5G Network Resources Requirements for Mobile Immersive Digital Environments

Experimental Validation of Mobile Virtual Reality Network Requirements in Unity 3D

Danaisy Prado-Álvarez, David Garcia-Roger, Jose F. Monserrat
iTEAM Research Institute
Universitat Politècnica de València, Camino de Vera, s/n
46022 Valencia, Spain
email: dapraal@teleco.upv.es, dagarro@iteam.upv.es, jomondel@iteam.upv.es

***Abstract*—The accelerated increase in the adoption of immersive digital technologies like virtual reality and 360-degree video escalates the pressure on mobile cellular networks. Its higher bandwidth demands and minimum latencies are crucial for enjoying the contents with a satisfactory quality and comfort. This paper describes a study that estimates the minimum critical bandwidth, and latency prerequisites, as well as video resolutions and bitrate needed in a mobile network so as to support immersive applications with a specific subjective quality of experience, using the consumer-ready hardware platforms Oculus Rift and Samsung Gear VR. One of the main conclusions is that the strict requirements of around 20 milliseconds of minimum latency highlights the important challenge that must be addressed by future 5G mobile cellular networks, but this is far from some target values discussed in the literature.**

Keywords-immersive technology; virtual reality; 5G networks; latency; quality of experience.

I. INTRODUCTION

In recent years, the improved technological development has made the dream of immersive digital environments like virtual reality (VR) and 360-degree video come true. In fact, VR headsets are now a tangible, consumer-ready visualization platform option available for PCs (Oculus Rift, HTC Vive), video game consoles (Sony Playstation VR) and smartphones (Samsung Gear VR, Google Cardboard, Google Daydream). VR is a computationally created environment that mimics reality, where the user enjoys an immersive experience and interacts with the virtual world that surrounds her/him as if it were real.

VR has three key characteristics: i) immersion, in which the user is only able to perceive the stimuli generated by the virtual environment; ii) interaction, in which the user is able to interact with the virtual environment in real time; iii) responsiveness, in which the user is able to sense simulated realities reacting quickly and positively. VR as a concept initially popularized by video games has gained relevance in sectors, such as medicine, archeology, artistic creation, military training, flight simulation, or virtual offices, among others.

On the go, wireless, cloud-powered VR via mobile cellular networks could be a future emerging trend, where the VR image is transmitted to the viewer from a remote entity in the cloud. In such scenarios, it is imperative to guarantee a superb user experience. Bandwidth demands in VR are remarkable because of the responsiveness trait. With tens of cameras capturing a scene, not even dynamic caching and multicast would be able to reduce the load. Because users should be able to select their individual point of view dynamically delivering content to thousands from a single feed is an unfeasible approach. Therefore, the feed from all of the cameras needs to be available almost instantly and the cellular network might be overwhelmed.

Together with bandwidth bottlenecks, coding artifacts, resolution degradation, and latency spikes are enough to trigger motion sickness. While VR technologies can handle slight impairments, many users will feel motion sick if they spend too much time in the headset perceiving a degraded experience [1]. With VR-induced motion sickness, the effect starts subtle, but is cumulative. What begins as slight discomfort or even a feeling of unease will progress into full-blown nausea. It is not something that the users can push through or become acclimated to. Once it starts, their discomfort will not end until they remove the headset.

Therefore, in order to guarantee a good user experience in the abovementioned scenarios, it is vital to consider which requisites need to be fulfilled by the oncoming the future fifth-generation (5G) mobile cellular network, and the quantitative values for such parameters.

