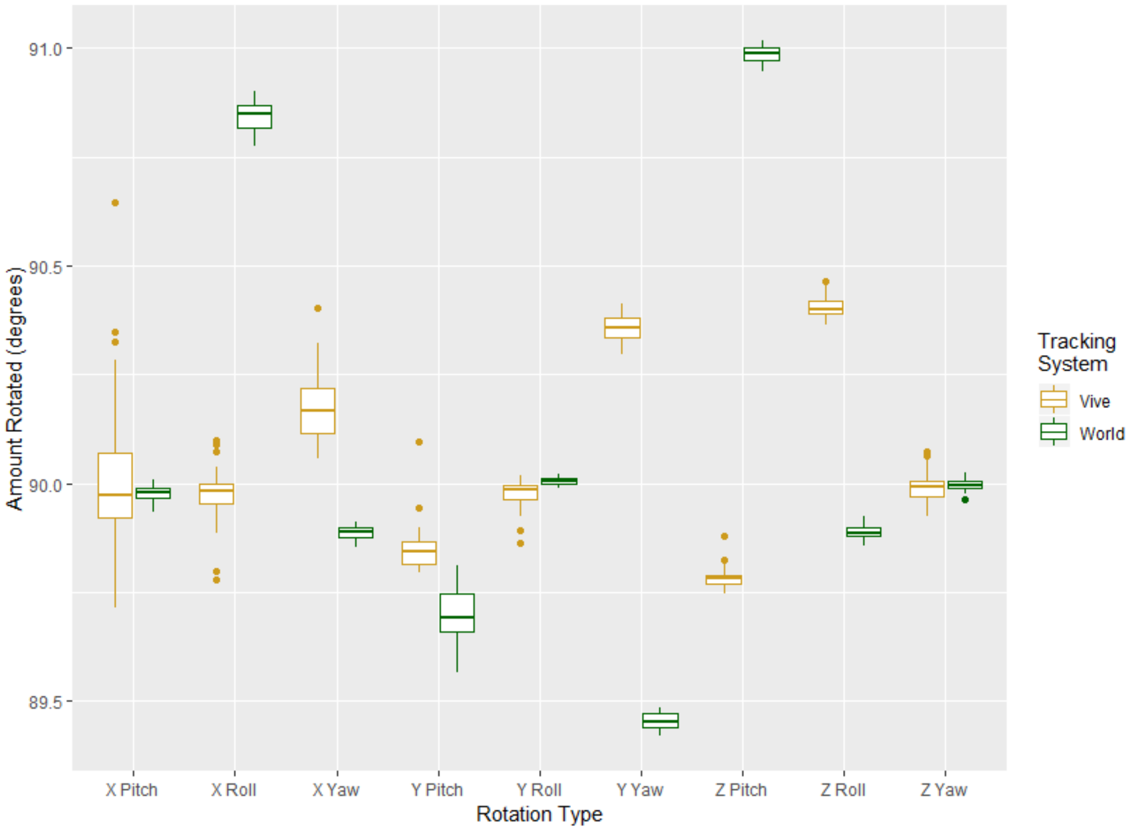


Figure 3.5 Amount rotated from a room-scale calibration across conditions.

One notable difference is that the standing calibration method yielded much more accurate and precise results in the Vive system than the Vicon system, as shown in Figure 3.5.

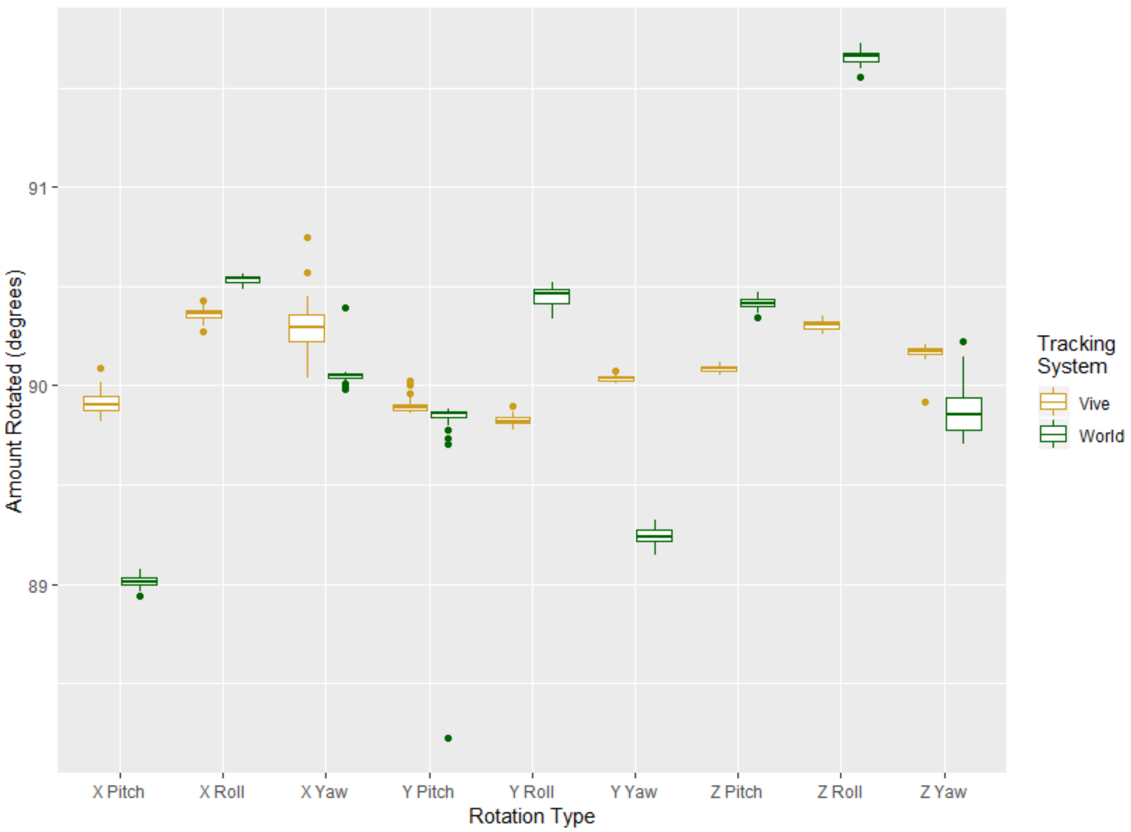


Figure 3.6 Amount rotated from a standing calibration across conditions.

This is likely due to the methodology required to calculate a rotation matrix representing the world orientation of the controller. To calculate the rotation matrix, 3 points are required, meaning any perturbation in any point will affect the rotation matrix. Furthermore, rotation matrices (and to a lesser extent, quaternions and other orientation representation methods) are subject to rounding errors which could have potentially impacted the results. Finally, the movement of the robot is ultimately a question of uncertainty in measurement. Both Vicon and the Universal Robots' robotic arm report accuracies of better than 0.5 mm, thus making it difficult to know which system to trust. Ultimately the Vicon output was chosen because the robot was not able to output data in real time, other laboratories in this field are more likely to have access to a motion capture system, and future studies performed on humans could then be referenced back using a similar method. Figure 3.7 shows the difference over all conditions between the Vive and Vicon systems.

