Immersive Interconnected Virtual and Augmented Reality: A 5G and IoT Perspective

14 authors, including:

Some of the authors of this publication are also working on these related projects:

Project   Multi-objective broadcast scheduling techniques View project

Project   WiPLASH: Architecting More Than Moore – Wireless Plasticity for Heterogeneous Massive Computer Architectures View project

Immersive Interconnected Virtual and Augmented Reality: A 5G and IoT perspective

**Maria Torres Vega · Christos Liaskos ·
Sergi Abadal · Evangelos Papapetrou ·
Akshay Jain · Belkacem Mouhouche ·
Gökhan Kalem · Salih Ergüt ·
Marian Mach · Tomas Sabol ·
Albert Cabellos-Aparicio ·
Christoph Grimm · Filip De Turck ·
Jeroen Famaey**

Abstract Despite remarkable advances, current augmented and virtual reality (AR/VR) applications are a largely individual and local experience. Interconnected AR/VR, where participants can virtually interact across vast distances, remains a distant dream. The great barrier that stands between current technology and such applications is the stringent end-to-end latency requirement, which should not exceed 20 ms in order to avoid motion sickness and other discomforts. Bringing AR/VR to the next level to enable immersive interconnected AR/VR will require significant advances towards 5G ultra-

M. Torres Vega · Filip De Turck
IDLab, Ghent University – imec, Belgium
E-mail: maria.torresvega@ugent.be

C. Liaskos · E. Papapetrou
University of Ioannina, Greece

S. Abadal · A. Jain · A. Cabellos-Aparicio
Universitat Politècnica de Catalunya, Spain

B. Mouhouche
Samsung Electronics, United Kingdom

G. Kalem · S. Ergut
Turkcell Technology, Turkey

C. Grimm
Technische Universität Kaiserslautern, Germany

M. Mach · T. Sabol
Technical University of Kosice, Slovakia

J. Famaey
IDLab, University of Antwerp – imec, Belgium

reliable low-latency communication (URLLC) and a Tactile Internet of Things (IoT). In this article, we articulate the technical challenges to enable a future AR/VR end-to-end architecture, that combines 5G URLLC and Tactile IoT technology to support this next generation of interconnected AR/VR applications. Through the use of IoT sensors and actuators, AR/VR applications will be aware of the environmental and user context, supporting human-centric adaptations of the application logic, and lifelike interactions with the virtual environment. We present potential use cases and the required technological building blocks. For each of them, we delve into the current state of the art and challenges that need to be addressed before the dream of remote AR/VR interaction can become reality.

Keywords Augmented and Virtual Reality · Tactile Internet · 5G · Internet of Things

1 Introduction

Despite the remarkable advances in augmented and virtual reality (AR/VR) technology, interconnected AR/VR remains confined to the world of science fiction. In interconnected AR/VR applications, participants can interact with each other and the environment across long physical distances [1]. The great barrier that stands between current technology and remote AR/VR presence is the extremely stringent motion-to-photon latency, which should not exceed 20 milliseconds (ms) in order to avoid motion sickness and enable lifelike experiences [2].

The motion-to-photon latency includes any delay incurred by sensor sampling (e.g., motion capturing, location estimation, interactions with the environment), data processing and fusion, image rendering and encoding, transmission, and decoding and displaying of each frame. Figure 1 show more details about the different tasks that lie in the critical path of the motion-to-photon latency, and round-trip time (RTT) of current state of the art and expected future high-end components [3]. In AR, processes such as object identification, registration, or retrieval of data already take considerable time; whereas, in VR, the sheer processing throughput required by the video stream is highly taxing. On top of them, the display scanning and photon launching further contribute to this delay. However, the main contributor to latency in current communications and computing systems is the network. As information traverses the network, it is queued, processed, and transmitted by an increasing number of routers. This procedure adds a delay that is both significant, generally in the order of tens of milliseconds per router, and dependent on the state of the network, which makes it hard to predict. These factors disrupt immersion and create discomfort in AR/VR systems and result in end-to-end latency in the order of tens to hundreds of ms. Even when considering high-end hardware and novel processing techniques, sampling, rendering and displaying will take 5 to 8 ms. To keep the full system RTT below 20 ms, this will leave about 7 ms for the transmission of information (Figure 1). However, in current
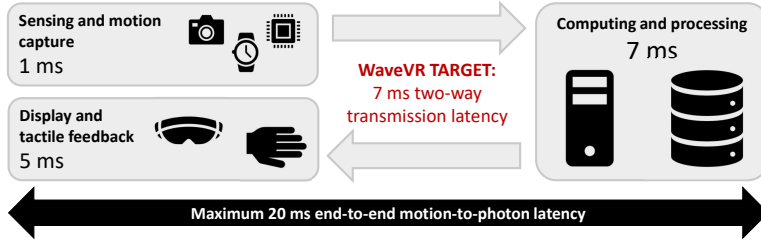
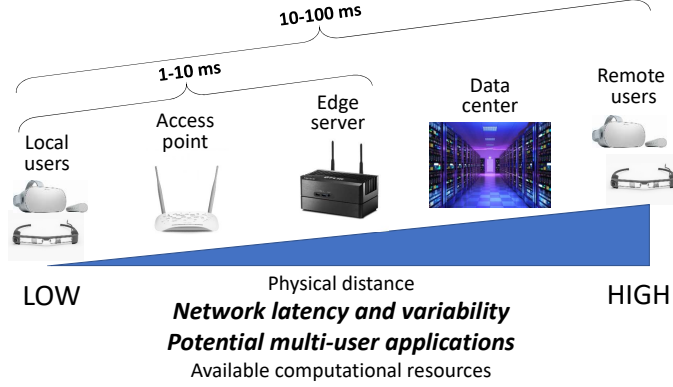**Fig. 1** Distribution of latency in the interconnected AR/VR system.

Fig. 2 In current computer networks, the latency increases significantly based on the distance between the transmitter and receiver, going up to 100 ms or more between remote end-devices

networks latency increases significantly with the distance, going up to 100 ms or more between remote devices (Figure 2). Therefore, it will not be enough to improve network infrastructures, but novel solutions are needed that will bring the latency down to the requirements of AR/VR remote applications.

Due to the high and unpredictable RTT of current computer networks, AR/VR is currently a largely individual experience, and at best allows participation of multiple locally connected users. Moreover, current AR/VR applications do not support interactions among participants or with the environment through sensors and actuators, and instead focus solely on auditory and visual signals. Bringing AR/VR to the next level to enable lifelike and human-centric remote interactions among users will require significant scientific and technological advances towards a Tactile Internet of Things (IoT) [2]. The Tactile IoT vision is the next breakthrough in technology enabling the manipulation of physical/biological objects from a distance. It envisions communication across the Internet at millisecond latency among users and heterogeneous IoT devices, such as AR/VR head-mounted displays (HMDs), wearables, touch sensors, radars, cameras, and haptic gloves. To achieve this, different network enablers need to be combined. 3GPP coined the term ultra-reliable low-latency communication (URLLC) as one of the 5G traffic classes. It requires a target

latency of 1 ms at a packet loss ratio of 10^{-5} for 32-byte packets on the wireless radio access network (RAN) segment [4]. Another enabler is deterministic networking (DetNet), coined by IETF. It is an extension of the time-sensitive networking (TSN) concept, and guarantees specific latency and jitter bounds for packets routed through the core network segment [5].

In this article, we describe the required components of the end-to-end transmission chain to enable multi-user immersive interconnected AR/VR applications. For each of the presented components, we articulate the challenges and provide with possible solutions. The end-devices are wirelessly connected via a heterogeneous URLLC RAN, that satisfies varying device requirements in terms of latency, throughput, reliability, and energy efficiency. The different RANs are connected via a deterministic core network, supporting Quality of Service (QoS) differentiated deterministic routing. A distributed IoT management framework ensures the interoperability among devices and secure dissemination of data. It interfaces with the network to reserve the required resources to guarantee the end-to-end latency and reliability requirements of the applications. Finally, the AR/VR applications interface IoT framework to obtain relevant context information about the environment and users, to allow human-centric adaptability of the application logic. The article provides an overview of the architecture, and for each component discussion about potential enabling technologies, the current state-of-the-art, and open challenges.

The remainder of the article is structured as follows. Section 2 presents the envisioned end-to-end architecture of the interconnected AR/VR concept. Four aspects of the chain are highlighted, namely the heterogeneous RAN, the deterministic core network, the IoT middleware infrastructure, and the AR/VR application. Furthermore, this Section provides a set of possible application examples with their requirements and characteristics. Sections 3 to Section 6 provide both state of the art and challenges for each of these four elements. Section 7 concludes this paper.

2 Overview of the interconnected AR/VR concept

This section first introduces our envisioned architecture and its components. Subsequently, more details are provided about application and vertical domains that may benefit from remote multi-user interconnected AR/VR.

2.1 Envisioned end-to-end architecture

Figure 3 summarizes the overall architecture conceived to enable immersive interconnected AR/VR. The presented architecture considers a number of locations, each of which hosts a number of physical users, sensors, actuators, and potentially interactive targets (i.e., objects the users can interact with, either physically or virtually). All devices are connected over the **heterogeneous RAN**, which provides local wireless connectivity with a variety of
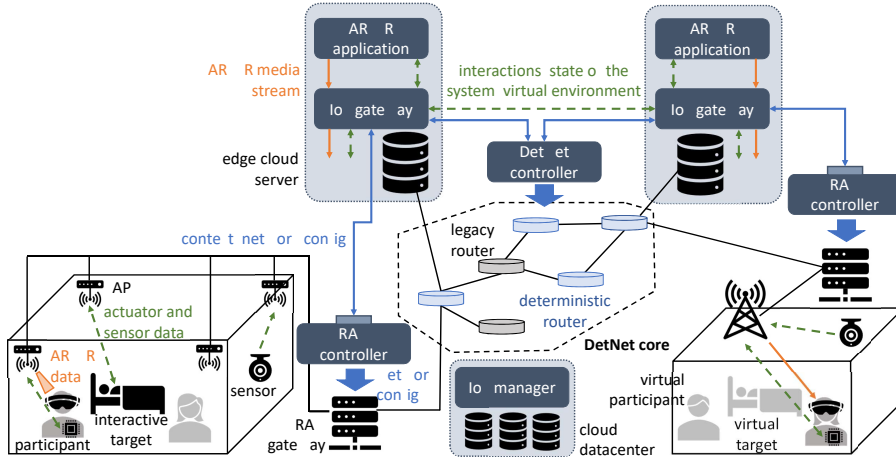
Fig. 3 The heterogeneous RAN and its devices are managed by a local controller, which interfaces with a local IoT gateway that ensures interoperability among IoT devices, and processes and disseminates data between locations over the low-latency DetNet core which supports end-to-end QoS and resource reservations to enable immersive interconnected AR/VR applications

QoS guarantees. The different locations are connected via the **deterministic networking (DetNet) core**, which serves the communication demands generated in all locations with a bounded and practically jitter-free latency. The distributed **IoT middleware** consists of a set of **IoT gateways** (one for each location) and a coordinating **IoT manager**. The middleware is responsible for managing the IoT devices and their data. Finally, the AR/VR applications use the IoT gateways to obtain and exchange context and application state information, and render the AR/VR data streams. The remainder of this section zooms in on each of the architectural building blocks and their function.

