Toward interactive multi-user extended reality using millimeter-wave networking

Reference:
Struye Jakob, Van Damme Sam, Bhat Nabeel Nisar, Troch Arno, Van Liempd Barend, Assasa Hany, Lemic Filip, Famaey Jeroen, Vega Maria Torres.- Toward interactive multi-user extended reality using millimeter-wave networking
IEEE communications magazine / Institute of Electrical and Electronics Engineers [New York, N.Y.] - ISSN 1558-1896 - 62:8(2024), p. 54-60
Full text (Publisher's DOI): https://doi.org/10.1109/MCOM.001.2300804
To cite this reference: https://hdl.handle.net/10067/2075790151162165141

Towards Interactive Multi-User Extended Reality using Millimeter-Wave Networking

Jakob Struye, *Graduate Student Member*, Sam Van Damme, Nabeel Nisar Bhat, *Graduate Student Member*, Arno Troch, Barend van Liempd, Hany Assasa, Filip Lemic, Jeroen Famaey, *Senior Member*, Maria Torres Vega, *Senior Member*

***Abstract*—Extended Reality (XR) enables a plethora of novel interactive shared experiences. Ideally, users are allowed to roam around freely, while audiovisual content is delivered wirelessly to their Head-Mounted Displays (HMDs). Therefore, truly immersive experiences will require massive amounts of data, in the range of tens of gigabits per second, to be delivered reliably at extremely low latencies. We identify Millimeter-Wave (mmWave) communications, at frequencies between 24 and 300 GHz, as a key enabler for such experiences. In this article, we show how the mmWave state of the art does not yet achieve sufficient performance, and identify several key active research directions expected to eventually pave the way for extremely-high-quality mmWave-enabled interactive multi-user XR.**

***Index Terms*—Extended Reality, Virtual Reality, Millimeter-Wave, Beamforming, Multimedia Streaming**

I. Introduction

Since the inception of modern Extended Reality (XR) (which comprises Augmented, Virtual, and Mixed Reality, or AR/VR/MR), Head-Mounted Displays (HMDs) have evolved from experimental, bulky, low-resolution devices to sleek, lightweight user-oriented peripherals. More and more applications of VR, where the user is transported to a fully artificial world, as well as its sibling technologies AR and MR, where virtual elements are overlaid onto, or integrated into the physical world, are being widely deployed. XR applications include employee training and education, sightseeing tours, and entertainment. We expect the recent release of the Apple Vision Pro MR HMD to lead to further mainstream acceptance and adoption of these technologies.

Traditionally, HMDs are connected by a wire to a powerful computer, which generates and renders the XR content. However, this tether inhibits the users' mobility and immersion, and can result in a tripping hazard. As an alternative, recent HMDs offer on-board processing capabilities. These are often focused on AR/MR, which does not require generating a full 360° environment (e.g., Apple Vision Pro). For VR, they are limited to rendering lower-quality content due to their constrained computational capabilities (e.g., Meta Quest 3). The obvious solution is to provide a high-data-rate wireless connection between the HMD and rendering location (e.g., a nearby computer or (edge) cloud server) [1], [2]. Several State of the Art (SotA) VR HMDs, such as the Meta Quest 3, offer wireless connectivity using Wi-Fi on the 5 GHz frequency band. However, as uncompressed XR content may require tens of gigabits per second, a high compression rate is applied, resulting in increased latency and visual artifacts. Achieving high Quality of Experience (QoE) in XR is further complicated by the fact that it requires a *motion-to-photon latency* of at most 20 ms, to avoid cybersickness [1]. This encompasses the total latency between a user's motion and the corresponding update of the visual image on the HMD. These requirements become even more stringent in interactive multi-user XR experiences, where users interact with each other, as well as with the virtual or hybrid environment. They may be co-located, or may participate from different physical locations. Enabling such seamless interactivity requires extremely low latency, alongside dense multi-user and wide-area connectivity.

XR thus requires a combination of high data rate, high reliability, and low latency network connectivity, known as High-Rate and High-Reliability Low-Latency Communications (HR2LLC) [3]. Due to the limited bandwidth and high congestion of the sub-6 GHz frequency bands, Millimeter-Wave (mmWave) wireless communications (i.e., 24–300 GHz) has been identified as a prime enabler of wireless XR [1]. The multi-gigahertz bands available in mmWave offer data rates of up to tens of gigabits at extremely low latency, but pose their own set of challenges to achieving consistent transmission quality. Notably, mmWave experiences high path and penetration loss. This hinders the establishment of consistently high-gain links and renders them prone to blockage, including by users themselves. Ensuring HR2LLC at mmWave frequencies requires a combination of large antenna arrays, directional beamforming, and multi-Access Point (AP) connectivity. This is especially challenging in multi-user interactive XR, featuring highly mobile users within a confined space.

