Evaluating the Applicability of Repurposed Entertainment Virtual Reality Devices for Military Training

Douglas Maxwell, Evan Oster, Spencer Lynch Aptima, Inc. Woburn, Mass. dmaxwell@, eoster@, slynch@aptima.com

ABSTRACT

Low cost virtual reality systems are currently available to consumers at a quality and price point never before offered. The military training community has begun to investigate these devices given the potential for immersive training. While these systems may be acceptable for training many tasks, their budget-focused quality introduces visual impairments that compromise performance on some critical tasks. We will demonstrate how users with normal vision may be limited to lower visual acuity within virtual environments due to the limitations of the VR device. Given the specific visual acuity needs of military training applications, this paper calls for caution when using low cost HMDs for AR/VR based training. This paper discusses how the quality of current virtual reality components affects critical training tasks and we mathematically demonstrate the issues with the typical COTS VR offerings. We discuss a strategy for determining if a specific training need is applicable for the VR systems by showing the calculations needed to pair the training requirements to the hardware specifications of the equipment. Lastly, we will discuss a use case that directly relates to the topic that requires a mounted soldier to identify signs of improvised explosive devices at distance. It is Aptima's intent to leave the reader with the knowledge needed to properly evaluate the training activity and pair it to the technical attributes and hardware specifications of the candidate virtual reality apparatus.

ABOUT THE AUTHORS

Dr. Douglas Maxwell is a Senior Research Engineer at Aptima, Inc., with a background in mechanical engineering and simulation-based training technology. Dr. Maxwell's expertise includes a practical background of custom virtual reality hardware interfaces, scientific visualization, games for training and simulation-based training systems. He has also conducted large scale training effectiveness studies with hundreds of participants in an effort to determine what key technical attributes were of value in a virtual simulator. He holds a Ph.D. and M.S. in Modeling and Simulation from the University of Central Florida, and a M.S. and B.S. in Mechanical Engineering from Louisiana Tech University.

Evan Oster is a Scientist at Aptima, Inc. with a background in courseware development, game-based training, and mixed reality technology. His current work centers on using cognitive science, instructional design, and analytics to increase training effectiveness in live, virtual, and constructive environments. He has experience managing interdisciplinary teams responsible for the development and integration of various immersive training products. Past work has involved designing game and simulation-based training using immersive virtual environments for the Navy. Mr. Oster has also created several augmented reality games with a university games lab while providing software development evaluation. Evan Oster holds an M.A. in Instructional Design and Technology and a B.S. in Human Development from Virginia Polytechnic Institute and State University as well as an M.S. in Curriculum and Instruction from Radford University.

Spencer Lynch is a Senior Software Engineer at Aptima, Inc., where he specializes in high-performance, web-based applications and full-stack development utilizing the latest web technologies such as WebGL, WebVR, WebRTC and WebSockets. He leads the Visualizations Community of Practice at Aptima, where colleagues meet to discuss web-based visualizations and related technologies. Mr. Lynch has experience at every stage of software development from requirements gathering to full deployment. Prior to Aptima, Mr. Lynch developed full-stack web applications for intelligence reporting and real-time simulator software for use in LVC training. Mr. Lynch is a member of SIGGRAPH and holds a B.S. in Computer Science from Bowling Green State University.

Evaluating the Applicability of Repurposed Entertainment Virtual Reality Devices for Military Training

Douglas Maxwell, Evan Oster, Spencer Lynch Aptima, Inc. Woburn, Mass. dmaxwell@, eoster@, slynch@aptima.com

INTRODUCTION

The entertainment industry is producing high quality virtual reality devices at an attractive price point. This price point achieves a long-anticipated ambition of the entertainment industry, making virtual reality accessible to the average consumer. This has subsequently gained the interest of the U.S. military for training applications. However, while the Head Mounted Displays (HMDs) currently available on the market are acceptable for many military training tasks, they are not suitable for all.