2.1.1 Heterogeneous RAN

The RAN provides last-hop wireless connectivity to the heterogeneous devices present in each location, including the participants' AR/VR HMDs, IoT sensors, (haptic) actuators, an (optional) application-dependent interactive target (e.g., a bed and mannequin in healthcare training), and wearables. The RAN consists of one or more access points (APs) and/or base stations (BSs) that support different wireless technologies, accounting for the various heterogeneous device requirements, including long-range low-power (e.g., IEEE 802.11ah, NB-IoT), traditional mid-band (e.g., IEEE 802.11ac, LTE), and millimetre-wave (mmWave) high-throughput (e.g., IEEE 802.11ay, 5G NR mmWave) connectivity. It uses URLLC principles to provide low-latency and reliability connectivity to all devices. By combining various wireless technologies, the network can cater to a variety of requirements, such as the long battery lifetime needs of IoT sensors, and the potentially extreme throughput

requirements, up to gigabits per second (Gbps), of high-definition AR/VR data streams [6]. The local network is managed by a **RAN controller** running on the gateway. This controller uses software-defined networking (SDN) principles to manage the flows going in and out of the local network. It enables handovers across APs/BSs and different technologies, QoS-aware flow routing and duplication, and device/network monitoring. Its northbound interface can be used by the **IoT gateways** to configure QoS requirements of flows, to obtain monitoring information, or to share context information (e.g., user locations) for use in managing the RAN.

2.1.2 Deterministic Networking core

The DetNet core offers a holistic approach towards ultra-low latency networking. It combines a deterministic networking protocol that offers jitter-free packet transfer across a network path with network virtualization capabilities to be able to offer isolated low-latency end-to-end network slices. Via the DetNet controller, the IoT gateways can reserve slices that provide guaranteed end-to-end latency, jitter and throughput between the locations participating in the same AR/VR session. It also allows the gateway and applications to monitor network performance.

2.1.3 IoT middleware

The IoT middleware provides an abstraction between network resources and devices (e.g., IoT sensors, haptic actuators, HMDs) on the one hand, and the AR/VR applications on the other. It consists of a centralized manager, and a set of distributed gateways. The gateway is envisioned to run on an edge cloud server, positioned close to the local network to minimize latency. It is responsible for managing the locally connected IoT devices (i.e., discovery, semantic interoperability, data collection and actuator control), and acting as a bridge between the applications and the network to translate application Quality of Experience (QoE) into network QoS requirements (for network configuration), expose processed sensor data as semantic context information, and generically control actuators, and implement mechanisms for self-awareness and resilience. The gateways at different locations can communicate with each other in a secure and privacy-ensured way. They are orchestrated by the manager, running on a remote cloud server. The middleware supports the integration of various micro-services, that add advanced data processing services on top of the default functionality (e.g., advanced localization algorithms, image processing, object recognition, text-to-speech).

2.1.4 Human-centric, context-aware AR/VR applications

The AR/VR applications are distributed among the different participating locations, running on the different edge cloud servers and communicating with the local and remote networks through the local IoT gateway. This will allow

the applications to obtain sensor data about the environment and users to enable context-aware and human-centric application adaptation, as well as allow them to influence the environment through (haptic) actuators. Rendering of AR/VR content is envisioned to be done locally, to be streamed towards the HMDs over the RAN, while the DetNet core is only used to exchange state information among the different local application instances. The applications interface with the middleware through a human-centric AR/VR application interface, which exposes all the functions necessary to implement context-aware interconnected remote AR/VR applications.

2.2 Potential use cases

AR/VR provides an immersive experience in environments which are either fictitious or dangerous-unlikely-unaffordable in real-life. Through interconnection, we open such an immersive world to collaboration, cooperation, and interactions across physically remote locations in real time. Aspects such as training, decision-making, or supervision can be easily done remotely with an unprecedented immersiveness and a huge impact across many application domains including industry, healthcare, entertainment, and security, among others. In addition to enabling multi-user interactions across physical locations, The proposed solution will greatly improve immersion and interactivity of the user with their (virtual) environment, by applying Tactile IoT principles and closely integrating sensors and actuators in the application loop. This will allow adapting the virtual environment to the actions, emotions, physiological parameters, and behavior of the users. Moreover, it will allow haptic feedback, resulting in a truly lifelike virtual presence. The incorporation of the emotional and physiological context information obtained from wearable sensors will result in a real human-centric AR/VR experience. We envisioned three main use cases for remote interconnected AR/VR, as described below. Figures 4 and 5 illustrate the possible cases in two different environments, i.e. Industry 4.0 and remote collaboration and social interaction.

2.2.1 Training

Interconnected AR/VR provides a unique opportunity to implement remote interactive training, where trainer and trainee(s) can interact without being in the same physical location. This relaxes scheduling constraints of high-demanded experts or, if taken to the extreme, of members of a team for team-work skills training. Other users, such as external observers, can participate in the activities through the eyes of the trainees. Potential vertical domains are healthcare (e.g., medical training), or industry (e.g., machine safety operations training).
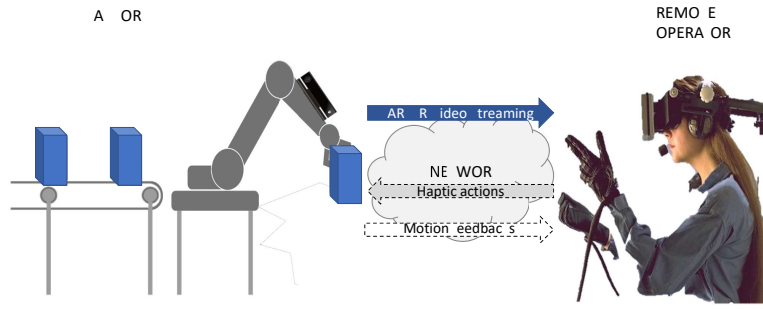
Fig. 4 Training/Collaborative decision-making use-case: A remote operator wears a HMD and haptic gloves and connects to the factory to control the quality of an assembly line. Other operators (in a collaborative decision-making case) or trainees (for training) could connect to the session and actively interact with the operator or follow their instructions.
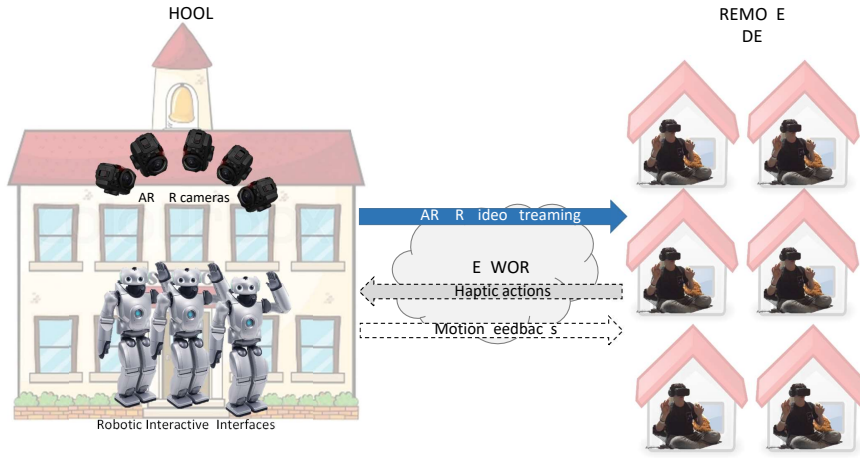

Fig. 5 Remote learning/social interaction: Due to a confinement, students must remain at home. At the school, a set of AR/VR cameras and robotic interfaces allow the remotely located students to interact with one another by means of the video feed and the robotic interfaces.

2.2.2 Collaborative decision-making

Rather than relying on one highly experienced individual, collaborative decision-making is based on a team of specialists working together. This approach is extremely beneficial in a variety of situations, but becomes a necessity in environments where, driven by high downtime costs, companies are forced to develop complex solutions with quick turnaround times. Moreover, specialists are often overburdened and expensive. Allowing them to participate remotely, would reduce their time wasted in transit, and enable them to quickly jump in to solve specific issues as the need arises. Potential vertical domains are healthcare (e.g., remote surgery, disease diagnosis), and industry (e.g., remote maintenance). Figure 4 illustrates an example of such remote maintenance.

A remote operator controls a robotic arm connected to an assembly line and makes different evaluations to test its performance.

2.2.3 Social interaction

Face-to-face communication and interaction among people is not always possible. It may not be desirable, because users may for example wish to interact with each other in a virtual rather than a physical environment. Alternatively, it may not be feasible due to mobility issues, or the physical distance just being too great. Potential vertical domains are healthcare (e.g., elderly participation), entertainment (e.g., multi-user AR/VR gaming), and media (e.g., holographic video call). Figure 5 presents an example of such a remote social interaction use-case. If due to some extreme conditions (for example, the COVID19 pandemic of 2020), students are required to stay at home and follow lectures remotely. Furthermore, they cannot interact with their school friends face-to-face. By means of an AR/VR infrastructure at home (HMD, headphones and haptic interfaces), the students could remotely interact with their friends (by means of some kind of robotic interfaces at the school) and follow the classes from the teachers. This could also be applied to seniors confined in their rooms of a nursing home. By menas of the interconnected AR/VR, they could interact with one other in a virtual reality environment and play games or talk without risking their health.

3 Heterogeneous RAN

As stated above, in order to support connecting heterogeneous IoT devices with vastly different requirements, the RAN is envisioned to offer different wireless network technologies using different frequency bands. A single AP or BS can offer one or multiple wireless technologies. Moreover, to achieve the high throughput requirements of future AR/VR applications, of up to several Gbps for a single data stream [6], there is a need to exploit the large amounts of uncongested bandwidth available above 6 Gigahertz (GHz). Especially the mmWave frequency bands have a lot of untapped potential in terms of available bandwidth.

To cover the wireless connectivity requirements of all types of devices present in the network, there is a need for several types of complementary radio access technologies (RATs) to co-exist: (i) low-power to connect battery-powered sensors and actuators (i.e., the GSM band or the 700 MHz band) , (ii) mid-band (i.e., 2 to 6 GHz for medium-throughput (up to hundreds of megabits per second (Mbps)) and (iii) mmWave (i.e., generally using frequency bands in the range of 24 to 70 GHz) to support Gbps throughput for high-definition AR/VR data streams. Two main alternative technological paths exist for offering these types of RANs: (i) the IEEE 802.11 family (i.e., Wi-Fi) of technologies using unlicensed spectrum, and (ii) the 3GPP mobile standards using licensed spectrum. The former has technologies available that cover low-power

connectivity (i.e., IEEE 802.11ah) [7], traditional mid-band connectivity (e.g., IEEE 802.11ax) [8], and mmWave connectivity (e.g., IEEE 802.11ay) [9]. The latter similarly covers low-power (e.g., Narrowband IoT (NB-IoT)) [10], mid-band (i.e., Long-Term Evolution (LTE) and 5G New Radio (5G NR)) [11], and mmWave (i.e., 5G NR) [12] connectivity.

