

Workgroup: MOPS

Internet-Draft: draft-ietf-mops-ar-use-case-10

Published: 13 March 2023

Intended Status: Informational

Expires: 14 September 2023

Authors: R. Krishna

A. Rahman

InterDigital Europe Limited Ericsson

## **Media Operations Use Case for an Extended Reality Application on Edge Computing Infrastructure**

### **Abstract**

This document explores the issues involved in the use of Edge Computing resources to operationalize media use cases that involve Extended Reality (XR) applications. In particular, we discuss those applications that run on devices having different form factors and need Edge computing resources to mitigate the effect of problems such as a need to support interactive communication requiring low latency, limited battery power, and heat dissipation from those devices. The intended audience for this document are network operators who are interested in providing edge computing resources to operationalize the requirements of such applications. We discuss the expected behavior of XR applications which can be used to manage the traffic. In addition, we discuss the service requirements of XR applications to be able to run on the network.

### **Status of This Memo**

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 14 September 2023.

### **Copyright Notice**

Copyright (c) 2023 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

- [1. Introduction](#)
- [2. Conventions used in this document](#)
- [3. Use Case](#)
  - [3.1. Processing of Scenes](#)
  - [3.2. Generation of Images](#)
- [4. Requirements](#)
- [5. AR Network Traffic](#)
  - [5.1. Traffic Workload](#)
  - [5.2. Traffic Performance Metrics](#)
- [6. Acknowledgements](#)
- [7. Informative References](#)
- [Authors' Addresses](#)

## 1. Introduction

Extended Reality (XR) is a term that includes Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR) [[XR](#)]. AR combines the real and virtual, is interactive and is aligned to the physical world of the user [[AUGMENTED\\_2](#)]. On the other hand, VR places the user inside a virtual environment generated by a computer [[AUGMENTED](#)]. MR merges the real and virtual world along a continuum that connects completely real environment at one end to a completely virtual environment at the other end. In this continuum, all combinations of the real and virtual are captured [[AUGMENTED](#)].

XR applications will bring several requirements for the network and the mobile devices running these applications. Some XR applications such as AR require a real-time processing of video streams to recognize specific objects. This is then used to overlay information on the video being displayed to the user. In addition XR applications such as AR and VR will also require generation of new video frames to be played to the user. Both the real-time processing of video streams and the generation of overlay information are computationally intensive tasks that generate heat [[DEV\\_HEAT\\_1](#)], [[DEV\\_HEAT\\_2](#)] and drain battery power [[BATT\\_DRAIN](#)] on the mobile device running the XR application. Consequently, in order to run applications with XR characteristics on mobile devices,

computationally intensive tasks need to be offloaded to resources provided by Edge Computing.

Edge Computing is an emerging paradigm where computing resources and storage are made available in close network proximity at the edge of the Internet to mobile devices and sensors [[EDGE\\_1](#)], [[EDGE\\_2](#)]. These edge computing devices use cloud technologies that enable them to support offloaded XR applications. In particular, the edge devices deploy cloud computing implementation techniques such as disaggregation (breaking vertically integrated systems into independent components with open interfaces using SDN), virtualization (being able to run multiple independent copies of those components such as SDN Controller apps, Virtual Network Functions on a common hardware platform) and commoditization (being able to elastically scale those virtual components across commodity hardware as the workload dictates) [[EDGE\\_3](#)]. Such techniques enable XR applications requiring low-latency and high bandwidth to be delivered by mini-clouds running on proximate edge devices

In this document, we discuss the issues involved when edge computing resources are offered by network operators to operationalize the requirements of XR applications running on devices with various form factors. Examples of such form factors include Head Mounted Displays (HMD) such as Optical-see through HMDs and video-see-through HMDs and Hand-held displays. Smart phones with video cameras and GPS are another example of such devices. These devices have limited battery capacity and dissipate heat when running. Besides as the user of these devices moves around as they run the XR application, the wireless latency and bandwidth available to the devices fluctuates and the communication link itself might fail. As a result algorithms such as those based on adaptive-bit-rate techniques that base their policy on heuristics or models of deployment perform sub-optimally in such dynamic environments[[ABR\\_1](#)]. In addition, network operators can expect that the parameters that characterize the expected behavior of XR applications are heavy-tailed. Such workloads require appropriate resource management policies to be used on the Edge. The service requirements of XR applications are also challenging when compared to the current video applications. In particular several QoE factors such as motion sickness are unique to XR applications and must be considered when operationalizing a network. We motivate these issues with a use-case that we present in the following sections.