In this article, we present our vision for future truly high-QoE interactive multi-user XR experiences. We present an overview of the scenario and the HR2LLC requirements it poses on the network. We expect that mmWave wireless communications will be an essential building block towards consistently fulfilling these requirements. We identify the shortcomings of the current mmWave SotA, and explore potential avenues towards addressing these, including XR-specific beamforming, Reflective Intelligent Surfaces, channel access optimizations, Integrated Sensing and Communication, low-latency real-time streaming protocols and multi-user human-centric perception for performance evaluations.

J. Struye, N. Nisar Bhat, A. Troch and J. Famaey are with the University of Antwerp and imec, Belgium. S. Van Damme is with Ghent University-imec and KU Leuven, Belgium. B. van Liempd and H. Assasa are with Pharrowtech, Belgium. F. Lemic is with i2Cat Foundation, Spain. M. Torres Vega is with Ghent University-imec and KU Leuven, Belgium.

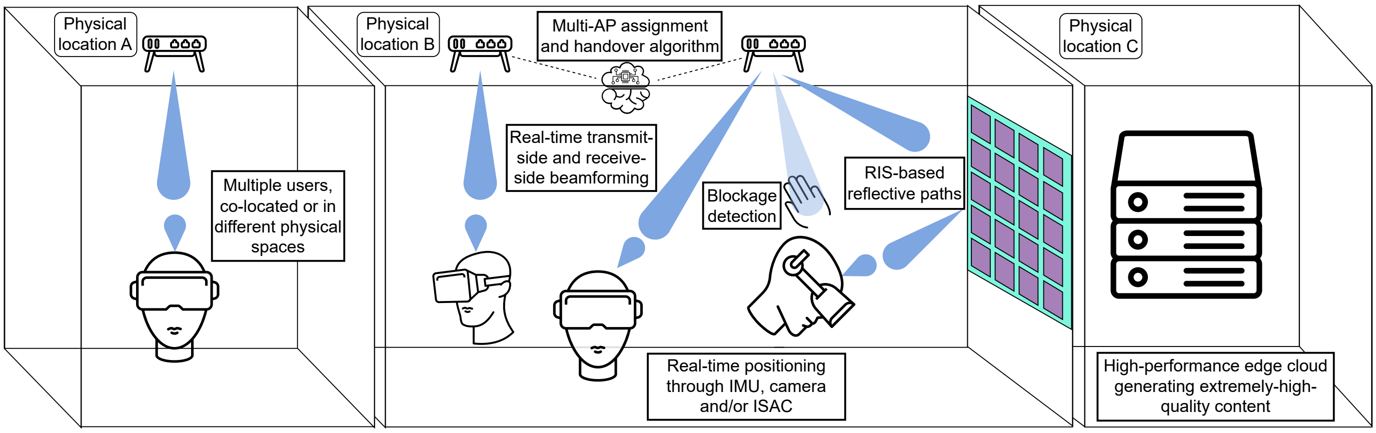


Fig. 1. Overview of the envisioned multi-user interactive XR scenario.

II. Collaborative Wireless XR

Arguably the most technically challenging form of XR is collaborative wireless XR. The inherent difficulty of achieving HR2LLC wirelessly is further amplified by the interactivity of such experiences. Common solutions such as content caching and heavy-duty compression become ineffective when content generation is dependent on several users' real-time actions [4]. Furthermore, wireless resource allocation when several users require an uninterrupted high-quality link is extremely challenging. We provide further details of the physical scenario, discuss the requirements, and summarize current solutions and their shortcomings.

A. System Architecture

Collaborative XR encompasses multi-user, interactive experiences, where users roam freely within a shared virtual (or mixed virtual-physical) environment, as shown in Fig. 1. Users may be separated across multiple physical locations (A and B in the figure), with each location hosting one or more users. Geographically separated users are projected into each other's spaces. Content is generated in real-time, likely at another physical location. In the figure, content is computer-generated in an edge cloud at physical location C. To facilitate realistic and unconstrained Six Degrees of Freedom (6DoF) motion, HMDs estimate their own pose through Internal Measurement Units (IMUs) and/or built-in cameras. Furthermore, they incorporate a mmWave antenna, and content is delivered by one or more mmWave APs, mounted on walls or the ceiling. HMDs and APs beamform towards each other, illustrated in blue in the figure. In multi-AP deployments, a centralized algorithm orchestrates AP assignment and handovers, with its objective commonly being to maximize the lowest QoE among all users. Crucially, when signal degradation indicates upcoming blockage, the algorithm should avoid quality degradation by reassigning the affected user. Finally, the figure incorporates several experimental systems for further improving the QoE, further covered in Sec. III, including novel approaches to real-time beamforming, establishing viable reflected links, and pose estimation.