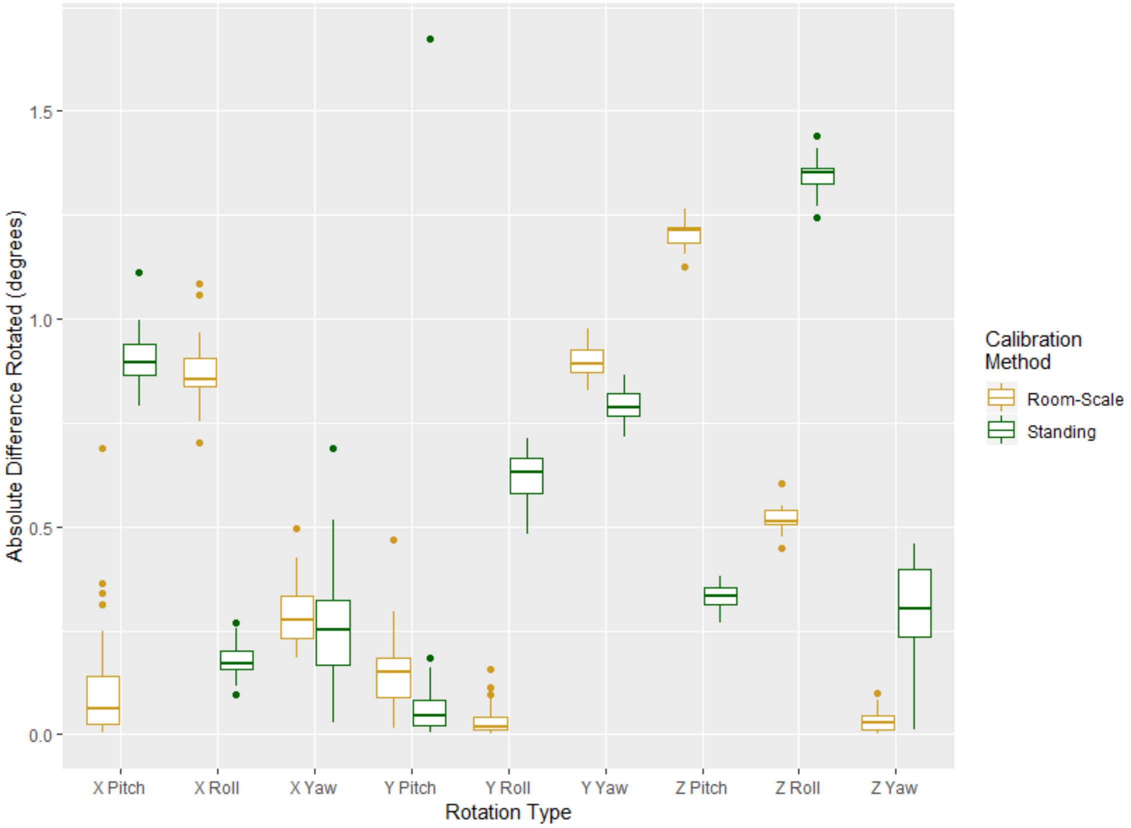


Figure 3.7 Difference in amount translated across conditions.

Similar to the translation data, there is no consistent condition which yields consistently superior results. The mean difference in angle rotated was 0.46° with an IQR of 0.81° for all room-scale calibration trials and 0.54° with an IQR of 0.62° for all standing calibration trials.

Results' Impact on Aims

Aims 2-5 are directly addressed by the results in this thesis. Aim 2, which tackles the translational accuracy of the HTC Vive, is addressed by Figure 3.4. It can be seen that the Vive does not track the controller better in any specific direction. While some directions have movement types that have significantly greater errors, no direction always had greater error. Aim 3, which addressed the rotational accuracy of the HTC Vive, had similar results to Aim 2 in that no condition yielded consistently better results. Aim 4 examined whether the Vive tracked better or worse in fast and slow conditions during translation. It can be seen in Figure 3.4 that moving the controller slowly did not yield better results than moving the controller quickly. Finally, Aim 5 addressed the calibration methods of the HTC Vive. It can be seen from Figures 3.4 and 3.7 that the calibration method did not have any consistent impact on the accuracy of the controller tracking in translation or rotation.

Human Movement Results

The matrix generated that theoretically related the two coordinate systems was:

$$\begin{matrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{matrix} \tag{14}$$

Figure 3.8 shows a representative trial where the controller is translated by a human.

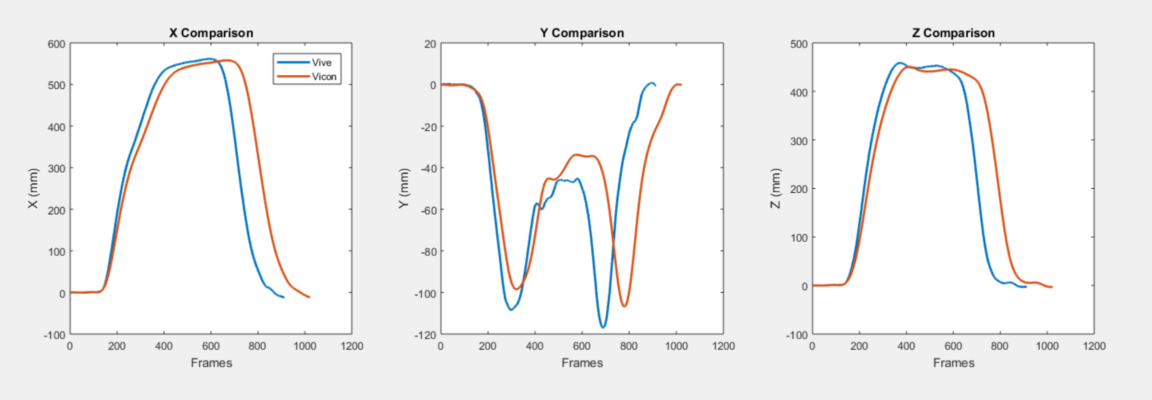


Figure 3.8 Rotated comparison of Vive and Vicon coordinates.

After rotating the Vive coordinates by the matrix in (14), the coordinates align giving Figure 3.8. It can be seen that the trials closely align in the beginning, but the Vive quickly lags behind in its framerate, but they are now in a unified coordinate frame. This alignment can be performed in any trial taken from multiple days and conditions. This demonstrates that the HTC Vive system is able to establish a consistent coordinate frame each time. Finally, human movement is shown in Figure 3.9. Each trial was cropped to start and stop at the same time. Then, the maximum Euclidean distance from the starting position was calculated for the movement reported by the Vive and Vicon system. The absolute difference between these values was then found. The average difference for the front raise exercise was 3.7 mm, and the average difference for the lateral raise exercise was 4.3 mm.