The military and entertainment industry have similar applications and challenges with this technology, but their goals and criterion can differ considerably. For example, the entertainment industry places a higher emphasis on cost and weight of HMDs over quality of optics to reach a higher number of buyers. When discussing visual fidelity, the goals and criterion for the military change depending on the end user. Consider the following use cases, "In a scenario designed to train military long-range snipers, gravitational influences on bullet trajectory may be an important variable to include in the simulation. Conversely, if the same scenario was used to train military medics in finding and treating casualties in an emergency, the gravitational dimension may not be as necessary as environmental metrics or treatment efficacy. Other characteristics, such as resolution and realistic texturing, may be noticed and appreciated by trainees, but they are directly impacted by these variables." (Alexander, Brunye, Sidman, & Weil, 2005). These characteristics matter from a military training measurement, assessment, and performance perspective and a one-size-fits-all approach to visual fidelity will not suffice.

It is the intent of this paper to highlight the necessity of evaluating virtual reality training device specifications based on training task attributes, safety severity, and asset criticality before the making a purchase. In order to accomplish this, a clear distinction must be made with our focus. We are not evaluating the user performance on awareness of objects in the virtual environment during training, but rather if the necessary fidelity exists for objects to be reasonably perceived. In essence, the focus is whether an HMD provides a clear enough representation of objects in the environment where a user with average visual acuity can successfully accomplish the training task. The ultimate goal is through this paper and with careful consideration the military can optimize the use of training funds and better serve its most valuable asset, the warfighter.

DISPLAY RESOLUTION AND OPTICS WORK TOGETHER

When evaluating a virtual reality device for a training task, there are a number of factors to consider. An examination of the tasks should produce a set of learning objectives and measures that can be evaluated to determine learner proficiency.

What Does 20/20 (6/6) Vision Mean?

According to the International Council of Ophthalmology, the specification of the measured visual acuity is determined by establishing the smallest figure or letters of different sizes (also known as optotypes) that can be identified correctly by a patient at a prescribed observation distance. (Colenbrander, 1988)

The Snellen Eye Chart (Figure 1) is perhaps the most familiar means of determining visual acuity for a human. The Snellen fraction is expressed as the test distance (m) divided by the distance at which the optotype can be correctly identified (M). 20/20 vision (6/6 Meter vision) is a term used to express normal visual acuity (the clarity of vision) measured at a distance of 20 feet, which also means they can discern a detail of one arc minute. (Colenbrander, 1988)

If a person has 20/20 vision, they can see clearly at 20 feet what should normally be seen by other people at that distance. If a person has 20/40 vision, then it means that person must be at a distance of 20 feet to see what people with normal vision may see at 40 feet.

The sizing and placement of the letters in the eye chart are purposefully designed to correspond to arc angles at the testing distance. The "E" on the 20/20 line corresponds to five arc minutes and each stroke of the letter corresponds to one arc minute at a distance of 20 feet (6 meters) from the test subject. (Figure 1) Snellen chose to baseline the size of the "normal" vision on one arc minute of resolution and calculate the testing distance through experimentation to find what statistically normal human vision would be. By using Snellen's notation, a person with 20/20 vision would have visual acuity of 1 and a fighter pilot with 20/8 vision would have exceptional acuity of 2.5.

