Using Virtual and Mixed Reality Technologies for Maintenance Training Applications

Purpose:
This paper describes how simulation and training system developers, acquisition agencies, and instructors can leverage mixed reality and virtual reality to support maintenance training.

For Support or Questions:
Email: sariotti@disti.com
Phone: +1-407-206-3390 ext 125
ABSTRACT

When Microsoft first introduced the HoloLens Developer Edition, the simulation and training industry scrambled to explore novel training applications. This new head wearable technology arrived at a time when the industry was still testing the bounds of virtual reality (VR) headset functionality available from products, such as, the Oculus Rift and HTC Vive. The HoloLens system attempts to transition mixed reality (MR) technology from the academic research community into practical real-world applications. Mixed reality, sometimes referred to as hybrid reality, merges real and virtual worlds to produce new environments and visualizations where physical and digital objects co-exist and interact in real time.

Simulation and training system developers, acquisition agencies, and instructors can leverage mixed and virtual reality to support maintenance training. Maintenance training systems, equipped with the new headsets, benefit from new functionality, but also suffer from new constraints. Maintenance training systems including MR or VR approaches employ different hardware to address key issues, including immersion, navigation, human interaction, simulator sickness, and physiology. Decision makers considering the use of MR or VR for maintenance training must consider the impact on training effectiveness and workflows, student throughput, and training facility space requirements.

This paper describes prototyping efforts conducted to better understand these issues and identifies current limitations and trade-offs when these technologies are used in maintenance training systems.

ABOUT THE AUTHORS

Mr. Christopher Van Duyne is the Chief Engineer at The DiSTI Corporation. Since 1993, Van Duyne has worked in the simulation and training industry starting at Environmental Tectonics Corporation where he was instrumental in the delivery of man-rated centrifuges, virtual fire fighting command and control trainers, and an integrated avionics maintenance trainer. He came to DiSTI in 1998 and oversaw much of the development of the GL Studio Toolkit, now an industry standard for object-oriented graphical application design. Later, as the lead engineer and system architect on numerous projects, Van Duyne shaped much of the design and implementation of the Virtual Maintenance Trainers (VMT) that have made DiSTI the leader in the industry. At present, Van Duyne is responsible for overseeing all the technical aspects of the projects and services, as well as, the overall technical direction of the company. Van Duyne holds a Bachelor of Science in Electrical Engineering from the University of Central Florida.

Mr. Scott Ariotti is a graduate from Embry-Riddle Aeronautical University with a Bachelor of Science in Aerospace Engineering. Ariotti began his career in the simulation and training industry in 1994 working as a 3D Database Engineer creating 3D visual content for simulators ranging from helicopter and aircraft flight trainers to urban and airport planning applications, and rail transportation to firearms training.
systems for companies such as Lockheed Martin, Real3D, Engineering and Computer Simulations (ECS), and ECC International. In 2000, Ariotti became a Systems Engineer at MultiGen-Paradigm providing support to customers interested in commercial tools and technology to build their own virtual worlds. Since 2005, he has worked for The DiSTI Corporation and currently serves as the company’s Director of Global Marketing. Ariotti has also served on various subcommittees for the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) from 2008 to 2015 and as the Chair of the Training Subcommittee in 2012.

**Dr. Lee Lacy** has over 30 years of experience supporting modeling, simulation, and training (MS&T) research and development efforts. Dr. Lacy attended the University of Central Florida (UCF) where he earned Bachelors and Masters degrees in Computer Science. After graduating from UCF, Dr. Lacy worked for MS&T companies on contracts for U.S. Department of Defense organizations. Dr. Lacy returned to UCF and earned a Ph.D. in Modeling and Simulation. He currently works at The DiSTI Corporation on U.S. and International Virtual Maintenance Trainer (VMT) contracts. Dr. Lacy has written and presented numerous technical papers for simulation conferences and published a book on OWL – the Web Ontology Language.
INTRODUCTION