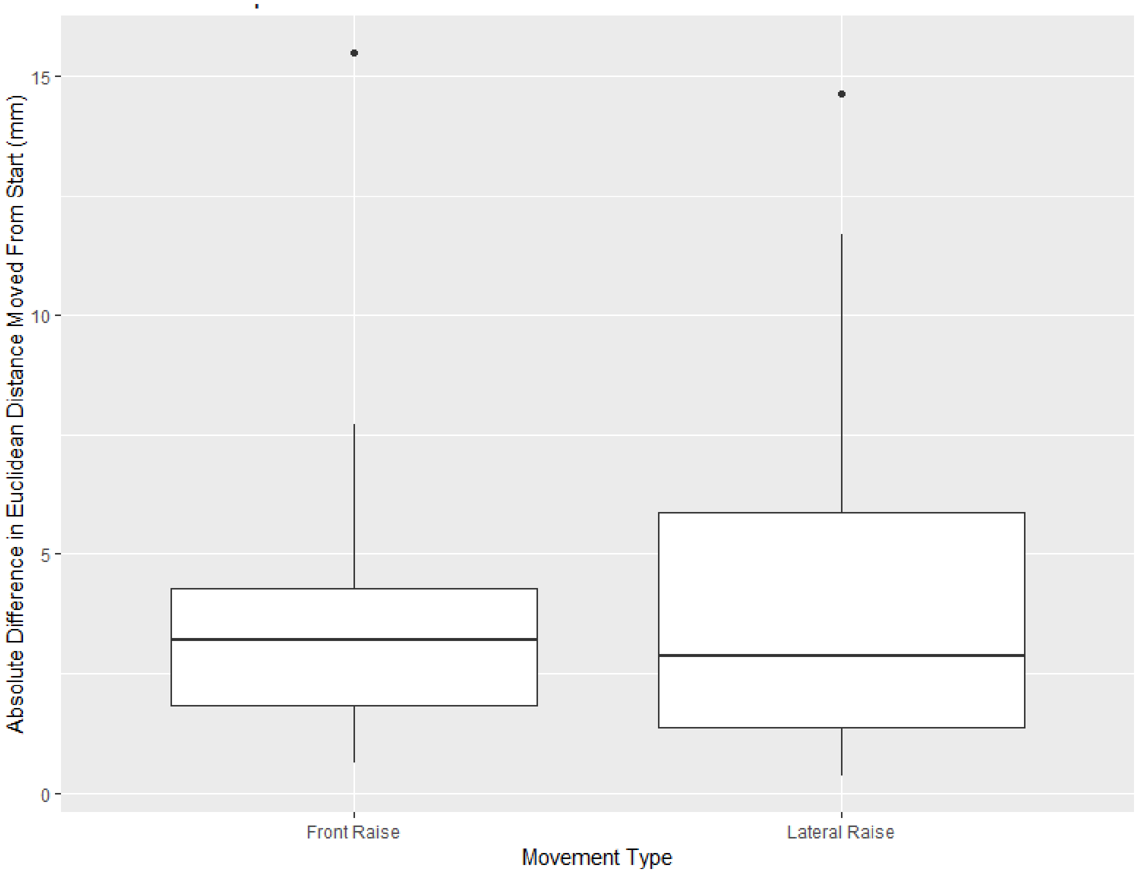


Figure 3.9 Human movement in two directions..

CHAPTER FOUR

Discussion

This thesis details a systematic methodology that can be used to evaluate the tracking accuracy and precision of 6 degree of freedom virtual reality based motion tracking systems such as the HTC Vive. Furthermore, it uses this methodology to investigate the translational and rotational accuracy and precision of the HTC Vive. This methodology was used to test whether different speeds, configurations, and directions had any consistent effects.

Aim 1: Establish a Framework for Testing Six Degree of Freedom (DOF) Virtual Reality Systems

A primary objective for this study was to establish a framework for testing six degree of freedom (DOF) virtual reality systems. The first step was to identify the different conditions which will affect the system. In this case, the calibration method was a unique condition because not all systems have multiple calibration methods. However, every system should be tested by translating the system in each direction (X, Y, Z) and rotating the system in 3 orthogonal directions around each axis (i.e. yaw, pitch, and roll around X, Y, Z). Furthermore, a slow and fast condition should be implemented to ensure that the system is able to track under moderately challenging conditions at least.

The second step is to use a method which is able to consistently move the system within defined parameters. In this case, the controller was moved 400 mm by an industry grade robot. However, further investigations could choose to move the system in a different way. The purpose is to move the system in a repeatable and consistent fashion. Movement of the system is important because it reflects intended use of the system. Simply keeping the system still will likely not yield optimum results because it is using the system in an unnatural way.

Finally, a great strength of this methodology is its use of accurate motion tracking. The method for tracking in this paper, optical motion capture, has been extensively used to validate prior systems efficacy[19], [20], [23], [25], [44]. Future investigations should aim to use optical tracking, but other methods can be used. However, this framework benefits most from tracking methods that are meant for motion capture. For example, a grid drawn on a floor would not be able to track the position precisely enough.

Aim 2: Translational Accuracy of the HTC Vive Controller

Translational accuracy is important to the HTC Vive because the primary use of the controller’s movement is to swing it around. It can be seen from figures 3.2 and 3.3 that the Vive produces different ratios and has different levels of error depending on direction and condition. However, no singular direction or condition yields better results. Figure 3.4 shows the total Euclidean distance moved as reported by the controller (after being converted using the average ratio) and the motion capture system. It can be clearly seen from this plot that the Vive, while close in average distance moved to the motion capture system, is less precise. When the difference between the Vive and Vicon is averaged, there is an IQR of 0.66 mm for the room-scale calibration trials with an average of 0.74 mm, and 0.29 mm for the standing calibration trials with an average of 0.63 mm. These ranges give an indication as to how the Vive will perform under serious use. When compared to translation tracking in similar systems such as the Xbox Kinect, this is a significant improvement[19].

Aim 3: Rotational Accuracy of the HTC Vive Controller

While not as initially useful as translation accuracy, validating the rotational accuracy of the HTC Vive controller has potential other uses. The primary potential use would be getting common biomechanical angles, such as those for hip, knee, and ankle. This would be relatively easy with the controllers, since they both output a rotation matrix. The rotation matrix relating the controllers' orientations could then be used to find the angle between them.

The total amount rotated by the controllers can be found in figures 3.5 and 3.6. It can be seen that, in the standing calibration method, there is a greater level of precision than the room-scale calibration in the Vive controller. This is notable because there was no remarkable increase in precision in the translation trials when comparing calibration methods. This could potentially be attributed to the closer base stations as well as the lack of movement. It is possible that, when the towers were moved closer together, the robot occluded them as it translated the controller. However, when the controller rotated the controller, it simply rotated its gripper instead of performing a more dynamic translation.

Aims 4 and 5: Effect of Velocity and Configuration on the HTC Vive Controller

Velocity and configuration (or calibration method) are specific conditions which needed to be tested in order to ensure that the HTC Vive could be used as a tool for clinic, research, and industry. For velocity, two speeds were chosen: 500 mm/s and 1000 mm/s. 1000 mm/s was chosen because it reflected the approximate speed of the controller during vigorous gameplay. 500 mm/s was used as a baseline, in order to see if the controller could track slow motions accurately as well as fast motions.

For the configuration, there are two internal calibration methods provided by the HTC Vive software: room-scale and standing. Room-scale is for users with a large environment, with the base stations having a distance of 16 ft 4 in[47]. Standing creates a virtual environment with a set space. Because the eventual goal of clinical research regarding the HTC Vive is to bring the HTC Vive into the home of patients, both configurations are important because it is unknown whether a subject will have an empty room with lots of space. However, utilizing a large capture volume in a clinical setting would enable more interactive and immersive games to be developed. Anecdotally, it is more pleasurable to play games in wide open spaces.

As reflected in figure 3.4, there was no consistent difference between configuration types and speeds when translating the controller. Similarly, rotating the controller did not cause any consistent difference between configuration types as shown in figure 3.7. As previously stated, ensuring the Vive operates in a similar manner will enable developers to have more freedom when creating games. For example, clinicians will be confident that games with lots of frantic action will still track the controller as well as games with more slow-paced movement. Likewise, the HTC Vive can be used in a larger variety of spaces because the base stations can always be moved closer together.

