O-RAN next Generation Research Group (nGRG)
Contributed Research Report

Research Report on Emerging Indoor Use Cases

Report ID: RR-2024-06

Contributors:

Apple
Mavenir

Release date: 2024.05

Authors

Company	Name	Email
Apple	Sharad Sambhwani Dirk Nickisch Wanlu Sun	ssambhwani@apple.com dnickisch@apple.com wanlu_sun@apple.com
Mavenir	Prerit Jain Sapna Sangal	prerit.jain@mavenir.com sapna.sangal@mavenir.com

Reviewers

Company	Name	Email
DTAG	Jan Plachy	Jan.Plachy@telekom.de
NICT	Gabriel Villardi	gpvillardi@nict.go.jp
Nokia	Daiju CA	daiju.ca@nokia.com
Reliance	Vikas Dixit	Vikas1.Dixit@ril.com
Sumitomo Electric Industries	Hiroshi Miyata	miyata-hiroshi@sei.co.jp
TechMahindra Ltd	Rajat Agarwal	Rajat.Agarwal@TechMahindra.com
TechMahindra Ltd	Sandeep Sharma	Sharma.Sandeep@TechMahindra.com
Capgemini	Lalit Garg	lalit.a.garg@capgemini.com
Verizon	Nirlay Kundu	nirlay.kundu@verizon.com

Disclaimer

Executive summary

Recognizing the critical role that indoor wireless communication plays in the current digital ecosystem, this technical report provides examinations of emerging and advanced indoor use case, and sheds light on the distinctive characteristics and requirements of these settings.

Each of the use cases detailed in the report is selected for its relevance to illustrate the potentials of 6G in indoor settings, e.g., either to enable the use cases that cannot be realized by 5G, or to bring the use cases that can to some extent already be achieved by 5G to the next level in terms of performance. These use cases include scenarios ranging from enhanced smart home functionalities and immersive entertainment experiences to efficient workplace communication systems. More specifically, the following eleven use cases are analysed:

- Mixed reality telepresence
- Immersive education
- Immersive gaming/entertainment
- Immersive exercise & fitness
- Multi-media streaming
- Real-time video/3D communication
- Immersive virtual shopping
- Personalized hotel experience
- Mega-events with high data rate
- Cooperative mobile robots
- Telemedicine and Remote Healthcare

By providing a detailed description of each use case, the report aims to highlight the significant opportunities that 6G technology offers for indoor environments. Moreover, the report outlines the potential requirements of the discussed use cases, emphasizing the need for advanced connectivity solutions to meet the expectations of the next generation of users and applications.

By analysing the potential use cases and their requirements, this technical report targets to contribute to the ongoing dialogue among stakeholders in the telecommunications industry, other related industries, policymakers, and technology developers. The ultimate goal is to pave the way for the successful integration of 6G technology into ~~our~~ indoor spaces, ensuring that the future wireless communication fulfils its promise of creating smarter, better connected, and more efficient indoor environments.

Table of Contents

List of abbreviations

3GPP	3rd Generation Partnership Project
4G	4th Generation Mobile Network
5G	5th Generation Mobile Network
6G	6th Generation Mobile Network
6DOF	Six degrees of freedom
AI	Artificial Intelligence
AP	Access Point
AR	Augmented Reality
DL	Downlink
E2E	End to end
IoT	Internet of Things
ML	Machine Learning
MR	Mixed Reality
MRTA	Multi-Robot Task Assignment
QoS	Quality of Service
TDD	Time Division Duplex
UE	User Equipment
UL	Uplink
VR	Virtual Reality
XR	Extended Reality

List of figures

List of tables

N/A

1 Introduction

Building on the foundation laid by 5G, 6G is expected to offer far more than just incremental improvements in speed and efficiency. It promises a transformation a new era of ultra-reliable, high-bandwidth, and low-latency connections. This transformation is not only confined to the outdoor environments but extends deep into the indoor spaces where we live, work, and play.

The significance of indoor use cases in the 6G era cannot be ignored. With most of the wireless communications already occurring indoors [1], the demand for seamless connectivity, capable of supporting an ever-growing ecosystem of smart devices and applications, is at an all-time high. From smart homes and offices to advanced manufacturing facilities and virtual spaces, the advent of 6G technology will unlock high levels of efficiency, safety, and convenience.

Compared to outdoor environments, indoor scenarios have distinct characteristics in their respective use cases. First, indoor environments often have higher signal attenuation compared to outdoor spaces. Walls, ceilings, and floors made of materials like concrete, steel, and glass can significantly reduce signal strength. Second, indoor wireless networks may serve a higher density of users compared to outdoor networks. For instances, in environments like offices, shopping malls, or stadiums, there could be a large number of devices concurrently accessing the network. Last but not least, indoor wireless networks may require higher levels of security due to the proximity of potential attackers. Unauthorized access points, rogue devices, and eavesdropping threats are more prevalent indoors, necessitating robust security measures and intrusion detection systems.

The report aims to provide a comprehensive overview of the emerging indoor use cases, framed through the lens of wireless communication engineering principles.

2 Emerging Indoor Use Cases

In this section, we describe eleven representative indoor use cases, including mixed reality (MR) Telepresence, immersive education, immersive gaming & entertainment, immersive exercise & fitness, multi-media streaming, real-time video/3D communication, immersive virtual shopping, personalized hotel experience, mega-events with high data rate, cooperative mobile robots, and telemedicine and remote healthcare. Each use case is examined through its importance, service flow, and potential requirements.

2.1 Use Case 1: Mixed Reality Telepresence

2.1.1 Description

Mixed reality (MR) is a technology that merges the real and virtual worlds to coexist with each other. Depending on the amount of real and virtual information, MR can be categorized as shown in *Figure 1: Mixed reality (MR)*[2], where two extremes are real environment and virtual environment. The former case defines environments consisting solely of real objects, while the latter case defines environments consisting solely of virtual objects. In this regard, MR refers to anywhere between the two extremes.

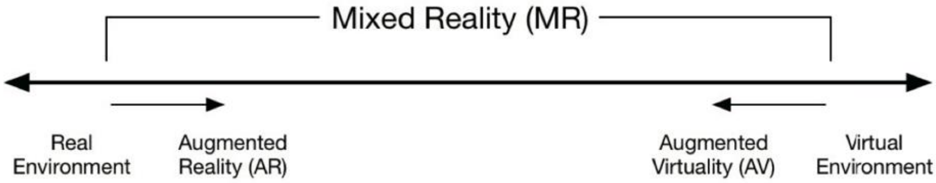

Figure 1: Mixed reality (MR)

According to [3], telepresence refers to the technologies which allow a person to feel as if they were present, to give the appearance or effect of being present, at a place other than their true physical location.

By combining MR and telepresence, virtual objects appear realistic, and remote users can get an experience of local and real-world interaction even though they are far away from each other.

MR telepresence has multiple applications and has the potential to revolutionize how we communicate, collaborate, play, and work.

