Navigating Augmented and Virtual Reality Hardware Choices

Gardner Congdon, John Williamson
SAIC
Seattle, WA
gardner.congdon@saic.com, john.e.williamson@saic.com

ABSTRACT

The interest in and use of augmented reality (AR) and virtual reality (VR) has grown significantly in the last several years as has the availability of commercial off the shelf hardware for AR and VR. Matching the hardware platform with the intent of an application is key to its adoption and success. We conducted a review of the most popular AR and VR hardware platforms and analyzed their strengths and weaknesses. Furthermore we developed several serious games applications for both AR and VR and analyzed the results we obtained.

ABOUT THE AUTHORS

Gardner Congdon is a Solutions Architect for SAIC and runs their Big Timber Games studio in Seattle, WA. Mr. Congdon has been developing serious games for the last 11 years and has shipped titles on PC, iOS, Android and console. He is particularly interested in applying game technology and methodology to education and healthcare markets.

John Williamson is a Senior Communications Analyst for SAIC. John has worked in games, both serious and commercial entertainment, for over 20 years shipping games in nearly every genre on nearly every platform. He lectures internationally on keeping students, especially young women, interested in STEM, by making their own video games.

Navigating Augmented and Virtual Reality Hardware Choices

Gardner Congdon, John Williamson
SAIC
Seattle, WA
gardner.congdon@saic.com, john.e.williamson@saic.com

A COMPLEX LANDSCAPE

With billions of dollars of investment and multiple commercial products, Augmented Reality (AR) and Virtual Reality (VR) are undergoing a second renaissance far bigger in scope than the last mainstream commercial push in the mid-1990s.

For the purposes of this paper, we will define Augmented Reality as the addition of Computer Generated Imagery (CGI) to a user's view of the real world environment. AR is more than just a simple Heads Up Display (HUD), as AR typically tracks some or all of the real-world environment in three dimensions in real time, giving the impression that the CGI is part of the real-world. Simple AR demos populate the app stores for all modern smart phones, but more immersive experiences are available through the use of a Head Mounted Display (HMD).

Using AR, a CGI cartoon character could appear to be in your living room. With certain AR hardware, all of the details in your living room would be scanned, allowing the CGI cartoon character to interact with the furniture; jumping from the floor to the couch, to the coffee table. You would see both your living room (and all the objects therein) and the cartoon character interacting together in three dimensions.

In contrast, Virtual Reality (VR) replaces your entire field of view with CGI. In your living room, in VR, you would no longer see your couch and coffee table. No matter where you looked, you would only see a virtual environment; the interior of a medieval castle or the stone monument of Stonehenge or the plains of Mars for example.

In the 1990s long forgotten companies such as Virtual IO, VictorMaxx and VPL tried to move VR from the pages of science fiction into American living rooms. While impressive technology for their time, the low resolution (320x200), flat shaded images rendered at less than 30 frames per second and high costs ($800 not adjusted for inflation) were not enough to satisfy consumers, and other than research institutions and a few high end markets, the products rescinded.

But in 2012, VR got a new lease on life via the intersection of high resolution smart phone screens, home computer power and crowdsourcing. Crowdsourcing meant that VR did not need approval from large corporations or venture capitalists, but money could be raised directly from consumers. The success of Oculus was quickly followed by Vive from HTC/Valve and HoloLens from Microsoft and a host of others. The net result is that within a relatively short timeframe many options have become available for experiencing AR and VR.

AR and VR options vary in price, processing power, features, and resolution. The variety of options and their features, strengths, and weaknesses create a challenge for businesses or individuals interested in using these tools; how do you choose the right platform for the problem you are trying to solve? In this paper we will share our observations gleaned from working with a variety of AR and VR platforms. We do not suggest that we can identify every relevant point, but will share what we have learned with the hope of spurring others to imagine new questions and use cases. We will not cover all of the AR/VR options available, but we will hit upon some of key platforms.

KEY PLATFORMS

Over the last year, SAIC's Big Timber Games has created a series of public facing AR and VR demos in a wide range of hardware and software solutions. The hardware used includes the Google Cardboard, Samsung Gear VR, Oculus Rift, HTC Vive and Microsoft HoloLens.