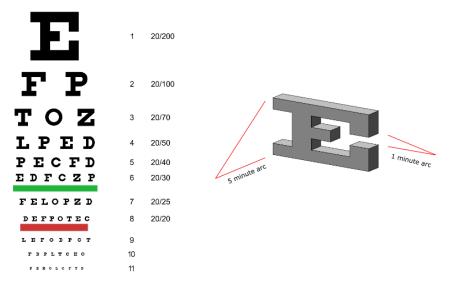


Figure 1. Snellen Eye Chart and Arcs

The Basic Visual Components of the Wearable Virtual Reality System

Two popular virtual reality devices were used as examples of COTS (commercially available off the shelf) components for this study. The Oculus Rift is a VR product developed by Oculus VR, now a division of Facebook, Inc. It's primarily audience is the entertainment industry and has been on the market to consumers since mid-2016. (Oculus VR, 2018) The Vive is a competing VR headset developed in cooperation by HTC and the Valve Corporation. (HTC Corporation, 2017) Like the Oculus, the Vive is also targeting the home VR entertainment market. Both have a robust catalogue of available games and are compatible with common game development platforms such as Unity. Table 1 provides a quick view of the specifications for each unit, relevant to this paper.

	HTC Vive	Oculus Rift
Display Type	Samsung OLED	Samsung OLED
Display Size	91.9 mm x 2, 447 ppi	90 mm x 2, 456 ppi
Resolution	1200 x 1080 per eye	1200 x 1080 per eye
Refresh Rate	90 Hz	90 Hz
FOV	110H x 113V, 8mm lens to eye	94H x 93V, 12 mm lens to eye

Much attention is paid to the screen resolution of the display technology used in the wearable VR headsets, however the optics are just as important. Humans cannot resolve a picture with clarity that is placed eight to 12 millimeters from the eye. (Mon-Williams, Warm, & Rushton, 1993) This distance is simply too short for the anatomy of our optics to focus on, without a little help. Both the Oculus and the Vive use Fresnel lenses to collimate light from their OLED image generators and focus that light onto the human eye. Fresnel lenses are a key component to low cost virtual reality devices. The lens design allows a substantial reduction in thickness (and thus mass and volume of material), at the expense of reducing the imaging quality of the lens, which is why precise imaging applications such as photography customarily use larger conventional lenses. Fresnel lenses are typically used in imaging, photography, illumination and projection. They are used in handheld magnifiers, automotive visual enhancers, short length telephoto lenses, lighthouses, and automotive headlamps. The design of the lens is efficient; however, it also introduces visual artifacts and distortions, so Fresnel lenses are typically not used when quality is a priority. Figure 2 shows how light enters a Fresnel lens, collimates to different focal lengths, thus causing visual distortions manifested as "wavy" concentric circles in a scene.

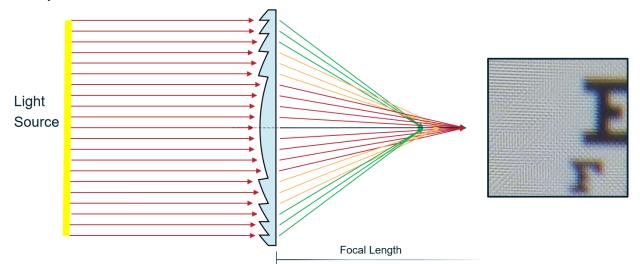


Figure 2. Fresnel Lens Collimation & Optical Distortion

The HTC Vive and Oculus Rift both use Fresnel lenses in the optics. The photograph in Figure 3 shows the distortion created by the lenses inside the Vive. The Fresnel lens creates a wavy effect that radiates outward from the center of the scene. Figure 4 shows a side by side comparison of the Vive (on the left) to the Rift (on the right) optics.



Figure 3. HTC Vive Fresnel Lens (left) and Oculus Rift Hybrid Fresnel Lens (right)

A second key component in the VR device is the image generator, usually an Organic Light Emitting Diode (OLED) array. This array is very similar to a television or computer monitor, but in this case much smaller and lower powered.

The device manufacturers are quick to advertise the resolution and pixel density, but what does this mean in terms of visual acuity?

Visual Acuity Inside a Virtual Reality Device

A common misconception is that there is some relationship between LED resolution or pixel density (or pixels per inch) to eye anatomy of rods & cones. Human visual resolution and acuity is calculated differently. When humans look at a display, there are many elements to take into consideration. The **size, field of view** of the display and the **distance** from which it is being viewed are key factors. The last element we will discuss is the **density** of the dots on the display. Referring to the product information in Table 1 and the previous acuity discussion, how does the presentation of each VR device compare to "reality"?