One compelling scenario is lifelike remote and immersive conference, where participants are located at different places. With telepresence, people feel that they are sitting in the same room. Moreover, participants are free to move around using Six degrees of freedom (6DOF) [4]. During the movement, the surrounding environments are processed at a server and sent back to the participants timely to create XR experience. In addition to the improved levels of video and audio clarity compared to conventional video/audio conference, a participant can also write or draw on paper or white/black board; then other participants can feel it as if they were just in front of it and can point to anywhere they want. This functionality is enabled by MR. In this way, not only much more efficient collaboration can be facilitated; but also, the costs and carbon emissions due to the traveling can be largely reduced. An example is illustrated in *Figure 2: Illustration of immersive conference*.

Figure 2: Illustration of immersive conference

2.1.2 Scenario illustration

1. Three users A, B, and C, who are located in their houses at different cities, are participating in an immersive conference.
2. Each user is equipped with an immersive conference capable device, e.g., Extended Reality (XR) headset or glasses.
3. Capture devices are available to record the movements and gestures of the three users, and to communicate to a server the recording outcome. Some examples of the capture devices include motion capture suits, which have sensors embedded throughout to capture full-body movement, and 3D cameras that are used to detect body positions and gestures.

2.1.3 Service flow

1. Users A, B and C select a virtual meeting room and start remote conference.
2. The three users connect to the conference server.
3. The conference server transmits the layout of the virtual meeting room to the three users. At this time, the three users have a common view of the virtual meeting room and can already talk to each other.
4. The three users select their preferred seats respectively within the meeting room and send the selections back to the conference server. Assume that user A selects a seat between users B and C's seats.
5. Based on the seat selections, the conference server assigns the three users to different positions. Then the view and sound perceived by each user will be updated accordingly and then sent to each user respectively.
6. User A first turns left to talk to user B. User A's movement is captured by the corresponding capture device and sent back to the conference server.

7. Based on user A's movement, the view and sound perceived by each user will be updated accordingly at the conference server, and then sent to each user respectively.
8. User A physically stands in front of his desk, writes down an equation on his notebook, and explains the equation. User A's gesture, the written outcome, and voice are captured by his local capture device and transmitted to the conference server.
9. After processing, the conference server sends the audio and video to users B and C, so that B and C perceive the real-time writing and talking behaviors of user A and they can also see what is being written on user A's notebook.
10. After user A explains the equation, the conference is finished, and all the three users disconnect with the conference server.

2.1.4 Service deliverable

Three users A, B, and C enjoy the remote but real-time, immersive, and lifelike discussions.

2.1.5 Potential technical requirements

This use case requires:

- ultra-high data rate to transmit and receive high-resolution videos.
- High reliability and extremely low End to End (E2E) latency and jitter for both uplink and downlink to provide seamless immersive experience.
- capabilities on capture devices to capture users' positions, movements and gesture variations with high accuracy
- Conference server is able to fuse the movements of different users to generate more appropriate views and sounds

Among these requirements, data rate and E2E latency will have impacts on standardization bodies, e.g., O-RAN ALLIANCE and/or 3rd Generation Partnership Project (3GPP). Demanding capturing and tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them. Fusion capability at the server will mostly like rely on custom implementation.

2.2 Use Case 2: Immersive Education

2.2.1 Description

Education is one of the most important drivers of societal development. High quality education should be widely and equally available for everyone all around the world. Accessibility and quality of education is, however, facing challenges such as lack of infrastructure, high costs, logistic challenges, or outdated and restrictive tools and means.

A clear example of such challenges is the COVID-19 pandemic that forced schools all around the world to physically shut down and moved the education process into homes to operate in a remote fashion. Lack of student-student and student-teacher interaction, difficulty in using traditional tools of education remotely, lack of a suitable learning environment, and low-quality internet connection made the process suffer both in scope and quality.

In general, there are two main purposes of immersive education.

1. Provide high-quality education for anyone and anywhere. In this case, some people have to attend the immersive classroom remotely due to various reasons. There are typically two scenarios: fully remote education where all students and instructors are located at different places; partially remote education where some students physically present in the classroom, while others join the session remotely.
2. Convey the knowledge in a more effective and attractive manner. For example, immersive education puts people/students in an environment where they are immersed in the material they are currently learning. In this way, people/students can be more engaged and interactive with the environment to gain realistic experience. And thus, they can understand the concepts better and more quickly via 3D visualization. Another example could be immersive museum which will be detailed later.

The service flow related to the first purpose is similar with the use case of mixed reality telepresence. Hence, in the following we consider a representative application for the second purpose only: immersive museum tour as illustrated in Figure *3: Immersive Museum tour*.

Figure 3: Immersive Museum tour

2.2.2 Scenario illustration

1. A group of people are visiting a Nature History Museum, and they are learning about dinosaurs.
2. Visitors are wearing Virtual Reality (VR) devices.

3. Capture devices are available in the museum to record the movements and gestures of the visitors, and to communicate to a learning server the recording outcome.

2.2.3 Service flow

1. Visitor A starts the dinosaur learning session on his XR device and connects to the learning server.
2. The learning server transmits a virtual scene of an island full of different categories of dinosaurs to visitor A's VR device. There is also a name on top of each dinosaur to show which species it belongs to.
3. Visitor A is interested in tyrannosaurus and thus taps his finger on it in the virtual scene.
4. The capture device captures visitor A's finger movement and sends the information to the learning server.
5. Based on the input, the learning server sends 3D video and audio to visitor A, which contains two tyrannosauruses. One is eating while the other one is walking. Visitor A is currently facing the tyrannosaurus who is eating.
6. Visitor A uses his fingers to zoom in to see more details on the tyrannosaurus's teeth.
7. The capture device captures visitor A's finger movement and sends the information to the learning server.
8. Based on the input, the learning server sends 3D video and audio to visitor A, which contains the enlarged area around the tyrannosaurus's teeth. The background audio is explaining the characteristics of the teeth.
9. Visitor A turns his head right.
10. The capture device captures visitor A's head movement and sends the information to the learning server.
11. Based on the input, the learning server sends 3D video and audio to visitor A, which contains the other tyrannosaurus who is walking and roaring. The background audio is explaining tyrannosaurus hunting's habits.
12. After learning, visitor A closed the session and disconnects with the server.

2.2.4 Service deliverable

Visitors have learnt more about dinosaurs in a more effective and interactive manner.

2.2.5 Potential technical requirements

This use case requires:

- ultra-high data rate especially downlink to transmit high-resolution videos;
- extremely low E2E latency and jitter for both uplink and downlink to provide seamless immersive experience;
- capabilities to capture users' positions, movements and gesture variations with high accuracy

Among these requirements, data rate and E2E latency will have impacts on standardization bodies, e.g., O-RAN ALLIANCE and/or 3GPP. Demanding tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them.