As a subject of research, cloud-powered VR is pushing the limits of current mobile cellular networks, and there is the need of specifying measurable bounds for the magnitude of enhanced capabilities of 5G that are required for VR to reach their full potential. 5G is expected to deliver a high speed network that will allow 10 Gb/s transmission speeds per device, as well as below 1 ms, both specs, 10 to 100 times better than current 4G/LTE networks [2]. Anticipating the impact of such technology on VR, well-known surveys on mobile and wireless technologies for VR like [3] on low bandwidth and high round-trip times in 3G networks have given way to the significant plethora of reviews of current trends towards 5G performed very recently. For example, the authors of [4] review the state of the art in virtual and augmented reality communications and highlights efforts for an operative, universal 5G network in niche markets like telepresence, education, healthcare, streaming media, and haptics. In addition, [5] describes the 5G low-latency applications business case and the potential market benefits for operators on other vertical industries such as automotive, public transport, infrastructure, entertainment, and manufacturing.

The purpose of this paper is describing a study that estimates the minimum critical bandwidth and latency prerequisites, as well as video resolutions and bitrate needed in a 5G mobile network to support immersive applications with a specific subjective Quality of Experience (QoE). This is made by using the consumer-ready hardware platforms Oculus Rift and Samsung Gear VR, with the objective of statistically ascertaining the weight that network induced effects can have on an adequate VR experience under such conditions of latency and bandwidth.

The remaining sections of the paper are structured as follows. Section II describes the relevant features of the immersive technologies addressed in the present work. Section III explains the set of specific performance goals that are evaluated. Section IV gives an overview of the experimental virtual reality testbed used. Section V presents the results of the performance assessment. Finally, Section VII draws the main conclusions of this work.

II. IMMERSIVE TECHNOLOGIES AND 5G

On the go, wireless, cloud-powered VR via mobile cellular networks has only just started to be developed as a concept. Conceived as a solution to the substantial local storage demands of traditional VR applications, cloud-powered VR streams the contents from storage resources located in a cloud, and playbacks it directly in the headset without the requirement of a previous download.

Prominent research initiatives on 5G like METIS-II European project [6] mention cloud-powered applications of telepresence like VR as a specific use case. Bidirectional flows with sustained transmission rates of 1 Gb/s, in addition to synchronization flows at 5 Gb/s and packet loss rates less than 5% are discussed. In this sense, there is a current application, 360-degree video, that although may be experienced in a VR headset, it is not technically VR. 360-degree video is recorded from every direction and users may observe all aspects of the virtual world that surrounds her, but cannot interact with it. It is expected that 360-degree video may escalate up to bandwidth demands similar to VR, if resolutions beyond 4K are considered in the future, but also research is being done at ways of minimizing the amount of data through user orientation prediction and transmitting only a given subset of cameras [7].

With respect to latencies, 60 ms has been mentioned as the absolute upper-bound, but uncomfortable delays are reported with technologies such as Oculus Rift when latency is larger than 40 ms [8][4]. Stringent requisites for 5G in 0, state that average packet-to-packet latency should less than 10 ms, and further extremely low latencies of 8 ms are desired in [4], or 15 ms to 7 ms application to application delay, i.e., action to reaction, is the threshold to provide a smooth action-reaction experience in [5]. Other studies claim that latencies less than 20 ms are imperceptible [9], and still several related works provide different estimations [1][10][11] With respect to latencies. Network latencies may be solved through processing at the edge or the user's device [12][13]. However, it must be taken into account that latencies in VR arise not only from the restrictions imposed by current mobile network technology but also from several sources, e.g., CPU-intensive tasks inherent to VR like real time capture and video stitching, and also sensing time (although recently reduced to an amount that is imperceptible by humans [4]), USB data speed delays, data crosschecking, game code execution time, frame rendering delay, video output delay, pixel switching time of a LCD, and frame buffering. Furthermore, individual VR experiences have unidirectional latency, but in games and telepresence latency is bidirectional, actually doubling the network infrastructure requirements [4].

Bandwidth and latency limitations still prevent current networks from achieving smooth telepresence and collaborative virtual and augmented reality applications [4]. Specifically, 4G/LTE is not able to support services requiring big data sizes, where the transmission of the gigabytes of data comprising a 3D model at a reasonable cost is a necessity. 4G/LTE networks also exhibit typical latencies of 50 ms [14] and are not able to satisfy instantaneous access to future cloud services.