Results' Impact on Clinic, Research, and Industry

The results presented will likely hold the most significance in those interested in using the HTC Vive for rehabilitation in the clinic. The accuracy presented for translation is notably better than the accuracy for the Xbox Kinect Systems[19], [20], [23], [25],

[26]. However, the HTC Vive is not necessarily a direct replacement for the Xbox Kinect. Despite the superior tracking, the Kinect has pose tracking out of the box, whereas the Vive will need modifications to get specific variables such as knee and hip flexion. The Kinect also has existing FDA approved software such as VERA (San Diego, CA) [58] which is specifically designed for physical therapy, as well as a lower initial startup cost when compared to the Vive. Finally, the Kinect avoids some of the pitfalls commonly associated with virtual reality. Most notably, users of the Kinect are free to play and interact with the game without use of an HMD or controllers, which frees them up to perform exercises in a more natural way. Use of an HMD makes virtual reality a markedly isolating experience. The advantage of this is that the user is drawn into an immersive experience, however this experience can usually only support the user. Thus, a clinician who wishes to use the HTC Vive must accept that their ability to help and instruct the patient will be more limited than it would be with the Kinect.

While the Kinect does have a number of advantages, the HTC Vive is still useful for numerous situations. The capture volume provided by the HTC Vive allows for greater depth and complexity when designing physical rehabilitation techniques. For example, moving objects will actually pass by the patient, as opposed to just being represented on a screen. The reaction to this will be far more dynamic than the Kinect, as the patient can reach anywhere around their body. Because the controllers are significantly harder to occlude, tracking range of motion with the HTC Vive will be much more reliable and accurate and allow for a wider range of exercises to be performed. Anecdotally, the Vive is also able to influence the way a user interacts based on the surrounding environment it creates. For example, an open warehouse environment can help the subject feel as if they are in a large room, and they may be more likely to perform exercises in a way that reflects this. In contrast, it might be useful for a clinician to place a subject in a more cramped environment to see how they interact.

Because industry and other fields of research do not have established guidelines for what constitutes “good” tracking, it is not as clear if the HTC Vive fits their needs precisely. In light of the submillimeter accuracy shown in this thesis, it seems reasonable that using the HTC Vive for research and industry purposes can support games designed for those purposes. For example, someone designing a game meant for exposure therapy may want to track the reaction of the user. They can be confident in knowing that the location of the users’ interactions are accurate. In industry, it may be of use to track where the user’s hands are during simulation to see where a user’s hands might naturally go, or to warn a user that their hands are in a dangerous spot. Overall, the HTC Vive can help reassure developers that their research and industry simulations are true to life representations.

While the Vive has shown itself to be a more than adequate tool for clinic, research, and industry, it still should not be considered a replacement for a research-grade system. For the most part, this is due to its high level of variance. Its translation tracking may be extremely close to the Vicon system, but it would likely struggle to keep up during dynamic gameplay given its high level of variance. This is not to diminish the capacity of the HTC Vive as a low cost tool, but to caution against using it as a ground truth over other more traditional methods of tracking motion.

Human Movement Data

While the tracking abilities of the controller under controlled conditions was the primary focus of this thesis, the incorporation of initial human data is an important corollary. The human data collected in this thesis was still taken under largely controlled circumstances. The subject performed 2 basic functional exercises in a slow, controlled manner. Although this does not fully reflect the frenzied motion that can be achieved with most VR video games, it does incorporate the stochastic fluctuations involved with human movement. It can be seen in Figure 3.9 that both movements had a mean error of under 5 mm. While this difference is not as low as the controlled movement, this is to be expected. It is possible that some difference comes from inaccuracy in cropping the trials. Despite this, the Vive shows promise as a tool which can gauge functional movements at a high degree of accuracy.

Differences in Conditions

The Vive controller was tested under a variety of conditions as defined by the sections “Translation Difference between Vive and World” and “Orientation Difference between Vive and World” in Chapter 2. The primary purpose in doing so was to evaluate its accuracy and precision to determine if any direction, speed, or calibration type produced consistently superior results. For example, if the standing calibration was shown to be more accurate than the room-scale calibration, a suggestion could be made to clinicians to only calibrate the system in the standing mode to optimize performance. Likewise, if the system was less favorable to a specific direction, it could be recommended that clinicians not use exercises that have primary movement in that direction. However, it should be reiterated that the X, Y, and Z directions shown in the results do not correlate to the X, Y, and Z directions in the Vive's internal coordinate system.

Overall, there were no consistent differences in any condition tested. This indicates that the HTC Vive can be used in a variety of conditions and situations. However, the difference in translation and rotation is not consistent across every condition, with some conditions producing lower errors. For example, moving the controller quickly in x-direction was, on average, less accurate than moving the controller in the z-direction quickly. One explanation could be the orientation of the robot relative to the controllers and base stations. The robot was meant to be oriented in a way that prevented it from occluding the controller, but it is possible that it still reflected some light or otherwise impeded upon the controller's ability to track. This is supported in part by the data because each direction is somewhat grouped. For example, the X direction room-scale trials are greater than both Y direction room-scale trials, but lower than the Z direction room-scale conditions. One improvement that could be using a series of stepper motors to translate the controllers. While the robot arm was useful due to its flexibility, a system which can move the controller without blocking its sensors would likely do a better job.

Alignment of Vive to World Coordinate System

By visual inspection, this rotation was successful. Unfortunately, the matrix in (14) was unable to be experimentally validated because of the dissimilar start and stop times between the Vive and Vicon systems, the unequal frame rates, and the lack of consistently labeled markers. However, the theoretically generated matrix is able to do a more than adequate job in aligning the two systems.

The primary significance of this alignment is that the output from the system will make more sense to clinicians and users. One hurdle experienced while developing this project is that the movement displayed by the Vive controller did not seem to easily correspond to the movement in the real world. This was mostly overcome by using relative measures such as Euclidean distance. However, this concern has been eliminated by being able to consistently align the coordinate systems. This will also enable any software developed for the Vive to output results that make sense in real-world terms. For example, arm movement can be described in anatomical terms such as adduction/abduction and flexion/extension. Easily interpreted data will also help the Vive to be more easily accepted by the community at large.

Comparison to Current Research

As previously mentioned, two prior studies[45], [51] examined the tracking accuracy of the Vive. The conclusions of those studies are in conflict with the results of the present thesis in many ways. Niehorster et al.'s conclusion that the plane orientation changed on loss of tracking was not supported by the results of this thesis, which found a rotation matrix that could consistently bring the controller into a world coordinate frame. These differences arise primarily because of differences in experimental design. The previous studies used measurement tools which were not equipped for dynamic motion, used the HTC Vive in an unnatural way (i.e. without motion), and had a capture volume which far exceeded recommended levels. This thesis had several methodological advantages over those previous studies. The primary advantage was the motion tracking system used to track the location of the controllers. The researchers' system was calibrated to have an error of less than 1mm, providing a significant boost in precision over the previous studies. Another significant advantage was the use of a robot arm, which enabled the controllers to be in movement instead of in a static pose. This allowed for the Vive to be tracked in a situation which is closer to actual gameplay and use.