B. Requirements

HR2LLC summarizes the network requirements of collaborative wireless XR [3]. The **high rate** is determined by the quality of the content, which is in turn limited by the hardware specifications of the HMD [5]. Without compression, the Meta Quest 3 requires between 15.75 and 26.25 Gbps depending on the refresh rate. While compression may reduce this staggering requirement, this introduces an additional (de)compression delay. This may impact the **low latency** requirements, with the *motion-to-photon* latency limit for XR being commonly defined as 20 ms, meaning the result of any user motion must be reflected in-experience within 20 ms to avoid nausea [1]. Depending on other factors, this may leave between 5 and 8 ms for one-way network transmissions. Any content not arriving on time is essentially lost, which is highly impactful given the **high-reliability** requirement. This reliability requirement is defined at two levels. The *intra-image reliability* determines the fraction of an image that needs to arrive on time in order to be considered as complete. The exact target fraction depends on redundancy in any compression algorithm, along any reconstruction algorithms extrapolating missing data from arrived content. The *inter-image reliability* defines how many images may be lost without unacceptable impact on QoE, both overall and within a single loss burst. The exact threshold depends on the specific experience, but may reach as high as 99.999 %, or roughly 1 missed image per 15 minutes. While there are many methods for measuring the QoE of interactive XR [6], fulfilling the HR2LLC requirements is always a necessary condition for achieving satisfactory results.

C. Existing Solutions and Shortcomings

Achieving HR2LLC wirelessly will require mmWave communications. Overcoming mmWave's high path loss and susceptibility to blockage demands a specialized approach. Through antenna arrays consisting of many elements, along with *beamforming*, a process in which energy is focused in a carefully selected direction, a sufficiently high gain can be achieved even with modest overall energy budgets. As the range of directions in which a mmWave antenna can beamform is limited, they effectively have a limited Field of View (FoV),

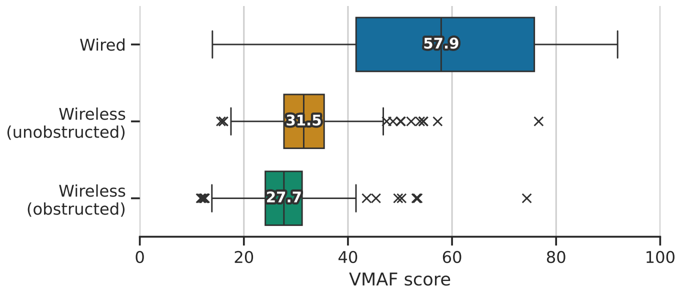


Fig. 2. Perceived video quality comparison for the HTC Vive in wired and (un)obstructed wireless scenarios. A higher VMAF score is better.

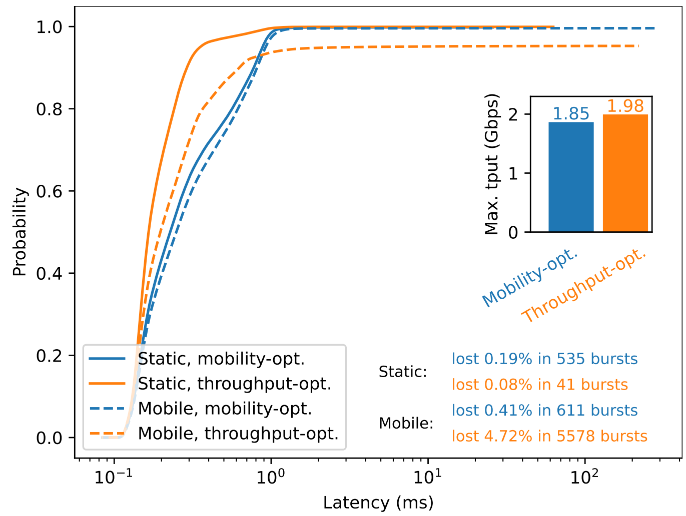


Fig. 3. Throughput, latency and loss with SotA mmWave hardware, for a single AP and user, both static and under moderate motion (45 °/s), with parameters optimized for either throughput or mobility.

usually between 60 and 120°. As such, a connection can easily be interrupted by the user walking around or even simply turning their head. Enabling consistent high-gain coverage in the face of blockage and limited FoVs therefore requires multiple, spatially separated APs.

At the protocol level, both 5G and Wi-Fi offer explicit mmWave support within their specifications. 5G New Radio (NR) supports several bands within the mmWave range, with most deployments using licensed bands between 24.5 and 29.5 GHz. Wi-Fi, in turn, supports mmWave using unlicensed bands between 57 and 70 GHz. Frequencies above 100 GHz are being explored for potential use in 6G. Concerning current-day protocols, both 5G and Wi-Fi consider XR as an explicit use case for their mmWave functionality in official documentation. However, 5G defines this as an outdoors use case with expected data rates in the tens of megabits per second, while Wi-Fi describes indoors experiences with multi-gigabit requirements. As the latter overlaps closely with the scenario described in this article, we adopt Wi-Fi's terminology (e.g., "AP") throughout the article, but note that all solutions described in Sec. III are protocol-agnostic.

