



POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Mixed Reality Applications for Industrial Maintenance and Special Education of ASPOC and ILSEME Association Students

TESI DI LAUREA MAGISTRALE IN
MECHANICAL ENGINEERING - INGEGNERIA MECCANICA

Author:

Kavin Prakasam

Co-Author:

Keerthivasan Dhandapani Ramesh

Author Student ID: 943906

Co-Author Student ID: 943908

Advisor: Prof. Mario Covarrubias Rodriguez

Academic Year: 2022-23

Abstract

In Our Contemporary, high-speed world, technology has become an irreplaceable cornerstone in numerous industries. This context delves into the profound impact of Mixed Reality (MR), an emerging technology that seamlessly blends the virtual and physical realms, on education and industry. MR offers immersive and interactive experiences, creating innovative solutions to complex problems. In industry, MR technology streamlines processes, boosts efficiency, and reduces costs. Remote Assistance applications enable real-time communication between the local maintenance group and the expert teams from the machine's manufacturer, irrespective of geographical distance. This accelerates troubleshooting, minimizes downtime, and enhances productivity. Mixed reality devices like the Hololens2 provide real-time guidance, improving precision and confidence in complex tasks. Immersive simulations in the industrial sector enhance training programs, ultimately leading to increased productivity. In the education sector, MR promises to revolutionize learning experiences, for all students, particularly in assisting students with cognitive disabilities. It enables both students to visualize complex concepts in three dimensions, making learning engaging and memorable. Utilizing devices like the Hololens2, students gain access to virtual demo applications, facilitating hands-on learning across subjects. These applications offer step-by-step instructions, allowing students to learn at their own pace and in their preferred environment. Furthermore, Organizations use MR, coupled with Vuforia Markers and the Hololens2, to guide individuals in learning basic day-to-day activities. This proactive approach empowers them to participate in work and self-care, fostering a sense of normalcy.

The context highlights the positive impact of MR in education and industry, emphasizing its potential to transform learning experiences and streamline processes. While challenges remain, the adoption of MR technology holds immense promise in enhancing education and technology across various sectors.

Keywords: Mixed Reality (MR), Hololens2, Industrial Remote Assistance, Cognitive Retardation pupils, Vuforia markers, System usability scale (SUS)

Abstract in lingua italiana

Nel nostro mondo contemporaneo ad alta velocità, la tecnologia è diventata un pilastro insostituibile in numerose industrie. Questo contesto approfondisce l'ampio impatto della Realtà Mista (MR), una tecnologia emergente che fonde in modo fluido i mondi virtuale e fisico, sull'istruzione e sull'industria. La MR offre esperienze immersive ed interattive, creando soluzioni innovative per problemi complessi. Nell'industria, la tecnologia MR semplifica i processi, aumenta l'efficienza e riduce i costi. Le applicazioni di assistenza remota consentono la comunicazione in tempo reale tra operatori ed esperti, indipendentemente dalla distanza geografica. Ciò accelera la risoluzione dei problemi, riduce i tempi di inattività e migliora la produttività. Dispositivi di realtà mista come l'Hololens2 forniscono assistenza in tempo reale, migliorando la precisione e la sicurezza nelle attività complesse. Simulazioni coinvolgenti nel settore industriale potenziano i programmi di formazione, portando infine ad un aumento della produttività. Nel settore dell'istruzione, la MR promette di rivoluzionare le esperienze di apprendimento, per tutti gli studenti, in particolare nell'assistenza agli studenti con disabilità cognitive. Consente a tutti gli studenti di visualizzare concetti complessi in tre dimensioni, rendendo l'apprendimento coinvolgente e memorabile. Utilizzando dispositivi come l'Hololens2, gli studenti hanno accesso ad applicazioni dimostrative virtuali, facilitando l'apprendimento pratico in diverse materie. Queste applicazioni offrono istruzioni passo dopo passo, consentendo agli studenti di apprendere al proprio ritmo e nell'ambiente preferito. Inoltre, le organizzazioni utilizzano la MR, in combinazione con i Marcatori Vuforia e l'Hololens2, per guidare le persone nell'apprendimento di attività quotidiane di base. Questo approccio proattivo permette loro di partecipare al lavoro e all'autonomia, promuovendo un senso di normalità. Il contesto evidenzia l'impatto positivo della MR nell'istruzione e nell'industria, sottolineando il suo potenziale per trasformare le esperienze di apprendimento e semplificare i processi. Nonostante le sfide rimangano, l'adozione della tecnologia MR offre un'immensa promessa nell'ambito dell'istruzione e della tecnologia in vari settori.