For a human with normal vision (visual acuity of 20/20), the smallest arc they can detect in visual degrees is 0.016667. Equation 1 is used to calculate this visual resolution:

Visual Resolution =
$$\frac{1}{Visual Acuity} \times \frac{1}{60} = \frac{1}{20/20} \times \frac{1}{60} = 0.016667 \, deg$$
 (1)

Equations 2 and 3 calculate an effective distance for the OLED to the eye (*as opposed to the lens to the eye, which is much shorter*). Using Width and Field of View, the optimal viewing distance can be calculated:

$$Eye \ Distance_{Vive} = \frac{\frac{OLED \ Width}{2}}{Tan(\frac{Viewing \ Angle}{2})} = \frac{\frac{91.9 \ mm}{2}}{Tan(\frac{110 \ degrees}{2})} = 32.2 \ mm$$
(2)

$$Eye \ Distance_{Rift} = \frac{\frac{OLED \ Width}{2}}{Tan(\frac{Viewing \ Angle}{2})} = \frac{\frac{90 \ mm}{2}}{Tan(\frac{94 \ degrees}{2})} = 42 \ mm$$
(3)

The distances calculated here represent values need *without* the Fresnel lenses. Since these distances would require the OLED displays to be placed an unacceptable distance from the face and interfere with the ergonomics of the head gear, Fresnel lenses are then added in front of the displays to collate the light and focus it on the eye at a much closer distance.

Equation 4 calculates the smallest pixel size that can be discerned at distance from eye to lens.

Smallest Pixel Size =
$$2 \times distance \times Tan(\frac{Visual Resolution}{2})$$
 (4)

For a human with normal vision and at the distances required, the smallest pixel sized they could discern inside the HTC Vive would be (5) and inside the Oculus Rift would be (6):

Smallest Pixel Size_{Vive} = 2 × 32.2 mm × Tan
$$\binom{0.016667 \ deg}{2}$$
 = 0.0094 mm ~ 2700 ppi (5)

Smallest Pixel Size_{Rift} = 2 × 42 mm × Tan
$$\binom{0.016667 \text{ deg}}{2}$$
 = 0.0122 mm ~ 2100 ppi (6)

Table 2 shows a comparison of the smallest perceptible pixel density for a human with normal vision and the pixel densities provided by the Vive and Rift. There is a significant difference between what the devices are capable of presenting versus what is needed to meet minimum normal human vision acuity requirements. This explains why individual pixels can be seen when wearing either device and why edges that should be sharp are seen as fuzzy.

 Table 2. PPI Delta Between Human Perception and Device Specification

	Device Specification (ppi)	Calculated Required Pixel Density (ppi)
HTC Vive	447	2700
Oculus Rift	456	2100

The device manufacturers are aware of this disparity and are working on new display solutions. As far back as 2013, Sony announced a 2098 ppi OLED display (0.7 inch @ 1280x720) for use in medical applications. (Sony, 2013) Samsung announced in 2015 a development effort for a new 11k Super-Resolution display (Sung, 2015) that it plans to eventually deploy to its mobile phones. This new super resolution display was aiming for a goal that would produce a 2250 ppi display. The reason we don't have these displays today (as of this writing) in commercial products is mostly costs associated with mass production. Sharp announced a 4k display with over 1000 ppi in 2016 (Byford, 2016) with the Aquos Crystal display that has panels that can be cut into various shapes. Late in 2017, Samsung announced new small display technology with a resolution of 2014x2200 resolution with a density of 858 ppi aimed at VR devices and is slated to be included in the new HTC Vive Pro and next generation Oculus Rift. (Martindale, 2017)

Proposed Approach to Testing Visual Acuity of Virtual Reality Hardware

The Snellen approach to testing human vision was used to assess real-world eyesight acuity within the HTC Vive and Oculus Rift virtual reality devices tested for this paper. A Snellen eye chart was prepared by placing a texture of the chart on a geometric primitive with dimensions 1-meter-wide by 1-meter-tall and ensuring the topmost "E" was 88.6 mm tall. We prepared a Unity environment and placed the eye chart in an empty scene, placing the viewport directly in front of the chart at a distance of 20 feet (approx. 6 meters). Figure 5 shows a photograph of the eyechart as seen through the HTC Vive.