3.1 State of the art

In order to achieve an end-to-end latency below 20 ms (cf. Figure 1), a latency in the order of at most a few milliseconds needs to be guaranteed in the RAN segment of the network. To achieve this, a variety of physical and MAC layer optimizations have been proposed in the context of 5G URLLC communications, such as novel packet structures, waveform designs, modulation and coding schemes, symbol detection methods, control signaling, and grant-free channel access solutions [13]. These methods are able to minimize the packet transmission and reception latency on the wireless link. However, unexpected packet loss may still result in significant drops in reliability and/or increases in latency and jitter. This is mainly due to retransmissions, which are traditionally used to improve reliability in wireless networks (e.g., using variants of Automatic Repeat Request (ARQ) protocols). However, retransmissions add additional latency (e.g., around 8 ms per retransmission in LTE [13]), which are not compatible with URLLC and Tactile Internet requirements. This problem is especially prevalent in mmWave networks due to the high propagation loss and attenuation by obstacles at high radio frequencies.

To enable URLLC in lossy conditions, researchers have been looking at alternatives to retransmissions to optimize latency and reliability. Such mechanisms include the use of higher frequency bands (i.e., mmWave), multi-access edge computing (MEC), and various forms of diversity [4]. While attenuation and propagation loss are major obstacles, mmWave frequencies offer the ability to reduce the on-air time, i.e., latency, via the ability to allocate large contiguous bandwidth segments as well as by allowing simultaneous transmission on an increased number of frequency channels. MEC brings computing and storage capabilities closer to the end user, supporting computational offloading and significantly reducing end-to-end latency [14]. Finally, diversity in terms of frequency (e.g., multi-channel), space (e.g., AP/BS densification, massive multiple-input and multiple-output (MIMO)), and interface (e.g., multiple radio access technologies (multi-RAT)) is seen as the main solution for the joint optimization of latency and reliability.

In terms of improving spatial diversity through densification, di Pietro *et al.* studied solving the intermittency problem in mmWave networks (i.e., link quality reduction caused by obstacles) by sending all packets over multiple links (i.e., APs) simultaneously [15]. They looked at MEC offloading tasks where the computational load is high compared to the communications load. This assumption, however, does not hold for AR/VR traffic, which has high bandwidth requirements. Moreover, their study was purely theoretical, assum-

ing perfect instantaneous knowledge about channel state information (CSI) and ignoring aspects related to AP selection, handovers, and packet repetition/duplication strategies.

Nielsen *et al.* studied the use of interface diversity (i.e., using multiple RATs simultaneously) [16]. They proposed a framework to theoretically evaluate the reliability and latency of various multi-RAT packet duplication strategies (i.e., full duplication, or redundant splitting). They showed that intelligent splitting provides the best results, with latency reductions up to 40% in low-throughput systems. As such, their conclusions may not be applicable to mmWave. Drago *et al.* did look at interface diversity for mmWave in combination with LTE [17]. They proposed a packet splitting mechanism that relies on network coding to enable packet reconstruction even if a certain percentage of packet fragments is lost. However, they focus mainly on reliability, reporting latencies between 10 and 20 ms. As such, their solution is not applicable to URLLC. De Schepper *et al.* proposed the ORCHESTRA framework as an enabler for interface diversity [18]. It supports packet duplication, splitting, load balancing, and seamless inter-AP/BS (i.e., horizontal) and inter-RAT (i.e., vertical) handovers at the network layer through a central controller and virtual medium access control (MAC) abstractions. The use of a virtual MAC makes the process programmable and transparent to the upper network layers.

In practice, performing seamless AP selection and handovers at mmWave frequencies is far from trivial owing to the need for directional connectivity (due to high propagation loss) and obstacle blockage (due to attenuation loss). This is especially challenging in dynamic environments, where users are mobile and obstacles may randomly appear. Sur *et al.* showed that this can lead to extreme spikes in latency (as high as multiple seconds) and drops in throughput [19]. To avoid this and to guarantee latencies below 1 ms, there is a need to perform proactive packet duplication and handovers between APs/BSs and technologies. Such proactive schemes require foresight into when signal quality will degrade and blockage will occur. An interesting research direction to tackle this is to use context information about the devices and their environment (e.g., location information). Palacios *et al.* recently evaluated the use of estimated location information to perform inter-AP handovers in IEEE 802.11ad mmWave networks [20]. However, their approach focused on throughput maximization and did not consider latency optimization.

3.2 Discussion and open challenges

Existing packet and flow management algorithms to enable millisecond URLLC in RANs focus on low-throughput applications. There is a clear lack of practical solutions that can guarantee reliable low-latency communications for high-throughput applications. Existing reliability techniques for high-throughput mmWave communications are not able to guarantee latencies below tens of milliseconds or even seconds when devices are mobile and blockage of the signal may occur. In order to bring down this latency to the order of 1 ms, there

is a need to perform predictive management of traffic flows based on accurate context information about the device and environment and exploit multiple forms of diversity and redundancy simultaneously. An interesting approach in this direction is the use of location information, such as studied by Palacios *et al.* [20]. Future research in this direction should go beyond pure throughput maximization and study how the location and movement of users and potential obstacles can be used for proactive handovers that guarantee a bounded latency. This requires studying the usefulness of different types of information (e.g., location, movement, mobility patterns), their needed accuracy, and the potential sources from which it can be obtained (e.g., wearables, environmental sensors).

4 Deterministic Core Network

The methodology for developing the core DetNet is based on a novel deterministic networking protocol that offers jitter-free packet transfer across a network path. It offers global timing and synchronized scheduling of packet traversal through router buffers. On top of that, the architecture envisions three mechanisms to further ensure deterministic low-latency routing: (a) a network coding mechanism for latency minimization, (b) an AI-driven mechanism for Application-Network collaboration, and (c) network slicing. Mechanism (a) enforces an adaptive network coding scheme for the AR/VR data flow that should be integrated into the underlying network technology. Mechanism (b) constitutes the interface between any application and the network. Finally, mechanism (c) allocates appropriate resources to meet latency requirements. In the remainder of this section, more details about the jitter-free routing protocol and the three extension mechanisms are provided.

4.1 State of the art

Time Sensitive networking (TSN) over large-scale and heterogeneous networks such as the Internet is receiving considerable research and industrial attention. On the one hand, the Resource reSerVation Protocol (RSVP) [21] is a protocol standardized by the Internet Engineering Task Force (IETF), which aims at providing guaranteed transmission performance for network flows by reserving forwarding resource on every forwarding hop along flow's path. This provides enough methods for network resource reservation for specific flows. However, as RSVP is a resource consumption method, the resource reservation on each node limits its scalability. On the other hand, the TSN standard is the IEEE 802.1Q standard technology to provide deterministic messaging on standard Ethernet [22]. TSN is centrally managed and provides guarantees of delivery and minimized jitter using (i) tight clock synchronization across the network and (ii) time scheduling methods for those real-time applications that require determinism. However, given that TSN is a Layer 2 technology,

it provides satisfactory performance in local networks, but cannot be applied to large-scale deployments.

To solve this, the DetNet Working Group (WG) of the IETF[1] aims to extend deterministic networking to forwarding at the Internet Protocol (IP) layer. To that end, the WG has proposed an overall architecture that standardizes the data plane and has discussed how aspects such as flow synchronization, management, control, and security should be rethought to enable multihop routing with low latency and virtually zero jitter. These initial discussions have indeed acknowledged the challenge of scaling TSN concepts to large-scale networks, i.e. involving many network devices, long distances, and many flows [22]. For instance, it is stated that per-flow queuing schemes that may be used in TSN are not applicable here due to the large state space requirements. Also, single-queue cyclic queuing and forwarding may suffer when frames arrive in the wrong cycle. To cope with the requirements of remote AR/VR applications, there is a need for the deterministic network research to go beyond those efforts.

The required flow synchronization and resource allocation functionalities can be modified to improve the scalability in terms of number of hops/flows. TSN envisioned two key enablers to achieve this: cyclic queuing and forwarding (CQF) and scalable deterministic forwarding (SDF) [23]. First, using CQF with three cyclic queues instead of one will mitigate the problems arising when frames arrive in the wrong cycle. Employing SDF will allow to rely on frequency synchronization rather than on the costly time synchronization across all routers. These methods, together with specific per-flow status maintained at the control plane, allow to make a fine-grained reservation of resources at every hop from source to destination ensuring minimal queuing and, thus, minimal jitter. Furthermore, to guarantee the scalability of the approach, the proposed methods can be implemented in a set of routers where nodes condition the traffic flows and ensure synchronization among them. The other routers can rely on legacy protocols as they only see the aggregate flows.

4.1.1 Network coding for low-latency

Current networking infrastructures, such as the Internet, fall short of the stringent requirements detailed by remote multi-user AR/VR, especially with respect to throughput and latency. In a best-effort infrastructure, such as the Internet, packet loss, especially occurring at core routers due to the network status, is one of the major factors that not only hampers throughput but most importantly contributes to the increase of delay through the retransmission policies that transport protocols use. One strategy to mitigate the problem is the use of source coding. However, this approach is bandwidth consuming, computationally intensive and allows little room for dynamic adaptation to the network status as well as the user requirements. A more distributed coding approach, such as network coding (network coding (NC)), provides a better
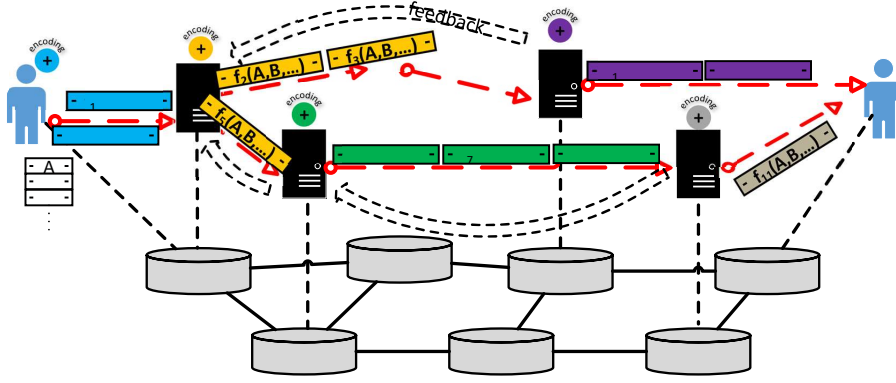
---

[1] https://datatracker.ietf.org/wg/detnet/about/

Fig. 6 Visualization of network coding to support low-latency end-to-end interconnected AR/VR.

fit for the context of remote AR/VR. In this case, coding is performed synergistically by all network nodes in the communication path. This essentially transforms those nodes into information regeneration points. This strategy comes with several advantages that can be beneficial for remote AR/VR:

1. Reaction to packet losses is prompter since the upstream node can react immediately, thus significantly reducing latency.
2. Rate adaptation based on the network status can be performed in a modular fashion by each intermediate node, thus economizing on bandwidth consumption. In other words, it is possible to use coding redundancy only where needed and not in the entire communication path.
3. It is possible to sustain high throughput even during congestion through multipath routing of the coded streams.
4. Combining network coding with multipath routing also adds security to remote AR/VR application, since it is virtually impossible to intercept the entire exchanged traffic, whereas any intercepted part cannot be decoded.