Recent technological advances have enabled Virtual Reality (VR) devices to generate realistic virtual environments. Most VR approaches utilize headsets to block off the user’s real surroundings and immerse them in stereoscopic, 360-degree worlds, with the goal of making the user forget where they are and feel like they are someplace else. While VR technology provides amazing opportunities for interactive training, full VR devices may not be the best choice for supporting maintenance training systems. Today’s best VR headsets are tethered to high-powered gaming PCs with cables, while less powerful systems use game consoles or an inserted smartphone (Shanklin, 2016). Examples of current VR headsets include the Oculus Rift™, HTC Vive™, Sony PlayStation VR™, and Samsung Gear VR™.

Augmented Reality (AR) combines a live view of the user’s surroundings with virtual content. The user’s real-world environment is obtained either through a camera feed or, preferably, directly through clear lenses. The virtual content may include generated objects, characters, and animations that are layered on top of the real-world view (Shanklin, 2016).

Mixed Reality (MR) occurs when an AR system is capable of scanning and mapping the immediate surroundings and enabling virtual objects to seemingly interact with the real world. Figure 1 contrasts VR, AR, and MR.

MAINTENANCE TRAINING APPLICATIONS

Unlike entertainment applications, training systems are focused on meeting instructional objectives. Maintenance training application developers can leverage MR technologies in many ways including supporting equipment familiarity training and just-in-time training on actual equipment. One
maintenance training system use case that employs MR technology involves an empty room. In this “empty room” scenario, fully virtual content is placed in the space without physical objects, where the trainee is located. Because there are no real-world objects, this scenario provides a very similar comparison to a VR-enabled training system approach.

The next two figures illustrate the use of the VR and MR devices in maintenance training scenarios discussed in this paper. Figure 2 represents a trainee using a VR-enabled maintenance training system. The left side of the figure depicts an observer’s view of the trainee using the VR-enabled system and the right side shows the virtual world seen by the trainee.

![Observer View](image1.png)

**Figure 2. Observer’s view of a VR trainee**

Figure 3 represents a trainee using an MR-enabled system in the “empty room” scenario with a full sized virtual automobile. Two perspectives of the trainee wearing the MR device are shown on the left and the trainee’s MR views of the room are shown on the right. The primary difference between using the room-scale VR approach and the MR-enabled system in the empty room is that the real world is still fully visible to the trainee using the MR system (Rouse & Haughn, 2016).
HARDWARE CHARACTERISTICS AND LIMITATIONS

The HTC Vive room-scale virtual reality system and the Microsoft HoloLens mixed reality device can support discussions comparing VR and MR (see Figure 4).

The HTC Vive uses a full capability desktop PC. The PC enables the Vive to provide much higher graphical fidelity and resolution. The current Vive headset is tethered to the PC through a long cable limiting the user’s range of motion. Also, the Vive does not have any room sensors of its own, it relies on a cluster of external devices to help calculate its position in space. Training environment installers must position, wire, protect, and calibrate these components.
The HoloLens is a completely self-contained device with no external PC, cords, external sensors or wires. This makes the HoloLens ideal for freely moving around an arbitrary sized area. However, this freedom comes at a cost. Since the HoloLens runs all graphics locally on a small embedded computer, there are severe restrictions on the complexity of the display content.

The hardware supporting the HTC Vive room-scale VR system and the HoloLens MR device are significantly different. Table 1 identifies key system hardware characteristics for the Vive and the HoloLens (Rubino, 2016).