2.3 Use Case 3: Immersive Gaming/Entertainment

2.3.1 Description

Gaming and entertainment are the first industry to reap the benefits and tap into the potential of digital reality, transforming into an interactive and immersive form through technologies like Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR). AR/VR/MR-based gaming may evolve beyond the traditional online gaming to immersive gaming parties, where users at different locations can join a Gaming Party either physically or virtually. Game players can interact with each other without feeling a boundary between the virtual and the physical worlds.

Furthermore, conventional games typically only provide users with what to see and what to hear. With immersive gaming, haptic feedback, such as forces or vibrations, can be offered as well to further improve the gaming experiences.

More and more gaming companies started to recognize the potential immersive gaming market. There are several key finds provided by [5]:

- 58% already pay a premium to their provider to enjoy the best gaming experience possible.
- 79% would consider upgrading their current home broadband and existing 4G mobile connectivity for better gaming experience with 5G and beyond.
- 95% would pay more for this improved experience, with 60% willing to pay 50% more (or $126 per month compared to the current monthly average of $84);
- 58% would switch connectivity provider as soon as they could if a competitor offered a high-quality gaming service.

If gaming companies manage to provide smooth immersive gaming experiences, it is expected they can further augment their revenues. This clearly requires high-performance wireless connectivity.

In the following scenario illustration, we will explain an example where two players who locate at different locations play virtual boxing game with each other, as illustrated in *Figure 4: Two players play virtual boxing game.*

Figure 4: Two players play virtual boxing game

2.3.2 Scenario illustration

1. Two players A and B, who are located in different locations, are participating in a virtual boxing game.
2. Each user is equipped with haptic feedback kit, including vest, arms, gloves, and facial interface.
3. Capture devices are available to record the movements and gestures of the two players respectively, and to communicate to a server the recording outcome.

2.3.3 Service flow

1. Players A and B join a virtual box gaming session.
2. The two players connect to the gaming server.
3. Both players see the menu in their virtual scene.
4. Both players punch the start button.
5. The capture devices capture the punch of the respective player and send the information to the gaming server.
6. The gaming server sends haptic feedbacks to both players so that they can feel the force from the button.
7. Player A can see a virtual player B through his facial interface.
8. Player A punches player B in his chest.
9. The capture device captures the punch from player A and send the information to the gaming server.
10. The gaming server processes the information and then sends the corresponding haptic feedback to player B.
11. Player B can see a virtual player A and also his punch via player B’s facial interface. Moreover, Player B can feel the punch in his own chest.
12. After one game session, both players disconnect with the server.

2.3.4 Service deliverable

Players enjoy a virtual boxing gaming without getting actual hurt.

2.3.5 Potential technical requirements

- High data rate to receive and transmit high-resolution videos.
- High reliability and extremely low E2E latency and jitter for both uplink and downlink to provide seamless immersive experience.
- Capabilities to capture users' positions, movements and gesture variations with high accuracy.
- Gaming server is able to fuse the movements of different players to generate more appropriate views, sounds, and haptic feedbacks

Among these requirements, data rate and E2E latency will have impacts on standardization bodies, e.g., O-RAN ALLIANCE and/or 3GPP. Demanding tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them. Fusion capability at the server will mostly like rely on custom implementation.

2.4 Use Case 4: Immersive Exercise & Fitness

2.4.1 Description

Online fitness exercises, e.g., Yoga, Zumba etc. are becoming very popular due to its immense health benefits. However, doing fitness exercises alone is boring and with time people lose interest. Conducting any fitness exercise in a group of friends along with an instructor can help in both correctness of activity as well as in keeping the interest of participants. E.g., a yoga instructor can constantly guide and correct postures of participants.

Using Mixed Reality (MR), people located at different places can form a group. A theme like beach, park, mountains etc. can be selected. MR helps these people to feel as if fitness activity is performed together in the environment selected by an individual.

Here, we are taking the example of MR based Online Yoga Session as illustrated in *Figure 5: Online Yoga with MR.*

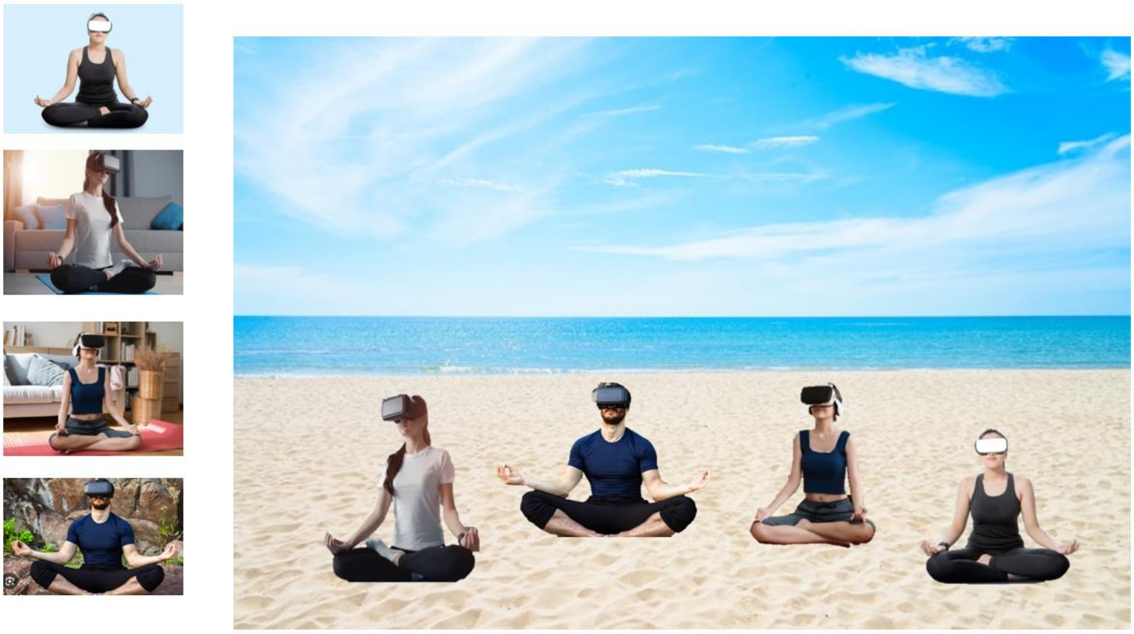

Figure 5: Online Yoga with MR

2.4.2 Scenario illustration

- Four users A, B, C and D are sitting in their houses at different cities and are participating in Immersive Yoga class.
- Each user is equipped with a VR headset.
- 360 degree camera that captures 6DoF movements of the four users and send it to a server for compiling all four user videos together in same environment.

2.4.3 Service flow

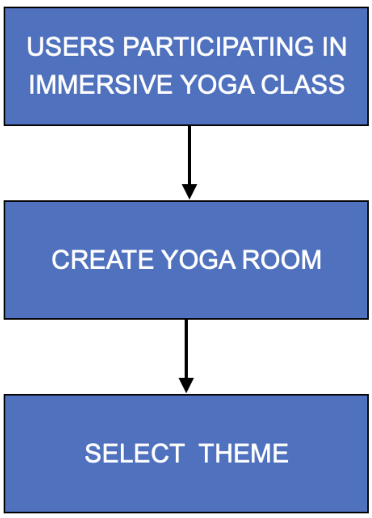

Figure 6: Online Yoga Service flow

1. As shown in *Figure 6: Online Yoga Service flow* user A, B and C are sitting in their houses in different cities and participating in immersive yoga class.