Implementing the NC approach in the interconnected AR/VR context requires a novel architectural approach that will allow flexibility. This can be done using an overlay network implemented with a set of cloud servers. The principle is shown in Figure 6. The end-to-end communication is established as a set of consecutive transport links to provide transparency to the underlying oblivious best effort infrastructure where NC is performed not only at the source, but also in every cloud server (intermediate node). These actively participate in the coding process with two critical functions. First, they provide feedback about the coding state to the upstream server in order to allow very fast reaction to packet losses. Second, they actively fine-tune the coding rate to the downstream server to adapt to the underlying infrastructure performance and at the same time economize on the bandwidth.

4.1.2 AI-driven application-network collaboration

The network effect in delay-sensitive applications has been studied mostly from a single perspective, i.e., either from the network side or the application side. Moreover, very few existing studies have covered and evaluated audio and video streaming, whereas the AR/VR case remains largely unexplored.

From the application side, studies on adaptive-quality media streaming for real-time communication are well known and have had a huge impact on the way we use the Internet [24]. In these cases, the application continuously monitors the path quality (bandwidth, latency) on its own end and subsequently reverts to a better or worse video/audio quality accordingly. This effect is noticeable, since short pauses or even connection drops are very common. Such approaches may be conditionally acceptable for every-day web browsing or tele-conferencing, but not for an AR/VR setting [25].

From the network side, deploying the shortest paths between connectivity end-points and using protocols that are naturally forwarded within a network with high priority cover most of the related studies [26]. Furthermore, the few existing solutions that consider a synergy between the network and the applications have been studied in vitro, i.e., in controlled network environments only, such as local area networks (LANs), leased lines or experimental testbeds (such as SDN-based [27]). Even in these limited exploratory studies, the security aspect is ignored completely, raising questions on their applicability to real-world conditions. Moreover, they do not study AR/VR applications either. Finally, it is worth noting that the use of AI/ML methods for computer networking has been proposed in multitude of occasions [28], but with limited success thus far. The main reason is that the widespread algorithms (e.g., fully connected, convolutional, or recurrent neural networks) are not amenable to the graph-structured nature of computer networks. Techniques applicable to such problems exist (such as [29]), but have not been applied successfully in networks.

With perfect and fast knowledge exchange between the network and the application, network resource allocation is no longer a matter of converging heuristics, but rather a deterministic process that results into final data routes and constant bit rates over them. To that end, modern network architectures, such as SDN [30] or knowledge-defined networking (KDN) [31], which have reverted to a centralized network monitoring and control, can be used. Monitoring, resource allocation and user access control are therefore modeled as applications, which are subsequently translated to network routing table rules and packet scheduling programs. Both aspects can be effectively driven by Artificial Intelligence (artificial intelligence (AI)) methods.

An AR/VR-to-network mediation service, based on a predefined logic [32], will (i) maintain an updated view of the application requirements and the network state in real-time, (ii) adaptively match application requirements to network configurations, and (iii) if a match is not attainable, it will automatically negotiate application modifications to resolve the issue in minimal time. This logic decides the deployment of proactive policies, such as over-

provisioning resources to AR/VR users, and performs traffic duplication to account for unforeseen network events, such as failures and congestion owed to sudden link-flooding. Reactive policies are instead based on foreseeable application or network events (e.g., an increase in AR/VR users, network traffic increase expectation), or events that have just occurred. A crucial point here is the methodology to, first, maintain the view of the network state and, then, make accurate predictions of the outcome of the decisions taken. In this regard, AI techniques can be applied [33] that will receive as input the current state of the network or application and yield an occurrence probability of events that require handling. One possible approach could be to implement it taking ground on the recent KDN approach and particularly the use of graph neural networks (GNNs) for network-level modeling and reinforcement learning (RL) to obtain optimal decision policies [34].

4.1.3 Reinforcement learning-based network slicing

Network slicing is considered key in enabling network operators to run overlay networks over radio, transport and core domains. Network slicing will allow automatic configuration of resources based on demands from application providers without the need for manual intervention and provide isolation and special service treatment for applications requiring predictable data rates and low latency [35]. Concepts such as tuning the underlying network for specific use cases has been used in existing network topologies. However, network slicing separates itself from such inflexible and dedicated assignments by providing elastic and dynamic set up of system resources. 3GPP Release 15 has already defined basic operation of end-to-end network slices, and with upcoming Release 16 and 17 enhancements such as integration with ETSI network function virtualization (NFV) management and orchestration (MANO) framework, RAN slicing specification to enable vendor-agnostic implementation are expected. Slices need to be monitored and resources should be adjusted using AI algorithms continuously to ensure that the specified service level agreements (SLAs) are met.

Recent research has shown the potential to use of RL to automate the configuration of network slices [36] [37] [38]. An RL agent interacts with its environment to identify the optimal policy through observing the state transitions and perceived rewards based on its actions. However, it will be challenging to implement RL-based algorithms on a commercial large-scale operator network due to the existence of large number of state and action spaces, which takes too long to reach to an optimal policy. Deep reinforcement learning (DRL) is preferred in such complex networks to overcome the long convergence times. [39] gives a survey of latest research on DRL applications to communication and networking domains and discusses complications that can arise from operating a multi-agent RL system where the learning is slowed down due to interaction between agents and resulting increased state space.

Distribution based RL algorithms improved the estimation of action values [40]. Generative adversarial network (GAN) is a class of machine learning

algorithm known to generate new data by approximating the distributions of the underlying training set. [41] proposes combining GAN with deep distributional RLs to learn the optimal policy for network slicing resource management that takes into account the demands.

Even though they provide very accurate predictions, highly non-linear nature of deep learning (DL) algorithms make it hard to interpret the proposed decisions. Automated decisions regarding critical network management for mobile operators will be required to be explainable in case of system failures or simply to ensure system integrity. A method relying on Taylor expansions without the need for retraining of the neural network is proposed to decompose the input output relationship for image classification[42].

Both DL and DRL are vulnerable to intentionally designed attacks to fool the machine learning models [43]. Automated configuration and resource management of network slices should take into account such vulnerabilities and introduce fail-safe mechanisms.

4.2 Discussion and open challenges

The network latency is a major factor affecting the performance of AR/VR applications for true remote and real-time use. This section outlined a set of promising mechanisms towards minimizing the network latency and mitigating its negative effects on overall AR/VR experience. At the application layer, a close application-to-network collaboration was proposed, based on successful existing outcomes [27, 44]. At the network layer, a latency reduction oriented network coding scheme was presented and the interplay between the two mechanisms was outlined. An AI system was also proposed, which will monitor and tune the whole system based on prior observations.

However, true low-latency, real-time performance remains an open issue that can only be studied via real-world system try-out. While any new mechanism introduces potential latency mitigation, their cross-collaboration and interplay can be a source of latency on its own. Regarding the use of NC, the idea of utilizing it to reduce delay and increase throughput has been proposed in the past, e.g. [45, 46]. However, the use of NC has never been tested over long distance communication links and under the stringent requirements of URLLC. The challenge here is designing the architecture that will enable an efficient deployment of NC over the long communication distances required by a remote AV/VR application. Moreover, the computational overhead induced by NC and AI techniques must also be taken into account in terms of potential latency introduction. Setting up pilot testbeds to develop and tune the whole system is an imminent challenge, in order to evaluate real-world application prospects.

5 IoT middleware

The RAN controller and gateways provide the technical basis for communication. To manage the technical connections and the information transmitted at different layers, complementary services are needed. These services have to deal with the identification of different resources, the management of connection at higher levels, the management of data and access to it (privacy, security), and finally data analytics and information extraction. These tasks are handled in IoT middleware and IoT cloud platforms, which aim at managing and hiding the complexity and heterogeneity of distributed infrastructures. This type of software has developed from early transaction monitors into general purpose service oriented platforms of today.

5.1 State of the art

Given the increasing complexity of variability among IoT platforms, the current tendency is to opt for a layered architecture. One of the most common architectures is the one defined in [47] and [48]. In it, the control is distributed in three layers: (1) Things; (2) Network; and (3) Application. The Things layer takes care of receiving the data from and providing control of the connected devices. The Network layer is in charge of the translation between different protocol as well as mediating data exchanges. Finally, the application layer processes the received data and converts it to useful information in an application context.

Also driven by complexity, the IoT architectures can be built on top of pre-existing platforms, using middleware infrastructure to provide the desired communication between the Things and the Network layers. A significant attention has been given to designing middleware, which aims to facilitate the efficient management and control of active infrastructure elements and their optimal coordination during system operation. This results in enhancement of the systems' interoperability in terms of quality and reliability. Furthermore, the design of middleware is envisioned to support features such as heterogeneity, mobility, scalability, multiplicity, and security [49]. IoT middleware systems can be classified according to their placement. For example, at gateway layer, platforms such as KURA, OpenHAB or AllJoyn [50], have the objective of collecting information from the things layer and to adapt it to the network layer at the gateway. On the other side of the gateway, IoT middleware focuses on gathering data, managing the connections, gathering and representing information; furthermore, privacy and security aspects are addressed. Modern IoT platforms add functionality to the hardware and application layer like capabilities for edge data processing or complex data analytic algorithms. One clear example of such platforms is FIWARE [2]. Finally, IoT cloud platforms such as Amazon Web Services or Microsoft Azure Cloud[51] are targeting the

---

[2] https://www.fiware.org/

collection of data in the cloud and data analytics at different levels. Moreover, they often provide overall-frameworks including templates for gateway adapters, middleware, and cloud services.

Independently from its placement, one key aspect of the design of middleware is to ensure the dynamic identification and seamless dynamic integration of a large number of heterogeneous data sources and sinks [52]. This is usually done by means of adding a semantic layer into the IoT reference architecture [53]. Semantic modelling for IoT middleware comes in two main flavors which complement one another. First, semantic device resolution models identify the device type and all related properties (e.g., services description, profiles, etc). This model is tied to the physical device. It does thus enable to query for all device properties. Second, semantic contextual models contain knowledge related to application context. This information includes but it is not limited to the inter-device relations, usually in the form of domain or application specific rules. This model is tied with context-aware reasoning used to make decisions depending on the real-time situation. This type of semantic models automate the interoperability across different existing standards and device Application Programming Interfaces (API). Furthermore, it extends traditional object discovery methods (based on entity type or entity state in static environments) with the possibility to discover objects in dynamic context environments [54]. Several semantic based IoT middleware systems have been developed, an example of an earlier system being Linksmart [55] while VICINITY [56] is an example of more recently developed platform.

5.2 Discussion and open challenges

In order to combine the benefits of IoT and AR/VR, we see the two following challenges: Integration and management of direct connections between two AR/VR devices and the semantification of AR/VR data for its interoperability as offered by appropriate IoT platforms.

Direct communication between devices, has been addressed by IoT platforms that permit the management of peer-to-peer (P2P) connections. Examples for P2P capable platforms are Nabto[3] and VICINITY[56]. However, they are far from allowing the management of both IoT and AR/VR devices.

