Creating the Perfect Illusion : What will it take to Create Life-Like Virtual Reality Headsets?

Eduardo Cuervo
Microsoft Research
cuervo@microsoft.com

Krishna Chintalapudi
Microsoft Research
krchinta@microsoft.com

Manikanta Kotaru
Stanford University
mkotaru@stanford.edu

ABSTRACT

As Virtual Reality (VR) Head Mounted Displays (HMD) push the boundaries of technology, in this paper, we try and answer the question, “What would it take to make the visual experience of a VR-HMD Life-Like, i.e., indistinguishable from physical reality?” Based on the limits of human perception, we first try and establish the specifications for a Life-Like HMD. We then examine crucial technological trends and speculate on the feasibility of Life-Like VR headsets in the near future. Our study indicates that while display technology will be capable of Life-Like VR, rendering computation is likely to be the key bottleneck. Life-Like VR solutions will likely involve frames rendered on a separate machine and then transmitted to the HMD. Can we transmit Life-Like VR frames wirelessly to the HMD and make the HMD cable-free? We find that current wireless and compression technology may not be sufficient to accommodate the bandwidth and latency requirements. We outline research directions towards achieving Life-Like VR.

KEYWORDS

Virtual reality, Computer Graphics, Mobile devices, Video compression, wireless networking

ACM Reference Format:
Eduardo Cuervo, Krishna Chintalapudi, and Manikanta Kotaru. 2018. Creating the Perfect Illusion : What will it take to Create Life-Like Virtual Reality Headsets?. In *Proceedings of 19th International Workshop on Mobile Computing Systems and Applications (HotMobile'18).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3177102.3177115

1 INTRODUCTION

The ultimate goal of VR Head Mounted Displays (HMD) is to provide an experience of virtual worlds that are indistinguishable from physical reality – deemed *Life-Like VR* in this paper. Today’s HMDs such as Oculus Rift [42], HTC Vive [23] etc. remain far from providing this perfect illusion [10]. In order to achieve Life-Like VR, displays must match the limits of human visual perception specifically w.r.t to display properties such as resolution, field of view and frame refresh rates. In this paper we ask the question, *“What will be the specifications of a HMD that provides a Life-Like VR experience and what is the path to such a device being built?”*

Based on the limits of human visual perception, we first show that Life-Like VR headsets will require $6-10\times$ higher pixel densities and $10-20\times$ higher frame rates of that of existing commercially available VR headsets. We then ask, “Will it be possible to create Life-Like headsets in the near future, given this gap in specifications to be bridged?” We examine various technological trends such as pixel densities, frame rates, computational power of GPUs etc. Our analysis indicates that while displays are likely to achieve Life-Like specifications by 2025, the computational capacity of GPUs will remain a bottleneck and not be able to render at the high frame-rates required for achieving a Life-Like VR experience.

We discuss various approximation techniques that can be used to bridge the computational gap and the open research challenges. Offloading computation to a powerful computer (not part of the headset) will provide better quality experience than a standalone headset. However, this means that the rendered frames need to be transmitted to the HMD using cables connected to the headset. In order to allow unrestricted mobility and avoid tripping hazards, a key requirement for HMDs is to be cable-free. We then ask the question, “What is the path to making the Life-Like VR headsets cable-free?” We find that the required data rates will be so high that existing as well as upcoming wireless standards (802.11ad, 802.11ay) will find it hard or even impossible to transmit rendered video to the headset. We examine techniques to understand the path to reducing this extremely high data rates that will have “acceptable” degradation to the Life-Like video quality and discuss open research challenges.

In summary, we make the following contributions,

- Drawing on fundamentals of human perception, we posit that a Life-Like VR-HMD will require $6-10\times$ higher pixel density and $20-30\times$ higher frame rate compared to current day VR-HMDs.
- We analyze various technological trends we show that the key bottleneck to realizing Life-Like VR will be the significantly high demand on rendering computation.
- We show that the wireless bandwidth needed to transfer rendered Life-Like VR frames from a rendering machine to the HMD will be extremely high and no wireless standard in horizon will be able to accommodate it.
- Finally, we outline possible solutions and research directions that might help in realizing Life-Like untethered VR.

2 SPECIFICATIONS FOR LIFE-LIKE VR

In this section we try to establish the specifications of a Like-Like VR headset based on the limits of human perception.