One note from Peer et al. which is not acknowledged by Niehorster et al. is that the reported height errors, at times up to around 15 centimeters, are largely unnoticed by users. An error this great seems like it would be noticeable and interfere with gameplay or normal use. However, anecdotal use of the system indicates that this does not occur during normal gameplay. One similarity between Niehorster et al. and Peer et al. is their method of importing the Vive's positional and rotational data. Both used a 3rd party resource called Vizard by Worldviz (Santa Barbara, CA) to import data. Vizard is essentially a Python IDE which enables use of proprietary Python libraries to develop environments for research involving virtual reality. The researchers attempted to use Vizard, but it was unsuccessful in picking up the location of the Vive. There are several less straightforward but easier to understand methods to record Vive positional data. Both involve using the developer's kit provided by Valve. Using this, it is possible to get both positional data as well as rotation data (in the form of a rotation matrix). It is not stated how Vizard determines the Vive's location, although it would be reasonable to assume that it uses a modified version of the developer's kit. Additionally, the outputted positional information from the developer's kit does not have a given unit, and one of the key findings from this study was determining the ratio between the Vive and world movement to determine its unit. It is not clear whether or not Vizard informs users of the unit it outputs or whether this unit has been converted at all. Additionally, it is uncertain whether or not the researchers from the Niehorster paper were able to access the orientation data. Browsing through Niehorster et al.'s GitHub repository, there does not seem to be a function which records orientation data. It is possible that Vizard does not provide tools for accessing orientation, although the researchers are unable to verify this claim. Since neither paper utilized orientation data, only 3D positional data, their methods were greatly complicated in determining orientation accuracy.

Finally, the methods utilized in this study were intentionally used to eliminate the need to unify the coordinate system of the Vive with the real world coordinate system. That is to say, a movement in the x-direction in the real world did not need to correspond to a movement in the x-direction in the Vive's space. Likewise, a rotation in the yaw direction in the real world did not need to correspond to a rotation in the Vive's space. The original reason for this was because the researchers were unable to successfully align the coordinate systems. By using relative motion (such as Euclidean distance), the researchers did not have to align the coordinate systems.

CHAPTER FIVE

Limitations and Future Works

Limitations

There are a number of limitations to this study which could be improved upon. While the experiments were performed under meticulous conditions with a variety of factors, there are still ways to potentially improve upon the results. A primary purpose of this thesis was to investigate the capabilities of the HTC Vive, but there are some technical limitations that inhibit the Vive from being researched to its full extent. With this in mind, several changes could be made in future studies to account for shortcomings in the HTC Vive's capabilities as well as changes in the methods used to collect and test the HTC Vive.

Unnatural Conditions

For a portion of the study, the primary limitation is that the HTC Vive research was performed under ideal rigid conditions. However, it is not abundantly clear how this impacts the results. The first criticism of the rigid motion is that it is not an accurate reflection of what gameplay would be like in clinic, research, and industry. There is inherent movement in human gameplay, even when holding the controller completely still. The unnatural position that the controller was placed in likely caused the results to be better than if a human were to move the controller. However, it has been previously stated that a completely static position is unnatural for the Vive controller due to its IMU experiencing inertial drift. This could have negatively impacted the results because the controller will jitter about when placed on the ground or in an otherwise static position.

Controller Jitters

When a controller “jitters”, it is visibly reflected in the HMD of the HTC Vive. The controller will begin to drift and often moves very rapidly from where it is meant to be. When the controller jitters, it actually is not sending a signal, which means that the data acquisition program records a 0 for every position and orientation value. While the static position of the controller was likely a contributing factor to the jittering of the controller, there are still unknown factors that influenced this jittering. Initial experiments had the controller held such that the trigger touched the bottom of the robot’s gripper. In contrast, the final experiment made it so that the tip of the controller touched the bottom of the robot’s gripper. This was because the controller would eventually begin to jitter so badly that data could not be taken. During the initial experiments, the controller would work fine for roughly 60 trials and begin to jitter. To reset the controller, a researcher had to remove the controller and shake it by hand. This would “reset” the controller, but only for about 10 more trials. Eventually the controller could not be reset and put back in the gripper without losing its tracking ability. However, researchers discovered that this could be greatly mitigated by putting the controller in the new position where the tip touched the bottom of the gripper. While some jittering was still experienced, the controller could be reliably reset for about 30-60 trials.

There were many efforts to mitigate the jittering experienced by the controller because data could not be reliably taken as described in the methods section of the thesis. While occlusions are very common in motion capture due to markers being hidden by other body parts, spline fitting or interpolating the data from the HTC Vive would have lessened the results from this study. Furthermore, the researchers had never experienced jittering like that seen with the robot during normal gameplay, thus it was hypothesized that there was an outside factor influencing the controller. The first factor considered was the motion capture system's IR interfering with the controller's IR receivers. However, the controller would still experience drift and rapid movements when the system was turned off. The second factor considered was the robot. If the controller was placed in the robot's gripper and the system was turned off, the controller would still jitter. The final factor considered was the orientation of the towers and controllers relative to the robot. The researchers hypothesized that the IR sensors on the controller were unable to properly receive light from the base stations. To account for this, the researchers first rearranged the base stations and the orientation of the robot so that the base stations had a clear line of sight to the controller. However, this did not have any effect.

Because the researchers knew the controller worked fine under normal conditions, they also tried to get the robot to reset the controller by attempting to simulate human movement. However, the rigid grip of the robot likely was the issue because no combination of swinging or shaking by the robot was able to reset the controller like a human. Finally, the researchers tried different orientations of the controller and noticed that having the tip of the controller touch the base worked best. While it was not perfect, this was able to alleviate the jittering of the controller enough that the researchers were able to take data.

Further Limitations

One notable further limitation that could have impacted results is the use of the motion capture laboratory. The cameras used in this study use IR light at a frequency of 850 nm. While this caused no noticeable change when playing games with the cameras on, it is still possible that this IR light caused interference with the controllers and headset. The initial setup for this study took place in a room with a much smaller capture volume than the motion capture lab. When calibrating the HTC Vive in this room, the controllers would drift off and were completely incapable of tracking properly. It is highly probable that the large capture volume of the motion capture laboratory enable the IR light to be diluted such that it did not noticeably interfere with the controllers. However, other Vive products, such as the trackers, are extremely susceptible to IR interference and are incapable of working with the cameras on. It has been previously suggested by Vive developers[8] that because the trackers use different, "light to digital [converters]" than the controllers or headset.

A major limitation is the sampling rate of the Vive acquisition code. Some tests were performed to determine the sampling rate, and it was found to be around 230 Hz. However, this framerate is likely variable. When examining a complicated movement's trajectory, it is clearly seen that peaks from the HTC Vive will move in and out of phase with the signal from the Vicon system. This discrepancy in sampling rate can largely be ignored when comparing rigid motion, because the controller is still at the beginning and ending of the movement. However, this makes it difficult to validate the controller during complicated motion because the signals cannot be properly aligned. While understanding how the controller tracks under ideal conditions is useful, it is equally useful to understand how the controller tracks under stress.

Future Works

The HTC Vive has shown great promise as a tool for tracking its controllers. Because of this potential, it can now be used in a number of different applications. One question that still remains is how effective the Vive will be under non-ideal conditions. Additionally, because virtual reality shows such potential for physical rehabilitation, a small video game could be developed to see if virtual reality is as effective as hypothesized. Finally, the Vive's tracker accessories show potential for developing a full body motion capture system.