To solve the second problem, we argue that there is need for a holistic solution that combines data from traditional IoT devices (sensors) with AR/VR data in a single platform, combining semantic interoperability with state-of-the-art techniques for achieving a high quality of service at the connection layer. Existing combinations of the IoTand AR/VR solutions such as [57] adapt some features of the IoT, i.e. direct connections, but fail to provide a comprehensive IoT platform with all its services, in particular, the semantification of data and interoperability at semantic layer. This has been addressed by some, rather research driven, IoT platforms, e.g. VICINITY[56].

---

[3] https://www.nabto.com/

6 Interactive, personalized and human-centric AR/VR applications

Haptic feedback and sensor monitoring enhance the immersiveness and interactivity of AR/VR applications. However, due to the early stages of research of immersive technologies, the successful integration of haptics, sensor data within AR/VR applications is still far from unlocking its full potential. Three elements currently drive the technology. These are the low latency transport and application layer protocols, the personalization of content and the perception-driven performance evaluation of applications. In this Section, first state of the art on these three aspects is provided. Further, discussion on the open challenges and hints to a new architecture are provided.

6.1 State of the art

In order to ensure interconnected and interactive AR/VR applications, three different aspects of the problem have been addressed on the state of the art. First and foremost, there is a need for low latency transport and application protocols that can cope with the future DetNet, the Heterogeneous RAN and the middleware IoT. Second, mechanisms for adapting and personalizing the experiences would enhance the experience of a broad range of users, thus making the applications suitable for the use-cases presented in Section 2. Finally, there is a need for mechanisms to assess the perception of the users of the applications so that changes can be fastly implemented. In the following three subsections, state of the art on these three aspects is provided.

6.1.1 Low latency transport and application protocols for AR/VR

The most common protocols used in the research literature for haptic communication over the Internet are the TCP and UDP protocols. TCP minimizes the effects of jitter, but its reliability mechanisms prevent it from being applicable as real-time protocol. On the other hand, UDP, despite minimizing the network delay, does not meet the reliability requirements of most haptic applications, especially in heterogeneous networks under packet congestion [58]. To fill the gap between the two, several approaches have appeared in literature. The Smoothed Synchronous Collaboration Transport Protocol (S-SCTP) uses a receiver buffer and time-stamps on the packet headers to cope with the jitter. The Interactive Real-Time Protocol (IRTP) prioritizes messages to send them either through TCP or UDP according to their significance. A hybrid solution, a protocol that tries to leverage the advantages of other protocols is the Hybrid Multicast Transport Protocol (HMTP) and is mainly used for realizing haptic collaboration in virtual environments. However, with the exception of TCP and UDP, which represent maximum reliability and minimum packet header overhead respectively, all other possible protocols need to balance the trade-off between reliability and latency.

This trade-off is also found in the case of the transmission of AR/VR media. AR/VR streaming has been performed almost exclusively by TCP-based solutions, where resiliency is prioritized over latency [59]. UDP solutions are better suited. However, these solutions have shown problems to maintain quality levels (which is necessary for immersiveness in VR). Approaches such as QUIC (Quick UDP Internet Connections) and WebRTC (Web Real-Time Communication) are improving the performance of UDP-based solutions [60]. Furthermore, given the stringent latency requirements, there is a need to reduce the AR/VR content compression time to its limits. Therefore, one possible solution would be to stream the content frame-by-frame and only encode at the destination [61]. The applications of these protocols and algorithms to the case of AR/VR are still open for exploration.

Apart from only streaming haptic and AR/VR data through the communication channel, the QoE requirements demand the synchronized transmission of both audio and AR/VR video data. Therefore, a fundamental aspect of the application layer protocols is the temporal management of the data streams. Several attempts have been made in this direction. In [62], haptic communication is achieved by employing the Session Initiation Protocol (SIP) to establish a teleoperation session and manage haptic transport streams. The multimodal communication framework of PAHCP [63] synchronizes data by using the Network Time Protocol (NTP) while graphics and haptic data are transmitted with Virtual Network Connection (VNC) and PAHCP respectively. The MPEG Media Transport, is an application layer transport protocol used in [64] for the purpose of multi-modal data transmission on 3D tele-immersion environments (3DTI). While multiplexing, if no force data are to be sent then the video data are prioritized. Then, by assuming a constant bitrate connection, packet delay can be computed and used for correctly demultiplexing the data stream. These approaches have provided promising results for simulations or experimental evaluations where the network conditions were mostly assumed constant. However, they have not been applied to real applications and have not dealt with the stringent requirements of AR/VR applications.

6.1.2 Personalization of AR/VR applications

A self-adaptive system requires a control loop which allows it to adapt to its context. Based on this feedback, the system makes intelligent decisions. Previous research has shown that machine learning (ML) techniques can be used for creating self-adaptive AR/VR systems [65]. Each session type belonging to an AR/VR system is designed for one task that can be altered by setting a fixed set of parameters, i.e., parameterized exercises. A self-adaptive, personalization system needs to understand the effect of each exercise and parameter on every user profile. Two different paths to build such self-adaptive system have been found in literature [66]: (1) knowledge-driven; and (2) data-driven. On the one hand, a knowledge-driven design requires a human expert to define the influence of each parameter and each context variable on the user. For example, by formulating a set of rules, the expert decides the behavior of the learning

system. On the other hand, a data-driven approach requires a lot of data and results in actions that are difficult to interpret by a human. Thus, enabling effective personalized immersive AR/VR applications require for methods able to map the sensor data information to the virtual environment (VE).

Most HMD systems often work with head motion tracking sensors or eye tracking for enhanced realism [67]. This sensor information can be used to provide accurate renderings of images on the displays. One clear example of using the HMD sensors is for adaptive tile-based streaming, where the sensor information is utilized to predict where the user will look next [59]. The analysis of the sensor information and haptics can be used both as input and output of the AR/VR application. On the one hand, it can help in assessing the user during the execution of the exercise. On the other hand, this assessment can, in turn, help with adaptations to the environment (content) for the following sessions. Analysis of motion is particularly useful for assessing motor skills. It enables to provide feedback on specific motoric tasks. For instance, by providing auditory or haptic feedback to notify the user of bad posture or wrong movements. However, the integration of haptics within the VE further that for entertainment is still largely unexplored. Some works have shown the application of ML techniques to automate the interaction between haptics and the visual feed. Such are the neural network-based data-gloves for automatic gesture detection of Luzanin et al. [68] or the predictive models enabling realistic resistance on airplane controls of Yamashita et al. [69]. These approaches, although promising, are still far from the full haptic-AR/VR automatic interaction.

6.1.3 Human-centric Quality Evaluation

Interconnected AR/VR experiences are built from multiple input and feedback channels to improve the perception of the user, in terms of feeling of immersion and interactivity. In order to make this experience as natural as possible, however, the system's QoS parameters need to be kept within the boundaries of acceptance of human perception. Delay, jitter and data loss, for example, are perceived rather differently depending on the sensorial type. When requirements are not fulfilled, the user experience will feel less authentic and the QoE decreases, possibly even inducing cybersickness in the most severe cases [70].

Currently, most research is limited to delay while jitter and data loss are barely researched [71]. In addition, most test setups include only one or two modalities, while studies on truly multi-modal systems are rather scarce. As a result, the synchronization and prioritization of the simultaneously transmitted signals is barely investigated within the light of QoE maximalization. Thus, it is not hard to imagine that (a lack of) synchronization between the multiple feeds will have a highly influence on the user experience. As most of the current network infrastructure consists of reliable, high-speed connections, synchronization mistakes typically arise within the local network. Especially when using wireless connections (which is preferred to enable maximal free-
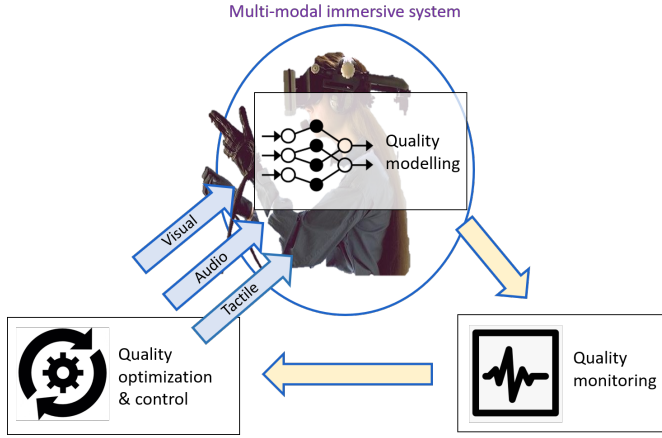
Fig. 7 QoE management loop for interconnected AR/VR applications [70].

dom of movement), the above QoS parameters might become more stringent. Managing these applications will require to shift the focus from the network (QoS) to the human. Therefore, the expectations of the users, i.e. the QoE, will drive the application and network decisions.

QoE is defined as the degree of delight or annoyance of the user of an application or service, based on both objective and subjective psychological measurements [72]. QoE management is performed by means of three processes as shown in Figure 7. First, quality modelling provides an estimation of the client-side quality as perceived by the end-user. Second, quality monitoring focuses on analyzing the services and gaining understanding of the different factors that influence the quality of the application. Third, the quality optimization and control part aims to optimize the QoE over time.

The most straight-forward procedure to assess the effects of multimedia feeds on the user's perception has traditionally been subjective evaluations [73]. Subjective studies and evaluations of multimedia services are typically performed by means of experimental setups in laboratory environments, where a few dozen people with varying demographic backgrounds and limited knowledge about signal processing and encoding rate some multimedia content. For each presented sequence, users grade the quality on a certain scale (usually between 1 and 5). The average score over all subjects is called the Mean Opinion Score (MOS) [73].

As the degree of immersiveness of the application rises assessing the QoE of the end-user by means of MOS becomes more difficult. In 360° videos, for example, one must take into account that a user only sees a limited part of the 360° hemisphere (the viewport) at each instant. Therefore, users might watch different portions of the video during playback, which makes it rather difficult to compare quality scores among users and to combine them to a single average score per video. At this point in time, no standardized assessing methodology exists to solve this issue.

It becomes even more complex when additional sensory inputs, and tactile feedback in particular, are added to the experience. Due to the complex combination of sensorial data types that influence the user, MOS becomes rather infeasible to define the quality. Therefore, it is beneficial to define the effectiveness of the interactive system in terms of the ability to perform certain tasks, e.g. the ability to pick up an object, localizing an object etc. [72] and the feeling (or rather the absent) of cybersickness. Research towards the actual design of such performance tests is currently scarce, however.

Most studies within existing, scientific literature are limited to subjective studies on limited groups of users. Although such studies provide an accurate view on the end-user quality perception, they are rather limited in scalability and inefficient in terms of time and money. In addition, multimedia systems benefit from real-time quality assessment to allow for dynamic adaptation of the system parameters to optimize end-user experience. Therefore, objective metrics are more tailored for this task. However, the amount of research concerning overarching objective QoE metrics for AR/VR experiences is limited.