Field of view (FOV). A Life-Like VR headset’s FOV should ideally be same as that of the human eye i.e. 210^o horizontally and 135^o vertically [4] as depicted in Figure 1. For stereoscopic vision there

---

HotMobile'18, February 12–13, 2018, Tempe, AZ, USA

ACM ISBN ISBN 978-1-4503-5630-5/18/02…$15.00
https://doi.org/10.1145/3177102.3177115

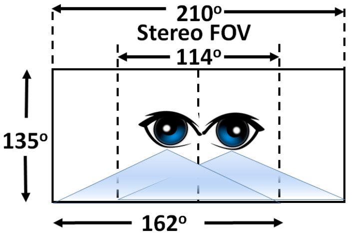

Figure 1: Field of View for a Life-Like VR-HMD

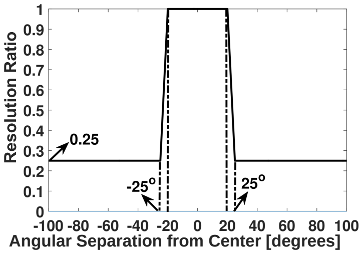

Figure 2: The foveation function

Display Property	Value
Field of View	$210^o \times 135^o$
Angular Pixel Density	>60 pixels/degree
Frame Rate	>1800 Hz
Dynamic Range	$1 : 10^9$
Colors	10 Million
Latency	> 7-20 ms

Figure 3: Specification for a Life-Like VR-HMD

is an overlap of 114^o degrees between the two eyes in the center. Thus, each individual eye has a view of 162^o ($\frac{210}{2} + \frac{114}{2}$).

Angular Pixel Density (APD) 20/20 vision allows a person to resolve two contours separated by 1 arc-minute (1.75mm at a distance of 20 feet) or 60 pixels/degree of angle subtended at the eye [43]. This ability to distinguish between two closely placed lines or contours is called *Resolution acuity.* Most people with corrected vision have better than 20/20 vision and human visual acuity ranges between 0.3-1 arc-minutes or (60-200 pixels/degree) [9].

Another important measure of acuity is *detection acuity* – the ability to detect tiny objects in the scene e.g. small spots. Detection acuity of the human eye far exceeds resolution acuity, especially in the periphery and in high contrast scenes. It has been measured that it is possible to detect a fine black line subtending only 0.5 arc-seconds [39] against an illuminated white background. In order to accommodate such a high level of detection acuity, the display intuitively should have a pixel density of 7200 pixels/degree. However, detection acuity works by averaging over large spatial extents and a lower acuity may suffice.

Today, HMDs typically offer a pixel density 10 pixels/degree and up to 13 pixels (e.g. Vive Pro [24]) – about 6× lower than the basic 20/20 vision standard. Thus, there is an immense gap between current day VR-HMDs and what the human eye is capable of. In the absence of user studies that establish what pixel densities will provide a Life-Like experience, based on the above discussion we speculate that *Life-Like VR will require somewhere between 60-120 pixels/degree (resolution of 1-0.5 arc-minutes).*

Foveal and Peripheral Vision. Broadly human vision can be classified into *foveal* and *peripheral* vision. Foveal, around center of the eye (fovea), has high resolution acuity (Section 2) and is able to discern finer details in the scene. Peripheral vision extends far from the center and while it cannot discern fine details, it can detect and help the eye focus its attention on important parts of the scene.

A VR-HMD with built-in eye tracking capabilities, e.g. the Fove VR Devkit [17], could potentially use lesser resolutions away from the person's center of the eye. Guided by user studies, [31] proposed a *foveation function* that does not "significantly" affect the users' perception. Figure 2 [19] depicts resolution as a function of angular separation (eccentricity) from the center of the eye. Resolution is 100% between -20^o to 20^o, then it is linearly reduced to 25% within 5^o and finally kept constant at 25% elsewhere. For a $210^o \times 135^o$ FOV, foveation reduces the overall number of pixels by 70%.

Frame Rate (Refresh Rate) While most current day VR headsets support frame rates of 90Hz, some commercial gaming monitors support frame rates of 240Hz (Foris FG2421 [16]) and 480Hz (Zisworks [47]). Nvidia recently debuted a working version of a 1700Hz display [26]. Why are such high frame rates essential for VR? While objects in the real world move continuously, in a video, pixels corresponding to moving objects will jump from a location to another in discrete steps across frames. This affects human visual perception due to two characteristics of the human eye.

Eye's Tracking Movements : As we follow objects with our gaze, human eyes reflexively tend to predict and track relative movement at high angular speeds in order to maintain gaze on the object. With *smooth pursuit movements* [33], the eye keeps a moving stimulus in view of the *fovea* – the area of the retina with the highest resolution acuity. These movements are continuous and can be as fast as 30 degrees/second [8]. When objects in view move faster, eyes try and follow them with *saccades*, which are extremely rapid changes in the eye viewpoint, in jumps from one position to another, as fast as 180 degrees/second or higher. Finally, the eyes move involuntarily to track objects if the head moves through the vestibulo-ocular reflex [33], matching the yaw and pitch angular velocities of the head of approximately 17 degrees/second [30].