Still, some argue that even 5G could face latency and bandwidth challenges in more remote or crowded locations [4], or 5G is unlikely to deliver the resolution and responsiveness requirements of some high-end applications of VR [12]. A detailed list of potential solutions in RAN and core network is [16], including software defined network (SDN), network function virtualization (NFV), caching, and mobile edge computing (MEC), will allow 5G networks to meet latency and other 5G requirements. Consequently, it is essential to evaluate the QoE of available VR applications over current mobile networks to find out if the limiting factor is the network or video processing in real-time.

There are previous works that address network latency in immersive applications, but either only non-VR cloud-gaming [16], or actual VR but without taking into account the network effects [17]. However, so far, to the best of the authors' knowledge, there has been no exhaustive study of a cloud-VR network scenario in 5G, ascertaining the value of the limiting latency for a sample of population.

The purpose of the work described in this paper is to create a VR testbed assisted by the 3D game engine Unity 3D, to experimentally determine the critical values of bandwidth and latency for a mobile cellular network supporting immersive applications, expressed with respect to the QoE perceived by its users, with the main goal of statistically profile the ratio of users that will enjoy a satisfactory VR experience under such conditions of latency and bandwidth.

III. VIRTUAL REALITY TESTBED

The VR testbed is shown in Figure 1. It consists of a Samsung Gear VR headset and a Samsung Galaxy S7 (4 GB of RAM, Android Nougat 7.0, and a screen resolution of 2560x1440 pixels at 60 frames per second). There is also an Oculus Rift consumer version (headset with resolution 2160x1200 pixels, 1080x1200 per eye at 90 fps, sensor, Xbox controller and Oculus Remote controller), and workstation with the following specs: graphics card GeForce GTX 1070 G1 Gaming 8GB GDDR5, Intel i7-6700K 4.0Ghz processor, 32 GB of RAM, Windows 10 Enterprise.

Figure 1. Virtual reality testbed.

The VR was used to carry two experiments on latency and display resolution QoE, experiment I and experiment II, respectively. Experiment I consists of evaluating the QoE perceived by a random group of persons enjoying a VR experience with respect to different values of latency. Experiment II uses the same population sample and several 360-degree test videos with known bitrate and resolutions to estimate the minimum bandwidth required for the satisfactory transmission of the 360-degree video.

During experiment I, the volunteer used the Oculus Rift headset to visualize a 3D rendering of Madrid with increasing degrees of latency. The planning of Experiment I made use of the Unity 3D game engine and required building on the code of the Madrid Grid visualization tool of METIS-II [18]. This platform simulates the shortcomings of 4G and introduces candidate 5G solutions for improved network traffic in a typical urban environment.

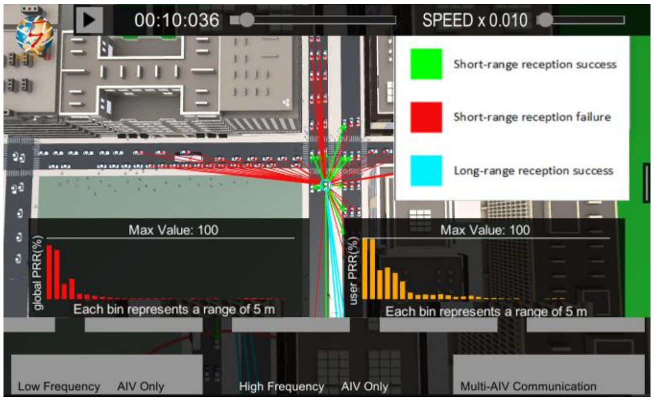

Figure 2. METIS-II Madrid Grid visualization platform.

The Madrid Grid visualization platform interface, is shown in Figure 2. which depicts a scenario with multi air interface variant communications for V2V. The interface includes the street traffic simulation and an overlay layer that provides information to the and allows interaction through a flexible, but traditional, mouse-based, graphical user interface (GUI).