### Table 1 – Vive and HoloLens System Specifications

<table>
<thead>
<tr>
<th></th>
<th>HTC Vive Minimum System Specifications</th>
<th>Microsoft HoloLens System Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU</strong></td>
<td>Intel Core i5-4590 or AMD FX 8350, equivalent or better</td>
<td>Intel Atom x5-Z8100</td>
</tr>
<tr>
<td><strong>GPU</strong></td>
<td>NVIDIA GeForce GTX 1060 or AMD Radeon RX 480, equivalent or better</td>
<td>GPU/HPU HoloLens Graphics</td>
</tr>
<tr>
<td><strong>Resolution</strong></td>
<td>1080x1200 per eye</td>
<td>1268x720 per eye</td>
</tr>
<tr>
<td><strong>Frame Rate</strong></td>
<td>90 FPS</td>
<td>240 FPS, allocating 60 FPS each to a red, green, blue, and a second green layer</td>
</tr>
<tr>
<td><strong>Connection Type</strong></td>
<td>Tethered</td>
<td>Untethered</td>
</tr>
<tr>
<td><strong>Other Notes</strong></td>
<td>“Chaperone” forward facing camera. Approximately 110 degree FOV.</td>
<td>GPU draws at up to 60 FPS, and the HPU draws the colors sequentially while automatically correcting the image placement for the current head position</td>
</tr>
</tbody>
</table>

The HoloLens graphics hardware is much less powerful than the dedicated wired computer used for the HTC Vive. However, the HoloLens’ less powerful graphics and dedicated Holographic Processing Unit (HPU) enable it to be untethered and therefore, more portable. Although this tradeoff currently exists, the quality of both VR and MR displays is improving rapidly (Yang, 2016).

Another direction for VR is to simply replace the cord with wireless technology. There are some ongoing attempts to create a wireless VR experience. One option is to perform all processing on the headset, like the HoloLens. Qualcomm has demonstrated this technology (Segan, 2017).

The challenge for wireless VR is keeping any latency as low as possible. The latency between head movement and graphical output is critical for reducing simulator sickness issues. In either case, the VR market recognizes the limitations of a tethered unit and knows they need to remove the wire as soon as possible (McCarthy, 2016).

Although the Microsoft HoloLens is often used to represent MR technology, there are several companies working on competing MR solutions. Two MR companies are Meta and Magic Leap. As of this writing, Meta is accepting pre-orders for their device and Magic Leap is still reportedly working on a device of their own.
USER EXPERIENCE FEATURES

**Visual Interaction**

In both MR and VR worlds, trainees must interact with virtual elements. Training application developers should support comfortable and streamlined interactions to avoid distracting the student and slowing them down as they proceed through their tasks.

The Vive has two wireless motion tracked controllers from SteamVR with multiple buttons, track pads, and pressure sensitive grips. These are fully tracked in the designated space. This allows them to be used to interact with the virtual world in many ways that are suitable for multiple types of interactions. A common user interface pattern is bringing up 3D option “menus” that appear to be attached to one of the two controllers. The user selects an option by pointing with the other controller. This allows users to quickly make selections in the scene (see Figure 5) (Duffy, 2016).

HoloLens users move their heads to point and gesture with their fingers to select interactions with virtual elements (see Figure 6). These types of interactions are more limiting than manipulating two handed VR controllers. Interacting with just one selection (finger gesture) and one cursor (head controlled) can be very limiting and impede training. HoloLens also supports voice commands. Voice recognition can alleviate having to use gestures for modal commands (e.g., select, remove, install) but may have limited use in some spatial interactions.

**Immersion**

Both VR and MR help support a user’s sense of immersion. However, MR is not as immersive as VR because the user still sees objects present in their environment. MR does not attempt to transport the user to another world, it lets the user experience additional content in their own environment. MR virtual elements are typically partially transparent and most of the surrounding environment is simply the existing real world. Also, current MR hardware is not as capable as the hardware that supports current VR systems. For example, the field of view on current MR hardware is quite narrow and the MR GPU and CPU are much less capable than tethered VR systems.
Although immersion is an important feature for applications such as first person shooter games, it is not a critical feature for maintenance training using virtual environments (Van Duyne, Giordano, & Jackson, 2014). VR may actually be a distraction from the “strategic immersion” needed for virtual maintenance training and might impair the trainee’s ability to follow the real-world workflow used for tasks such as fault isolation.

**NAVIGATION**

Trainees must be able to move around in their environment. Moving throughout the VR or MR world is natural while the trainee stays within the “designated space.” The designated space is the area in which the VR or MR technology is capable of tracking the trainee’s movements and inputs. The difference occurs in what happens at the edges of the designated space.