Eye's Continuous Integration of Visual Stimuli : Processing in human eyes is not a discrete but a continuous analogue process where photo-receptors continuously accumulate light energy falling on them, which is in turn continuously processed by the brain to draw inferences about the world. When pixels corresponding to an object remain stationary in the video frame for a complete refresh cycle while the eyes move in anticipation of motion, they accumulate light energy from different parts of the image rather than the intended moving object. This causes a smearing and strobing effect called *judder*, reducing the eye's ability to detect objects clearly and inducing *simulator sickness* [3].

To avoid the above effects, frame rate of the HMD must be fast enough to avoid any abrupt discontinuous jumps in pixels corresponding to a moving object. Consider the eyes following an object in view, moving smoothly at an angular velocity of up to $30^o/sec$ (corresponding to vestibulo-ocular reflex and smooth pursuit movements). For an HMD with resolution of 60 pixels per degree, the pixels corresponding to this object will displace by 1800 pixels each second. In other words, a frame rate of about 1800Hz would ensure that there are no discontinuous pixel jumps. Note, that given eye's ability to follow objects at speeds up to $180^o/sec$ (due to saccades), higher frame rates may be required for a Life-Like experience. In the absence user studies, in this paper, we conservatively assume that frame rates of at least 1800Hz are required for a Life-Like experience.

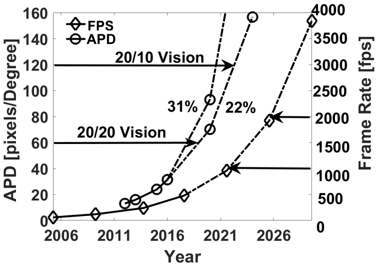

Figure 4: Trends for frame rates and pixel/degree.

Dynamic Range, Colors and Bits Per Pixel

Human eyes have a fantastic ability to *adapt* to lighting conditions, allowing us to see dim stars in low starlight as well as objects in bright sunny conditions – a contrast range (ratio of minimum to maximum intensities) of up to $1:10^9$. Further, the eye can discern about 10 million colors. These two factors in combination require a wider color space than the one offered by the standard 24-bit RGB space. It is for this reason that modern displays, especially OLED technology, provide high dynamic range color (HDR), often represented using a 96 bits per pixel floating point representation. [6]

3 TRENDS AND PATH TO LIFE-LIKE VR

Based on Section 2, Figure 3 summarizes the specs for a Life-Like VR HMD. In this section, we ask the question, “is it likely that Life-Like VR HMDs will be technologically feasible in the near future?” We try to answer this question by examining various relevant technological trends for both displays and GPUs.

Field of View (FOV). VR headset with a Life-Like FOV of $210^o \times 130^o$ is already available today (StarVR [38]). At a slightly lower FOV (200 degrees horizontally) but higher resolution, the Pimax 8K also aims to cover most of the human visible field of view [32]

Angular Pixel Density (APD). Today, the availability, specification and prices of displays for screens under 10 inches is driven by smartphones and tablets. Since VR is an upcoming technology with much smaller adoption, displays in VR headsets lag behind those of smartphone screens even though the underlying LCD/OLED technology is fundamentally the same. VR manufacturers today are forced to either rely on displays designed for smartphones (GearVR) or rely on smaller batches of specialized displays relying in slightly older technology (Vive, Rift). As VR gains popularity and the adoption gap between smartphones and VR headsets become narrower, the technological lag currently affecting VR displays will become narrower or disappear. Consequently, in this analysis of the evolution of screen pixel densities, we include all screens under the 10 inch form factor i.e. smartphones, tablets and VR-HMDs.

Table 2 provides the progression of highest pixels-per-inch (PPI) of commercial displays seen since 2012. In Table 2, we also provide the equivalent APD if this display technology were to be used in a 210^o Field of View. In order to compute the equivalent APD, we use as reference the StarVR [38] HMD, that offers a pixel density of 533 PPI and a corresponding APD of 15.8 pixels/degree (2560 pixels over 162^o). Further, we provide the percentage increase in APD per year during the last interval. The trend indicates between 22-31% increase in PPI per year. Figure 4 speculates the PPI based on the maximum and minimum growth trends that displays with 60 pixels/degree (20/20 vision) will be available by 2018 and 120 pixels/degree (20/10 vision) between 2021-2023. Indeed a display with PPI 2250 (APD of 67 pixels/degree) has been announced by Samsung to be released in 2018 [27]. We speculate it will take roughly two-three years for these announced displays to materialize in commercial devices and so VR-HMD displays with Life-Like pixel densities will probably appear between 2021-2025.