More research has been conducted on each of the individual feeds of the system. Especially for traditional, 2D video applications a wide variety of objective metrics exist. Some of them are a pure mathematical comparison of signals (e.g. Peak Signal-to-Noise Ratio (PSNR)), while others take the Human Visual System (HVS) into account (e.g. Video Multimethod Assessment Fusion (VMAF)[4]). The latter are often based on ML techniques such as Support Vector Machines (SVMs) to combine multiple metrics. Limited attempts exist to expand this towards 360° and holographic content, such as the Point cloud Quality Rating (PQR) for point clouds [74] or Weighted to Spherically uniform PSNR (WS-PSNR) [75] and the gaze driven model for adaptive tile-based streaming of van der Hooft et al. [76] for omnidirectional video content. Alexiou et al. [77] presented an study of the correlation between objective and subjective evaluation for the case of volumetric media. They showed that while some correlation can be found, traditionally used objective metrics need to be heavily tuned to make them ready for the analysis of volumetric media.

For auditory feeds, there also exist a handful of metrics such as ViSQO-LAudio, ViSQOL Speech and AMBIQUAL [78] for traditional audio, speech and ambisonics (i.e. a full sphere audio surrounding technique) respectively.

The haptic feed is the least explored path of the three senses. The limited amount of haptic-related, objective metrics is based on generic metrics for evaluation of signal quality in general, e.g. Mean Squared Error (MSE) or Signal-to-Noise Ratio (SNR). Haptic Perceptually Weighted PSNR (HPWP-SNR) [79] is probably the most known example of a haptic objective metric. Another, more recent one is Haptic SSIM (HSSIM) [80], which is an adaptation of the classic Structural Similarity Index (SSIM) [81] for video quality estimation to the specific case of haptic feedback signals.
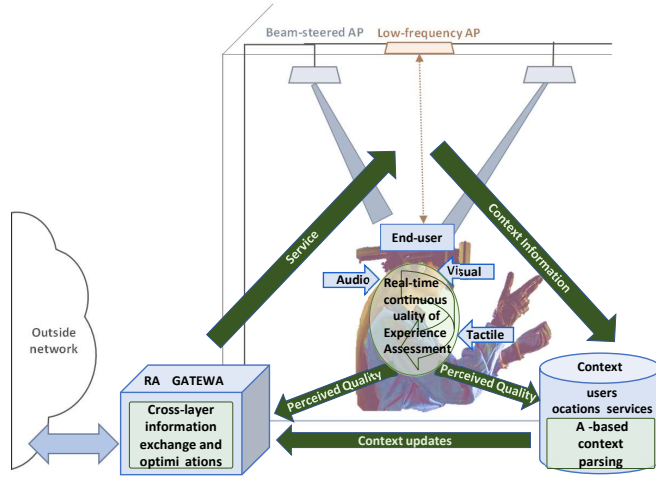
---

[4] https://github.com/Netflix/vmaf

Fig. 8 End-user application challenges to enable interconnected AR/VR applications.

6.2 Discussion and open challenges

In order to enable full interactivity and interconnected AR/VR applications, we find that the three elements of the system presented in the previous section (streaming, personalization and quality) need to be thoroughly improved. Furthermore, they need to be integrated into a close architecture, devising novel algorithms and methods. Figure 8 presents a proposal for this management loop of future human-centric AR/VR applications.

1. **Cross-layer architecture enabling information exchange and low latency protocols:** Cross-layer information exchange breaks the boundaries that traditionally exist among the protocol layers in the OSI stack. This allows higher layer protocols to take into account and anticipate the effects of beam training/steering and MAC scheduling on end-to-end performance. For example, at transport layer, smart flow and congestion control (by means of smart retransmissions, buffering and packet reordering) can be performed. Also, the network layer could benefit from cross-layer information to decide on the best route for the packets given the multipath, the context information and the required quality. Transport and application layer protocols need to become faster where UDP-based solutions are aided by control mechanisms to provided the required resilience. One promising approach is QUIC, which could be used both to provide a per-frame transmission [61] as well as for leveraging the multiple paths available for transmissions [82].
2. **Context awareness and personalization:** Context awareness enables the algorithms to take into account additional context information about the users, applications, and network (e.g., user location and speed, required throughput, detected obstacles). However, the beam steering mechanism in

beam-steered networks is unpredictable in terms of link establishment latency and achievable throughput. This unpredictability in large part comes from the obstacles in an environment, as well as due to mobility of the clients. These open a plethora of research questions. First, given the massive amount of potentially generated context, it is first fundamental to understand which sources of information are needed. Second, given the possible inaccuracies of information sources (e.g. location estimates are burdened with localization errors, clients' mobility patterns can become obsolete), it is necessary to learn how to cope with such context errors. Finally, based on the selected and processed context, there is a need to create an estimation quality model for each potential device-AP pair. Furthermore, full awareness of the context will aid to provide solutions to enable the full personalization of the environments, both to the user as well as to the conditions.

3. **Quality of Experience management:** Satisfying the needs of users of immersive media services requires not only to manage the network Quality of Service (QoS), but also to address the user's QoE. Nowadays, the perception of users of multimodal immersive applications (visual, audio and tactile) is only measured by means of subjective evaluation (a group of users rate the performance of the systems using a rating system called Mean Opinion Score). Despite providing the most accurate feedback, they are time-consuming (human testers need to be involved), and expensive to conduct. Thus, even if they are necessary as a benchmark, there is a need for novel objective evaluation models, that can provide a measurement of the user's QoE in real-time. These methods need to be valid not only for the AR/VR multimedia feedback but also need to include the haptic communications as well, which is widely unaddressed in the literature.

7 Conclusions

Enabling interactivity and interconnectivity for AR/VR applications imposes immense challenges on the currently available technologies, networks and applications. The purpose of this paper has been to bring forward four areas of the transmission chain that will need deep analysis. These are the heterogeneous radio area network, the deterministic core network, the IoT middleware layer and the end-user device transmission protocols and perception modelling. For each of them, we delve into the current state of the art and challenges that need to be addressed before the dream of remote AR/VR interaction can become reality.

References

1. Bastug, E., Bennis, M., Medard, M., Debbah, M.: Toward interconnected virtual reality: Opportunities, challenges, and enablers. IEEE Communications Magazine **55**(6), 110–117 (2017). DOI 10.1109/MCOM.2017.1601089
2. Aijaz, A., Dohler, M., Aghvami, A.H., Friderikos, V., Frodigh, M.: Realizing the tactile internet: Haptic communications over next generation 5G cellular networks. IEEE Wireless Communications **24**(2), 82–89 (2017)
3. Mangiante, S., Klas, G., Navon, A., GuanHua, Z., Ran, J., Dias Silva, M.: VR is on the edge: How to deliver 360-videos in mobile networks. In: Workshop on Virtual Reality and Augmented Reality Network (VR/AR Network), pp. 30–35 (2017). DOI 10.1145/3097895.3097901
4. Elbamby, M.S., Perfecto, C., Bennis, M., Doppler, K.: Toward low-latency and ultra-reliable virtual reality. IEEE Network **32**(2), 78–84 (2018). DOI 10.1109/MNET.2018.1700268
5. Finn, N.: Introduction to time-sensitive networking. IEEE Communications Standards Magazine **2**(2), 22–28 (2018)
6. Zhang, H., Elmokashfi, A., Yang, Z., Mohapatra, P.: Wireless access to ultimate virtual reality 360-degree video at home. In: International Conference on Internet of Things Design and Implementation, pp. 271–272 (2019)
7. Baños-Gonzalez, V., Afaqui, M., Lopez-Aguilera, E., Garcia-Villegas, E.: IEEE 802.11ah: A technology to face the IoT challenge. Sensors **16**(11) (2016). DOI 10.3390/s16111960
8. Khorov, E., Kiryanov, A., Lyakhov, A., Bianchi, G.: A tutorial on IEEE 802.11ax high efficiency WLANs. IEEE Communications Surveys and Tutorials **21**(1), 197–216 (2019). DOI 10.1109/COMST.2018.2871099
9. Zhou, P., Cheng, K., Han, X., Fang, X., Fang, Y., He, R., Long, Y., Liu, Y.: IEEE 802.11ay-based mmWave WLANs: Design challenges and solutions. IEEE Communications Surveys & Tutorials **20**(3), 1654–1681 (2018). DOI 10.1109/COMST.2018.2816920
10. Beyene, Y.D., Jantti, R., Tirkkonen, O., Ruttik, K., Iraji, S., Larmo, A., Tirronen, T., Torsner, J.: NB-IoT technology overview and experience from cloud-RAN implementation. IEEE Wireless Communications **24**(3), 26–32 (2017). DOI 10.1109/MWC.2017.1600418
11. Lien, S.Y., Shieh, S.L., Huang, Y., Su, B., Hsu, Y.L., Wei, H.Y.: 5G new radio: Waveform, frame structure, multiple access, and initial access. IEEE Communications Magazine **55**(6), 64–71 (2017). DOI 10.1109/MCOM.2017.1601107
12. Lopez, A.V., Chervyakov, A., Chance, G., Verma, S., Tang, Y.: Opportunities and challenges of mmWave NR. IEEE Wireless Communications **26**(2), 4–6 (2019). DOI 10.1109/MWC.2019.8700132
13. Parvez, I., Rahmati, A., Guvenc, I., Sarwat, A.I., Dai, H.: A survey on low latency towards 5G: RAN, core network and caching solutions. IEEE Communications Surveys and Tutorials **20**(4), 3098–3130 (2018). DOI 10.1109/COMST.2018.2841349
14. Barbarossa, S., Ceci, E., Merluzzi, M.: Overbooking radio and computation resources in mmW-mobile edge computing to reduce vulnerability to channel intermittency. In: European Conference on Networks and Communications (EuCNC) (2017). DOI 10.1109/EuCNC.2017.7980746
15. di Pietro, N., Merluzzi, M., Calvanese Strinati, E., Barbarossa, S.: Resilient design of 5G mobile-edge computing over intermittent mmWave links (2019)
16. Nielsen, J.J., Liu, R., Popovski, P.: Ultra-reliable low latency communication using interface diversity. IEEE Transactions on Communications **66**(3), 1322–1334 (2018)
17. Drago, M., Azzino, T., Polese, M., Stefanović, C., Zorzi, M.: Reliable video streaming over mmWave with multi connectivity and network coding. In: International Conference on Computing, Networking and Communications (ICNC), pp. 508–512 (2018)
18. De Schepper, T., Bosch, P., Zeljkovic, E., Mahfoudhi, F., Haxhibeqiri, J., Hoebeke, J., Famaey, J., Latre, S.: ORCHESTRA: Enabling inter-technology network management in heterogeneous wireless networks. IEEE Transactions on Network and Service Management **15**(4), 1733–1746 (2018)