The visualization platform required an extension to so as to enable VR output compatible with the Oculus Rift headset. However, the traditional GUI could not be easily translated in to a VR-friendly paradigm and therefore it was necessary to create also a proper VR GUI as well as the definition of a set of ways to interact with the GUI through the Oculus Remote controller instead of the mouse. The status bar was relocated as a head-up display (HUD) inside world-space and the lack of mouse in VR was solved with a fixed reticle assisted by the buttons of the Oculus Remote as shown in Figure 3.

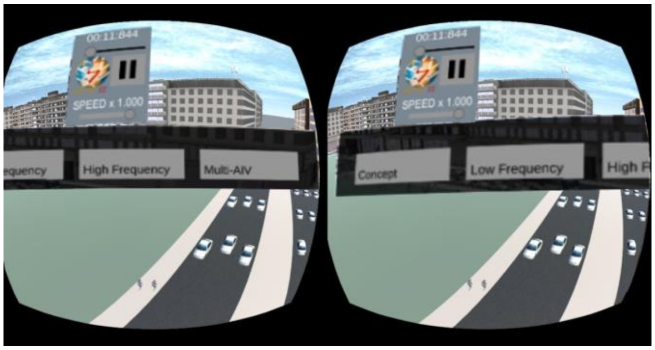

Figure 3. Visualization platform VR GUI reticle and head-up display.

Experiment I also required the configuration of the Oculus Debug Tool to force a range of different latency values. This is a debugging tool provided by Oculus that allows the study of real-time performance of a VR experience in runtime. By using the performance HUD, and selecting the option *latency timing*, the performance in ms of several contributions to delay may be observed, as shown in Figure 4. Specifically, the parameter *App Tracking to Mid-Photons* summarizes all the contributions to latency, and it may be impacted by changing the *Pixels per display pixel override* parameter in the Oculus Debug Tool.

Specifically, the parameter App Tracking to Mid-Photons summarizes all the contributions to latency, and it may be impacted by changing the Pixels per display pixel override parameter in the Oculus Debug Tool as shown in Figure 5.

Increasing the pixel density from 1.0 to 2.0 improves in the image resolution but after 2.0 there is no discernable effect and further increasing comes at a cost: latency increase, which may be handily used to adjust the latency experienced during the experiment as shown in Table I.

 ISBN: 978-1-61208-596-8

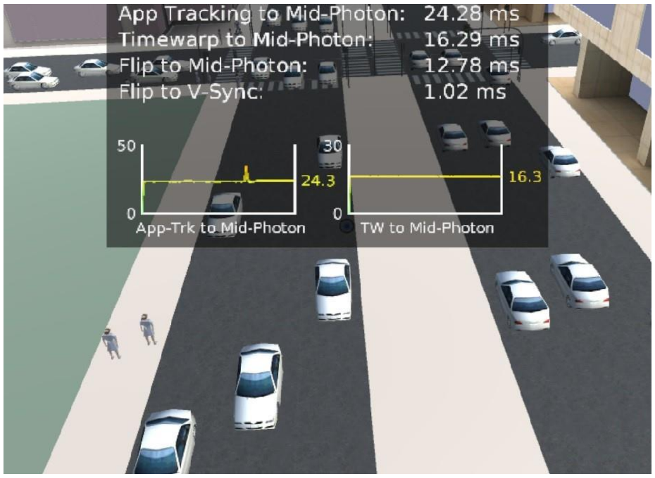

Figure 4. Oculus Debug Tool, displaying latency contributions.