Table 1: Trends in Pixel Densities

Year	Max PPI	Equiv. APD	Increase/Year	Device
2012	440	13		LG [22]
2013	538	16	22.7%	LG [35]
2015	806	24	22.4%	XperiaZ5 [37]
2016	1058	31	31.3%	SEL [46]

Frame Rate. Framerates for LCD/OLED displays in the gaming industry have doubled steadily almost every 4 years over the last 12 years as depicted in Table 3. Since inherent switching times of OLEDs are under 0.1ms, the frame rate is fundamentally limited by speed of the digital circuits that control the OLEDs. Given this trend, it is reasonable to expect displays like the 1800Hz prototype presented by Nvidia [26] to be commercially available before 2025. In Figure 4 we speculate that we will see commercial monitors with frame rates of around 1000Hz by 2021 and 1800Hz before 2025. Commercial VR HMDs will likely see these frames rates two-three years later, depending on the adoption gap between VR headsets and smartphones at the time. This means that by 2028 VR-HMDs will probably start having refresh rates that offer a Life-Like experience.

Table 2: Trends in Frame Rates

Year	Max Frame Rate (fps)	Device
2005	60	Laptops
2009	120	Samsung 2233RZ
2013	240	Foris FG2421
2017	480	Zisworks [47]

Computation Capability for Rendering. In VR, the frames to be displayed are typically computed by a powerful computer based on physics, geometry and other logic related to the virtual scene. A higher pixel density implies that computation needs to be performed for a larger number of pixels and in a finer detail. A higher frame rate implies that the computation for each frame needs to be performed faster. Even though displays might have the pixel density and refresh rates that are required for a Life-Like experience, the computers generating these frames must be able to keep up.

VR headsets today, typically use powerful GPUs to perform the computations required for rendering in VR. Table 3 depicts the trend in increase in GPU computational power with the progress of years. As seen from Table 3, GPU computational power in GFlops has been increasing at an average of 34% per annum. These increases have been primarily due to improvements in the fabrication processes over the years from 40nm to 12nm technology. IEEE’s International Roadmap for Devices and Systems (IRDS) [25], predicts improvements in the fabrication processes up to 5nm over the next four years. The IRDS foresees complications in shrinking beyond 5 nm, and focuses on alternative devices such as neuromorphic circuits, quantum "‘qubits"’, spintronics, and others [25]. Relying on process shrinking and related incremental improvements thereafter, we speculate that the average growth in computational capacities at 34% to continue until 2025 resulting in about 140TFlops.

In Table 3 GTX780 , is comparable to what Vive and Oculus consider VR-ready, the GTX 970. Both operating at about 4TFlops [1] are capable of rendering 2.6 Million Pixels (2160×1200) at a frame rate of 90Hz. A VR HMD with $210^o \times 135^o$ FOV ($162^0 \times 135^o$ per eye 2) with APD of 60 pixels/degree operating at 90Hz refresh rate will have to render 157 Million Pixels ($162 \times 135 \times 60 \times 60 \times 2$) and will require the equivalent of about 244TFlops. If we allow for the VR-HMD to make use of foveation technology, then this requirement will drop to about 73TFlops. If GPUs are able to render at the equivalent of 140TFlops by 2025, then it will be possible to have VR-HMDs that operate offer 20/20 vision at 180Hz refresh rate. Note that this refresh rate is still 10× less than the minimum of 1800Hz required to avoid judder completely (Section 2).

Table 3: Trends in GPU Computational Power

Date	GFlops	Fab	Rate of Increase	Device
2010-03	1345	40nm		GTX480
2010-10	1581	40nm	27%/yr	GTX580
2012-12	3090	28nm	65%/yr	GTX680
2013-05	3977	28nm	24%/yr	GTX780
2014-07	4612	28nm	11%/yr	GTX980
2016-05	8228	16nm	39%yr	GTX1080
2017-12	13800	12nm	41%yr	Titan V

Summary. We summarize our analysis of trends as follows,

- We speculate that between 2025-2028, display technology will be Life-Like offering $210^o \times 135^o$ FOV, APDs of 60 pixels/degree (20/20 vision) and refresh rates of up to 1800 Hz.
- We also speculate that GPU technology will lag behind and will only be able to render 180 frames/sec at that resolution. This implies that intermediate frames will have to be generated using cheap approximate techniques such as image-based rendering (Section 4).

4 UNTETHERING THE HMD

As we saw in Section 3, the key bottleneck in VR is the computation involved in rendering frames. Given space, power and heat dissipation constraints, mobile GPUs suitable for standalone HMDs are typically $6 - 10\times$ less capable, than state-of-the-art GPUs found in desktop/gaming PCs or server class machines. For example, mobile GPUs today operate between 500GFlops (Adreno) - 1500GFlops (Tegra X2) while the best desktop GPUs operate at 13000GFlops. Similarly, graphics intensive content is often ported to mobile devices 7-10 years after debuting on full-sized devices (e.g. Doom 3, GTA Vice City, and Skyrim). Consequently, most high-end HMDs today, rely on a powerful GPU enabled PC (separate from the HMD) to render frames. These frames are then transmitted to the HMD via cables that connect the HMD to the PC. However, these cables are undesirable since they not only restrain free-motion but also might result in a hazard. In this section we explore the question "What is the path to un-tethering Life-Like VR-HMDs?"