2. User A opens the Immersive Fitness application and creates the New Yoga Room - YogaInBeach and sets the theme on it as per the Service Flow shown in ***Error! Reference source not found.***
3. User A sends the URL link to User B and User C.
4. User B and User C user opens the Immersive Fitness application using the URL provided and enters the room –"YogaInBeach"
5. Three Users connect to same Yoga room of "YogaInBeach" of the Immersive Fitness application.
6. The Immersive Fitness application capture the video of all the three users and send to the server for processing.
7. After processing the server will send video back to all the users such that that every user will have same view.
8. Immersive Fitness application is connected to both VR and 360-degree camera.

2.4.4 Service deliverable

Three users A, B, and C enjoy the fitness activity in real-time, immersive, and at the location of their choice.

2.4.5 Potential technical requirements

This use case requires:

- ultra-high data rate to transmit high-resolution videos in both UL and DL direction;
- extremely low E2E latency and jitter for both uplink and downlink to provide seamless immersive experience.
- capabilities to capture users' positions, movements and gesture variations with high accuracy;
- server is able to fuse the movements of different users to generate more appropriate views and sounds;
- instructor and participants can view the 360-degree view of each other and can correct the posture in real time.

2.5 Use Case 5: Multi-media Streaming

2.5.1 Description

Envisioning the future of multi-media streaming, it can be considered as highly immersive, interactive, and personalized digital content delivery service. We can expect several ground-breaking enhancements and novel applications due to the advancements in technology and network capabilities. Three representative categories are provided below.

- **Ultra-high resolution streaming**
  Ultra-high resolution streaming in the future will likely represent a significant leap forward from today's standards. It is potentially going beyond 8K resolution and offering even higher and seamless frame rates. It could allow for crisp, clear, and highly detailed imagery, making distant textures and subtleties as clear as if they

were right in front of you. Whether it is film, sport, or live event, the streaming quality would feel like looking through a clear window rather than watching a screen. Moreover, after recognizing the device and context of each user, the future streaming technologies can adapt the resolution and streaming quality in real time. For instance, as the user moves closer to a screen, the resolution might increase to maintain the impeccable detail. Last but not the least, the high-quality streaming should be provided ubiquitously, irrespective of time, user locations, or user densities.

- **Holographic display**
  Generally speaking, holographic displays will project lifelike 3D images into space, allowing viewers to see and interact with them. These images, known as holograms, can be viewed from different angles and distances, offering a realistic depth and perspective that mirrors real objects or scenes. Holographic display can potentially redefine the way we consume content, providing a 3D, interactive, and highly immersive experience. The advanced holographic display will likely have a profound impact on entertainment, education, healthcare, and communication.

- **Real-time translation and dubbing**
  Real-time translation and dubbing are to convert spoken or written text from one language to another instantaneously, allowing viewers to understand and engage with content regardless of their original language. This goes beyond the simple subtitles to include voice dubbing, where the original speech is replaced with another language, synchronized with the speaker’s lip movements and expressions. Therefore, users can watch shows, movies, or live broadcast in their preferred language without delays. Real-time translation and dubbing can mark a substantial advancement in making content universally accessible and enjoyable, breaking down language barriers.

We consider an example of an explorer sharing his wildlife adventures of Amazon rainforest in a video show. Although the explorer narrates his journey in English, a Japanese viewer is watching the show on her phone and trying to understand it in Japanese. This is illustrated in *Figure 7: Video streaming with real-time translation and dubbing.*

Figure 7: Video streaming with real-time translation and dubbing

2.5.2 Scenario illustration

1. The explorer is recording the surrounding environment of Amazon rainforest using his phone with ultra-high resolution. His phone can connect to a cloud server in a real-time manner.
2. The cloud server has data processing capabilities in terms of translation and voice synthesis.
3. Application on the viewer's phone can support the service of multi-media streaming with real-time translation and dubbing.

2.5.3 Service flow

1. As the explorer records his show in the rainforest, the camera on his phone captures video in ultra-high resolution. The raw footage is compressed using advanced algorithms designed to retain quality while reducing file size.
2. The compressed video is uploaded in real-time to the cloud server via advanced wireless connection.
3. The translation software in the cloud performs real-time voice translation from English into Japanese. This includes not only language conversion but also cultural adaptation and lip-syncing adjustments. The translated text is then converted into speech, matching the explorer voice and intonation as closely as possible.
4. The Japanese viewer open her app, which then connects to the cloud server.
5. The now translated and compressed video stream in the cloud server is sent to the phone of the reviewer, via either broadcast or unicast.
6. After receiving the data, the app on the viewer's phone adjusts the resolution, frame rate, and compression, accordingly, ensuring the best possible experience for the viewer.

2.5.4 Service deliverable

Through the combination of the ultra-high resolution and real-time voice translation, the Japanese viewer can feel a seamless and deeply engaging experience without being blocked by the language barrier.

2.5.5 Potential technical requirements

The evolution of multi-media streaming will require wireless communication technologies that are faster, more reliable, and broader reaching than ever before.

1. High data rate to transmit ultra-high-resolution videos.
2. Low E2E latency and jitter to provide seamless experience.
3. Uniform QoS guarantee to be able to have high-quality streaming at anytime and anywhere.
4. Accurate and context-aware voice translation, voice synthesis, and real-time content adaptation.

Among these requirements, data rate, E2E latency, and uniform QoS guarantee impact technology, network architecture and standards defined by bodies like 3GPP and O-RAN ALLIANCE. Demanding tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them. Accurate translation will mostly like rely on custom implementation.

2.6 Use Case 6: Real-time Video/3D Communication

2.6.1 Description

Real-time video calls have become a standard and widely adopted form of remote communications. They are of great interests at people’s daily life. Even though the popularity of 2D video calls has already been significant and widespread, the qualities and user experiences are not always satisfied. As technology advances, the motivation for real-time video calls in the future lies in an increasing emphasis on fostering genuine human connection, which includes further enhanced 2D video calls as well as 3D video calls which create a more immersive and engaging environment.

- In 2D video calls, participants appear as digital personas on the flat canvas of screens. It’s a shared window into each other’s worlds, where the depth is limited to the visual plane, creating an intimate yet inherently two-dimensional exchange reminiscent of peering through a digital looking glass.
- 3D video calls are experiences that transcend the flatness of screens. Here, participants morph into vibrant, three-dimensional avatars navigating a digital space that extends far beyond the confines of 2D calls. Envision a virtual realm where interactions possess depth, and the shared environment becomes a dynamic stage for communication. It's as if participants step into a holographic landscape, where the nuances of spatial awareness amplify the richness of conversation, creating a multisensory rendezvous that mirrors the complexities of face-to-face encounters.