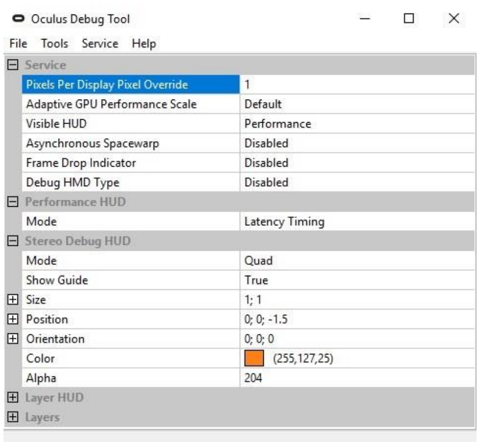

Figure 5. Oculus Debug Tool, Pixels per display override.

TABLE I. LATENCIES PRODUCED BY SEVERAL PIXEL DENSITIES

Pixel density	Associated latency
0	22
1	22
2	25
3	30
4	40
5	50
5.2	60

During experiment II, the volunteer used the Samsung Gear VR to watch a 360-degree video depicting a location surrounded by African wild elephants. The preparation of Experiment II required the selection and download of a 4K 360-degree video of excellent quality from Youtube [19]. The video was re-encoded with the FFmpeg tool [20] to generate 2K, 1080p, 720p, 480p, 360p, 240p, and 144p versions of the same video to deduce estimations of the required streaming bandwidth required to transmit them, as shown in Table II, assuming that 8 camera rigs have been used for the recording of the video. The FFmpeg tool *ffprobe* was used to analyze and may be also used to live-stream the videos to the Samsung Galaxy S7 smartphone inside the Samsung Gear VR headset.

IV. PERFORMACE ASSESSMENT

The experiments were performed on 25 volunteers, a typical value similar to previous studies [18][21]. Participants ranged from 25 to 40 years old, and were initially surveyed about their previous experiences with VR headsets just in case that would bias their opinion about the experience; 20% of them had tested either Samsung Gear VR or Oculus Rift.

A. *Experiment I*

During experiment I on latency, the Oculus Rift headset was used. The experimental results are shown in Figure 6. It should be noted that QoE rating by users was worsened if they were subjected to degraded quality for too much time. Evaluation criteria are linked to the tolerance of the subjects. The latency achieved with the best configuration is 22 ms, and is used as the reference value. The rating method is based in degradation category rating (DCR), which evaluates the QoE according to a metric mean opinion score (MOS) [22]. As shown, 22 ms of latency seems to be imperceptible for 76% of the subjects; it seems reasonable that reducing latency by some seconds would achieve almost 100% of agreement on its imperceptibility. At 25 ms the noticeable effect of latency is not only perceptible for someone but unpleasant for 32% of the subjects. At 30 ms latency is clearly perceptible for 84% of the subjects. As latency is increased, subjects agree on the distressing nature of the experience, and at 60 ms 80% of the surveyed persons expressed their intention to stop the experiment.

TABLE II. 360-DEGREE VIDEO DETAILS

Video ID	Resolution	Bitrate (kb/s)	Bandwidth (Mb/s)
1	3840x2048	12672	101.7
2	2560x1440	8217	65.7
3	1920x1080	3458	27.7
4	1280x720	1572	12.6
5	854x480	767	6.2
6	640x360	459	3.7
7	426x240	296	2.4
8	256x144	161	1.3

The percentage of surveyed persons that report motion sickness for different latencies is shown in Figure 7.

As shown, the percentage of subjects that report motion sickness with respect to latency follows an exponential law. Note that subjects reported 60 ms of latency as unbearable, because they were visually exhausted and dizzied.

B. *Experiment II*

During experiment II on display resolution, the Samsung Gear VR headset with a Samsung Galaxy S7 smartphone was used and eight 360-degree videos coded at diminishing resolutions were displayed in succession. The rating method is based in absolute category rating (ACR), which evaluates with simple stimuli the QoE according also to a MOS metric

 ISBN: 978-1-61208-596-8

[22]. The QoE rating of each of the eight videos is shown in Figure 8. The subjects had to stand while evaluating all videos with audio enabled except the one at 2K which was watched while sitting on a chair with audio at minimum. Note that although the video coded at 4K has the best resolution the resolution of the Samsung S7 used in the experiment is limited to 2K.