Some mmWave solutions for wireless XR have been brought to market. Most notably, the HTC VIVE Wireless Adapter replaces the usual cable with a wireless mmWave bridge, running a custom protocol. To understand the performance of these wireless options, we compared the performance in wired and wireless mode, additionally considering the scenario in which the wireless link is obstructed. We transmitted the same content, computer-generated in real-time, in the three scenarios, calculating the Video Multimethod Assessment Fusion (VMAF) score to assess video quality objectively. The VMAF score (0-100) is a per-frame quality similarity metric compared to a reference recording of the content, with 100 indicating no quality reduction compared to the reference. Fig. 2 shows the mean VMAF scores, along with box plots of their per-frame values. Due to frame rate instability, the VMAF score of the wired scenario is already suboptimal despite there being no compression or data loss. The VMAF drops significantly with a clear wireless link, indicating significant compression occurs as to not saturate the wireless link capacity. Blockage causes an additional reduction in perceived quality.

More recently, the Meta Quest line of HMDs offers a wirelessly streamed solution through users' existing 5 GHz Wi-Fi deployments rather than requiring additional hardware. Even with the recently introduced H.264+ compression option, the bitrate is still limited to 400 Mbps, leading to readily noticeable compression artifacts, substantiating the need for mmWave.

To assess the capabilities of SotA mmWave hardware, we evaluated the data rate, latency and loss under motion of a mmWave link provided by Pharrowtech. Here we evaluated two sets of Medium Access Control (MAC)-level parameters, aimed at optimizing data rate and adaptation to mobility, respectively. This showed that optimizing MAC-level parameters boosts throughput, and that moderate motion can lead to an increase in packet loss, even with an unobstructed Line-of-Sight (LoS) and parameters optimized for mobility. This is illustrated in Fig. 3, where, in each case, there is room for improvement in reducing loss to meet high-QoE XR standards. This indicates the need for developing proactive beamforming approaches. Overall, SotA solutions are unable to fulfill the HR2LLC requirements of current-day HMDs. Future HMDs are expected to drastically increase specifications, enabling truly realistic experiences but also further inflating HR2LLC requirements, mainly in terms of data rate.

Even experimental mmWave Wi-Fi hardware currently fails to achieve the maximal Modulation and Coding Scheme (MCS) supported by the specification, which would offer a data rate of 8.5 Gbps. In parallel, two optional features of the specification can improve the data rate orthogonally. Through channel bonding and aggregation, up to four channels may be combined to linearly increase the data rate. This is partially supported by the devices above, and we consider this a viable way to increase effective data rate for interactive XR. Next, Multiple-Input Multiple-Output (MIMO) can increase Wi-Fi data rate by, theoretically, a factor 8, through multiple antenna arrays leveraging multipath propagation. However, to increase data rate towards a single XR device, this would require equipping those with multiple (or larger, logically sub-divisible) antenna arrays. Given cost and physical size constraints, we do not consider this to be viable for wearable XR devices. Furthermore, based on several measurement campaigns, power consumption of a simple mmWave Wi-Fi chip is, at worst, 1 W higher than with sub-6 GHz Wi-Fi during active transmission. As overall power consumption of modern HMDs is around 10 W, the reduction in battery life from incorporating mmWave is modest. However, MIMO would further increase this consumption, with noticeably reduced battery life eventually leading to a reduction in longer-term QoE. We do note that Multi-User MIMO (MU-MIMO), in which the multiple streams are sent to *different* receiving XR devices, would only require additional hardware support at the APs, making this a cost-effective way of supporting more users within a physical environment.

Independent of the networking approach, XR's strict latency requirements are often alleviated through Asynchronous Time-Warp (ATW) [7]. With this algorithm, images generated based on a stale orientation measurement are perturbed to offset for recent rotational motion, reducing the effective motion-to-photon latency in some use cases. However, this can only address the motion of the viewpoint; other visible physical motions (e.g., in tele-operation) are largely unaffected by ATW, meaning it can not generally reduce effective latency.

III. Open Challenges and Way Forward

Above, we argued that mmWave networking is necessary for high-QoE interactive multi-user wireless XR, but showed that SotA performance of mmWave solutions does not yet suffice. In this section, we discuss several avenues where XR-centric research is ongoing, but more efforts are needed to achieve truly immersive experiences.

A. Beamforming for mobile users

In many present-day mmWave deployments, transmit-side beamforming suffices to achieve a performant link. The HR2LLC requirements of XR, however, may necessitate additional receive-side beamforming at the HMD. Beamforming at this side is inherently more challenging in mobile XR; most user rotation changes the Angle of Arrival (AoA) at the HMD drastically, but has minimal impact on the Angle of Departure (AoD) at the static AP. As such, mobile XR necessitates the development of receive-side beamforming approaches which adapt to user rotation with minimal latency, or even proactively. To this end, algorithms can leverage the plethora of sensors already available on modern-day HMDs [8]. Through IMUs and on-device cameras, HMDs can accurately estimate their own position and orientation in real-time. By combining this *context information* with the fixed location of APs, the HMD can beamform towards the AP directly, foregoing the time-consuming beam searching algorithms prevalent in beamforming approaches. In addition, the HMD could predict its upcoming motion and form a receive beam that will consistently cover the AP during this rotation, as shown in Fig. 4 [9].