Bandwidth Needed for Loss-less Wireless Transmission of Life-Like Rendered Video A straightforward way to un-tether an HMD is through wireless transmission of rendered frames. A few virtual reality systems and headset manufacturers today, use wireless connectivity to un-tether the HMD [1, 41]. Will it be possible to transmit a Life-Like rendered video wirelessly in the future?

As depicted in Figure 1 in a Life-Like, a stereo HMD with $210^o \times 135^o$, each eye individually has a $162^o \times 135^o$ field of view. For an APD of 60 pixels per degree, with 96 bits per pixel and 1800 frames/sec refresh rate the raw bit rate required to transmit frames is approximately 27 Tbps ($2 \times 162 \times 135 \times 60 \times 60 \times 96 \times 1800$). If the HMD includes foveation, the resulting bit-rate will be about 8Tbps. The state of the art hardware accelerated loss-less compression (DSC Display Stream Compression) [44] provides a compression ratio of 3:1, reducing this data rate to about 2.7 Tbps.

Current and Future Wireless Standards. 60GHz wireless standards promise the highest wireless bandwidths for wireless transfer. IEEE 802.11ad (WiGig) [21] prototypes today provide around 3 Gbps and promise up to 7Gbps in the future. The Wireless HD standard (UltraGig) [45] was created for HDMI cable replacement and products today provide data rates up to 4 Gbps while promising a theoretical of 25 Gbps. The 802.11ay standard in the making aims to provide a peak theoretical bandwidth of 176 Gbps. All these rates are at least an order of magnitude less than the required 2.7Tbps. This indicates that *there is no wireless standard in the horizon capable of accommodating lossless Life-Like VR video data rates.*

Free-space Optics. Optical transceivers transmit data at extremely high data rates over optical fibers in data centers. Ranovus [34] has a commercially available 200Gbps optical transceiver and often multiple transceivers are combined to provide an aggregate bandwidth of several Tbps. High bandwidth free-space-optics (FSO) based communication systems have been considered for stationary links [14, 20] where the transmitter and receiver are at fixed locations. Recently, a FSO system [15] has been proposed which allows for the transceivers to have micro-motions, for example, tiny movements of a building. However, to the best of our knowledge no FSO system exists that can maintain a robust communication link when the transceiver is subject to large and quick motions such as those caused by head and body motions on a VR-HMD.

Partial Computation at HMD. As described in Section 3 GPUs will only be able to compute at an order of magnitude lesser frame rate than the Life-Like requirement of 1800Hz and consequently, the remaining frames will be generated using computationally cheap approximation *image-based rendering* (IBR) [29] techniques such as timewarp and spacewarp [5]. One way to reduce the bandwidth dramatically is transmit only the rendered frames and let a mobile GPU in the HMD generate the remaining frames through local render [13] or IBR [36]. This can reduce the bandwidth requirement by at-least an order of magnitude to about 270Gbps. While this is a significant improvement, this is still much higher than what any wireless standard in the horizon can hope to provide. Further there could be more sophisticated schemes to partition computation between the HMD and the rendering machine that allows fewer frames to be transmitted, thus alleviating the bandwidth bottleneck.

Lossy Compression and Latency Constraints. State-of-the-art lossy compression schemes like H.265 or VP9 can provide an average compression ratio of 200:1 - 1000:1 [40], but may provide much

---

[1] Graphics performance depends on many more factors other than Flops, including number of shading units, bandwidth, amount of memory and scene complexity. We used Flops because we consider it is a simple yet effective predictor of GPU growth.

higher compression at very high framerates where temporal locality is stronger. However, these schemes were designed for displays that are located farther from the human eye than an HMD. When used on a stereo VR display that is located close to the human eye, they can cause discomfort, especially when one or more visible artifacts are present in one eye and not in the other, inducing double vision, as we noticed during studies for previous VR systems we built [7, 12]. There is no user study to date that we know of that studies the effect of various compression levels on VR.

Latency constraints : In response to humans actions such as head movements, the VR system must render frames and transmit to the HMD. VR and neuroscience experts have found through through user studies that a latency greater than 20ms causes motion sickness and discomfort, and have projected that it may be necessary to reduce it to 15ms or even 7ms to fully eliminate them [2]. Thus, this end-to-end latency of up to 20ms must include various latencies such as, HMD tracking latency [28], transmission latency of new HMD pose to the rendering machine, rendering computation latency on the GPU, latency of compression of the video frames, latency of transmission of the compressed frames to HMD, decompression latency at HMD and finally display latency of the frames.