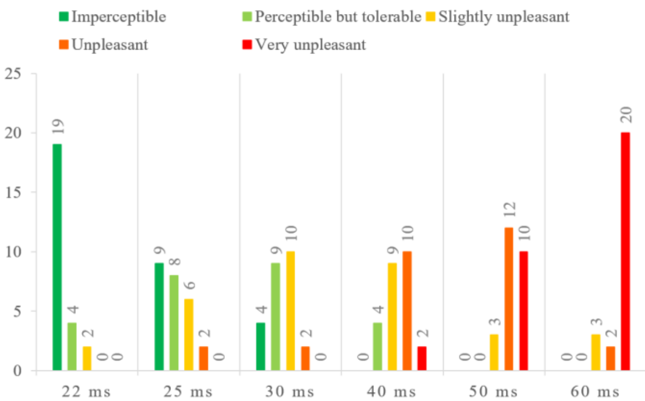

Figure 6. Experiment I on latency: QoE results.

As shown, reducing the display resolution negatively affects QoE. Even at 4K, 52% of the subjects did not considered the video quality to be excellent, and 36% of them were able to detect the pixelization on the smartphone screen. At 1080p, 64% of the users expressed their disagreement (average, poor) with the quality of the visualized video. Compared to conventional applications, VR requires higher resolution in order to perceive a similar QoE mainly because of the magnification optics lenses that bend the light inside the headset in a way that helps the user to see clearly.

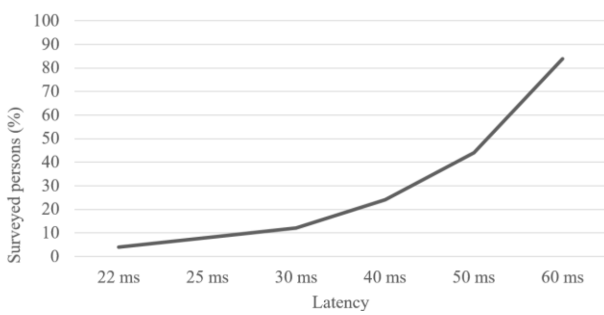

Figure 7. Experiment I on latency: motion sickness results.

The percentage of surveyed persons that report motion sickness for different display resolutions is shown in Figure 9. As shown, as resolution is reduced, the extra visual effort required to distinguish the features in the video yields motion sickness. From 720p to lower resolutions, motion sickness increases logarithmically. At 480p, 52% of the subjects report vertigo, and 60% at 144p. Subjects that did not experienced dizziness reported that were forcing themselves to tell apart the objects in the video.

The percentage of surveyed persons that report the experience as immersive is shown in Figure 10. The surveyed persons reported that as resolution decreased, the scene depicted was less immersive because of the degrading visual depth perception, and turned into an unpleasant experience. Also of note is the fact that although the view shown was all a straight capture from a GoPro Omni camera, even the 2K resolution induced certain participants to wonder whether the elements where real or computer-generated imagery. In addition, the video at 2K, received lower rates than the 2K resized version of the 4K video even though both had the same resolution. A reason could be that the lack of audio and the sitting position worsens the immersive experience, as notified by most of the subjects. This hypothesis is confirmed because the 1080p video (again with audio) produces better perception for the viewers than the 2K version. When resolution is 480p video quality is unacceptable and 92% of the participants thought that the experience was not immersive. Note finally that in order to transmit 360-degree 4K videos, it would be necessary a bandwidth of roughly 11 Mbps, but it is foreseeable that for future applications, even 8K per eye would not be enough for a perfect VR experience.