## **2. Conventions used in this document**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

### 3. Use Case

We now describe a use case that involves an application with AR systems' characteristics. Consider a group of tourists who are being conducted in a tour around the historical site of the Tower of London. As they move around the site and within the historical buildings, they can watch and listen to historical scenes in 3D that are generated by the AR application and then overlaid by their AR headsets onto their real-world view. The headset then continuously updates their view as they move around.

The AR application first processes the scene that the walking tourist is watching in real-time and identifies objects that will be targeted for overlay of high resolution videos. It then generates high resolution 3D images of historical scenes related to the perspective of the tourist in real-time. These generated video images are then overlaid on the view of the real-world as seen by the tourist.

We now discuss this processing of scenes and generation of high resolution images in greater detail.

#### 3.1. Processing of Scenes

The task of processing a scene can be broken down into a pipeline of three consecutive subtasks namely tracking, followed by an acquisition of a model of the real world, and finally registration [[AUGMENTED](#)].

Tracking: This includes tracking of the three dimensional coordinates and six dimensional pose (coordinates and orientation) of objects in the real world [[AUGMENTED](#)]. The AR application that runs on the mobile device needs to track the pose of the user's head, eyes and the objects that are in view. This requires tracking natural features that are then used in the next stage of the pipeline.

Acquisition of a model of the real world: The tracked natural features are used to develop an annotated point cloud based model that is then stored in a database. To ensure that this database can be scaled up, techniques such as combining a client side simultaneous tracking and mapping and a server-side localization are used [[SLAM 1](#)], [[SLAM 2](#)], [[SLAM 3](#)], [[SLAM 4](#)]. Another model that can be built is based on polygon mesh and texture mapping technique. The polygon mesh encodes a 3D object's shape which is expressed as a collection of small flat surfaces that are polygons. In texture mapping, color patterns are mapped on to an object's surface. A third modelling technique uses a 2D lightfield that describes the intensity or color of the light rays arriving at a single point from

arbitrary directions. Assuming distant light sources, the single point is approximately valid for small scenes. For larger scenes, a 5D lightfield is used which encodes separate 2D lightfields for many 3D positions in space [[AUGMENTED](#)].

Registration: The coordinate systems, brightness, and color of virtual and real objects need to be aligned in a process called registration [[REG](#)]. Once the natural features are tracked as discussed above, virtual objects are geometrically aligned with those features by geometric registration. This is followed by resolving occlusion that can occur between virtual and the real objects [[OCCL\\_1](#)], [[OCCL\\_2](#)]. The AR application also applies photometric registration [[PHOTO\\_REG](#)] by aligning the brightness and color between the virtual and real objects. Additionally, algorithms that calculate global illumination of both the virtual and real objects [[GLB\\_ILLUM\\_1](#)], [[GLB\\_ILLUM\\_2](#)] are executed. Various algorithms to deal with artifacts generated by lens distortion [[LENS\\_DIST](#)], blur [[BLUR](#)], noise [[NOISE](#)] etc are also required.

### **3.2. Generation of Images**

The AR application must generate a high-quality video that has the properties described in the previous step and overlay the video on the AR device's display- a step called situated visualization. This entails dealing with registration errors that may arise, ensuring that there is no visual interference [[VIS\\_INTERFERE](#)], and finally maintaining temporal coherence by adapting to the movement of user's eyes and head.