Tracking under Non-Ideal Conditions

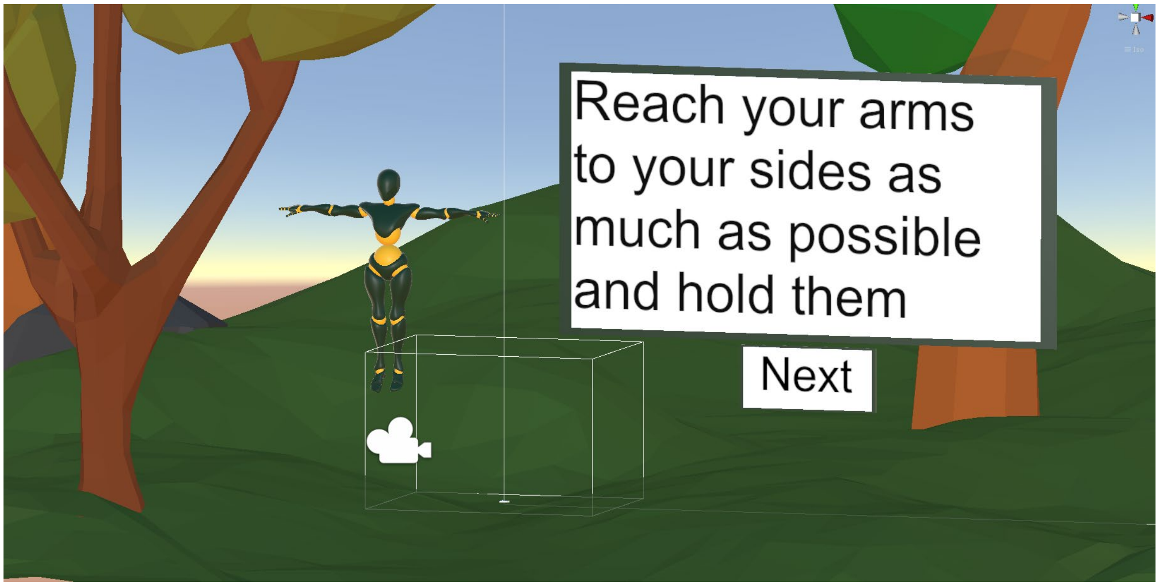
One condition not addressed in this study is a combination of both translation and rotation. It was hypothesized by the researchers that because the HTC Vive controller uses an IMU to report orientation and position (when combined with the base stations) it would track equally as well when doing a combination of translation and rotation. However, performing a combination of translation and rotation under a variety of circumstances would give further confirmation of the HTC Vive's tracking accuracy and precision.

As previously mentioned, exploring the movement of the Vive under non-ideal conditions will give greater insight into how useful the Vive is in tracking human movement. There are several different levels on which this can be achieved, although it is likely that the results will not be quite as promising as those under the robot. The first level is having a human perform clinical movements. These are rigid movements that are the first step up from the robot. They are highly controlled, and slow in nature. Some preliminary data for clinical movements has been taken in this thesis, but exploring more complex clinical movements will yield greater insight into how the HTC Vive will perform in a clinical setting. The second level is subject performance under normal gameplay. This is the least ideal situation but will give useful insight as to the reliability of the system under gameplay that can become intensely erratic. The main barrier to understanding the Vive's error when performed under non-controlled conditions is the temporal syncing of the start and stop times. Currently, a manual sync combined with a manual cropping of the signals is the only available method to sync the two systems.

Development of a Rehabilitative Video Game

Currently no existing physical therapy system utilizes the HTC Vive. As previously stated, home therapy is capable of enhancing home exercise programs through gamification. Virtual reality should be no different, and it could be hypothesized that it will be more effective due to its immersive environment. While full body motion tracking is not currently capable of being tested, it might be possible to utilize the controllers in a few different fashions.

Assessment of range of motion is crucial for stroke recovery patients[59], and is a key area in which the HTC Vive can utilized. Because the HTC Vive is able to generate objects which can be interacted with at any point around the subject, a game could be developed which selectively places targets relative to the patient's range of motion. A custom functional assessment could be created which establishes a baseline for a user in terms of the Vive's frame of reference. Over a period of weeks, the Vive could slowly move the targets further and further away, causing the user to have to reach out and increase their range of motion. One nice feature about this is that it has an obvious game connotation instead of being just a series of exercises. However, the Vive could also be utilized so that the subject is instructed as to how to properly perform exercises. It would be feasible to use the motion capture laboratory to import the actual motion of a physical therapist into the game world. Based on a static pose, the game could set regions of interest where they want the subject to move the controllers. These regions could even implement haptic feedback where the controllers vibrate when the subject has reached the correct pose. While the pose tracking is not as complex as the Xbox Kinect's, it could still be effective. A potential implementation of this game is currently being developed by the laboratory, as shown in Figure 5.1.



Reach your arms to your sides as much as possible and hold them

Figure 5.1 A sample of the video game.

Range of motion assessment is the most obvious and straightforward implementation of the HTC Vive's controllers because they naturally reside in the user's hands. However, creative placement of the controllers could yield different results. For example, the controllers could be strapped to the thigh and shank of a user. Because each controller generates a rotation matrix, the angle between them could be determined which would yield the knee angle. Just like the hands, range of motion assessment and exercise instruction could then be implemented to coach patients back to full health.

Improvements in Data Collection

Any further controlled tests should ideally be collected using Unity or Unreal Engine 4 instead of the Python scripts included in this thesis. The primary motivation for this is to move the lab forwards in terms of technology. The reason Python was used is because it is an easy to use programming language that had a neatly written wrapper for Valve's developer's kit. However, a suspected disadvantage in using Python is that the inefficiency of the programming language when compared to Unity or Unreal Engine prevented the sampling of the signal from being optimal. This was most clearly seen when attempting to manipulate the data while the script was collecting data. Anything not essential would cause the script to collect at an extremely low rate (around 10 Hz), so all calculations or data manipulations were offloaded to occur after collection finished. This was acceptable for the present thesis, but further collections will be greatly improved by using a more direct method for accessing position and rotation (i.e. Unity or Unreal Engine). Another improvement that should be made for any collection method is to implement a global time stamp for the beginning and end of the trial. This can be implemented simply in both C# and Python. Collecting the time collected is more convoluted for Vicon Nexus. The “Created” tab in

Joint Kinematics

As previously mentioned, there is an accessory to the primary HTC Vive system called Vive trackers. These trackers are small pucks which use the same technology as the controllers and HMD, albeit with a few minor differences. Additionally, these trackers have trouble in the current motion capture laboratory setup due to IR interference. However, there are several solutions which can potentially counteract these issues. The first solution is to use an active motion tracking system. The active motion tracking system will be able to use a variety of wavelengths, enabling wavelengths which do not inhibit the Vive trackers to be used. A second solution is to physically build the light emitting portion of a motion tracking system at the correct wavelength, or manually filter some of the light being cast by the cameras. The tracking will not be as reliable, but the IR being emitted from the cameras can be turned off so that they do not interfere with the trackers.