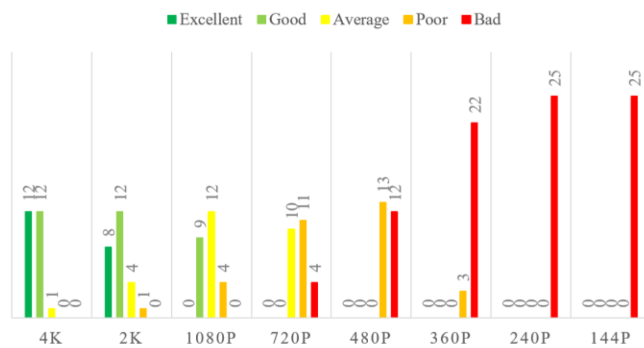

Figure 8. Experiment II on display resolution: QoE results.

Experiment I on latency shows that, by default and without taking into account communication network effects, the baseline latency of a consumer-ready headset like Oculus Rift is 22 ms. Since up to 25 ms is a tolerable latency, there is only 3 ms of buffer for network induced latency for a cloud-powered VR experience. Although reduced latencies are expected from headsets yet to come, current LTE networks have a latency of 50 ms, and therefore reducing latency by an order of magnitude in upcoming 5G cellular networks is also of utmost importance.

Experiment II demonstrates that even 360-degree videos at the 2K native resolution of smartphones do not satisfy 50% of the participants. Consequently, to enhance the VR experience, future devices should have screens with higher resolutions. In this sense, Samsung has announced a 11K smartphone in 2018; with such higher resolutions, network capacity demands for cloud-based VR are expected to climb sharply.

V. CONCLUSIONS

The higher bandwidth demands and minimum latencies of wireless, cloud-powered VR will escalate the pressure on future 5G mobile cellular networks. This paper provides estimates of the minimum critical bandwidth (4K great quality lossy compression at 80 Mb/s), and latency prerequisites (less than 25 ms) needed in a mobile network to support immersive applications with a specific subjective quality of experience Potential solutions that will allow 5G networks to meet the requirements in RAN and core network, include SDN, NFV, caching, and MEC.

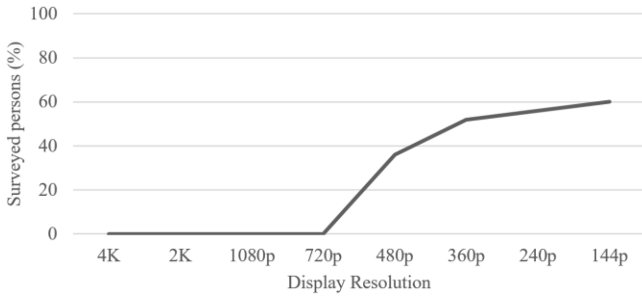

Figure 9. Experiment II on display resolution: motion sickness results.

ACKNOWLEDGMENTS

Part of this work has been performed in the framework of the H2020 project METIS-II (H2020-ICT-2014-2, topic ICT-14-2014) co-funded by the European Commission. The authors would like to acknowledge the contributions of their colleagues. The views expressed are those of the authors and do not necessarily represent the project. Part of this work has been also supported by the Ministerio de Economía y Competitividad, Spain (TEC2014-60258-C2-1-R), by the European FEDER funds.

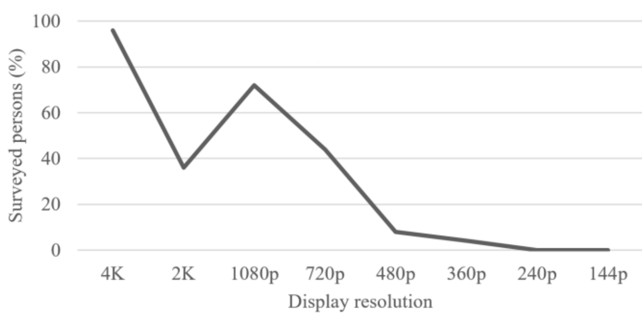

Figure 10. Experiment II on display resolution: degree of immersiveness.

REFERENCES

[1] M. Meehan, S. Razzaque, M. C. Whitton and F. P. Brooks, "Effect of latency on presence in stressful virtual environments," IEEE Virtual Reality, 2003 pp. 141-148.