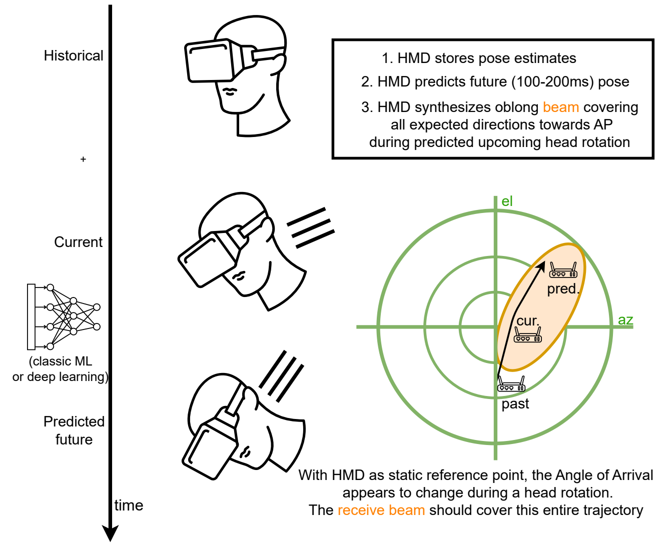


Fig. 4. HMD-side beamforming should proactively adapt to expected upcoming rotations, such that receive gain remains consistently high during rapid user motion.

B. Reconfigurable Intelligent Surfaces

Reconfigurable Intelligent Surfaces (RISs) are passive wall-mountable *metasurfaces* whose reflective properties can be altered in real-time [10]. By intelligently controlling the reflection angle of incident beams, high-gain reflected paths can be established where these would otherwise not be viable. An intelligent, advanced resource scheduler could take these into account to maximize the QoE in a large XR deployment. As the passive elements that comprise a RIS are low-complexity and therefore low-cost, RIS is a promising avenue towards increasing coverage in XR scenarios at a low cost. Existing prototypes at both mmWave and lower frequencies have proven the viability of the concept. However, practical and experimental works on mmWave RIS are still scarce, and none consider the extremely high data rates required for XR. Prototypes convincingly demonstrating the viability of RIS-enhanced interactive XR would aid in paving the way towards affordable consumer-grade solutions.

Furthermore, we identify a need for **real-time** RIS configuration algorithms. SotA configuration algorithms are already well-performing, but rely on high-complexity optimizers, leading to runtimes of minutes to hours [11]. These are focused on fully static environments, where a single iteration of such algorithms leads to a permanent configuration. However, deploying RISs in interactive XR scenarios necessitates heuristic approaches with runtimes in the range of milliseconds. We identify Deep Reinforcement Learning as a potential enabler to this end.

C. High-Data Rate Low-Latency Channel Access

Especially in multi-user scenarios, interactive XR poses its own challenges and opportunities in terms of scheduling channel access. Most traffic is downstream, meaning orchestration can easily occur in a centralized controller. Even in multi-AP deployments, most of the scheduling control overhead can occur over the wired, fully reliable network. The main challenge of XR traffic is the strict per-image deadline. An image *must* arrive fully at the HMD before the time it is intended to be displayed, or else it is fully lost. This makes efficient interweaving of traffic towards multiple HMDs highly challenging. Modern solutions such as channel aggregation and bonding, allowing for dynamic bandwidths, further complicate the resource allocation challenge. In addition, mmWave transmission schedules often incorporate repeating periods for MAC-layer tasks such as device discovery, association and beamforming, during which no application traffic can be transmitted. For example, the Beacon Header Interval (BHI) for mmWave Wi-Fi reoccurs every 100 ms and may take several milliseconds each time. Traffic must be scheduled around this carefully, taking deadlines into account [12]. In addition, an optimal scheduler would be aware of each HMD's refresh cycle, maximizing the percentage of images arriving on time. This requirement could be alleviated somewhat if HMDs support a variable refresh rate, with which a screen update could be briefly postponed until an image is fully received. Fig. 5 shows a schedule taking MAC-layer overheads, image deadlines and sudden blockage into account.