Wireless communications are prerequisites to enable calls over mobile phones. To support higher resolution and smoother 2D video calls or even more interactive and immersive 3D video calls, requirements on wireless communications become more demanding.

We will consider an example from the perspective of a user story.

2.6.2 Scenario illustration

1. User A and user B live in different cities, and they want to have a 3D video call using their phones. It is illustrated in *Figure 8: Example of 3D video call (from user A's perspective).*
2. Both of their phones have 3D communication capabilities.
3. Both of their phones can connect to a remote cloud in a real-time manner. The cloud has data processing capabilities.

Figure 8: Example of 3D video call (from user A's perspective)

2.6.3 Service flow

1. User A opens the 3D video call application on her phone to call her user B.
2. The application establishes a secure connection to user B's phone via wireless communication protocols.
3. User B receives a notification from his phone and answers the call.
4. As soon as the call starts, both phones begin to scan the user's surroundings and their movements using advanced sensors and cameras. Real-time renderings are used to create 3D model for both user A and user B.

5. Virtual spaces can also be created for both user A and user B. The space can be a virtual replication of one of their actual rooms, or a completely imaged environment like a virtual garden or beach.
6. User A and user B engage in a conversation. Their voices are transmitted in high-quality audio via wireless communications protocols.
7. Cameras and sensors capture their respective facial expressions and gestures.
8. After potential internal processing within the phones, the phones send this information to the remote cloud via wireless communication protocols in a real-time manner.
9. The cloud processes the received data, e.g., using AI/ML and rendering techniques.
10. The cloud transmits the relevant processed data to the corresponding users respectively.
11. Both user A and user B can see a 3D visualization of the other person with an enhanced sense of depth and presence.
12. After the conversation, either user A or user B can end the call. Their 3D models will disappear from the virtual space, and the connection is closed.

2.6.4 Service deliverable

enjoy a highly immersive and interactive way to communicate, bridging the gap created by physical distance. The call is also secure in the sense that all personal data and call records are securely stored and offers options for data deletion for privacy reasons.

2.6.5 Potential technical requirements

1. High DL and UL data rate to transfer high-resolution videos.
2. Extremely low E2E latency and jitter for both uplink and downlink to provide seamless experiences.
3. Capabilities to capture users' positions, movements, facial expressions and gesture variations with high accuracy.
4. Stable connection to avoid dropouts and to ensure a consistent user experience without interruptions.
5. Uniform QoS guarantee to be able to have high-quality calls at anytime and anywhere.
6. Security and privacy protocols to protect the data privacy and integrity of the user.

Among these requirements, data rate, E2E latency, and uniform QoS guarantee impact technology, network architecture and standards defined by bodies like 3GPP and O-RAN ALLIANCE. Demanding tracking capabilities and security as well as privacy can be achieved by custom implementations, standardized design, or a combination of them.

2.7 Use Case 7: Immersive Virtual Shopping

2.7.1 Description

Immersive virtual shopping transforms the standard retail experience, by blending the convenience of online shopping with the sensory engagement of in-store visits. The potential key characteristics are summarized as below.

- Customized spaces and items: shoppers enter personalized virtual stores designed to suit their tastes and preferences. As they step into these 3D environments from the comfort of their homes, they are greeted by a layout that mirrors their favourite physical stores or entirely new, imaginative setting. Moreover, the customers can choose to be offered with items that are selected based on their shopping histories and preferences.
- Interactive displays: products are displayed in high-resolution 3D, allowing customers to examine items for from every angle. Also, they can virtually touch, feel the texture, and even try on clothes or accessories using their digital avatars.
- Virtual assistants: there can be virtual assistants who guide the shoppers through the virtual space. These assistants can answer questions, provide additional product information, or even suggest complete outfits.
- Social experience: shoppers can invite friends to join their virtual shopping session. Together, they can explore stores, try on clothes side by side, or share opinions in real-time, making shopping a social and fun experience.
- Augmented reality trials: with AR technologies, shoppers can see how furniture looks in their actual home space or how a dress looks on them. This feature helps in making more informed decisions by providing a realistic preview of the products.
- Instant purchase: once a decision is made, customers can purchase items instantly with secure integrated payment systems. The products are then delivered to their home or prepared for quick pick-up at a local store.

In fact, immersive virtual shopping is not just about purchasing goods; it is an engaging, social, and sensory-rich experience that blurs the lines between physical and digital retail. It offers convenience, personalization, and a new realm of interactive possibilities, reshaping the way we typically think about shopping.

We consider an example of shopper A buying a dress through immersive virtual shopping from her living room, as illustrated i*n Figure 9: Illustration of immersive virtual shopping.*

s

Figure 9: Illustration of immersive virtual shopping

2.7.2 Scenario illustration

1. Shopper A wears a VR headset which connect to the virtual shopping cloud server.
2. Capture device is available to record the movements and gestures of shopper A, and to communicate to the cloud server the recording outcome.

2.7.3 Service flow

1. Via the VR headset, shopper A starts a session of virtual shopping and steps into a new world.
2. In the virtual world, shopper A selects several brands and details her wish, e.g., buying a dress. She also selects an avatar to represent herself and provides appropriate height and weight. The selection is done by shopper A taping her finger in the air.
3. After the setting, the space that shopper A sees through the VR headset transforms into a personalized paradise of plenty of dresses reflecting her selection (e.g., the selected brands).
4. A friendly avatar, the store's virtual assistant, greets shopper A. The assistant also recommends items to shopper A based on her taste and previous purchases. The assistant also offers to try them on a virtual fitting session.
5. Shopper A selects 'yes' for the fitting session.
6. In a virtual room, shopper A tries on various outfits. Each piece adapts to her avatar in real-time, allowing her to see how they look from every angle. She mixes and matches, playing with colours and styles – all without the physical hassle of changing clothes.
7. Shopper A decides on a beautiful dress and starts with the purchase. As the virtual store already has the payment and address information of shopper A, the purchase is completed by a simple gesture of shopper A.
8. After shopping, shopper A ends the shopping session and exits the virtual world.
9. The cloud server records the purchase outcome of shopper A, and sends the processed information to the actual store.

10. The purchased dress will be delivered to shopper A’s home.

2.7.4 Service deliverable

Shopper A not only bought what she wants in a time-efficient and cost-effective manner, but also went through a fun, interactive and personalized experience.

2.7.5 Potential technical requirements

Immersive virtual shopping places several critical demands on wireless communication.

- High DL and UL data rate to transmit high-resolution graphics, 3D models, and user interactions;
- Extremely low E2E latency and jitter for both uplink and downlink to provide seamless immersive experience;
- Capabilities to capture users’ positions, movements and gesture variations with high accuracy;
- Provision of recommended items based on user’s preferences, history purchases, and current selections.

Among these requirements, data rate and E2E latency will have impacts on standardization bodies, e.g., O-RAN ALLIANCE and/or 3GPP. Demanding tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them. Suitable provisions for users will mostly like rely on custom implementation.