19. Sur, S., Venkateswaran, V., Zhang, X., Ramanathan, P.: 60 GHz indoor networking through flexible beams: A link-level profiling. SIGMETRICS Perform. Eval. Rev. **43**(1), 71–84 (2015)
20. Palacios, J., Casari, P., Assasa, H., Widmer, J.: LEAP: Location estimation and predictive handover with consumer-grade mmWave devices. In: IEEE Conference on Computer Communications (INFOCOM), pp. 2377–2385 (2019)
21. Braden, B., Zhang, L., Berson, S., Herzog, S., Jamin, S.: Resource ReSerVation Protocol (RSVP) – Version 1 Functional Specification. RFC 2205. (1997)
22. Nasrallah, A., Balasubramanian, V., Thyagaturu, A., Reisslein, M., ElBakoury, H.: TSN Algorithms for Large Scale Networks: A Survey and Conceptual Comparison (2019)
23. Messenger, J.L.: Time-sensitive networking: An introduction. IEEE Communications Standards Magazine **2**(2), 29–33 (2018)
24. Kua, J., Armitage, G., Branch, P.: A survey of rate adaptation techniques for dynamic adaptive streaming over http. IEEE Communications Surveys & Tutorials **19**, 1842–1866 (2017)
25. He, D., Westphal, C., Garcia-Luna-Aceves, J.: Network support for ar/vr and immersive video application: A survey. pp. 359–369 (2018). DOI 10.5220/0006941703590369
26. Abdallah, M., Griwodz, C., Chen, K.T., Simon, G., Wang, P.C., Hsu, C.H.: Delay-sensitive video computing in the cloud: A survey. ACM Trans. Multimedia Comput. Commun. Appl. **14**(3s) (2018). DOI 10.1145/3212804. URL `https://doi.org/10.1145/3212804`
27. Lakiotakis, E., Liaskos, C., Dimitropoulos, X.: Improving networked music performance systems using application-network collaboration. Concurrency and Computation: Practice and Experience (2018). DOI 10.1002/cpe.4730
28. Wang, M., Cui, Y., Wang, G., Xiao, S., Jiang, J.: Machine learning for networking: Workflow, advances and opportunities. IEEE Network **PP** (2017). DOI 10.1109/MNET.2017.1700200
29. Battaglia, P.W., Hamrick, J.B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V.F., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., Gülçehre, Ç., Song, H.F., Ballard, A.J., Gilmer, J., Dahl, G.E., Vaswani, A., Allen, K.R., Nash, C., Langston, V., Dyer, C., Heess, N., Wierstra, D., Kohli, P., Botvinick, M., Vinyals, O., Li, Y., Pascanu, R.: Relational inductive biases, deep learning, and graph networks. CoRR **abs/1806.01261** (2018). URL `http://arxiv.org/abs/1806.01261`
30. Davie, B., Koponen, T., Pettit, J., Pfaff, B., Casado, M., Gude, N., Padmanabhan, A., Petty, T., Duda, K., Chanda, A.: A database approach to sdn control plane design. SIGCOMM Comput. Commun. Rev. **47**(1), 15–26 (2017). DOI 10.1145/3041027.3041030. URL `https://doi.org/10.1145/3041027.3041030`
31. Mestres, A., Rodriguez-Natal, A., Carner, J., Barlet-Ros, P., Alarcón, E., Solé, M., Muntés-Mulero, V., Meyer, D., Barkai, S., Hibbett, M.J., et al.: Knowledge-defined networking. SIGCOMM Comput. Commun. Rev. **47**(3), 2–10 (2017). DOI 10.1145/3138808.3138810. URL `https://doi.org/10.1145/3138808.3138810`
32. Lakiotakis, E., Liaskos, C., Dimitropoulos, X.A.: Improving networked music performance systems using application-network collaboration. CoRR **abs/1808.09405** (2018). URL `http://arxiv.org/abs/1808.09405`
33. Jiang, C., Zhang, H., Ren, Y., Han, Z., Chen, K., Hanzo, L.: Machine learning paradigms for next-generation wireless networks. IEEE Wireless Communications **24**(2), 98–105 (2017). DOI 10.1109/MWC.2016.1500356WC
34. Suárez-Varela, J., Mestres, A., Yu, J., Kuang, L., Feng, H., Cabellos-Aparicio, A., Barlet-Ros, P.: Routing in optical transport networks with deep reinforcement learning. J. Opt. Commun. Netw. **11**(11), 547–558 (2019). DOI 10.1364/JOCN.11.000547. URL `http://jocn.osa.org/abstract.cfm?URI=jocn-11-11-547`
35. Afolabi, I., Taleb, T., Samdanis, K., Ksentini, A., Flinck, H.: Network slicing and softwarization: A survey on principles, enabling technologies, and solutions. IEEE Communications Surveys Tutorials **20**(3), 2429–2453 (2018). DOI 10.1109/COMST.2018.2815638
36. Zhang, H., Liu, N., Chu, X., Long, K., Aghvami, A., Leung, V.C.M.: Network slicing based 5g and future mobile networks: Mobility, resource management, and challenges. IEEE Communications Magazine **55**(8), 138–145 (2017). DOI 10.1109/MCOM.2017.1600940

37. Li, R., Zhao, Z., Sun, Q., I, C., Yang, C., Chen, X., Zhao, M., Zhang, H.: Deep reinforcement learning for resource management in network slicing. IEEE Access **6**, 74429–74441 (2018). DOI 10.1109/ACCESS.2018.2881964

38. Li, R., Zhao, Z., Zhou, X., Ding, G., Chen, Y., Wang, Z., Zhang, H.: Intelligent 5g: When cellular networks meet artificial intelligence. IEEE Wireless Communications **24**(5), 175–183 (2017). DOI 10.1109/MWC.2017.1600304WC

39. Luong, N.C., Hoang, D.T., Gong, S., Niyato, D., Wang, P., Liang, Y.C., Kim, D.I.: Applications of deep reinforcement learning in communications and networking: A survey. IEEE Communications Surveys & Tutorials **21**(4), 3133–3174 (2019)

40. Bellemare, M.G., Dabney, W., Munos, R.: A distributional perspective on reinforcement learning. In: Proceedings of the 34th International Conference on Machine Learning-Volume 70, pp. 449–458. JMLR. org (2017)

41. Hua, Y., Li, R., Zhao, Z., Chen, X., Zhang, H.: Gan-powered deep distributional reinforcement learning for resource management in network slicing. IEEE Journal on Selected Areas in Communications (2019)

42. Montavon, G., Lapuschkin, S., Binder, A., Samek, W., Müller, K.R.: Explaining non-linear classification decisions with deep taylor decomposition. Pattern Recognition **65**, 211–222 (2017)

43. Zhang, C., Patras, P., Haddadi, H.: Deep learning in mobile and wireless networking: A survey. IEEE Communications Surveys & Tutorials **21**(3), 2224–2287 (2019)

44. Liaskos, C., Tsioliaridou, A., Ioannidis, S.: The socket store: An app model for the application-network interaction. In: IEEE ISCC 2017 (2017)

45. Cloud, J., Leith, D., Médard, M.: A coded generalization of selective repeat arq. In: 2015 IEEE Conference on Computer Communications (INFOCOM), pp. 2155–2163 (2015). DOI 10.1109/INFOCOM.2015.7218601

46. Papadopoulos, I., Papanikos, N., Papapetrou, E., Kondi, L.: Network-wide md and network coding for heterogeneous video multicast. In: 2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), pp. 3578–3582 (2013). DOI 10.1109/PIMRC.2013.6666770

47. Zhu, Q., Wang, R., Chen, Q., Liu, Y., Qin, W.: IOT Gateway: Bridging Wireless Sensor Networks into Internet of Things. In: IEEE/IFIP International Conference on Embedded and Ubiquitous Computing, pp. 347–352 (2010)

48. Chen, H., Jia, X., Li, H.: A brief introduction to IoT gateway. In: Communication Technology and Application (ICCTA 2011), IET International Conference on, pp. 610 – 613 (2011)

49. Chellough, S.A., El-Zawawy, M.A.: Middleware for internet of things: Survey and challenges. Intelligent Automation & Soft Computing **24**(2), 309–318 (2018)

50. Dickerson, K., Heinz, C., García-Castro, R., et al.: Analysis of Standardisation Context and Recommendations for Standards Involvement (2016). URL https://vicinity2020.eu/vicinity/sites/default/files/documents/vicinity_d2.1_analysis_of_standardisation_context_and_recommendations_for_standards_involvement.pdf

51. Nakhuva, B., Champaneria, T.: Study of various internet of things platforms. International Journal of Computer Science & Engineering Survey **6**(6), 61–74 (2015)

52. Gomes, P., Cavalcante, E., Batista, T., Taconet, C., Conan, D., Chabridon, S., Delicato, F.C., Pires, P.F.: A semantic-based discovery service for the internet of things. Journal of Internet Services and Applications **10** (2019). DOI 10.1186/s13174-019-0109-8

53. Song, Z., Cardenas, A., Masuoka, R.: Semantic middleware for the internet of things. In: 2010 Internet of Things (IOT), pp. 1–8 (2011). DOI 10.1109/IOT.2010.5678448

54. Wang, W., Lee, K., Guo, J., Murray, D.: Discovering objects and services in context-aware iot environments. International Journal of Services Technology and Management **25**(3/4), 326–347 (2019). DOI 10.1504/IJSTM.2019.10021608

55. Kostelnik, P., Sarnovsky, M., Furdík, K.: The semantic middleware for networked embedded systems applied in the internet of things and services domain. Scalable Computing: Practice and Experience **12**(3), 307–315 (2011)

56. Guan, Y., Vasquez, J.C., Guerrero, J.M., Samovich, N., Vanya, S., Oravec, V., García-Castro, R., Serena, F., Poveda-Villalón, M., Radojicic, C., Heinz, C., Grimm, C., Tryferidis, A., Tzovaras, D., Dickerson, K., Paralic, M., Skokan, M., Sabol, T.: An

open virtual neighbourhood network to connect iot infrastructures and smart objects
— vicinity: Iot enables interoperability as a service. In: 2017 Global Internet of Things
Summit (GIoTS), pp. 1–6 (2017). DOI 10.1109/GIOTS.2017.8016233

57. Alam, M.F., Katsikas, S., Beltramello, O., Hadjiefthymiades, S.: Augmented and vir-
    tual reality based monitoring and safety system: A prototype iot platform. Journal of
    Network and Computer Applications **89**, 109 – 119 (2017). DOI https://doi.org/10.
    1016/j.jnca.2017.03.022. URL http://www.sciencedirect.com/science/article/pii/
    S1084804517301315. Emerging Services for Internet of Things (IoT)

58. Antonakoglou, K., Xu, X., Steinbach, E., Mahmoodi, T., Dohler, M.: Toward haptic
    communications over the 5g tactile internet. IEEE Communications Surveys Tutorials
    20(4), 3034–3059 (2018)

59. Hooft, J.V.d., Torres Vega, M., Petrangeli, S., Wauters, T., Turck, F.D.: Tile-based
    adaptive streaming for virtual reality video. ACM Trans. Multimedia Comput. Com-
    mun. Appl. **15**(4) (2019). DOI 10.1145/3362101. URL https://doi.org/10.1145/
    3362101

60. Clemm, A., Torres Vega, M., Ravuri, H.K., Wauters, T., Turck, F.D.: Toward Truly
    Immersive Holographic-Type Communication: Challenges and Solutions. IEEE Com-
    munications Magazine **58**(1), 93–99 (2020)

61. Palmer, M., Krüger, T., Chandrasekaran, B., Feldmann, A.: The quic fix for optimal
    video streaming. In: Proceedings of the Workshop on the Evolution, Performance, and
    Interoperability of QUIC, EPIQ'18, p. 43–49. Association for Computing Machinery,
    New York, NY, USA (2018). DOI 10.1145/3284850.3284857. URL https://doi.org/
    10.1145/3284850.3284857