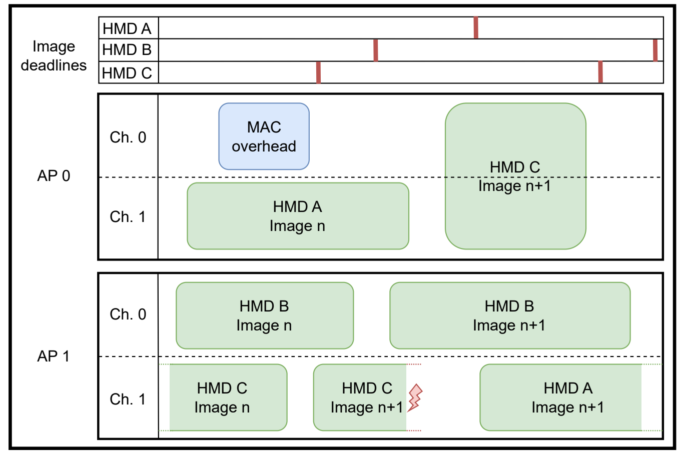


Fig. 5. Adaptive scheduling for a multi-AP, multi-user, multi-channel deployment. During transmission of image $n+1$ to HMD C, the connection is suddenly interrupted (e.g., due to hand blockage), after which HMD C is rapidly moved to another AP and given two channels to ensure the image deadline is met. To facilitate this, HMD A switches to the other AP.

D. Integrated Sensing and Communication

XR applications not only require low latency and high-speed communication, but also accurate and real-time sensing of user pose. While the HMD pose is needed for rapid beamforming, the continuous and accurate estimation of the *full body pose* is needed for applications where physical actions are translated into the XR environment. While current SotA solutions for XR pose estimation often rely on cameras or hand-held controllers, these approaches come with significant drawbacks. Camera-based methods infringe the privacy of users, require a well-lit environment, and are not scalable. Also, hand-held controllers can be unintuitive to users. Therefore, the overall setup becomes expensive and complex.

Instead, pose estimation can also occur through the use of wireless signals. In particular, Wi-Fi signals have been used for many *sensing* applications such as pose estimation, gesture recognition, and localization [13]. This concept of re-using communication signals for sensing is known as Integrated Sensing and Communication (ISAC). The standout advantage of the sensing approach is its limited additional cost. While Wi-Fi sensing has primarily focused on sub-6 GHz signals, the constrained bandwidth at these frequencies limits sensing resolution. In contrast, we identify an untapped potential in mmWave, where vast bandwidth significantly enhances sensing accuracy. For instance, leveraging a 2 GHz bandwidth at 60 GHz can yield an impressive 15 cm raw resolution in localization applications. Initial results on mmWave ISAC are promising, however some open challenges remain on the path towards real-time body pose estimation through mmWave. Specifically, current estimation performance is highly sensitive to both the body shapes of individuals and the surrounding environment [13]. As such, we identify the need for more robust estimation algorithms, and expect future Deep Learning-based solutions to fulfill this need.

E. Low Latency End-to-End Streaming

To ensure consistently high QoE, special care must also be taken at the application streaming level, where the system must adapt in real-time to context changes. With a context such as mmWave, given its volatile nature, special care is needed to ensure the user's perception of 6DoF video remains consistent [14]. The system must incorporate measurements from client and server to improve the real-time reactiveness of the streaming, such as on-the-fly parameter tuning related to the viewport of the user. Although low-latency video streaming protocols such as Low-Latency Dynamic Adaptive Streaming Over HTTP (LL-DASH) and Web Real-Time Communication (WebRTC) are already in used for traditional 2D video, their translation to immersive 3D content, especially when transmitted over mmWave, is not straightforward, as these protocols are constrained in terms of processing power, throughput, and latency. Recently, the Internet Engineering Task Force (IETF) initiated the Media over QUIC (MOQ) working group aiming to develop a simple low-latency media delivery solution addressing use cases including live video streaming, gaming, and media conferencing at scale. Currently, only early results on MOQ are available, and solutions for immersive media use cases are still to be developed [15].

F. Multi-User, Human-Centric Perception

In order to align the aforementioned technologies with human perception, experience and interaction (i.e. human-centric multimedia) and to assess the effectiveness of the above approaches to QoE improvement, we need detailed, accurate and expansive evaluations. Traditionally, this human perception evaluation has been performed by means of subjective methodologies, relying on user feedback gathered through questionnaires and prompts. However, these suffer from individual biases, are not scalable and may disrupt a person's experience [14]. Such disruptions impact the evaluation, as breaks in presence and immersion tend to significantly alter the experience. Moreover, subjective evaluations are a posteriori, meaning the full experience is rated at once. Given the volatile nature of mmWave, it is fundamental to move towards less intrusive and more real-time assessments of perception.

Ideally, objective metrics, driven by physiological data, would provide a more immediate result without requiring conscious assessment. These methodologies are still at their infancy and must thus be investigated and integrated into immersive media experience evaluation.