Parole chiave: Realtà mista (MR), Hololens2, Assistenza remota industriale, Studenti con ritardo cognitivo, Marcatori Vuforia, Scala di utilizzabilità del sistema (SUS).

Contents

Abstract	i
Abstract in lingua italiana	iii
Contents	v
Introduction	1
1 MIXED REALITY	3
1.1 Extended Reality (XR):	4
1.2 Augmented Reality (AR):	5
1.3 Virtual Reality (VR):	6
2 HOLOLENS 2	9
3 Developing Remote Assistance Application for FIMI-Machinery	15
3.1 Backend Network Development	16
3.1.1 Client-Server Architecture	17
3.1.2 WebRTC	21
3.2 Front end Application Development	35
3.2.1 Developing process of Mixed Reality application with Real-time Communication	35
3.2.2 Developing Mixed Reality environment for FIMI Simulations of Ma- chinery	44
3.3 Summary	48
4 Developing Educational Applications for Polimi and ASPOC Students	49
4.1 Students with Congenital Retardation:	49
4.1.1 Common Challenges faced by students with Cognitive Retardation in education programs:	50

4.1.2	Challenges faced by students with Cognitive Retardation in using Mixed Reality Technology:	52
4.1.3	Design strategies implemented for individuals with Cognitive disabilities:	53
4.2	Special Education Application development:	53
4.2.1	CHARLY 4U	53
4.2.2	UNITY	54
4.3	Summary	61
5	Developing Special Education Applications for the ILSEME Association	63
5.1	Special Education Application Development using Vuforia Engine	64
5.2	Summary	68
6	System Usability Scale (SUS)	69
6.1	SUS data Collected from FIMI Machinery	72
6.2	SUS data Collected from students of Politecnico Di Milano	73
6.3	SUS data collected from ASPOC students	74
6.4	SUS data collected from ILSEME Association	75
7	Conclusions	77
8	Future Improvements	79
	Bibliography	81
	List of Figures	85
	List of Tables	87

Introduction

In our fast-paced and technology-driven world, the role of technology has become indispensable in various sectors. Technological advancements have made it possible to find efficient solutions to complex problems and streamline various processes for time-saving benefits. One such groundbreaking technology that is revolutionizing multiple sectors is Mixed Reality. Mixed Reality, also known as MR, is an emerging technology that bridges the gap between the virtual and physical realms. This technology allows users to interact with digital content in a real-world environment, creating immersive and interactive experiences. With the rapid development of Mixed Reality technology, it has the potential to transform education and, industry sectors. Mixed Reality technology holds immense promise for the education sector, offering innovative ways to enhance learning experiences. Mixed Reality technology has the potential to revolutionize the way students learn by creating immersive and interactive experiences. It allows students to visualize complex concepts in a three-dimensional space, making learning more engaging and memorable. By using devices like the HoloLens 2, students can access virtual demo applications that facilitate hands-on learning in various subjects. These applications provide step-by-step instructions, allowing students to learn at their own pace and in their preferred learning environment. In the Industrial sector, By implementing Mixed Reality technology in the industrial sector, companies can streamline various processes, improve efficiency, and reduce costs. Remote Assistance applications using Mixed Reality technology enable the local maintenance group to communicate with the expert teams from the machine's manufacturer in real-time, regardless of geographical distance. This facilitates troubleshooting and problem-solving processes, leading to faster resolution times and minimizing downtime. By using mixed reality devices like the HoloLens2, the local maintenance group can receive real-time instructions, annotations, and guidance from expert teams of the machine's manufacturer, allowing them to carry out complex tasks with precision and confidence. Additionally, mixed-reality technology in the industrial sector can enhance training programs by providing immersive and interactive simulations [28]. These simulations can help employees develop practical skills and improve their understanding of complex procedures, ultimately leading to increased productivity and better

performance. Virtual and Mixed Reality technologies have the potential to transform the education sector by providing immersive and engaging learning experiences. In Education Sector, There are several organizations that have taken the initiative to engage individuals with cognitive retardation in basic day-to-day activities with the help of Mixed Reality technology, which acts as a proactive approach to enable students with Cognitive Retardation to participate in work and self-care, fostering a sense of normalcy and inclusion. The Hololens2 recognizes Vuforia Markers, which are utilized as a means to guide individuals with Cognitive disabilities in observing and learning basic activities that they can subsequently replicate in a real-world setting. This method allows them to learn at their own pace, repeating the process as many times as necessary until they feel confident performing these activities in a real-world scenario.