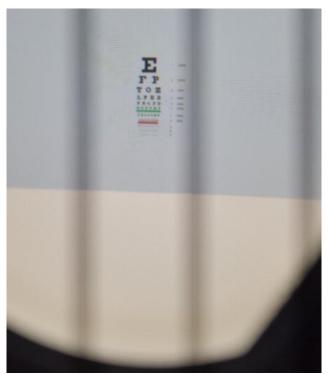


Figure 4. Photograph of Snellen Eye Chart within HTC Vive

Next, we viewed the chart in the HTC Vive and Oculus Rift using test subjects with normal vision. We then recorded the smallest line on the eye chart to determine the best visual acuity that each device could provide each viewer. "Real world" is a benchmark of a presentation for a person with 20/20 vision and results in an acuity of 1. The viewer in the Oculus Rift reported being able to read the 20/40 line on the chart, resulting in an acuity score of 0.5. Lastly the viewer in the HTC Vive reported being able to read the 20/50 line in the chart, resulting in an acuity score of 0.4. The results are shown in figure 6.

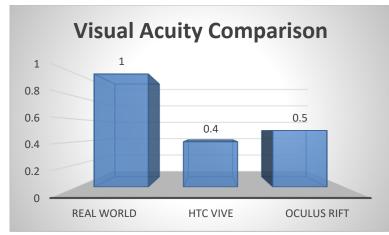


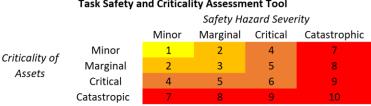
Figure 5. Visual Acuity Comparison

DETERMINE WHICH TRAINING TASKS ARE APPROPRIATE

Knowing the limitations of the VR devices allows for military trainers and course managers to determine which training tasks are appropriate for virtual reality training and which should be excluded. Military instructional designers start with doctrine documentation when constructing a training regimen in any military simulator, virtual reality and virtual environments are no exception. The entertainment industry has great leeway to create content and game levels within their platforms that highlight strengths and minimize weaker aspects of their commercial VR devices. The military does not have that luxury and must construct realistic scenarios with all of the training objectives represented. This in part explains why the military has so many simulators in a given domain as "one size" does not fit all and multiple kinds of simulators are sometimes included in a curriculum to cover all the desired training tasks. The types of training tasks of interest in this paper are ones requiring environments with high visual fidelity where performance must be demonstrable "at a distance" with average visual acuity. In order accurately identify training tasks vital to this performance, it is recommended tasks are evaluated across two dimensions (1) task attributes and (2) assessment of safety hazard severity and criticality of assets.

A literature review revealed the task attribute of primary importance for this device evaluation is a cognitive task requiring visual or perceptive skill. While cognitive tasks require a dynamic mental process and exert a load on the working memory, when it is highly dependent on sharp visual or perceptual skills, an additional layer of complexity is added. Placing trainees in virtual environments has many advantages, but as Baum concluded, "Very few data exist to describe the relationship between training device fidelity and training effectiveness for maintenance tasks. In particular, insufficient research has been conducted on the fidelity requirements for training cognitive (i.e., nonprocedural troubleshooting) or perceptual-motor maintenance tasks" (Baum, Riedel, Hayes, Mirabella, 1982). Even though this statement was made over three decades ago, unfortunately it remains relevant today given the rapid development of technology used for training and a lack of focus on simulator effectiveness. Overcoming the challenges of these complex visual and perceptive tasks is of critical importance for the warfighter today, but the same does not necessarily hold true for the entertainment industry developing the devices.