Moreover, these assessments are currently performed on the individual, meaning the individual perception and immersion are assessed. However, in multi-user (collaborative) environments, the multiple users share a common environment and goal. Therefore, a focus shift is required from the level of the individual user to the collective perception and experience of the group. This group cohesiveness, or the extent to which group members are attracted to the group and its goals, is affected by a plethora of factors. These include shared cognition (the relationship between group and application/system), shared awareness (the mutual relationship between the group's individual members), shared engagement (the relationship between group and its common goal) and the individual perception and immersion. The level in which they affect the perception will depend on context factors such as hardware, network, use case, subject and emotional state. As such, this focus shift from individual to collective, multi-user, human-centric perception poses an important and interesting direction for further research.

IV. Conclusion

In this article, we presented our vision for future deployments of extremely-high-quality interactive multi-user XR experiences. Multiple users can roam freely in a shared, collaborative experience, while they may be co-located or in different locations physically. A high-performance edge cloud handles real-time content-generation, while wireless mmWave links provide the last-hop connection to HMDs. Evaluation of several SotA hardware solutions showed that current mmWave technology cannot fulfill the extreme requirements of the envisioned scenario. We then identified several key active research tracks towards achieving consistently high-QoE interactive multi-user XR, discussing ongoing work and open challenges. We are confident that additional effort along these tracks will eventually lead to the realization of this scenario, enabling a plethora of novel applications and experiences.

Acknowledgments

This research was partially funded by the ICON project INTERACT and Research Foundation - Flanders (FWO) project WaveVR (Grant number G034322N). INTERACT was realized in collaboration with imec, with project support from VLAIO (Flanders Innovation and Entrepreneurship). Project partners are imec, Rhinox, Pharrowtech, Dekimo and TEO. This work is partially supported by the European Commission through the Horizon Europe JU SNS project Hexa-X-II (Grant Agreement no. 101095759). Nabeel Nisar Bhat and Sam Van Damme are supported by an FWO SB PhD fellowship (Grant numbers 1SH5X24N and 1SB1822N respectively). The work of Filip Lemic was supported by the Spanish Ministry of Economic Affairs and Digital Transformation and the European Union – NextGeneration EU, in the framework of the Recovery Plan, Transformation and Resilience (Call UNICO I+D 5G 2021, ref. number TSI-063000-2021-7); the European Union's Horizon Europe's research and innovation programme under grant agreement nº 101139161 — INSTINCT project; and MCIN / AEI / 10.13039 / 501100011033 / FEDER / UE HoloMit 2.0 project (nr. PID2021-126551OB-C21).

References

[1] M. S. Elbamby, C. Perfecto, M. Bennis, and K. Doppler, "Toward low-latency and ultra-reliable virtual reality," *IEEE Netw.*, vol. 32, no. 2, pp. 78–84, 2018.

[2] F. Hu, Y. Deng, W. Saad, M. Bennis, and A. H. Aghvami, "Cellular-connected wireless virtual reality: Requirements, challenges, and solutions," *IEEE Commun. Mag.*, vol. 58, no. 5, pp. 105–111, 2020.

[3] C. Chaccour, M. N. Soorki, W. Saad, M. Bennis, and P. Popovski, "Can terahertz provide high-rate reliable low-latency communications for wireless vr?" *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9712–9729, 2022.

[4] Y. Sun, Z. Chen, M. Tao, and H. Liu, "Communications, caching, and computing for mobile virtual reality: Modeling and tradeoff," *IEEE Trans. Commun.*, vol. 67, no. 11, pp. 7573–7586, 2019.

[5] R. Radoeva, E. Petkov, T. Kalushkov, D. Valcheva, and G. Shipkovenski, "Overview on hardware characteristics of virtual reality systems," in *2022 Int. Congr. Human-Comput. Interact., Optim. Robot. Appl.*, 2022, pp. 01–05.

[6] S. Vlahovic, M. Suznjevic, and L. Skorin-Kapov, "A survey of challenges and methods for quality of experience assessment of interactive vr applications," *J. Multimodal User Interfaces*, vol. 16, pp. 257–291, 2022.

[7] J. M. P. van Waveren, "The asynchronous time warp for virtual reality on consumer hardware," in *Proc. 22nd ACM Conf. Virtual Reality Softw. Technol.*, 2016, p. 37–46.

[8] A. M. Nor, S. Halunga, and O. Fratu, "Survey on positioning information assisted mmwave beamforming training," *Ad Hoc Netw.*, vol. 135, p. 102947, 2022.

[9] J. Struye, F. Lemic, and J. Famaey, "Covrage: Millimeter-wave beamforming for mobile interactive virtual reality," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4828–4842, 2023.

[10] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.

[11] C. Liaskos, L. Mamatas, A. Pourdamghani, A. Tsioliaridou, S. Ioannidis, A. Pitsillides, S. Schmid, and I. F. Akyildiz, "Software-defined reconfigurable intelligent surfaces: From theory to end-to-end implementation," *Proc. IEEE*, vol. 110, no. 9, pp. 1466–1493, 2022.

[12] T. Azzino, T. Ropitault, and M. Zorzi, "Scheduling the data transmission interval in ieee 802.11ad: A reinforcement learning approach," in *2020 Int. Conf. Comput. Netw. Commun.*, 2020, pp. 602–607.