In a typical VR-HMD today, at a refresh rate of 90Hz, a newly rendered and transmitted frame will have to wait up to 11ms before it is displayed. Further, rendering frames, having a variable cost, often takes close to the time given by the refresh rate (in this case 11 ms). Including other latencies, the end-to-end latency often exceeds the prescribed 20ms. As a result, HMDs employ IBR techniques to hide latency [5] and minimize motion sickness. Unfortunately, these techniques are only effective to hide small amounts of latency, producing increasingly distracting artifacts as latency grows. This also means that even hardware-accelerated compression such as H.265 or VP9 may not be suitable given their relatively high compression and decompression latencies.

As the refresh rates of displays increase to 1800Hz and computational capacity of GPUs increase to 180Hz, both rendering and display latencies will reduce, leaving more time (up to 15ms) for compression and decompression. This opens up the possibility for lossy compression schemes that provide near Life-Like quality within the latency constrains of VR while reducing the bandwidth requirements for transmitting frames enough to meet the capabilities of wireless radios. For example, a compression scheme that provides a lossy compression of 10:1, coupled with partial computation at HMD can reduce the bandwidth requirement to 27Gbps which will be accommodated by the upcoming 802.11ay standard.

Standalone Headsets. Given that mobile GPU performance lags behind desktop GPUs by a factor of $6 - 10\times$, it is not immediately clear, whether Life-Like standalone VR-HMDs will be feasible by the time displays are. In this section we consider power consumption, and ask the question, “will power consumption be a bottleneck if mobile GPUs are powerful enough to render Life-Like VR?”

For a typical VR-HMD, the battery should last long enough (at least 3 hours to be able to experience full-length VR movies, gaming sessions [18], but perhaps more as line of business applications in VR become common). The energy density of a Li-Ion battery has been doubling once every 20 years [11] and is unlikely to increase significantly in the next few decade unless there is a disruptive breakthrough. Table 4 depicts the improvement in power efficiency of GPUs with time. As seen from Table 4, in the last 6 years GPU power efficiency in GFlops/Watt has increased by a factor of 8.5×. However, in the same time as seen from Table 3, the computational power increased by a factor of 6×. This indicates that GPUs power efficiency is increasing faster than their computational power, growth difference that may grow more pronounced as it becomes harder to shrink the fabrication process of GPUs. Thus, we believe that it is unlikely that GPUs power consumption will be a bottleneck for mobile GPUs.

Table 4: Trends in Gigaflops per watt

Date	GFlops/watt	Device
2010-03	5.4	GTX480
2010-10	6.5	GTX580
2012-12	15.8	GTX680
2013-05	15.9	GTX780
2014-07	27.9	GTX980
2016-05	45.7	GTX1080
2017-12	55.2	Titan V

Summary We now summarize our findings in this section,

- Wireless transmission of loss-less Life-Like VR frames with foveation will require about 2.7Tbps - no wireless standard in the horizon will be able accommodate this data rate.
- Partial rendering at the HMD coupled with the invention of “near-Life-Like” lossy compression techniques have the potential of alleviating the bandwidth bottleneck and making it accessible to standards such as 802.11ay.
- Since GPU power efficiency is growing faster than their computational capacity, power consumption is unlikely to become the key bottleneck for standalone VR-HMDs.

5 RESEARCH CHALLENGES REALIZING LIFE-LIKE UNTETHERED VR

Based on our study, we see two key bottlenecks in realizing Life-Like untethered VR. First GPUs will fall short of rendering computational capacity. Second, wireless bandwidth will be insufficient to transmit Life-Like rendered VR frames to the HMD. In this section we outline some of the research directions to overcome these bottlenecks.

User study based foveation research. Foveation is a relatively new concept and not yet fully understood, especially at high resolutions, framerates, and FOVs. Research to understanding foveation, guided by user studies, will help answer the crucial question of “How can foveation be effectively used to reduce computational and wireless transfer overheads for Life-Like VR?”

Novel computation sharing mechanisms between HMD and rendering machine. We believe that significant amount of research is required to develop novel techniques where the HMD can share the burden on rendering frames with the rendering machine to not only help make the VR experience Life-Like but also help in reducing the bandwidth required for wireless transfer.

Novel near Life-Like compression techniques ratified by user studies. There is very little understanding of the effects of lossy compression techniques on VR users. We believe that there is a significant open research questions in terms of designing compression techniques that can provide significant compression gains while preserving Life-Like quality of the rendered frames.