## **4. Requirements**

The components of AR applications perform tasks such as real-time generation and processing of high-quality video content that are computationally intensive. As a result, on AR devices such as AR glasses excessive heat is generated by the chip-sets that are involved in the computation [[DEV\\_HEAT\\_1](#)], [[DEV\\_HEAT\\_2](#)]. Additionally, the battery on such devices discharges quickly when running such applications [[BATT\\_DRAIN](#)].

A solution to the heat dissipation and battery drainage problem is to offload the processing and video generation tasks to the remote cloud. However, running such tasks on the cloud is not feasible as the end-to-end delays must be within the order of a few milliseconds. Additionally, such applications require high bandwidth and low jitter to provide a high QoE to the user. In order to achieve such hard timing constraints, computationally intensive tasks can be offloaded to Edge devices.

Another requirement for our use case and similar applications such as 360 degree streaming is that the display on the AR/VR device should synchronize the visual input with the way the user is moving their head. This synchronization is necessary to avoid motion sickness that results from a time-lag between when the user moves their head and when the appropriate video scene is rendered. This time lag is often called "motion-to-photon" delay. Studies have shown [[PER\\_SENSE](#)], [[XR](#)], [[OCCL\\_3](#)] that this delay can be at most 20ms and preferably between 7-15ms in order to avoid the motion sickness problem. Out of these 20ms, display techniques including the refresh rate of write displays and pixel switching take 12-13ms [[OCCL\\_3](#)], [[CLOUD](#)]. This leaves 7-8ms for the processing of motion sensor inputs, graphic rendering, and RTT between the AR/VR device and the Edge. The use of predictive techniques to mask latencies has been considered as a mitigating strategy to reduce motion sickness [[PREDICT](#)]. In addition, Edge Devices that are proximate to the user might be used to offload these computationally intensive tasks. Towards this end, the 3GPP requires and supports an Ultra Reliable Low Latency of 0.1ms to 1ms for communication between an Edge server and User Equipment(UE) [[URLLC](#)].

Note that the Edge device providing the computation and storage is itself limited in such resources compared to the Cloud. So, for example, a sudden surge in demand from a large group of tourists can overwhelm that device. This will result in a degraded user experience as their AR device experiences delays in receiving the video frames. In order to deal with this problem, the client AR applications will need to use Adaptive Bit Rate (ABR) algorithms that choose bit-rates policies tailored in a fine-grained manner to the resource demands and playback the videos with appropriate QoE metrics as the user moves around with the group of tourists.

However, heavy-tailed nature of several operational parameters make prediction-based adaptation by ABR algorithms sub-optimal [[ABR\\_2](#)]. This is because with such distributions, law of large numbers works too slowly, the mean of sample does not equal the mean of distribution, and as a result standard deviation and variance are unsuitable as metrics for such operational parameters [[HEAVY\\_TAIL\\_1](#)], [[HEAVY\\_TAIL\\_2](#)]. Other subtle issues with these distributions include the "expectation paradox" [[HEAVY\\_TAIL\\_1](#)] where the longer we have waited for an event the longer we have to wait and the issue of mismatch between the size and count of events [[HEAVY\\_TAIL\\_1](#)]. This makes designing an algorithm for adaptation error-prone and challenging. Such operational parameters include but are not limited to buffer occupancy, throughput, client-server latency, and variable transmission times. In addition, edge devices and communication links may fail and logical communication relationships between various software components change frequently as the user moves around with their AR device [[UBICOMP](#)].

Thus, once the offloaded computationally intensive processing is completed on the Edge Computing, the video is streamed to the user with the help of an ABR algorithm which needs to meet the following requirements [[ABR\\_1](#)]:

- \*Dynamically changing ABR parameters: The ABR algorithm must be able to dynamically change parameters given the heavy-tailed nature of network throughput. This, for example, may be accomplished by AI/ML processing on the Edge Computing on a per client or global basis.
- \*Handling conflicting QoE requirements: QoE goals often require high bit-rates, and low frequency of buffer refills. However in practice, this can lead to a conflict between those goals. For example, increasing the bit-rate might result in the need to fill up the buffer more frequently as the buffer capacity might be limited on the AR device. The ABR algorithm must be able to handle this situation.
- \*Handling side effects of deciding a specific bit rate: For example, selecting a bit rate of a particular value might result in the ABR algorithm not changing to a different rate so as to ensure a non-fluctuating bit-rate and the resultant smoothness of video quality . The ABR algorithm must be able to handle this situation.