2.8 Use Case 8: Personalized Hotel Experience

2.8.1 Description

Personalized hotel experience refers to a hospitality service that is tailored to meet the specific needs, preferences, and expectations of individual guests, rather than offering a standard, one-size-fits-all approach. This kind of service recognizes that each guest is unique and seeks to provide them with a stay that feels specially curated for them. Here are key elements that characterize a personalized hotel experience.

- Streamlined check-in and check-out processes: this can be done through a mobile app, allowing for direct room access and efficient billing. During the check-in process, the guest will also be informed when the room will be ready.
- Understanding guest preferences: this involves collecting information about the guest's likes, dislikes, and special requests, often before they arrive.
- Customized services and amenities: based on the collected preferences, the hotel customizes the guest's experience. This could range from the room's layout and amenities to personalized dining options that cater to dietary restrictions or preferences.
- Localized experiences: offering guest a local experience, such as recommendations for local attractions or events based on their interests, and even customizing room decorations with local art or products.
- Feedback utilization: using guest feedback to continuously improve and personalize the service. This can involve follow-up communications after their

stay to thank them and ask for feedback, which is then used to enhance future stays.

- Emphasis on privacy and security: while personalizing experiences, hotels also need to balance this with respect for guest privacy and data security, ensuring that personal information is handled sensitively and securely.

In essence, a personalized hotel experience is about making the guest feel valued and understood, creating a memorable and comfortable stay that feels distinctly tailored to them. This approach not only enhances guest satisfaction but also builds loyalty and a strong reputation for the hotel.
~~We~~ consider an example of a family (mom, dad and kids) travelling to another city. They have made a reservation of hotel A which provide personalized hotel experience. This is illustrated in *Figure 10: Family enjoying personalized hotel experience.*

Figure 10: Family enjoying personalized hotel experience

2.8.2 Scenario illustration

1. Both the hotel and the guest's mobile phone support the personalized hotel experience
2. A cloud server has collected the guests' personal data regarding their preferences and is able to process the data.

2.8.3 Service flow

1. The family step off the plane. As they turn on their mobile phones, they receive a welcome notification from their booked hotel. The message includes a friendly greeting and weather updates for the city.
2. While in the taxi from the airport to the hotel, the parents use the mobile app to complete their check-in process. They select their room preferences, request a play table for the kid.
3. Upon reaching the hotel, they bypass the traditional check-in desk. The app guides them directly to their room.
4. Via the sensor system in the hotel, the guests' locations and trajectories can be tracked accurately. Then the elevator to their room is ready and waiting for them. The elevator's door is open as if anticipating their arrival.
5. After reaching their room, the door unlocks with a simple tap using the digital key ready on their smartphones.
6. As they enter the room, they find it perfectly set to their preference.
7. After settling in, the family uses the mobile app to explore options for dinner and entertainment. Based on their profile – which includes the child's interests and dietary preferences – the app suggests a family-friendly nearby restaurant. The profile can be pre-stored in the cloud server.
8. Throughout their stay, the family enjoys the convenience of the hotel's personalized services. They order room service through the app and get daily itinerary suggestions. The hotel staff, aware of their preferences, are always on hand to offer personalized assistance.
9. On their final day, the family checks out via the app while still in their room. The bill is settled through a saved payment method, confirming their check-out in seconds.

2.8.4 Service deliverable

The family feel a sense of satisfaction. The personalized experience makes their trip not only enjoyable but also convenient.

2.8.5 Potential technical requirements

This use case requires the following.

- High data rate to transmit high-resolution images and videos especially for downlink.
- Low E2E latency and jitter to provide real-time communication.
- Capabilities to capture guests' locations and movements with high accuracy.
- Personalized assistances based on guests' preferences.

Among these requirements, data rate, E2E latency, and uniform QoS guarantee impact technology, network architecture and standards defined by bodies like 3GPP and O-RAN ALLIANCE. Demanding tracking capabilities can be achieved by custom implementations, standardized design, or a combination of them. Appropriate personalized assistance will mostly like rely on custom implementation.

2.9 Use Case 9: Mega-Events Requiring High Data Rate

2.9.1 Description

At mega-events such as the Olympic Games and sports games, many users may stay indoors to participate and share their experiences with friends through video or XR technology, as shown in *Figure 11: Video/XR sharing in mega-event.*

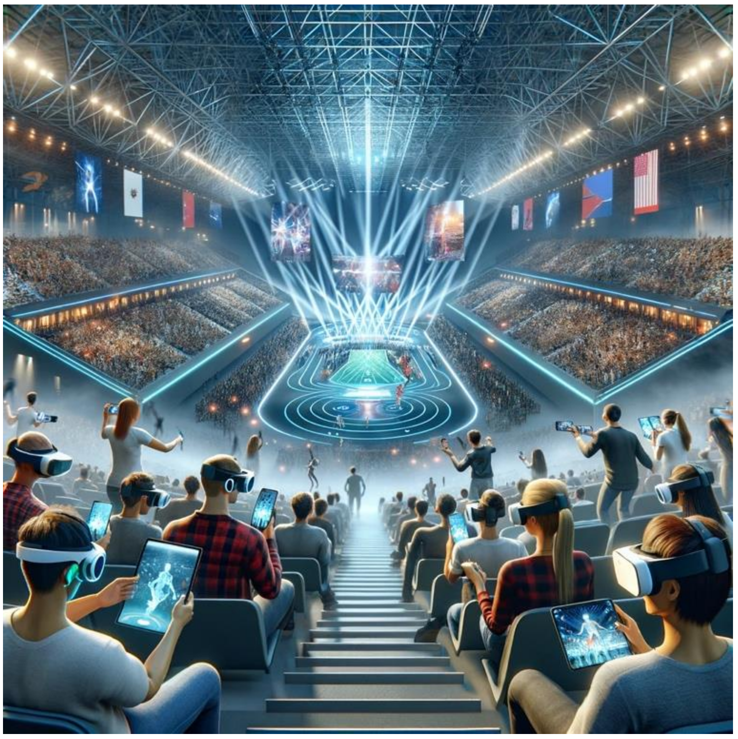

Figure 11: Video/XR sharing in mega-event

Very high data rate is required in such scenario. For example, if 5% of the 100,000 users at the Olympic Games opening ceremony were to share 4K video (around 25Mbps), the required data rate would be around 125 Gbps on the uplink. In the crowded indoor environment, frequency reuse is difficult. Without spectrum reuse, the required carrier bandwidth would be: Data Rate/ Spectral efficiency/UL ratio = 125/1.6/0.25 GHz = 312.5 GHz ~= 300 GHz, where UL spectral efficiency is assumed to be 1.6bps/Hz [6]; TDD DL/UL configuration is assumed to be 3:1. The required 300 GHz carrier frequency far exceeds the FR1 capacity and also FR1+FR2+FR3 total capacity.

Users may also watch videos shared by other people. DL data rate is usually much higher than uplink. User may also share the view and experience via XR which has even higher data rate and delay requirement than video. So, DL has similar capacity issue as UL.