High bandwidth Wireless technologies Given that even the upcoming wireless standards such as 802.11ay fall an order of magnitude shy of the bandwidth required to transfer Life-Like VR, the need of the hour is to invent new wireless technologies that can accommodate these high data rates. This might involve developing FSO based wireless communication techniques or novel wireless communication over hitherto unexplored higher frequency spectrum that allow for higher data rates.

Due the development and commercialization of high bandwidth FSO transceivers, and the availability of more spectrum for communication is infrared spectrum compared to radio waves spectrum, we believe that it will be an interesting and useful research direction to develop FSO based communication systems for mobile transceivers like virtual reality headsets.

User study based high framerate low-persistence perception research. Judder and low persistence at high framerates are still not well understood. As the framerate of displays increases, a key question to answer would be "What combination of frame rates and persistence will be sufficient to achieve uncompromising Life-Like VR?" This question can be answered through careful user studies and the development of experimentation techniques.

6 CONCLUSION

In this paper, guided by aspects of human vision we determined what would be the specifications of a VR headset that is indistinguishable from physical reality – Life-Like VR. We then examined the trends in various technologies to determine the bottlenecks in the path to Life-Like VR. We find GPU computational capacity to be the key bottleneck. We then considered the possibility of making VR headsets cable-free and find that none of the existing standards are poised to accommodate the extremely high bandwidth of transferring Life-Like rendered VR frames. Finally, we present open research challenges in the path to realizing Life-Like VR.

REFERENCES

[1] Abari, O., Bharadia, D., Duffield, A., and Katabi, D. Enabling high-quality untethered virtual reality. In *NSDI* (2017).

[2] Abrash, M. Latency, the sine qua non of ar and vr. http://blogs.valvesoftware.com/abrash/latency-the-sine-qua-non-of-ar-and-vr/.

[3] Abrash, M. Down the vr rabbit hole: Fixing judder. http://blogs.valvesoftware.com/abrash/down-the-vr-rabbit-hole-fixing-judder/, July 2013.

[4] Andersen, S. R. The history of the ophthalmological society of copenhagen 1900–50. *Acta Ophthalmologica 80*, s234 (2002), 6–17.

[5] Beeler, D., Hutchins, E., and Pedriana, P. Asynchronous spacewarp. https://developer.oculus.com/blog/asynchronous-spacewarp/.

[6] Boitard, R., et al. Evaluation of color encodings for high dynamic range pixels. In *Proceedings Human Vision and Electronic Imaging* (2015).

[7] Boos, K., Chu, D., and Cuervo, E. Flashback: Immersive virtual reality on mobile devices via rendering memoization. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services* (2016), ACM, pp. 291–304.

[8] Britannica. Sensory reception: Human vision: Structure and function of the human eye, 1987.

[9] Carney, T., and Klien, S. Resolution acuity is better than vernier acuity. *Vision Research 37*, 5 (1997), 525–539.

[10] Cuervo, E. Beyond reality: Head-mounted displays for mobile systems researchers. *GetMobile: Mobile Computing and Communications 21*, 2 (2017), 9–15.

[11] Cuervo, E., Balasubramanian, A., Cho, D.-k., Wolman, A., Saroiu, S., Chandra, R., and Bahl, P. Maui: Making smartphones last longer with code offload. In *MobiSys 2010* (2010).

[12] Cuervo, E., and Chu, D. Poster: Mobile virtual reality for head-mounted displays with interactive streaming video and likelihood-based foveation. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services Companion* (2016), ACM, pp. 130–130.

[13] Cuervo, E., et al. Kahawai: High-quality mobile gaming using gpu offload. In *MobiSys* (May 2015).

[14] Cui, Y., Xiao, S., Wang, X., Yang, Z., Yan, S., Zhu, C., Li, X.-Y., and Ge, N. Diamond: Nesting the data center network with wireless rings in 3-d space. *IEEE/ACM Transactions on Networking* (2017).

[15] Curran, M., Rahman, M. S., Gupta, H., Zheng, K., Longtin, J., Das, S. R., and Mohamed, T. Fsonet: A wireless backhaul for multi-gigabit picocells using steerable free space optics. In *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking* (New York, NY, USA, 2017), MobiCom '17, ACM, pp. 154–166.

[16] EIZO. Focus on the foris fg2421. http://gaming.eizo.com/news/the-gaming-monitor-weve-all-been-waiting-for/, June 2014.

[17] Fove-VR. Fove 0 eye tracking vr devkit. https://www.getfove.com/.

[18] GiantBomb. How long is your average gaming session? https://www.giantbomb.com/forums/general-discussion-30/how-long-is-your-average-gaming-session-1438129/.