## **5. AR Network Traffic**

### **5.1. Traffic Workload**

As discussed earlier, the parameters that capture the characteristics of XR application behavior are heavy-tailed. Examples of such parameters include the distribution of arrival times between XR application invocation, the amount of data transferred, and the inter-arrival times of packets within a session. As a result, any traffic model based on such parameters are themselves heavy-tailed. Using these models to predict performance under alternative resource allocations by the network operator is challenging. For example, both uplink and downlink traffic to a UE device has parameters such as volume of XR data, burst time, and idle time that are heavy tailed.

[Table 1](#) below shows various XR applications and their associated throughput requirements [[METRICS\\_1](#)]. Our use case envisages a 6DoF video or point cloud and so will require 200 to 1000Mbps of bandwidth. As seen from the table, the XR application such as our use case transmit a larger amount of data per unit time as compared to traditional video applications. As a result, issues arising out of heavy tailed parameters such as long-range dependent traffic

[METRICS 2], self-similar traffic [METRICS 3], would be experienced at time scales of milliseconds and microseconds rather than hours or seconds. Additionally, burstiness at the time scale of tens of milliseconds due to multi-fractal spectrum of traffic will be experienced [METRICS 4]. Long-range dependent traffic can have long bursts and various traffic parameters from widely separated time can show correlation. Self-similar traffic contains bursts at a wide range of time scales. Multi-fractal spectrum bursts for traffic summarizes the statistical distribution of local scaling exponents found in a traffic trace. The operational consequences of XR traffic having characteristics such as long-range dependency, and self-similarity is that the edge servers to which multiple XR devices are connected wirelessly could face long bursts of traffic. In addition, multi-fractal spectrum burstiness at the scale of milli-seconds could induce jitter contributing to motion sickness. The operators of edge servers will need to run a 'managed edge cloud service' [METRICS 5] to deal with the above problems. Functionalities that such a managed edge cloud service could operationally provide include dynamic placement of XR servers, mobility support and energy management [METRICS 6]. Providing Edge server support for the techniques being developed at the DETNET and RAW Working Groups at the IETF could guarantee performance of XR applications.

Application	Throughput Required
Image and Workflow Downloading	1 Mbps
Video Conferencing	2 Mbps
3D Model and Data Visualization	2 to 20 Mbps
Two way Telepresence	5 to 25 Mbps
Current-Gen 360 degree video (4K)	10 to 50 Mbps
Next-Gen 360 degree video (8K, 90+ FPS, HDR, Stereoscopic)	50 to 200 Mbps
6DoF Video or Point Cloud	200 to 1000 Mbps

Table 1: Throughput of some XR Applications

Thus, the provisioning of edge servers in terms of the number of servers, the topology, where to place them, the assignment of link capacity, CPUs and GPUs should keep the above factors in mind.

## 5.2. Traffic Performance Metrics

The performance requirements for AR/VR traffic have characteristics that need to be considered when operationalizing a network. We now discuss these characteristics.

The bandwidth requirements of XR applications are substantially higher than those of video based applications.



The latency requirements of XR applications have been studied recently [[AR TRAFFIC](#)]. The following issues were identified.:

- \*The uploading of data from an AR device to a remote server for processing dominates the end-to-end latency.
- \*A lack of visual features in the grid environment can cause increased latencies as the AR device uploads additional visual data for processing to the remote server.
- \*AR applications tend to have large bursts that are separated by significant time gaps.