After a workaround has been found which enables research grade motion tracking of the Vive trackers, work can be done to develop a system which is able to perform full body motion tracking using the Vive trackers. Computationally, the trackers are easy to implement because they report a rotation matrix. This means that joint angles could be found relatively easily. The Vive trackers will likely be subjected to more jiggling from fat and resting muscle than the markers, but the extent of this has yet to be researched. The precision reached by the Vive controllers under controlled conditions is a good sign for the Vive trackers, and proper implementation could lead to one of the cheapest, most accurate, and easily accessible motion tracking systems. Additionally, these joint kinematics can be output at a smaller scale by attaching the controllers to different body parts. However, implementation of the trackers is preferred because they are smaller and their number is limited by the hardware of the collection system as opposed to a 2-controller limit imposed by the HTC Vive software.

Extension of Methodology

A strength of this thesis is that it lays out a clear methodology for testing 6 DOF systems. Further studies can utilize this methodology in order to yield insight into the accuracy and precision of other systems. For example, the Oculus Rift is a competing virtual reality system. It is possible that it can achieve greater precision than the HTC Vive, therefore it would make sense to test it according to the methodology described in this thesis. Furthermore, this thesis only examined the controllers of the HTC Vive. Extending this methodology to include the HMD would expand the types of rehabilitation measures that can be addressed by the Vive. While previous systems (most notably the Microsoft Kinect) were capable of 6 DOF tracking, it would be sensible to expect the market to continue to expand in the direction of virtual reality. There are many improvements that can be made in slimming down the HMD, incorporating more players locally into the virtual world, and reducing the controllers into more natural apparatuses. Furthermore, it can be expected that the tracking accuracy will improve over time as well. Despite cosmetic or technological changes, the methodology described by this thesis will be able to determine accuracy and precision because it is applicable to any 6 DOF system which can be held by the robot.

Significance

This study quantitatively analyzed the orientation and tracking accuracy of the HTC Vive. In terms of clinic, the HTC Vive has been shown to be better than existing systems. For research and industry, the high level of precision can add confidence to experiments and simulations. Finally, a methodology was established for validating 6 degree of freedom virtual reality systems. This will enable researchers to have a rigid methodology when testing future HTC Vive iterations or systems.

BIBLIOGRAPHY

[1] I. E. Sutherland, "A head-mounted three dimensional display," in *Proceedings of the December 9-11, 1968, fall joint computer conference, part I on - AFIPS '68 (Fall, part I)*, San Francisco, California, 1968, p. 757.

[2] L. T. D. Paolis, "Virtual and Augmented Reality Applications," p. 23.

[3] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," vol. E77-D, no. 12, p. 15.

[4] B. Moon, "HTC VIVE and VIVE PRO comparison," *Best Buy Blog*, 13-Sep-2018. .

[5] HACKADAY, *Alan Yates on the Impossible Task of Making Valve's VR Work.* .

[6] "Oculus Rift CV1 Teardown," *iFixit*, 30-Mar-2016. [Online]. Available: https://www.ifixit.com/Teardown/Oculus+Rift+CV1+Teardown/60612. [Accessed: 05-Dec-2018].

[7] "VIVE™ | VIVE Virtual Reality System." [Online]. Available: https://www.vive.com/us/product/vive-virtual-reality-system/. [Accessed: 05-Dec-2018].

[8] "Re: VIVE Trackers conflict with cameras of motion capture system," *Vive Community*, 23-May-2017. [Online]. Available: https://community.viveport.com/t5/Developer-Support/VIVE-Trackers-conflict-with-cameras-of-motion-capture-system/m-p/7844#M149. [Accessed: 12-Dec-2018].

[9] P. Rego, P. M. Moreira, and L. P. Reis, "Serious games for rehabilitation: A survey and a classification towards a taxonomy," in *5th Iberian Conference on Information Systems and Technologies*, 2010, pp. 1–6.

[10] T. Susi, M. Johannesson, and P. Backlund, "Serious Games – An Overview," p. 28.

[11] C. Höchsmann, S. P. Walz, J. Schäfer, J. Holopainen, H. Hanssen, and A. Schmidt-Trucksäss, "Mobile Exergaming for Health—Effects of a serious game application for smartphones on physical activity and exercise adherence in type 2 diabetes mellitus—study protocol for a randomized controlled trial," *Trials*, vol. 18, no. 1, Dec. 2017.

[12] Ather Nawaz, Nina Skjæret, Jorunn Lægdheim Helbostad, Beatrix Vereijken, Elisabeth Boulton, and Dag Svanaes, “Usability and acceptability of balance exergames in older adults: A scoping review,” *Health Informatics J.*, vol. 22, no. 4, pp. 911–931, Dec. 2016.

[13] J. Sween, S. F. Wallington, V. Sheppard, T. Taylor, A. A. Llanos, and L. L. Adams-Campbell, “The Role of Exergaming in Improving Physical Activity: A Review,” *J. Phys. Act. Health*, vol. 11, no. 4, pp. 864–870, May 2014.

[14] W. Zhao, D. D. Espy, M. A. Reinthal, and H. Feng, “A feasibility study of using a single Kinect sensor for rehabilitation exercises monitoring: A rule based approach,” in *2014 IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE)*, 2014, pp. 1–8.

[15] M. E. Nixon, A. M. Howard, and Y. Chen, “Quantitative evaluation of the Microsoft KinectTM for use in an upper extremity virtual rehabilitation environment,” in *2013 International Conference on Virtual Rehabilitation (ICVR)*, 2013, pp. 222–228.

[16] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, “Evaluation of Pose Tracking Accuracy in the First and Second Generations of Microsoft Kinect,” in *2015 International Conference on Healthcare Informatics*, 2015, pp. 380–389.

[17] K. Otte *et al.*, “Accuracy and Reliability of the Kinect Version 2 for Clinical Measurement of Motor Function,” *PLOS ONE*, vol. 11, no. 11, p. e0166532, Nov. 2016.

[18] R. A. Clark *et al.*, “Validity of the Microsoft Kinect for assessment of postural control,” *Gait Posture*, vol. 36, no. 3, pp. 372–377, Jul. 2012.

[19] J. Shotton *et al.*, “Real-Time Human Pose Recognition in Parts from Single Depth Images,” p. 8.

[20] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, “Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson’s disease,” *Gait Posture*, vol. 39, no. 4, pp. 1062–1068, Apr. 2014.

[21] B. Sun, Z. Zhang, X. Liu, B. Hu, and T. Zhu, “Self-esteem recognition based on gait pattern using Kinect,” *Gait Posture*, vol. 58, pp. 428–432, Oct. 2017.

[22] A. Mobini, S. Behzadipour, and M. Saadat Foumani, “Accuracy of Kinect’s skeleton tracking for upper body rehabilitation applications,” *Disabil. Rehabil. Assist. Technol.*, vol. 9, no. 4, pp. 344–352, Jul. 2014.

[23] R. P. Kuster, B. Heinlein, C. M. Bauer, and E. S. Graf, “Accuracy of KinectOne to quantify kinematics of the upper body,” *Gait Posture*, vol. 47, pp. 80–85, Jun. 2016.

[24] Y. Blumberg, “The real reason medical care costs so much more in the US,” 22-Mar-2018. [Online]. Available: https://www.cnbc.com/2018/03/22/the-real-reason-medical-care-costs-so-much-more-in-the-us.html. [Accessed: 05-Sep-2018].

[25] S. M. Coppola and S. M. Collins, “Is physical therapy more beneficial than unsupervised home exercise in treatment of post surgical knee disorders? A systematic review,” *The Knee*, vol. 16, no. 3, pp. 171–175, Jun. 2009.