[19] Guenter, B., Finch, M., Drucker, S., Tan, D., and Snyder, J. Foveated 3d graphics. *ACM Trans. Graph. 31*, 6 (Nov. 2012), 164:1–164:10.

[20] Hamedazimi, et al. Firefly: A reconfigurable wireless data center fabric using free-space optics. In *ACM SIGCOMM Computer Communication Review* (2014), vol. 44, ACM, pp. 319–330.

[21] Hansen, C. J. Wigig: Multi-gigabit wireless communications in the 60 ghz band. *IEEE Wireless Communications 18*, 6 (2011).

[22] Hruska, J. Lg new 440 ppi display is way too much of a good thing. https://www.extremetech.com/computing/130051-lgs-new-440-ppi-display-is-way-too-much-of-a-good-thing, May 2012.

[23] HTC. Htc vive. https://www.htcvive.com/us/, Apr. 2016.

[24] HTC / Valve. Vive pro. https://www.vive.com/us/product/vive-pro/.

[25] IEEE. International roadmap for devices and systems 2016, more moore white paper. https://irds.ieee.org/images/files/pdf/2016_MM.pdf.

[26] Lang, B. Nvidia demonstrates experimental zero latency display running at 1,700hz. https://www.roadtovr.com/nvidia-demonstrates-experimental-zero-latency-display-running-at-17000hz/.

[27] Larsen, R. Review: Samsung 2233rz: First 120hz and 3d-monitor. http://www.flatpanelshd.com/review.php?subaction=showfull&id=1239184512, May 2009.

[28] LaValle, S. M., Yershova, A., Katsev, M., and Antonov, M. Head tracking for the oculus rift. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on* (2014), IEEE, pp. 187–194.

[29] Mark, W. R., McMillan, L., and Bishop, G. Post-rendering 3d warping. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics* (New York, NY, USA, 1997), I3D '97, ACM, pp. 7–ff.

[30] Moore, S. T., Hirasaki, E., Raphan, T., and Cohen, B. The human vestibulo-ocular reflex during linear locomotion. *Annals of the New York Academy of Sciences 942*, 1 (2001), 139–147.

[31] Patney, A., Salvi, M., Kim, J., Kaplanyan, A., Wyman, C., Benty, N., Luebke, D., and Lefohn, A. Towards foveated rendering for gaze-tracked virtual reality. *ACM Transactions on Graphics (TOG) 35*, 6 (2016), 179.

[32] PimaxTechnology. Pimax 8k. https://www.pimaxvr.com/en/8k/.

[33] Purves, D., Augustine, G., Fitzpatrick, D., et al. Neuroscience. 2nd edition. https://www.ncbi.nlm.nih.gov/books/NBK10991/, 2001. Types of Eye Movements and Their Functions.

[34] Ranovus. Ranovus announces availability of worlds first 200g cfp2 direct detect optical transceiver to enable 38.4 terabits per data center interconnect rack.

[35] Reed, B. Lg new 440 ppi display is way too much of a good thing. http://bgr.com/2013/08/21/lg-display-538-ppi/, May 2012.

[36] Reinert, B., Kopf, J., Ritschel, T., Cuervo, E., Chu, D., and Seidel, H.-P. Proxy-guided image-based rendering for mobile devices. *Computer Graphics Forum 35*, 7 (2016), 353–362.

[37] Sony. Xperia z5 premium. https://www.sonymobile.com/global-en/products/phones/xperia-z5-premium/.

[38] StarVR. Starvr. https://www.starvr.com/, July 2013.

[39] Strasburger, H., Rentschler, I., and Jüttner, M. Peripheral vision and pattern recognition: A review. *Journal of vision 11*, 5 (2011), 13–13.

[40] Sullivan, G., and Ohm, J.-R. Meeting report of the 13th meeting of the joint collaborative team on video coding (jct-vc). http://phenix.it-sudparis.eu/jct/doc_end_user/current_document.php?id=7746.

[41] TPCast. Tpcast vive. https://www.tpcastvr.com/.

[42] https://www.oculus.com/en-us/rift/. Oculus rift, Mar. 2016.

[43] Visual Functions Committee. *Visual acuity measurement standard.* May 2015.

[44] Walls, F., and MacInnis, S. Vesa display stream compression. http://www.vesa.org/wp-content/uploads/2014/04/VESA_DSC-ETP200.pdf.

[45] WirelessHD. Wirelesshd. https://www.wirelesshd.org/.

[46] Yokoyama, K., et al. Ultra-high-resolution 1058-ppi oled displays with 2.78-in size using caac-igzo fets with tandem oled device and single oled device. *Journal of the Society for Information Display 24*, 3 (2016), 159–167.

[47] Zisworks. Zisworks zws 480hz: Engineering sample purchase. http://www.zisworks.com/shop.html, 2017 Aug.