Table 1 summarizes the platforms we evaluated and developed upon:

Table 1. AR and VR Platforms

Platform	Cost	Primary Function	Add'l Equipment	Total Cost
Google Cardboard	$ 10	VR*	Smart phone	$ 310
Samsung Gear VR	$ 50	VR	Android phone	$ 350
Oculus Rift	$ 500	VR	PC	$ 2,000
HTC Vive™	$ 800	VR	PC	$ 2,300
Microsoft HoloLens	$ 3,000	AR	NA	$ 3,000

As inexpensive as less than $10 per viewer, the Google Cardboard (and similar products) is the most affordable option. These viewers are constructed from cardboard or other inexpensive materials, use plastic lenses, and require a smartphone running an application to generate the visual experience. Users insert the smartphone into the Cardboard and hold the device to their face to view the content. Typical applications provide the user a VR experience, but AR is possible using the smartphone's forward facing camera. Many Cardboard-like devices utilize a single button interface. The Cardboard uses the processing power of the smartphone and its built in accelerometer, magnetometer, and gyroscopic sensors to detect where the user is looking and draw the appropriate perspective based on this. The user perspective can be tracked in pitch, roll and yaw, but the user's actual position is not tracked. This means that locomotion though a virtual environment requires a user interface of some sort (as opposed to the user simply walking around). Google has reported that over a million Cardboard units have been sold and over 1,000 applications have been created (Wikipedia, 2017).

Similar to the Cardboard-like devises, enhanced smartphone viewer devices exist such as the Samsung Gear VR at a price point around $50 and up. These devices may have additional sensors integrated into the viewer for more accurate and responsive rotational tracking and often have a built in control pad or ship with a peripheral controller supporting more user interface options. In some cases, these devices work with only one brand of smartphone. Again, the user perspective can be tracked in pitch, roll, and yaw, but the user position is not tracked. Versions currently exist for Google (Daydream VR) and Samsung (Gear VR). We evaluated the Samsung Gear VR, which also can be used for AR, but not the Google Daydream.

Several platforms consist of Head Mounted Displays (HMD) coupled to personal computers (PC), such as the Oculus Rift and HTC Vive. These platforms cost in the neighborhood of $600 to $800 for the HMD and $1,500 and up for the PC. They require a relatively high-end PC to generate the graphics, which results in a higher fidelity VR experience. In addition to tracking user perspective in pitch, roll, and yaw, these platforms also track the user's location in a limited space utilizing infrared lasers and sensors on the HMD and controllers, an approach described as "outside-in" tracking. This tracking allows the user to walk around the interior of the capture area. As they do, their position and perspective in the virtual environment matches 1:1. Networking multiple units and utilizing careful design approaches can create experiences in which the user feels as though they are walking for much longer distances than they actually are. Tracking sensors on handheld controllers allow the user to interact with the virtual environment, such as to reach down to grab an object with their virtual hand or to interact with menus. Currently, these experiences require the use of a cable to connect the HMD to the PC, limiting the user's range of motion unless the PC may be run from batteries and carried by the user. These systems are designed to deliver predominantly VR experiences.

Augmented-reality specific systems like the Microsoft HoloLens consist of an HMD which is driven by an onboard computer, negating the need for cables. The HoloLens employs an "inside-out" tracking system, wherein the device uses infrared lasers and/or other means to map the local environment and support interaction of virtual and real objects, characters, etc. This device utilizes an HMD which is almost completely transparent, allowing the user to see their real physical environment unobstructed and see CGI objects making it specifically an AR platform. Being untethered and affording a view of the real world addresses many concerns about safety and privacy for users. The HoloLens is not without limitations however. The inside out approach works best in relatively small indoor spaces and does not perform as well outdoors or in large open venues. Additionally, the device is able to draw virtual objects in a relatively narrow field of view (FOV). At a cost of $3,000 this is the most expensive of the platforms we evaluated and, importantly, is still considered a development kit. No consumer version is available at this time. Although, anyone with the budget can purchase one.

SUMMARY OF APPLICATIONS USED TO DERIVE FINDINGS

We developed several applications using these technologies from which we developed our observations of the AR and VR platforms. While we will not describe the applications in depth, Table 2 summarizes these applications to offer insight into the source of our observations.