[26] “Physical Therapy Treatment Effectiveness for Osteoarthritis of the Knee: A Randomized Comparison of Supervised Clinical Exercise and Manual Therapy Procedures Versus a Home Exercise Program,” *Phys. Ther.*, Dec. 2005.

[27] Y.-J. Chang, S.-F. Chen, and J.-D. Huang, “A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities,” *Res. Dev. Disabil.*, vol. 32, no. 6, pp. 2566–2570, Nov. 2011.

[28] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: defining ‘gamification,’” p. 7.

[29] R. N. McLay *et al.*, “A Randomized, Controlled Trial of Virtual Reality-Graded Exposure Therapy for Post-Traumatic Stress Disorder in Active Duty Service Members with Combat-Related Post-Traumatic Stress Disorder,” *Cyberpsychology Behav. Soc. Netw.*, vol. 14, no. 4, pp. 223–229, Apr. 2011.

[30] Y.-H. Choi, F. Vincelli, G. Riva, B. K. Wiederhold, J.-H. Lee, and K.-H. Park, “Effects of Group Experiential Cognitive Therapy for the Treatment of Panic Disorder with Agoraphobia,” *Cyberpsychol. Behav.*, vol. 8, no. 4, pp. 387–393, Aug. 2005.

[31] A. Garcia-Palacios, H. Hoffman, A. Carlin, T. . Furness, and C. Botella, “Virtual reality in the treatment of spider phobia: a controlled study,” *Behav. Res. Ther.*, vol. 40, no. 9, pp. 983–993, Sep. 2002.

[32] H. G. Hoffman, “Virtual-Reality Therapy,” *Sci. Am.*, vol. 291, no. 2, pp. 58–65, 2004.

[33] X. Pan and A. F. de C. Hamilton, “Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape,” *Br. J. Psychol.*, vol. 109, no. 3, pp. 395–417, Aug. 2018.

[34] Open Science Collaboration, “Estimating the reproducibility of psychological science,” *Science*, vol. 349, no. 6251, p. aac4716, Aug. 2015.

[35] M. Zanon, G. Novembre, N. Zangrando, L. Chittaro, and G. Silani, “Brain activity and prosocial behavior in a simulated life-threatening situation,” *NeuroImage*, vol. 98, pp. 134–146, Sep. 2014.

[36] O. FeldmanHall, D. Mobbs, D. Evans, L. Hiscox, L. Navrady, and T. Dalgleish, “What we say and what we do: The relationship between real and hypothetical moral choices,” *Cognition*, vol. 123, no. 3, pp. 434–441, Jun. 2012.

[37] C. Fink, “Osso VR Surgical Training Makes Push Into Med Schools,” *Forbes*. [Online]. Available: https://www.forbes.com/sites/charliefink/2018/06/26/osso-vr-surgical-training-makes-push-into-med-schools/. [Accessed: 11-Dec-2018].

[38] D. Ota, B. Loftin, T. Saito, R. Lea, and J. Keller, “Virtual reality in surgical education,” *Comput. Biol. Med.*, vol. 25, no. 2, pp. 127–137, Mar. 1995.

[39] *Lowe’s Next-Generation VR Experience, Holoroom How To, Provides On-Demand DIY Clinics For Home Improvement Learning*. .

[40] N. Gillil, “Why are VR headset sales declining?,” *Econsultancy*, 13-Aug-2018. .

[41] X. Xu, K. B. Chen, J.-H. Lin, and R. G. Radwin, “The accuracy of the Oculus Rift virtual reality head-mounted display during cervical spine mobility measurement,” *J. Biomech.*, vol. 48, no. 4, pp. 721–724, Feb. 2015.

[42] D. C. Niehorster, L. Li, and M. Lappe, “The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research,” *-Percept.*, vol. 8, no. 3, p. 204166951770820, Jun. 2017.

[43] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-squares fitting of two 3-d point sets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, May 1987.

[44] “Tips for setting up the base stations.” [Online]. Available: https://www.vive.com/us/support/vive/category_howto/tips-for-setting-up-the-base-stations.html. [Accessed: 13-Oct-2018].

[45] “Floor Tilt a Hardware or Software Issue?,” 16-Aug-2016. [Online]. Available: https://steamcommunity.com/app/358720/discussions/0/353916981477560813/?tscn=1490201536. [Accessed: 16-Oct-2018].

[46] “Fix for Slanted Floor Issue - IMU Recalibration,” 09-Jul-2014. [Online]. Available: https://www.reddit.com/r/Vive/comments/6tzthx/fix_for_slanted_floor_issue_imu_recalibration/. [Accessed: 16-Oct-2018].

[47] “Tilted/Slanted/Sloped Floor Cause Determined!,” *reddit.com*, 22-Mar-2017. [Online]. Available: https://www.reddit.com/r/Vive/comments/60vhwi/tiltedslantedsloped_floor_cause_determined/. [Accessed: 16-Oct-2018].

[48] A. Peer, P. Ullich, and K. Ponto, “Vive Tracking Alignment and Correction Made Easy,” in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018, pp. 653–654.

[49] “What are the details of the Woltring Filter? | VICON.” [Online]. Available: https://www.vicon.com/faqs/software/what-are-the-details-of-the-woltring-filter. [Accessed: 02-Apr-2019].

[50] C. Bruns, “pyopenvr,” 10-Jun-2018. [Online]. Available: https://github.com/cmbruns/pyopenvr. [Accessed: 14-Jun-2018].

[51] TriadSemi, “triad_openvr,” 25-May-2018. [Online]. Available: https://github.com/TriadSemi/triad_openvr. [Accessed: 14-Jun-2018].

[52] Universal Robots, “The future is collaborative.” [Online]. Available: https://www.universal-robots.com/media/1801258/eng_199901_ur_main-product_brochure_web_1.pdf. [Accessed: 06-Nov-2018].

[53] B. Belousov, “Distance between rotations,” 01-Dec-2016. [Online]. Available: http://www.boris-belousov.net/2016/12/01/quat-dist/. [Accessed: 17-Oct-2018].

[54] J. J. Craig, *Introduction To Robotics: Mechanics And Control, 3/E*. Pearson Education, 2009.

[55] P. S. Smith, J. A. Hembree, and M. E. Thompson, “Berg Balance Scale and Functional Reach: determining the best clinical tool for individuals post acute stroke,” *Clin. Rehabil.*, vol. 18, no. 7, pp. 811–818, Nov. 2004.

[56] “Shoulder Mobility Screen – Performing and Scoring the FMS – Smart Group Training.” [Online]. Available: http://smartgrouptraining.com/shoulder-mobility-screen-performing-and-scoring-the-fms/. [Accessed: 05-Apr-2019].

[57] T. J. Stevenson, “Detecting change in patients with stroke using the Berg Balance Scale,” *Aust. J. Physiother.*, vol. 47, no. 1, pp. 29–38, 2001.

[58] R. Komatireddy, “Quality and Quantity of Rehabilitation Exercises Delivered By A 3-D Motion Controlled Camera: A Pilot Study,” *Int. J. Phys. Med. Rehabil.*, vol. 02, no. 04, 2014.

[59] “The Stroke Rehabilitation Assessment of Movement (STREAM): A Comparison With Other Measures Used to Evaluate Effects of Stroke and Rehabilitation,” *Phys. Ther.*, Jul. 2003.