For VR systems, the configuration of position sensors and the length of the tether explicitly define the designated space. Outside the designated space, the sensors are not guaranteed to work. For MR systems, the designated space can be arbitrarily large, including the complete room, if desired.

For a room-scale training application, the maintained item may be much larger than the designated space for the VR system. **Figure 7** shows the designated space requirements for various VR systems. The HoloLens does not have a predefined area requirement.

![Figure 7. VR System Designated Space Requirements](image)

With MR, the trainee can clearly see the real boundaries of the room (the walls) and real world objects (e.g., desks, other people). With VR, some sort of “in-game” element must indicate to the trainee that they are approaching the edge of the designated space. This can appear as a virtual fence or some other visual change as they approach the boundary. The HTC Vive uses a forward facing camera to show a version of the real world to the trainee. This “chaperone mode” will fade into the VR view when the trainee is about to bump into the tracking stations’ limits, showing blurry, perfectly aligned versions of walls and other real-life elements within your VR room (Machkovech, 2016).

If a user ignores a boundary warning, they can easily walk into real world obstacles such as, desks and walls. VR systems that warn users about obstacles remove some issues caused by having a completely enclosed headset. **Figure 8** shows the chaperone mode camera view on the HTC Vive.
Note that the view in Figure 8 is not a 3D view of the room, it is created from the system’s 2D camera and is therefore missing much of the normal depth cuing that a user gets from looking at the real world (like with MR).

When the trainee realizes that they need access to another area of the equipment outside the designated space, they need a way to access it. In VR, the most common way to achieve this is to teleport to the location. Teleporting involves selecting the new place, either by pointing and clicking, or by selecting the location from some user interface element. The eye point then materializes at the new position. Teleportation is preferred because shifting the eye point through space in room-scale VR could easily cause someone to lose their balance since it would appear as if the whole world just suddenly started shifting underneath them.

With MR, the world doesn’t move under software control, it is just the real world. The trainee can easily move the device using a drag mechanism, or a UI element to choose presets.

**INTERACTING WITH OTHER HUMANS**

Since VR completely occludes the trainee’s view of the real world, anything or anybody that is within the designated space must be either represented in the VR world or kept out of the designated space. Otherwise, the trainee in the virtual world could easily injure themselves or others.

Maintenance training systems may have four classifications for other humans besides the trainee. Each has a different interaction experience between MR and VR.

1. Instructor
2. Teammate
3. Other Trainee
4. Bystander

**1. INSTRUCTOR LEADING A CLASSROOM OF TRAINEES**

MR - The Instructor can wear an MR headset and observe a trainee interacting with the virtual content. The instructor can be in the designated space with the trainee (see Figure 9). The trainee will see them and naturally avoid physically colliding with them. The instructor can point with their real hands, and talk naturally to the trainee.
VR - Most VR systems do not allow multiple players in the same designated space, however, it is possible to have the instructor join the student remotely. The instructor would physically have their own designated space which would then allow the trainee and the instructor to work in the same virtual world. The trainee would see the instructor as some sort of simplified avatar (see Figure 10). Communication would have to be through electronic means since the designated spaces should be separated. The VR community is still trying to determine the best way to resolve cooperation between multiple sensor stations in a confined area (Quixotic7, 2015).

Desktop - In both cases, the Instructor can still observe the students’ virtual world remotely using simple desktop displays (see Figure 11).
2. **TEAMMATE - WORKING WITH THE TRAINEE ON THE SAME EQUIPMENT ON THE SAME TASK**

MR - The teammate can wear an MR headset to observe and interact with the same virtual content as the trainee similar to Figure 9. Both can communicate freely and can read and hold physical reference material, just like a real teammate.