The packet loss rates in wireless links between XR devices and the Edge server can be as high as 2% or more [[WIRELESS 1](#)].

Additionally, XR applications interact with each other on a time scale of a round-trip-time propagation and this must be considered when operationalizing a network.

The following [Table 2](#) [[METRICS 6](#)] shows a taxonomy of applications with their associated expected end-to-end latencies and bandwidths. Our use case requires an end-to-end latency of 20ms at most and preferably between 7-15ms as discussed earlier. The required bandwidth for our use case as discussed in section 5.2 is 200Mbps-1000Mbps. Since our use case envisages multiple users running the XR applications on their devices, and connected to an edge server that is closest to them, these latency and bandwidth connections will grow linearly with the number of users. The operators should match the network provisioning to the maximum number of tourists that can be supported by a link to an edge server.

Application	Expected End-To-End Latency	Expected Data Latency	Possible Implementations/ Examples
AR-based remote surgery with uncompressed 4K (3840x2160 pixels) 120 fps HDR 10-bit real time video stream	Less than 750 microseconds	Greater than 30 Gbps	World's first remote surgery over 5G
Mobile AR based remote assistance with uncompressed 4K (1920x1080 pixels) 120 fps HDR	Less than 10 milliseconds	Greater than 7.5 Gbps	Assisting maintenance technicians, Industry 4.0 remote maintenance, remote

Application	Expected End-To-End Latency	Expected Data Latency	Possible Implementations/ Examples
10-bit real-time video stream			assistance in robotics industry
Indoor and localized outdoor navigation	Less than 20 milliseconds	50 to 200 Mbps	Theme Parks, Shopping Malls, Archaeological Sites, Museum guidance
Cloud-based Mobile AR applications	Less than 50 milliseconds	50 to 100 Mbps	Google Live View, AR-enhanced Google Translate

Table 2: Traffic Performance Metrics of Selected XR Applications

## 6. Acknowledgements

Many Thanks to Spencer Dawkins, Rohit Abhishek, Jake Holland, Kiran Makhijani ,Ali Begen and Cullen Jennings for providing very helpful feedback suggestions and comments.

## 7. Informative References

- [ABR\_1] Mao, H., Netravali, R., and M. Alizadeh, "Neural Adaptive Video Streaming with Pensieve", In Proceedings of the Conference of the ACM Special Interest Group on Data Communication, pp. 197-210, 2017.
- [ABR\_2] Yan, F., Ayers, H., Zhu, C., Fouladi, S., Hong, J., Zhang, K., Levis, P., and K. Winstein, "Learning in situ: a randomized experiment in video streaming", In 17th USENIX Symposium on Networked Systems Design and Implementation (NSDI 20), pp. 495-511, 2020.
- [AR\_TRAFFIC] Apicharttrisorn, K., Balasubramanian, B., Chen, J., Sivaraj, R., Tsai, Y., Jana, R., Krishnamurthy, S., Tran, T., and Y. Zhou, "Characterization of Multi-User Augmented Reality over Cellular Networks", In 17th Annual

IEEE International Conference on Sensing, Communication, and Networking (SECON), pp. 1-9. IEEE, 2020.

[AUGMENTED] Schmalstieg, D. S. and T.H. Hollerer, "Augmented Reality", Addison Wesley, 2016.

[AUGMENTED\_2] Azuma, R. T., "A Survey of Augmented Reality.", Presence:Teleoperators and Virtual Environments 6.4, pp. 355-385., 1997.

[BATT\_DRAIN]

Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., Hassan, M., and A. Seneviratne, "A survey of wearable devices and challenges.", In IEEE Communication Surveys and Tutorials, 19(4), p.2573-2620., 2017.

[BLUR] Kan, P. and H. Kaufmann, "Physically-Based Depth of Field in Augmented Reality.", In Eurographics (Short Papers), pp. 89-92., 2012.