The second-dimension tasks were evaluated against the assessment of safety hazard severity and criticality of assets. A framework was developed as a decision aid to help answer the question, "How important is this task to the success of the mission in the operational environment?" while providing a way to classify and prioritize training tasks. Tasks are rated according to two criteria; safety hazard severity and criticality of assets. A four-tier scale from minor to catastrophic is used for both criteria and a numeric identifier is assigned to the output levels resulting in a ten-point assessment scale (Figure 6).



Task Safety and Criticality Assessment Tool

Figure 6. Criticality versus Safety Hazard Decision Support Tool

Decision making will vary and depend on the training provider's objectives and level of funding, but in general level 5 and above deserves consideration. This framework enables military trainers and course managers to quantifiably rate cognitive tasks requiring visual or perceptive skills so they can responsibly plan and spend training funds where it is most appropriate. The following brief case studies are examples of how training tasks are evaluated across these two dimensions to ensure they are appropriately presented and exercised within the virtual environment.

Case Study: Improvised Explosive Device Detection

There are practical advantages to using virtual simulators when training mounted and dismounted infantry for identifying signs of improvised explosive devices (IEDs). Given this task ranks as a 9 or 10 on the assessment of safety hazard severity and criticality of assets scale (Figure 6), it is imperative to receive high caliber training in a high-fidelity environment. Live training is considered the preferred option, but proxy devices are used instead as it is too dangerous and expensive to use real explosives. There is also limited availability of the live ranges, so virtual options such as the Army's Virtual Battle Spaces 3 and other simulators are considered.

For this example, we will use task number 052-COM-1271 "Identify Visual Indicators of an Improvised Explosive Device (IED)". Two of the performance measures supporting this task are to identify IED indicators and identify the IED casing. The indicators vary considerably given the environment, time of day, culture, population, and political climate. A few indicators that might identify a potential IED are freshly disturbed earth, concrete that does not match the surrounding areas, overloaded vehicles, a person on an overhead pass, and trash along roadways. A casing could be anything that can contain any or all of the components of the IED, may provide enhanced fragmentation, and may camouflage the IED. Examples of IED casings range in size from soda cans to 55-gallon barrels. For this example, we evaluated the ability to spot a soda can, a one-gallon jug and a 55-gallon barrel sized IED casings from 25 meters away (note: given the size of a soda can, it is rendered at the single pixel level and therefore impossible to perceive).

In order to determine if the VR devices are appropriate for this task, we can calculate if the object can be presented to the user in a way that can be seen. For example, the Oculus has a pixel density of 456 ppi and we previously calculated the optimum eye distance to the OLED is 42mm. A standard 55-gallon drum is 851 mm tall, so the image height in the device is calculated at 0.143 mm, which is 16 pixels tall, equations 7, 8, and 9.

$$Image Arc = atan\left(\frac{Drum Height}{Detection Distance}\right) = atan\left(\frac{0.851 m}{25 m}\right) = 0.034027 radians = 1.95 deg$$
(7)

Image Height = $tan(Image Arc) \times Eve Distance = tan(0.034027) \times 0.042 m = 0.00149 m$ (8)

Pix High = Image Height \times Pixel Density = 0.00149 m (0.0363 in) \times 456 ppi = 16.5 pix (9)

Similar calculations were performed for the 1-gallon jug and soda can, and repeated for the Vive. The results are summarized in table 3.