VR - Just as the instructor in Figure 10, with the current technology, the teammate needs to have their own designated space and will then appear as simply an avatar to the trainee. However, there are some interesting experiments for ways to allow the systems to be combined (see Figure 12) (Brekelmans, 2016). This one uses a Microsoft Kinect™ to show more detailed avatars and also is allowing multiple VR participants in the same designated space.

3. **OTHER TRAINEE - WORKING INDEPENDENTLY ON A DIFFERENT PIECE OF EQUIPMENT**

MR - Other trainees, working on separate instances of equipment, can have overlapping designated spaces (see Figure 13). Both trainees will see different content in different places, but their respective areas can overlap because they can still see each other, therefore naturally avoiding contact.
Figure 13. MR Trainees Sharing a Designated Space
VR - Designated areas need to be kept clear of other people, the areas cannot overlap.

4. – Bystander

MR - Bystanders can walk through the designated areas and engage in conversation with the trainees without any real consequences.

VR - Bystanders need to be kept out of designated areas for the safety of themselves and the trainees.

TRAINING IMPLICATIONS

Using VR or MR for maintenance training affects workflow, trainee physiology, physical space requirements, and trainee throughput.

Workflow
Real world maintenance training is often performed with physical reference materials in hand (see Figure 14). MR enables trainees to use tablets, paper, and laptops directly in their normal way. VR, on the other hand, would require them to use virtual versions of the materials with significant unnatural navigation and interaction challenges (mouse, keyboard, touch screen).

Physiological
VR has known issues with simulator sickness (Vincenzi, Blickensderfer, Deaton, Beker, & Pray, 2009). One of the biggest challenges with VR is updating the scene fast enough to make it seem real when the trainee moves their head. This latency between head movement and the corresponding refresh of the screens within the headset is critical. Technology has been rapidly advancing to help reduce this issue.
With most MR implementations (like the HoloLens), the real world is fully visible at all times as if looking through a pair of sunglasses. In the real world view, the virtual elements are projected. Having this live view of the real world may significantly reduce the chances for simulator sickness. Because the view of the area is real, to the student, it feels like the virtual elements have been added to the real area. Technology in the MR headset is used to make sure virtual elements appear to be anchored in the real world. The HoloLens does a very nice job of locking holograms into the real world, making them appear to be part of the room versus something attached to the trainee’s head.

The ability to use each system with eyeglasses varies, but because of the open nature of the HoloLens, it can accommodate any type of glasses, where VR systems have various restrictions on what size eyeglasses can be used. The comfort experience with VR headsets is hit or miss and largely dependent on the size and shape of the user’s glasses frame (Roston, 2016).

**Physical Space**

MR and room-scale VR both require enough space for each trainee to move around comfortably and perform the training task. For some training, it may be acceptable to reduce the scale of the content to something that fits on a desk. However, for many VMT types, the real world size is a desirable feature. With room-scale VR, there cannot be any obstructions, walls, or other non-players in the designated area. Trainees’ designated areas cannot overlap with each other or they could easily injure themselves or others. With MR, overlapping designated areas is acceptable. It may be slightly distracting having an unrelated trainee moving through the scene, but it doesn’t risk injury or stop training.

**Throughput**

Both MR and VR technologies require some teaching and assistance to use them. The headsets need to be fitted, and some sort of calibration sequence is necessary for each trainee to have the optimum experience.

With the prototype system, once trainees were setup with an MR device, they were able to continue on their own with minimal assistance from the instructor. In addition, the trainee can use the MR device for long periods of time without the onset of simulator sickness thus allowing them to meet their training objectives faster.

Some trainees that used the prototype were new to room-scale VR. They required a dedicated assistant to help them avoid getting tripped up on the tether, to help them understand how to deal with the virtual boundaries of the designated area, and to help keep them safe as they begin to use the device and deal with the (invisible) limitations. In addition, higher risk of simulator sickness with VR could limit the time scheduled for each training session, reducing overall training class throughput.

Because of these differences, throughput in a training class might be much higher for an MR solution vs. a room-scale VR solution. However, this hypothesis could be significantly affected by new hardware solutions and should be researched.