62. King, H.H., Hannaford, B., Kammerly, J., Steinbachy, E.: Establishing multimodal
    telepresence sessions using the session initiation protocol (sip) and advanced haptic
    codecs. In: 2010 IEEE Haptics Symposium, pp. 321–325 (2010)

63. Nasir, Q., Khalil, E.: Perception based adaptive haptic communication protocol (pahcp).
    In: 2012 International Conference on Computer Systems and Industrial Informatics, pp.
    1–6 (2012)

64. Venkatraman, K., Vellingiri, S., Prabhakaran, B., Nguyen, N.: Mpeg media transport
    (mmt) for 3d tele-immersion systems. In: 2014 IEEE International Symposium on
    Multimedia, pp. 279–282 (2014)

65. Rossol, N., Cheng, I., Bischof, W.F., Basu, A.: A framework for adaptive training and
    games in virtual reality rehabilitation environments. In: Proceedings of the 10th Interna-
    tional Conference on Virtual Reality Continuum and Its Applications in Industry, VR-
    CAI '11, p. 343–346. Association for Computing Machinery, New York, NY, USA (2011).
    DOI 10.1145/2087756.2087810. URL https://doi.org/10.1145/2087756.2087810

66. Power, D.J., Sharda, R.: Model-driven decision support systems: Concepts and re-
    search directions. Decision Support Systems **43**(3), 1044 – 1061 (2007). DOI https:
    //doi.org/10.1016/j.dss.2005.05.030. URL http://www.sciencedirect.com/science/
    article/pii/S0167923605000953. Integrated Decision Support

67. Vaughan, N., Gabrys, B., Dubey, V.N.: An overview of self-adaptive technologies within
    virtual reality training. Computer Science Review **22**, 65 – 87 (2016). DOI https://
    doi.org/10.1016/j.cosrev.2016.09.001. URL http://www.sciencedirect.com/science/
    article/pii/S1574013716300259

68. Luzanin, O., Plancak, M.: Hand gesture recognition using low-budget data glove and
    cluster-trained probabilistic neural network. Assembly Automation **34**(1), 94–105
    (2014). DOI 10.1108/AA-03-2013-020

69. Yamashita, M.: Assistive driving simulator with haptic manipulator using model predic-
    tive control and admittance control. Procedia Computer Science **39**, 107 – 114 (2014).
    DOI https://doi.org/10.1016/j.procs.2014.11.016. URL http://www.sciencedirect.
    com/science/article/pii/S1877050914014343. The 6th international conference on
    Intelligent Human Computer Interaction, IHCI 2014

70. Van Damme, S., Torres Vega, M., De Turck, F.: Human-centric Quality Management
    of Immersive Multimedia Applications. In: in proceedings of the fourth Quality of Ex-
    perience Management Workshop, collocated with NetSoft 2020. Ghent, Belgium (2020)

71. Simsek, M., Aijaz, A., Dohler, M., Sachs, J., Fettweis, G.: 5g-enabled tactile internet.
    IEEE Journal on Selected Areas in Communications **34**(3), 460–473 (2016)

72. Skorin-Kapov, L., Varela, M., Hoßfeld, T., Chen, K.T.: A survey of emerging concepts and challenges for qoe management of multimedia services. ACM Trans. Multimedia Comput. Commun. Appl. **14**(2s) (2018). DOI 10.1145/3176648. URL `https://doi.org/10.1145/3176648`
73. ITU-T: Recommendation P.910 (09/99) ITU-T RECOMMENDATION P.910: Subjective video quality assessment methods for multimedia applications (1999)
74. Alexiou, E., Ebrahimi, T.: On subjective and objective quality evaluation of point cloud geometry. In: 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–3 (2017)
75. Tran, H.T.T., Ngoc, N.P., Bui, C.M., Pham, M.H., Thang, T.C.: An evaluation of quality metrics for 360 videos. In: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), pp. 7–11 (2017)
76. van der Hooft, J., Torres Vega, M., Petrangeli, S., Wauters, T., De Turck, F.: Quality Assessment for Adaptive Virtual Reality Video Streaming: A Probabilistic Approach on the User's Gaze. In: 2019 22nd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN), pp. 19–24 (2019)
77. Alexiou, E., Viola, I., Borges, T.M., Fonseca, T.A., de Queiroz, R.L., Ebrahimi, T.: A comprehensive study of the rate-distortion performance in MPEG point cloud compression. APSIPA Transactions on Signal and Information Processing **8**, e27 (2019). DOI 10.1017/ATSIP.2019.20
78. Narbutt, M., Allen, A., Skoglund, J., Chinen, M., Hines, A.: AMBIQUAL - a full reference objective quality metric for ambisonic spatial audio. In: 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6 (2018)
79. Sakr, N., Georganas, N.D., Zhao, J.: A Perceptual Quality Metric for Haptic Signals. In: 2007 IEEE International Workshop on Haptic, Audio and Visual Environments and Games, pp. 27–32 (2007)
80. Hassen, R., Steinbach, E.: HSSIM: An Objective Haptic Quality Assessment Measure for Force-Feedback Signals. In: 2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6 (2018)
81. Zhou Wang, Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004)
82. De Coninck, Q., Bonaventure, O.: Multipath quic: Design and evaluation. In: Proceedings of the 13th International Conference on Emerging Networking EXperiments and Technologies, CoNEXT '17, p. 160–166. Association for Computing Machinery, New York, NY, USA (2017). DOI 10.1145/3143361.3143370. URL `https://doi.org/10.1145/3143361.3143370`

Biographies

Maria Torres Vega received her M.Sc. degree in telecommunication engineering from the Polytechnic University of Madrid, Spain, in 2009 and her Ph.D. from Eindhoven University of Technology, The Netherlands, in 2017. She is a postdoctoral researcher on multimedia delivery with the Department of Information Technology, Ghent University - imec, Belgium. Her research interests include quality of experience in immersive multimedia systems and autonomous management of future networks.

Christos Liaskos received the Diploma in Electrical and Computer Engineering from the Aristotle University of Thessaloniki (AUTH), Greece in 2004, the MSc degree in Medical Informatics in 2008 from the Medical School, AUTH and the PhD degree in Computer Networking from the Dept. of Informatics, AUTH in 2014. He is currently an assistant professor at the University of Ioannina and an affiliated researcher at the Foundation of Research and Technology, Hellas (FORTH).

Sergi Abadal is the director of research at the NaNoNetworking Center in Catalunya (N3Cat) at UPC-BarcelonaTech and coordinator of the WIPLASH FET-OPEN project. He has published 70+ articles in top-tier journals and conferences.

He is member of the IEEE, ACM, and HiPEAC. His research interests lie in the area of extreme communication networks.

Evangelos Papapetrou received the Diploma in Electrical and Computer Engineering and the Ph.D. degree in Telecommunications from the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki. He is an Assistant Professor in the Department of Computer Science and Engineering, University of Ioannina. His research interests include networking in wireless and mobile environments, such as mobile ad hoc, opportunistic, and satellite networks.

Akshay Jain is a Researcher at the NaNoNetworking Center in Catalunya (N3Cat) at UPC-BarcelonaTech and a PhD Candidate with the Network Engineering department at UPC. He received MSCA fellowship from EU and the KVPY fellowship from DST, Govt. of India. His research interests lie in next generation wireless networks.

Belkacem Mouhouche is a Chief Standards Engineer at Samsung Research UK. He holds a Ph.D. degree in Signal Processing from Ecole Nationale Superieure des Telecoms (Telecom Paristech) in France. Dr. Mouhouche has been involved in 3GPP standard since the early days of HSDPA and later LTE. He was also involved in FP7 European Project "End to End Efficiency" (E3) where he worked on cognitive radio in heterogeneous systems. His research interest are in the area of advanced techniques for the physical layer of future communication systems including MIMO, beamforming and coding applied to wireless systems. Dr. Mouhouche was also heavily involved in the development of the latest broadcast terrestrial standard ATSC3.0 and led the broadcast activity within the 5G-PPP project "Fantastic-5G". Recently, Was the technical manager of 5G-PPP project "5G-Xcast" focused on new point to multipoint techniques for 5G networks. He is currently the technical manager of the ICT-19 project 5G-TOURs focused on implementing 13 vertical use cases on top of 5G infrastructures.

Gökhan Kalem has around 10 years of experience in the telecommunications industry. He is currently working as an Expert 5G R&D Engineer at Turkcell, and pursuing his PhD in Engineering Management at Marmara University (Turkey). His research interests include 5G and beyond mobile network technologies.

Salih Ergüt has many years of experience in the telecommunications industry working at leading vendors and operators in the US and Turkey. He is currently working at Turkcell, and serving as a vice chairman at ITU-T Focus Group on "ML for Future Networks including 5G". He received his MS from Northeastern University, and PhD from University of California San Diego, all in Electrical and Computer Engineering.

Marian Mach is an associate professor at the University of Kosice, Slovakia. He received his PhD in Technical Cybernetics in 1993 and habilitation in Artificial Intelligence in 2001. His professional interests include IoT, knowledge representation and processing, machine learning and nature-inspired optimization. He is the author and co-author of six books and over 150 research papers in journals and conference proceedings.

Tomas Sabol is a professor of artificial intelligence at Technical University of Kosice, Slovakia. He has a professional background on Artificial Intelligence, Internet of Things, Data analytics and knowledge management. He has published more than 160 papers in books, journals, conference proceedings.

Albert Cabellos-Aparicio received a PhD (2008) degree in Computer Science Engineering from the Universitat Politècnica de Catalunya (www.upc.edu). He is the

co-founder of the NaNoNetworking Center in Catalunya (http://www.n3cat.upc.edu), the Open Overlay Router initiative (http://openoverlayrouter.org) and the Barcelona Neural Networking Center (https://bnn.upc.edu).

Christoph Grimm is professor for computer science at the Technische Universität Kaiserslautern. He studied electrical engineering at TU Darmstadt (1994), holds a Doctor in computer science from Frankfurt University (1999). He was professor for embedded systems at TU Vienna (2006-2012). He is active in standardization and development of modeling languages where he contributed to the development of VHDL-AMS and SystemC AMS. He also works in the field of ontological modeling, e.g. in the VICINITY project.

Filip De Turck received his M.Sc. and Ph.D. degrees in electronic engineering from Ghent University in 1997 and 2002, respectively. He is a full professor with the Department of Information Technology, Ghent University-imec. There he leads the Network and Service Management Research Group. His research interests include network and service management, and design of efficient virtualized networks. He is the Chair of the IEEE Technical Committee on Network Operations and Management and Editor-in-Chief of IEEE TNMS.

Jeroen Famaey received the M.Sc. and Ph.D. degrees in computer science from Ghent University, Belgium, in 2007 and 2012, respectively. Since 2016, he is an Assistant Professor associated with imec and the University of Antwerp, Belgium. He has co-authored over 120 articles published in international peer reviewed journals and conference proceedings. His research focuses on performance modelling and optimization of wireless networks.