[2] H. Tullberg et al., “The METIS 5G system concept: Meeting the 5G requirements”, IEEE Communications Magazine, vol. 54, no. 12, pp. 132-139, 2016.

[3] G. Papagiannakis, G. Singh, and N. Magnenat-Thalmann. “A survey of mobile and wireless technologies for augmented reality systems.” Comput. Animat. Virtual Worlds, 19(1), pp. 3-22, Feb. 2008.

[4] J. Orlosky, K. Kiyokawa, and H. Takemura, “Virtual and Augmented Reality on the 5G Highway”, Journal of Information Processing, vol 25. pp. 133-141, 2017.

[5] M. A. Lema et al., "Business Case and Technology Analysis for 5G Low Latency Applications," in IEEE Access, vol. 5, no. , pp. 5917-5935, 2017.

[6] S. E. El Ayoubi, F. Pujol, and M. Fallgren (ed.), METIS-II/D1.1 Refined scenarios and requirements, consolidated use cases, and qualitative techno-economic feasibility assessment. (2016).

[7] F. Qian, B. Han, L. Ji, and V. Gopalakrishnan, “Optimizing 360 video delivery over cellular networks”. In ACM MobiCom All Things Cellular Workshop, pp. 1-6, Oct. 2016.

[8] S. LaValle, “The latent power of prediction”, *Oculus Developers Blog*. 2013.

[9] J. Carmack, “Latency mitigation strategies,” *Twenty Milliseconds*, 2013.

[10] B. D. Adelstein, T. G. Lee, and S. R. Ellis, (2003, October). Head tracking latency in virtual environments: psychophysics and a model. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, No. 20, pp. 2083-2087). Sage CA: Los Angeles, CA: SAGE Publications.

[11] K. Mania, B. D. Adelstein, S. R. Ellis, and M. I. Hill, “Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity.” In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization* (pp. 39-47). ACM, August 2004.

[12] C. Westphal, “Challenges in networking to support augmented reality and virtual reality”, ICNC 2017, Silicon Valley, California, USA, 26-29 January, 2017.

[13] P. Mach, and Z. Becvar, "Mobile Edge Computing: A Survey on Architecture and Computation Offloading," in IEEE Communications Surveys & Tutorials, vol. 19, no. 3, pp. 1628-1656, thirdquarter 2017.

[14] 3GPP TR 36.913, “Requirements for further advancements for E-UTRA (LTE-Advanced)”.

[15] I. Parvez, A. Rahmati, I. Guvenc, A. I. Sarwat, and H. Dai, “A Survey on Low Latency Towards 5G: RAN, Core Network and Caching Solutions,” arXiv:1708.02562v1 Aug 2017.

[16] T. Kämäräinen, M. Siekkinen, A. Ylä-Jääski, W. Zhang, and Pan Hui, “A Measurement Study on Achieving Imperceptible Latency in Mobile Cloud Gaming”, *Proceedings of the 8th ACM on Multimedia Systems Conference*, MMSys'17, pp. 88-89, 2017.

[17] M. Nabiyouni, S. Scerbo Siroberto, D. A. Bowman, and T. Höllerer, “Relative Effects of Real-world and Virtual-World Latency on an Augmented Reality Training Task: An AR Simulation Experiment”, Frontiers in ICT, vol 3, pp 34, 2017.

[18] M. Maternia, and J. F. Monserrat (ed.), METIS-II/D2.1 Performance evaluation framework.

[19] IM360, “Surrounded by Wild Elephants in 4k 360” goo.gl/NdP1WN [retrieved: Sep. 2017]

[20] FFmpeg.org [retrieved: Sep. 2017]

[21] R. Schatz, A. Sackl, C. Timmerer, and B. Gardlo, “Towards subjective quality of experience assessment for omnidirectional video streaming,” QoMEX, Erfurt, 2017, pp. 1-6.

[22] Recommendation ITU-T P.913 Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment, 03/201