Table 2. Projects Developed on AR and VR Platforms

Platform	Project Description
Google Cardboard	AR + VR enabled maintenance training, 360-degree video
Samsung Gear VR	VR introduction to the MQ-9 Reaper UAV
Oculus Rift	VR weapons course familiarization
HTC Vive	VR intelligence site exploitation
Microsoft HoloLens	AR instructor / student networked maintenance training

For the Google Cardboard we used the Unity engine to develop a VR and AR enabled maintenance training application. Users select a vehicle from a library, then select a system (such as engine, transmission, suspension), after which they select the specific maintenance operation on which they wish to train. The user then is able to step through the procedure at their own pace while manipulating a 3D, animated model of the assembly. The animated 3D-model may be viewed on the smartphone screen with or without the Google Cardboard in AR or VR mode. We have recorded and delivered several 360 degree-video segments for this platform as well.

For the Samsung Gear VR, we developed a virtual tour of an MQ-9 Reaper Unmanned Aerial Vehicle (UAV) using Unity. (It also will run on the Google Cardboard.) The user can traverse around an articulated/animated model of the Reaper UAV, using a UI system we call teleportation, in order to review the preflight inspection sequence.

For the HTC Vive, we developed a training system built in Unreal Engine 4 that allows a student to explore a geo-typical residence for valuable intelligence (laptops, thumb drives, maps, documents, etc.). The controllers are used as virtual hands to interact with the environment, such as to open a drawer see what is in it (or taped to the bottom). As well as being scored by the application, the user's performance can be evaluated in real time or played back via video capture of the player perspective.

On the Oculus Rift, we developed a weapon and range familiarization course using the Unreal 4 engine. The application allows a soldier to rehearse the movements and firing positions found at an actual training range.

For the Microsoft HoloLens, we developed a collaborative student/teacher system that allows two geographically separate users to perform a repair procedure on a suspension assembly of an amphibious vehicle. We reused much of the content we developed for the Google Cardboard application.

COMPARISONS

Each of the platforms have specific advantages and disadvantages that should be taken into account when designing an AR or VR experience. As designers and developers of serious games, we focused on aspects critical to our user populations such as training, education and simulation value versus entertainment value. Table 3 summarizes our findings, followed by supporting and elaborating details.

Table 3. Platform Comparisons

	Collaboration	Cables	Cost	Processing Power	Portability	3D Volume	Inside-Out Outside-In
Cardboard	*	No	$	*	****	None	No
Enhanced Smart Phone	**	No	$$	*	****	None	No
PC HMD	***	Yes	$$$	****	*	4 meters	Outside-In
HoloLens	****	No	$$$$	**	***	3+ meter	Inside-Out
Ratings run from * = relatively weak to **** = relatively strong							

Cost

Cost is often a primary concern for any customer. A smart phone at roughly $300 paired with a $5 Cardboard is the most affordable option, but it offers the lowest fidelity and processing power of the systems we evaluated. Enhanced smartphone viewer-devices cost a bit more and provide some enhancements in sensors as well as user input options. The PC/HMD options at approximately $2,000 represent a significant step up in price, but they also provide the most processing power and opportunities for immersive, high fidelity environments. At $3,000, the HoloLens offers processing power higher than the phone based options but not as robust as the PC based options nor as immersive due to the narrower field of view in which CGI may be drawn.

Input Schemes

While almost every VR/AR device can be modified to include a wide range of input devices through hardware and software add-ons, our discussion will focus on the default user input for each system.

A common user interface for the Google Cardboard is simply “gaze select.” The user superimposes a crosshair User Interface (UI) element on an icon and holds it there for several seconds. Typically the UI element appears as an animated circle filling in, signifying the action is selected. This interface is common for these devices, in large part, because users need to hold the cardboard devices to their face with both hands. A single button tap or finger tap directly on the screen can also be used, but it can be unreliable. Voice commands may also be used with varying degrees of success depending on implementation, ambient noise, user accent, etc.

The Samsung Gear VR works with both a touchpad built into the side of the HMD or a Bluetooth controller. The Google Daydream has a single button peripheral which connects wirelessly to the viewer. As with the other smartphone based platforms, gaze select and voice may also be used.

Both the Vive and Oculus are compatible with game console controllers, but both offer better experiences with their propriety 3D tracking hand held controllers; one for each hand. Each support multiple buttons and trackpad/joystick.

The HoloLens has the widest range of built in UI Options. A single button “clicker” can be used to select and rotate objects (not unlike the Google Daydream device). Hand gestures are recognized by the device such as “Select” via index finger and thumb pinching together and “back” via expanding fist gesture. Voice recognition using Cortana (Microsoft’s version of voice recognition interface) is also supported.