[CLOUD]

Corneo, L., Eder, M., Mohan, N., Zavodovski, A., Bayhan, S., Wong, W., Gunningberg, P., Kangasharju, J., and J. Ott, "Surrounded by the Clouds: A Comprehensive Cloud Reachability Study.", In Proceedings of the Web Conference 2021, pp. 295-304, 2021.

[DEV\_HEAT\_1] LiKamWa, R., Wang, Z., Carroll, A., Lin, F., and L. Zhong, "Draining our Glass: An Energy and Heat characterization of Google Glass", In Proceedings of 5th Asia-Pacific Workshop on Systems pp. 1-7, 2013.

[DEV\_HEAT\_2] Matsushashi, K., Kanamoto, T., and A. Kurokawa, "Thermal model and countermeasures for future smart glasses.", In Sensors, 20(5), p.1446., 2020.

[EDGE\_1] Satyanarayanan, M., "The Emergence of Edge Computing", In Computer 50(1) pp. 30-39, 2017.

[EDGE\_2] Satyanarayanan, M., Klas, G., Silva, M., and S. Mangiante, "The Seminal Role of Edge-Native

Applications", In IEEE International Conference on Edge Computing (EDGE) pp. 33-40, 2019.

[EDGE\_3] Peterson, L. and O. Sunay, "5G mobile networks: A systems approach.", In Synthesis Lectures on Network Systems., 2020.

[GLB\_ILLUM\_1] Kan, P. and H. Kaufmann, "Differential irradiance caching for fast high-quality light transport between virtual and real worlds.", In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 133-141, 2013.

[GLB\_ILLUM\_2] Franke, T., "Delta voxel cone tracing.", In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 39-44, 2014.

[HEAVY\_TAIL\_1] Crovella, M. and B. Krishnamurthy, "Internet measurement: infrastructure, traffic and applications", John Wiley and Sons Inc., 2006.

[HEAVY\_TAIL\_2] Taleb, N., "The Statistical Consequences of Fat Tails", STEM Academic Press, 2020.

[LENS\_DIST] Fuhrmann, A. and D. Schmalstieg, "Practical calibration procedures for augmented reality.", In Virtual Environments 2000, pp. 3-12. Springer, Vienna, 2000.

[METRICS\_1] ABI Research, "Augmented and Virtual Reality: The first Wave of Killer Apps.", <https://gsacom.com/paper/augmented-virtual-reality-first-wave-5g-killer-apps-qualcomm-abi-research/>, 2017.

[METRICS\_2] Paxon, V. and S. Floyd, "Wide Area Traffic: The Failure of Poisson Modelling.", In IEEE/ACM Transactions on Networking, pp. 226-244., 1995.

[METRICS\_3] Willinger, W., Taqqu, M.S., Sherman, R., and D.V. Wilson, "Self-Similarity Through High Variability: Statistical Analysis and Ethernet LAN Traffic at Source Level.", In IEEE/ACM Transactions on Networking, pp. 71-86., 1997.

[METRICS\_4] Gilbert, A.C., "Multiscale Analysis and Data Networks.", In Applied and Computational Harmonic Analysis, pp. 185-202., 2001.

[METRICS\_5] Beyer, B., Jones, C., Petoff, J., and N.R. Murphy, "Site Reliability Engineering: How Google Runs Production Systems.", O'Reilly Media, Inc., 2016.

**[METRICS\_6]**

Siriwardhana, Y., Porambage, P., Liyanage, M., and M. Ylianttila, "A survey on mobile augmented reality with 5G mobile edge computing: architectures, applications, and technical aspects.", In IEEE Communications Surveys and Tutorials, Vol 23, No. 2, 2021.

**[NOISE]**

Fischer, J., Bartz, D., and W. Straßer, "Enhanced visual realism by incorporating camera image effects.", In IEEE/ACM International Symposium on Mixed and Augmented Reality, pp. 205-208., 2006.

**[OCCL\_1]**

Breen, D.E., Whitaker, R.T., and M. Tuceryan, "Interactive Occlusion and automatic object placement for augmented reality", In Computer Graphics Forum, vol. 15, no. 3, pp. 229-238, Edinburgh, UK: Blackwell Science Ltd, 1996.