IED Casing	HTC Vive	Oculus Rift
Soda Can	1 pixel height	2 pixels height
1 Gallon Jug @ 25m	3 pixels height	5 pixels height
55 Gallon Drum @ 25m	12 pixels height	16 pixels height

Figure 7 shows what a 55-gallon drum would look like if it were scaled down to just 16 pixels from a recognizable texture. It could be argued that it would not be reasonable to expect a student to be able to recognize this object at the required distance and is therefore not recommended to use the Oculus Rift for this training task. This example also does not take into account any distortions introduced by the Fresnel lenses. It could be argued that additional blur or "wavy" artifacts would exacerbate the image quality problems. This illustrates the coupling of optics and OLED displays, showing that in this case even a higher pixel density display may not help image quality unless the lens quality was increased as well.



Figure 7. Comparison of 55-Gallon Drum Texture (left) Scaled to 16 Pixels Tall (right)

CONCLUSIONS

This paper outlines a proposed approach to testing visual acuity within commercially available virtual reality devices. The purpose was to assist decision makers who are interested in using these devices for military training applications. We discussed how size, field of view of the display and the distance from which it is being viewed, and pixel density are key presentation factors to the user. We discussed how to calculate the difference between device parameters and normal human vision and how the delta impacted the selection of proper training tasks to include in the simulation.

The issue of matching training tasks to VR gear can be addressed in two directions. A cost-effective method of acquiring training devices is to purchase COTS equipment, so examining the device specifications is very important. The training tasks and measures for military training is clearly defined in the doctrine and accessible via Army Training Network. With the equations used in this paper, we demonstrated it is possible to calculate if training objectives can reasonably be expected to be achieved in a given apparatus. Conversely, using the equations in this paper, the minimum specifications for a VR training device can be calculated and the used to generate requirements for the acquisition community.

The next step is to prepare an experiment with a statistically significant number of human test subjects with normal vision to conduct a series of tests on these and new devices. With this information, decision makers may gain a better understanding of which training tasks most appropriate for candidate COTS virtual reality entertainment gear.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Zach Shields of Aptima, Inc. for his part in test scenario setup and data collection.

REFERENCES

- Alexander, A., Brunye, T., Sidman, A., Weil, S. (2005). From Gaming to Training: A Review of Studies on Fidelity, Immersion, Presence, and Buy-in and Their Effects on Transfer in PC-Based Simulations and Games. DARWARS Impact Training Group.
- Baum, D., Riedel, S., Hayes, R., Mirabella, A. (1982) Training Effectiveness as a Function of Training Device Fidelity. Final Report March 1981-August 1982. http://www.dtic.mil/docs/citations/ADA140997
- Byford, S. (2016). Now I know what a tiny 1,000-ppi 4K display looks like. Retrieved February 5, 2018, from https://www.theverge.com/circuitbreaker/2016/10/7/13197252/sharp-vr-4k-display-1000-ppi
- Colenbrander, A. (1988). Visual Acuity Measurement Standard. *Published in the Italian Journal of Ophthalmology II*, 1–15.
- HTC Corporation. (2017). HTC Vive. Retrieved February 16, 2018, from https://www.vive.com/us/
- Martindale, J. (2017). Samsung's next-gen VR display doubles pixel density of Rift, Vive screens. Retrieved February 7, 2018, from https://www.digitaltrends.com/virtual-reality/samsung-vr-display-resolution/
- Mon-Williams, M., Warm, J. P., & Rushton, S. (1993). Binocular vision in a virtual world: visual deficits following the wearing of a head-mounted display. *Ophthalmic and Physiological Optics*, *13*(4), 387–391. https://doi.org/10.1111/j.1475-1313.1993.tb00496.x
- Oculus VR. (2018). Oculus Research. Retrieved from https://www.oculus.com/blog/
- Sony. (2013). Sony Introduces "head-mount image processing unit" for endoscopic image display. Retrieved February 9, 2018, from https://www.sony.net/SonyInfo/News/Press/201307/13-085E/index.html
- Sung, H. (2015). Samsung Develops "11K" Super-Resolution Display Along With 13 Companies...Putting 26.5 Mil USD For 5Years. Retrieved February 7, 2018, from http://english.etnews.com/20150710200002