In our work, we have found the user input schemes of the PC HMD systems (Vive, Oculus) to be the most reliable, easiest to train, and best suited for fine-tuned gestures and manipulation. The HoloLens’ natural input gestures can take a few minutes to learn, and the voice recognition, while very reliable even in noisy conditions, can be tricky due to the nuances of language and wide range of synonyms and homonyms. Grab/Pinch/Hold/Grasp/Clutch/Seize/Take can all mean the same thing to different users as can Drop/Place/Release/Let Go/Place Here. Keeping vocabulary to a minimum while building in as much redundancy as possible is the best design choice. In all cases, user fatigue should be a design consideration as well.

Multiplayer

All the systems we evaluated can allow for multiple users to interact with the same CGI either in the same physical room or remotely, thereby supporting networked multiplayer experiences. Uniquely, the HoloLens allows a user to see both the CGI and a remote networked user's environment. This may be of particular use in a training environment.

Cables

Only the HoloLens currently allows the user to move naturally through 3D space without the use of cables. While the smartphone-based platforms are untethered, they do not track the user in 3D space. With these systems, navigation through the virtual world is accomplished via the user interface. Cable free options are in early prototype stages for PC HMDs and can be accomplished now through the user carrying the PC or laptop in a backpack or other mount.

Portability

The smartphone-based platforms can fit inside a coat pocket. The PC HMDs require a PC, the HMD itself, as well as several tracking devices and stands but can fit into a carry on suitcase. The HoloLens can fit into commuter backpack.

3D Volume

For volume, we are referring to the amount of physical space that can be tracked and traversed in 3D space. The smartphone-based platforms cannot track 3D space, but they can make use of integrated GPS receivers for some location functionality. The PC HMDs can track within a roughly 4 meter cube. The HoloLens can perform spatial mapping (different terminology due to the inside-out vice outside-in approach) to a distance of 3.1 meters, but it can also cache scanned areas to build out larger and connected spaces like corridors and connected rooms.

Inside-Out vs. Outside-In

Both inside-out and outside-in approaches have their merits. Current implementations allow outside-in to appear to support higher accuracy tracking. However, outside-in requires additional hardware to perform this tracking. This must be set in place and is static. It also requires setup time and power. Only the HoloLens currently accomplishes inside-out.

DEVELOPMENT TOOLS

One of the reasons that AR and VR has seen such amazing growth is the unheralded availability of low-cost, high-quality development tools. In the first commercial generation of VR content, everything had to be programmed from scratch. Now, the same game engines used to power Halo, Call of Duty®, and Angry Birds can be used to create AR and VR content with the flip of a single toggle- All of which are available for free for non-commercial and limited commercial use. These include the Unreal and Unity game engines and web-based systems such as CoSpaces.

While game engines can be used to create virtual interactive 3D VR-experiences, cameras can be used to create 360-degree video VR-experiences with live actors and real environments. This content can be shared via YouTube and Facebook and viewed on all of the above platforms.

VR cameras are available in a wide range of prices from $250, for easy-to-use consumer systems, to $70,000 and up, for professional-grade systems. The inexpensive consumer models render a 360-degree image or video from two 180-degree fish-eye lenses. Image processing is done in the camera, which means a user with a smartphone can edit and upload the stills or videos to social media while they are on location. These cameras are lower resolution and offer only 2D VR-stills and -videos. The higher-end systems can shoot in higher resolution and in stereoscopic 3D VR but require off line processing and video compositing.

ADDITIONAL THOUGHTS

Each of these platforms has strengths and weaknesses. We found that all the platforms integrated well with the game-based development tools most developers will be familiar with. In practice, there are details associated with all the platforms that should be considered at the design phase. For example, how many different smartphones will you support (of using a smartphone based platform) and is your user-interface responsive to different screen sizes in VR? What application distribution system do you intend to use? For your use case, what are the dangers of immersing your user in VR?

Novel hardware and software solutions, as well as updates to existing platforms, are being brought to market at an incredibly fast pace. Please keep in mind that information in this document is likely already dated.

We believe strongly in the value of VR and AR for a multitude of uses. Ultimately it is up to the developer to identify how the use of these technologies will improve an application beyond what can be achieved with common two dimensional displays.

REFERENCES

Wikipedia. (2017). Google Cardboard — Wikipedia, The Free Encyclopedia. Retrieved April 3, 2017, from https://en.wikipedia.org/w/index.php?title=Google_Cardboard&oldid=769625464