**Summary**

Table 2 summarizes the capabilities of the VR and MR systems discussed in this paper.
Table 2. VR MR Summary Table

<table>
<thead>
<tr>
<th></th>
<th>VR (HTC Vive)</th>
<th>MR (MS HoloLens)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware</td>
<td>Desktop PC</td>
<td>Fully Self-Contained</td>
</tr>
<tr>
<td></td>
<td>Wired Head Mounted Display</td>
<td>Power-Optimized Performance</td>
</tr>
<tr>
<td></td>
<td>High Performance</td>
<td></td>
</tr>
<tr>
<td>Virtual Interaction</td>
<td>Two hand controllers</td>
<td>Single-hand gestures</td>
</tr>
<tr>
<td></td>
<td>Multiple buttons and track pads</td>
<td>Hands are free to interact with real world</td>
</tr>
<tr>
<td>Immersion</td>
<td>Very immersive visuals</td>
<td>Strategic immersion may be higher when</td>
</tr>
<tr>
<td></td>
<td></td>
<td>working with real world reference materials</td>
</tr>
<tr>
<td></td>
<td></td>
<td>due to the lack of distraction</td>
</tr>
<tr>
<td>Navigation</td>
<td>Movement constrained to predefined space</td>
<td>Nearly unlimited space supported</td>
</tr>
<tr>
<td>Human Interaction</td>
<td>More difficult because of the lack of</td>
<td>Natural</td>
</tr>
<tr>
<td></td>
<td>awareness of the real world</td>
<td></td>
</tr>
<tr>
<td>Workflow</td>
<td>All reference materials and equipment must</td>
<td>Can use real reference material and</td>
</tr>
<tr>
<td></td>
<td>be virtual</td>
<td>equipment</td>
</tr>
<tr>
<td>Physiological</td>
<td>Potential for simulator sickness may</td>
<td>Can probably use the system much longer</td>
</tr>
<tr>
<td></td>
<td>limit useful training time</td>
<td>due to the solid connection to the real</td>
</tr>
<tr>
<td></td>
<td></td>
<td>world</td>
</tr>
<tr>
<td>Physical Space</td>
<td>Need to keep designated areas clear of</td>
<td>Can use in areas overlapping with other</td>
</tr>
<tr>
<td></td>
<td>people, and place and maintain position</td>
<td>trainees and with arbitrary obstacles</td>
</tr>
<tr>
<td></td>
<td>markers</td>
<td></td>
</tr>
<tr>
<td>Throughput</td>
<td>Especially when inexperienced, a dedicated</td>
<td>Once the basics have been explained, the</td>
</tr>
<tr>
<td></td>
<td>helper is needed to keep the trainee safe</td>
<td>trainee can operate on their own without</td>
</tr>
<tr>
<td></td>
<td></td>
<td>concern for their safety</td>
</tr>
<tr>
<td>Hardware Limitations</td>
<td>Performance tracks PC</td>
<td>Limited field of view can be distracting</td>
</tr>
<tr>
<td></td>
<td>performance as graphics cards can be</td>
<td>and the possible quality of content is</td>
</tr>
<tr>
<td></td>
<td>upgraded</td>
<td>limited by the hardware compute power</td>
</tr>
</tbody>
</table>

CONCLUSIONS

With the current state of technology, maintenance training in the near future may benefit most from MR technology versus VR technology, provided there are improvements made to the methods of interacting with the virtual world in MR. While the fidelity of the visuals is much more compelling on the VR system, the practical use for everyday training is severely hindered by the limited designated area, the tethered umbilical, the necessity to keep obstacles out of the way, and the higher risk of simulator sickness.

As hardware progresses, fidelity and performance will continue to improve for both MR and VR systems, eventually bringing them to an equal footing. At this point, the primary decision for training developers will be whether to bring the virtual content to the classroom (MR), or bring the trainee to the virtual content (VR). Each has their own positive and negative tradeoffs to consider and should fall to Instructional Systems Designers to consider these during their training task analysis.
REFERENCES