In summary, meeting super high data requirement during mega-event is very difficult (if not impossible) by current 4G/5G and current spectrum.

2.9.2 Scenario illustration

Users are connected to internet and many of them are using video/XR sharing and downloading.

2.9.3 Service flow

1. User connects to internet from indoor wireless coverage
2. User connects to the video/XR server
3. User shares its video/XR to others via the server
4. User downloads and views the video/XR of other users.

2.9.4 Service deliverable

Users use high speed internet for bidirectional video/XR sharing.

2.9.5 Potential requirements

Same as use case *Use Case 1: Mixed Reality Telepresence* and *Use Case 3: Immersive Gaming/Entertainment.*

2.10 Use Case 10: Cooperative Mobile Robots

2.10.1 Description

Collaborative robots constitute a pivotal technology towards the materialization of novel industrial and domestic use cases. For example, Industry 4.0 is a paradigm where the demands of the changing market can be rapidly met by enabling the flexible rearrangement of production flows [7]. The role of collaborative robots in Industry 4.0 is to form a dynamic ecosystem that realizes the required motion and actuation according to the desired manufacturing goal, which is communicated to them through a multi-robot task assignment (MRTA) controller [7]. A wireless network facilitates the transmission of task assignments as well as the exchange of information between robots, sensors, edge devices and a control center.

While collaborative robots have been widely discussed in the context of Industry 4.0, their reach can extend beyond industrial settings towards domestic environments, where they can procure convenience, safety and accessibility by exploiting the aforementioned task MRTA and the available wireless network and existing Internet of Things (IoT) setting. For instance, when cooperative robots are used at home as illustrated in ***Error! Reference source not found.***, they can greatly enhance the efficiency and quality of our daily life. Potential applications of collaborative robots in domestic settings include [8]: assisting in chores such as folding laundry; home healthcare for individuals with mobility issues or vulnerable members of the population; educational support for tutoring, assistance with homework and homeschooling; companionship for the elderly; and safer home maintenance.

Figure 12: Applications of cooperative robots at home

For example, consider the scenario illustrated in *Figure 12: Applications of cooperative robots at home*, where two collaborative robots aid in the tasks of elderly persons at an assisted living facility. The domestic environment includes a wireless network with multiple access points (APs), edge devices and UEs, and smart sensors.
One of the edge devices serves as a control center and transmits the MRTA instructions to the agent depending on the time of the day and the status of the environment, as determined through information from the sensors which is either post-processed locally by the edge devices or streamed to the cloud. The *Figure 13: Collaborative robots in domestic scenarios* illustrates three situations where the robots fulfil different roles and tasks:

1) Robot 1 assists subject 1 in opening a jar in the kitchen and robot 2 assists subject 2 in dressing.
2) Robot 1 and robot 2 assist a disabled subject in getting out of bed and getting to a different location, respectively.
3) Robot 1 is idle while Robot 2 displays the news to subject 3.

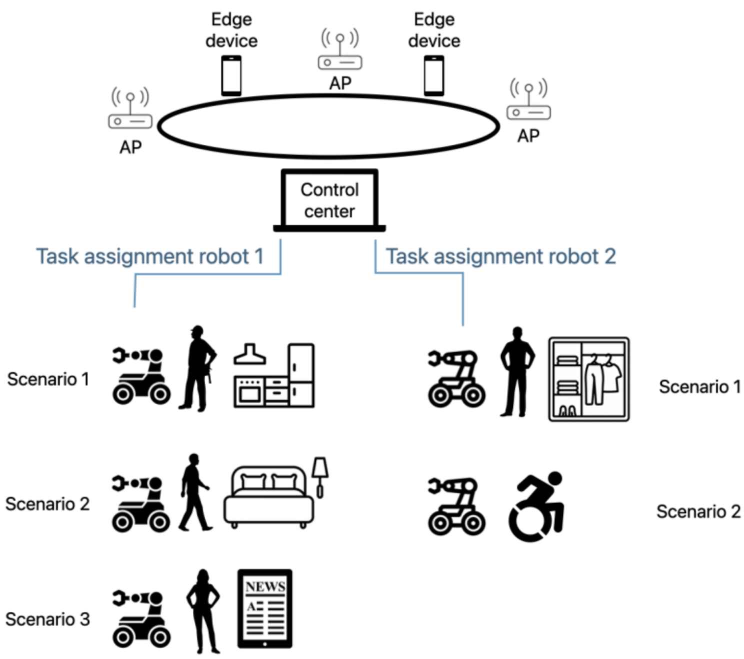

Figure 13: Collaborative robots in domestic scenarios

2.10.2 Scenario illustration

- The robots are connected to the local network and their location is known through post-processing of their onboard sensors or other environmental sensors.
- Environmental sensors (cameras, temperature sensors, etc) are available and their status must be delivered to one of the edge devices or the centralized controller to track the location of the human subjects in the environment.
- At least one of the robots is equipped with actuators (e.g., robotic hands) with a very low latency for real time.
- Edge devices are available for post-processing of some of the sensor data, while the network is also able to send heavy computational workloads to the cloud if needed (e.g., for computer vision or sensor post processing tasks).
- One of the edge devices can generate and send the Multi-Robot Task Assignments.
- The robots and environment are equipped with sensors, whose status must be delivered to a centralized controller and actuators with a very low latency for real time control.

2.10.3 Service flow

1. The sensors identify the status and location of the human subjects. Sensor information can directly be sent to the command center or through edge devices available in the environment.
2. The subjects provide voice commands or hand/arm gestures to indicate which task they need assistance with.
3. The control center processes sensor information and subject commands to generate the task assignments for the robots.

4. The robots execute the desired task and transmit their current state and control actions to the factory controller. Available sensors keep sending information to edge devices and the control center about the status of the robots, the environment and the subjects.
5. Steps 3 and 4 are executed simultaneously until the task is completed, as indicated by the human subjects.

2.10.4 Service deliverable

The task is executed successfully ensuring safety for the human subjects involved.

2.10.5 Potential technical requirements

- A centralized controller that is capable of transmitting and receiving sensor information and control commands, respectively with very low latency (e.g., < 1 ms) [9].
- A variety of robot mounted and environmental sensors such as lidar and cameras.
- Robots with high precision actuation capabilities.
- Environmental sensors.

Among these requirements, data rate, required low latency will have impacts on standardization bodies, e.g., O-RAN ALLIANCE and/or 3GPP. Robots with high precision actuation capabilities will mostly like rely on custom implementation.

2.11 Use Case 11: Telemedicine and Remote Healthcare

2.11.1 Description

The use case of Telemedicine and Remote Healthcare envisions a future where healthcare delivery transcends physical boundaries to offer comprehensive, patient-centric and high-quality medical services remotely as well as country-wise.
In this use case, telemedicine encompasses a wide array of applications: from virtual consultations and remote patient monitoring to sophisticated telesurgery. Doctors and medical professionals can engage with patient in real-time, despite being miles apart, ensuring timely medical advice, diagnosis, and treatment.