[13] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless sensing for human activity: A survey," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1629–1645, 2019.

[14] J. van der Hooft, H. Amirpour, M. Torres Vega, Y. Sanchez, R. Schatz, T. Schierl, and C. Timmerer, "A tutorial on immersive video delivery: From omnidirectional video to holography," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, p. 1336–1375, 2023.

[15] Z. Gurel, T. Erkilic Civelek, A. Bodur, S. Bilgin, D. Yeniceri, and A. C. Begen, "Media over quic: Initial testing, findings and results," in *Proc. 14th Conf. ACM Multimedia Syst.*, 2023, p. 301–306.

Jakob Struye is a Ph.D. researcher in the field of wireless networking at the IDLab research group (University of Antwerp) and imec research institute, Belgium. He obtained his B.Sc (2015) and M.Sc. (2017) in Computer Science at the University of Antwerp. His current research focuses on leveraging extremely-high-frequency wireless networks in the millimeter-wave bands to improve the performance of truly wireless Extended Reality experiences, and has experience in applying Artificial Intelligence to time series prediction problems.

Sam Van Damme is a Joint PhD Researcher affiliated with the Internet Technology and Data Science Lab (IDLab) at the Department of Information Technology (INTEC), Ghent University, Belgium and e-Media Research Lab, KU Leuven, Belgium. He obtained his B.Sc. (2016) and M.Sc. (2019) in Computer Science Engineering at Ghent University, Belgium. He has expertise in Quality-of-Experience modelling and assessment, Human-Computer Interaction, and Haptic Feedback with a main focus on immersive multimedia technologies such as Virtual and Augmented Reality.

Nabeel Nisar Bhat is a Ph.D. researcher in the field of Joint Communication and Sensing at the IDLab research group (University of Antwerp) and imec research institute, Belgium. He obtained his M.Sc. (2021) in Communications and Computer Networks Engineering at Politecnico di Torino. His current research focuses on leveraging mmWave communication signals for pose estimation in Extended Reality applications. He has experience in signal processing, wireless communications, and deep learning.

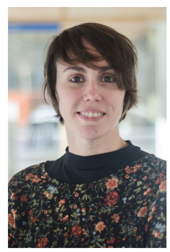
Arno Troch is a Ph.D. researcher in the field of wireless and cellular networking at the IDLab research group (University of Antwerp) and imec research institute, Belgium. He obtained his B.Sc. (2021) and M.Sc. (2023) in Computer Science at the University of Antwerp. His current research focuses on mesh networking at millimeter-wave and sub-terahertz frequencies for future cellular networks. His main interests are wireless and cellular networking, network modelling, and vehicular communications.

Barend van Liempd received the B.Sc. and M.Sc. degrees in EE from Eindhoven University of Technology (2009, 2011), and Ph.D. from Vrije Universiteit Brussel (2017), with a focus on tunable RF front-end circuits. He worked at imec, Belgium (2011-2022), first as R&D Engineer, and later Senior Researcher, Program Manager and R&D Manager. Then, he joined Pharrowtech as Manager Hardware Engineering. He published over 60 articles and patents and received the 2015 NXP Prize and 2019 Lewis Winner Award.

Hany Assasa is a senior system engineer at Pharrowtech, Belgium. He leads software and firmware activities for the Pharrowtech 60 GHz RFIC module. Previously, he worked at IMDEA Networks Institute, Spain (2015 - 2020), first as a Ph.D. researcher and then as a Postdoctoral researcher. While working at IMDEA Networks, he focused on building efficient, robust, and reliable millimeter-wave wireless networks. His main interests are wireless networking, prototyping, wireless communications, and signal processing.

Filip Lemic is a senior researcher at the i2Cat Foundation. He held positions at the University of Antwerp, imec, Universitat Politecnica de Catalunya, University of California at Berkeley, Shanghai Jiao Tong University, FIWARE Foundation, and Technische Universitat Berlin. He received his M.Sc. and Ph.D. from the University of Zagreb and Technische Universitat Berlin, respectively.

Jeroen Famaey is an associate research professor at the Department of Computer Science of the University of Antwerp and imec, Belgium. His research focuses on performance modeling and optimization of wireless network protocols. He has published over 170 peer-reviewed journal articles and conference papers, and 8 granted patents. He was listed as one of the top 2% most cited scientists worldwide by Stanford University both in 2022 and 2023.



Maria Torres Vega is a tenure track assistant professor at KU Leuven (Belgium), where her research focuses on devising human-driven control and management mechanisms for enhancing the perception of immersive systems. She received her M.Sc. degree in Telecommunication Engineering from the Polytechnic University of Madrid, Spain, in 2009 and her Ph.D. from the Eindhoven University of Technology, The Netherlands in 2017. Her research interests include quality of service and QoE in immersive multimedia systems and autonomous networks management.