**[OCCL\_2]**

Zheng, F., Schmalstieg, D., and G. Welch, "Pixel-wise closed-loop registration in video-based augmented reality", In IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 135-143, 2014.

**[OCCL\_3]**

Lang, B., "Oculus Shares 5 Key Ingredients for Presence in Virtual Reality.", <https://www.roadtovr.com/oculus-shares-5-key-ingredients-for-presence-in-virtual-reality/>, 2014.

**[PER\_SENSE]**

Mania, K., Adelstein, B.D., Ellis, S.R., 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., 2004.

**[PHOTO\_REG]**

Liu, Y. and X. Granier, "Online tracking of outdoor lighting variations for augmented reality with moving cameras", In IEEE Transactions on visualization and computer graphics, 18(4), pp.573-580, 2012.

**[PREDICT]**

Buker, T. J., Vincenzi, D.A., and J.E. Deaton, "The effect of apparent latency on simulator sickness while using a see-through helmet-mounted display: Reducing apparent latency with predictive compensation..", In Human factors 54.2, pp. 235-249., 2012.

**[REG]**

Holloway, R. L., "Registration error analysis for augmented reality.", In Presence:Teleoperators and Virtual Environments 6.4, pp. 413-432., 1997.

**[RFC2119]**

Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

**[SLAM\_1]**

Ventura, J., Arth, C., Reitmayr, G., and D. Schmalstieg, "A minimal solution to the generalized pose-and-scale problem", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 422-429, 2014.

**[SLAM\_2]**

Sweeny, C., Fragoso, V., Hollerer, T., and M. Turk, "A scalable solution to the generalized pose and scale problem", In European Conference on Computer Vision, pp. 16-31, 2014.

**[SLAM\_3]**

Gauglitz, S., Sweeny, C., Ventura, J., Turk, M., and T. Hollerer, "Model estimation and selection towards unconstrained real-time tracking and mapping", In IEEE transactions on visualization and computer graphics, 20(6), pp. 825-838, 2013.

**[SLAM\_4]**

Pirchheim, C., Schmalstieg, D., and G. Reitmayr, "Handling pure camera rotation in keyframe-based SLAM", In 2013 IEEE international symposium on mixed and augmented reality (ISMAR), pp. 229-238, 2013.

**[UBICOMP]**

Bardram, J. and A. Friday, "Ubiquitous Computing Systems", In Ubiquitous Computing Fundamentals pp. 37-94. CRC Press, 2009.

**[URLLC]**

3GPP, "3GPP TR 23.725: Study on enhancement of Ultra-Reliable Low-Latency Communication (URLLC) support in the 5G Core network (5GC).", <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3453>, 2019.

**[VIS\_INTERFERE]**

Kalkofen, D., Mendez, E., and D. Schmalstieg, "Interactive focus and context visualization for augmented reality.", In 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 191-201., 2007.

**[WIRELESS\_1]**

Balachandran, A., Voelker, G.M., Bahl, P., and P.V. Rangan, "Characterizing user behavior and network performance in a public wireless LAN.", In Proceedings of the 2002 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, pp. 195-205., 2002.

[XR]

3GPP, "3GPP TR 26.928: Extended Reality (XR) in 5G.",  
[https://portal.3gpp.org/desktopmodules/Specifications/  
SpecificationDetails.aspx?specificationId=3534](https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3534), 2020.

#### Authors' Addresses

Renan Krishna  
InterDigital Europe Limited  
64, Great Eastern Street  
London  
EC2A 3QR  
United Kingdom

Email: [renan.krishna@interdigital.com](mailto:renan.krishna@interdigital.com)

Akbar Rahman  
Ericsson  
8275 route Transcanadienne  
Montreal H4S 0B6  
Canada

Email: [rahmansakbar@yahoo.com](mailto:rahmansakbar@yahoo.com)