- In virtual consultations, patients can engage with healthcare providers from the comfort of their own homes or any location with internet access. Through a secure digital platform, they can schedule appointments, share health records, and receive medical consultations via video conferencing. This setup not only saves time and reduces the need for travel but also minimizes exposure to potential pathogens in clinical settings—a particularly valuable benefit in the wake of global health crises.
- Remote patient monitoring becomes more efficient with advanced wireless technology, allowing for continuous, non-invasive monitoring of patients' health indicators through wearable sensors and embedded devices. These devices can measure vital signs such as heart rate, blood pressure, and glucose levels, transmitting data seamlessly to healthcare providers for ongoing assessment. This capability is particularly beneficial for managing chronic conditions, post-operative care, and elderly care, offering a proactive approach to healthcare that can prevent hospital readmissions and improve quality of life.

- At its core, telesurgery involves a surgeon operating robotic instruments remotely to perform surgical interventions on a patient who is not in the same physical location as the surgeon. This is facilitated through a real-time, high-definition video feed that allows the surgeon to see the surgical site as if they were present in the operating room. The key to telesurgery's success lies in the seamless integration of robotics, advanced imaging, and ultra-fast, reliable communication networks.

By harnessing the power of advanced wireless connectivity, the Telemedicine and Remote Healthcare use case promises to revolutionize the healthcare industry. It aims to overcome geographical barriers, enhance the efficiency and effectiveness of healthcare delivery, and ultimately improve patient outcomes. This futuristic vision of healthcare is not only about embracing technological advancements but also about reimagining the possibilities for care in the digital age.

We consider an example of remote surgery as described in the following. In a remote village, miles from the nearest city hospital, Patient A is diagnosed with a condition that requires specialized surgery. The nearest surgeon with the necessary expertise is Dr. B, who is located in a state-of-the-art medical center in a bustling metropolitan city thousands of miles away. With traditional surgery, Patient A would face significant challenges, including the risk of long-distance travel and the potential wait times for surgery. However, the advent of remote surgery offers a different path, as illustrated in *Figure 14: Remote surgery.*

Figure 14: Remote surgery

2.11.2 Scenario illustration

1. Around Patient A, there is local hospital equipped with remote surgery capabilities. The operating room is fitted with advanced robotic arms, high-definition cameras, and a large display screen.

2. At Dr. B's office or operating room, there are remote surgery capabilities with controls that are directly linked to the robotic surgical instruments in the above operating room in Patient A's local hospital.

2.11.3 Service flow

1. As the surgery date approaches, Patient is admitted to the local hospital with remote surgery capabilities.
2. Meanwhile, Dr. B prepares for the surgery at his location.
3. On the day of the surgery, Dr. B and his team, along with Patient A's local medical team, connect via a secure, ultra-high-speed wireless network.
4. With the real-time communication provided by the wireless network, every movement Dr. B makes with the controls is instantly mirrored by the robotic arms performing the surgery on Patient A.
5. The high-definition cameras provide Dr. B with a crystal-clear, 3D view of the surgical site, allowing for precision that rivals, and in some cases surpasses, traditional hands-on surgery.
6. Throughout the procedure, vital signs and other critical health data from Patient A are continuously monitored and shared in real time, ensuring that Dr. B and both medical teams are fully informed and can quickly respond to any situation.
7. As Dr. B completes the final stitches, the surgery is successful.
8. Patient A's recovery begins under the careful watch of her local healthcare providers, with Dr. B available for follow-up consultations via virtual meetings.

2.11.4 Service deliverable

Through the power of remote surgery, Patient A receives the best possible care, despite the physical distance from Dr. B.

2.11.5 Potential technical requirements

This use case requires the following.

- Extremely high data rate to transmit and receive high-definition live video.
- High reliability and extremely low E2E latency and jitter to transmit and receive haptic feedback controls of robotic arms.
- High-accuracy and resolution sensing for imaging, motion tracking, and biometric monitoring.

Among these requirements, data rate and E2E latency will impact technology, network architecture and standards defined by bodies like 3GPP and O-RAN ALLIANCE. Demanding sensing capabilities can be achieved by custom implementations, standardized design, or a combination of them.

3 Conclusion

The exploration of emerging indoor use cases within this report unveils a future where the physical and digital worlds are seamlessly integrated, redefining the boundaries of communication, entertainment, education, and industry. As we have seen, each use case brings new potentials of immersive experiences and operational efficiencies.

The importance of these use cases cannot be overstated. They promise to enhance human interactions, expand educational horizons, revolutionize entertainment, transform fitness regimes, streamline media consumption, enable new dimensions in communication, reimagine retail, personalize hospitality, enrich mega-events, and automate industry with unprecedented precision. Through detailed examination, it is clear that the service flows of these applications demand not only high data rates and low latency but also reliability, positioning and sensing accuracy, ubiquity, and security. Compared to outdoor environments, indoor use cases have their distinct features, e.g., higher signal attenuation, higher user density, and more stringent security requirement. These factors add layers of difficulty to wireless communication within indoor setting.

In conclusion, the emerging indoor use cases present an appealing vision of the future. By addressing the associated challenges and utilizing the collective expertise around the whole world, we can unlock a new dimension of connectivity that enriches the human experience in all ways.

References

[1] ABI Research Anticipates In-Building Mobile Data Traffic to Grow by More Than 600% by 2020, https://www.abiresearch.com/press/abi-research-anticipates-building-mobile-data-traf/

[2] Paul Milgram and Fumio Kishino, "A taxonomy of mixed reality visual displays. " *IEICE Transactions on Information System*, Vol EE77-D, No.12 December 1994.

[3] Telepresence, https://en.wikipedia.org/wiki/Telepresence

[4] Six degrees of freedom, https://en.wikipedia.org/wiki/Six_degrees_of_freedom#:~:text=Six%20degrees%20of%20freedom%20(6DOF,body%20in%20three%2Ddimensional%20space

[5] Ribbon Research on Cloud Gaming Reveals Revenue-Generating Opportunities for 5G Networks, https://www.prnewswire.com/news-releases/ribbon-research-on-cloud-gaming-reveals-revenue-generating-opportunities-for-5g-networks-301128554.html

[6] The economics of mmWave 5G (gsmaintelligence.com), https://data.gsmaintelligence.com/api-web/v2/research-file-download?id=59768858&file=210121-Economics-of-mmWave.pdf

[7] Chen, Kwang-Cheng, et al, "Wireless networked multirobot systems in smart factories." *Proceedings of the IEEE* 109.4 (2020): 468-494.

[8] Cobots at Home, "The Future of Collaborative Robots in Household Settings. " Global Robot Marketplace. https://www.linkedin.com/pulse/cobots-home-future-collaborative-robots-household-settings-robotmp/

[9] Yang Yulin, "What are Cooperative Robots and Collaborative Robots?" https://www.idtechex.com/en/research-article/what-are-cooperative-robots-and-collaborative-robots/28376