1 | MIXED REALITY

A **Mixed Reality (MR)** Experience is one that unifies the user's real-world environment and virtual world, where both environments can co-exist and interact with each other. It allows users to interact with digital content in a way that feels natural and immersive, thereby enhancing their perception and understanding of the surrounding environment. This technology enables users to manipulate and interact with virtual objects using natural gestures and movements, further blurring the line between the physical and virtual worlds. With the assistance of hand tracking or motion controllers, users can seamlessly grab, move, rotate, and resize virtual objects as if they were tangible physical objects. Mixed Reality depends on the evolving relationship between humans and machines. For efficient work, an MR solution needs to understand the various actions of a person in a specific space, as well as the surrounding landscape. For that, Mixed Reality uses a series of cameras, sensors, and often AI-enhanced technology to process data about space and use that information to create digitally-enhanced experiences. It can scan the space and project a mesh-like structure as a safety standard. By tracking the environment, the software creates a virtual map, where the MR technology can add holographic images and content into the world through the use of image projections.

With this Mixed Reality technology, we could:

Innovate faster: Construction and design teams can use Mixed Reality to easily create digital twins of products and explore creative ideas with minimal waste or expense. With MR, you can easily experiment with different materials and ideas in seconds.

Collaborate: Imagine being able to build a holographic version of a new car with your colleagues, who are all working hundreds of miles away from each other. Similar to VR and AR, Mixed Reality allows for the creation of more immersive collaborative experiences.

Support workers: Speaking of collaboration, cameras on Mixed Reality headsets can stream content to colleagues in real-time, so specialists can annotate what they see (like the image of a machine in front of another employee) and provide guidance in real-time.

Improve training: Extended Reality has incredible potential as a training and development tool. With Mixed Reality, companies can create immersive employee training experiences in all environments. Not only does this help with building muscle memory, but it can also reduce the risk often associated with some training environments.

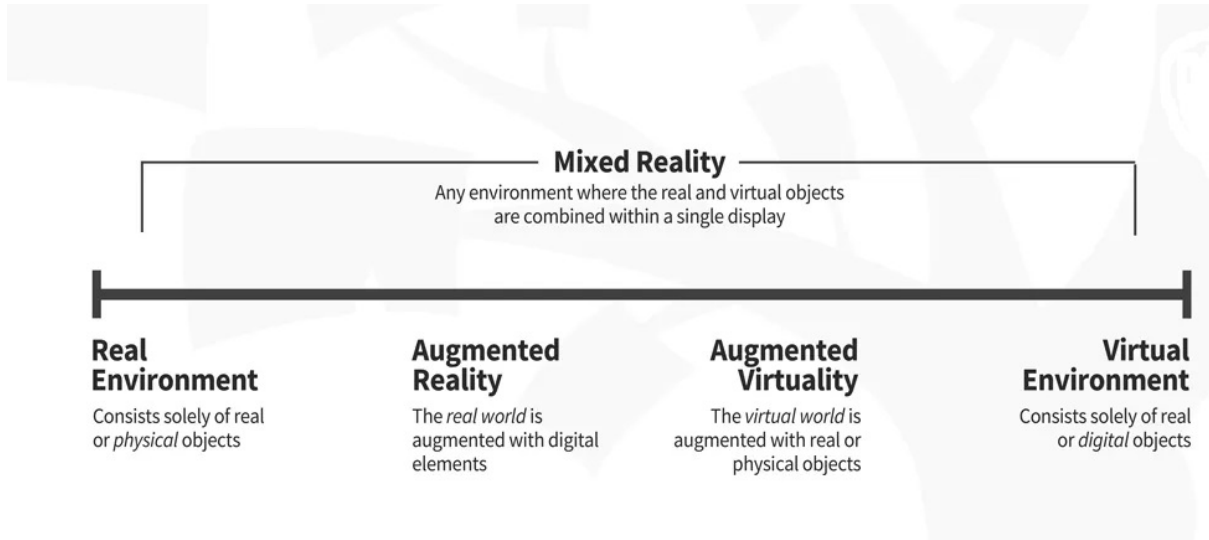


Figure 1.1: Mixed Reality branching [5]

1.1. Extended Reality (XR):

Extended Reality (XR) is a collective term that holds any sort of technology that alters reality by adding digital elements to the physical or real-world environment to any extent, blurring the line between the physical and the digital world. XR is a library where AR, MR, VR, and any future technology related to the virtual continuum is born.

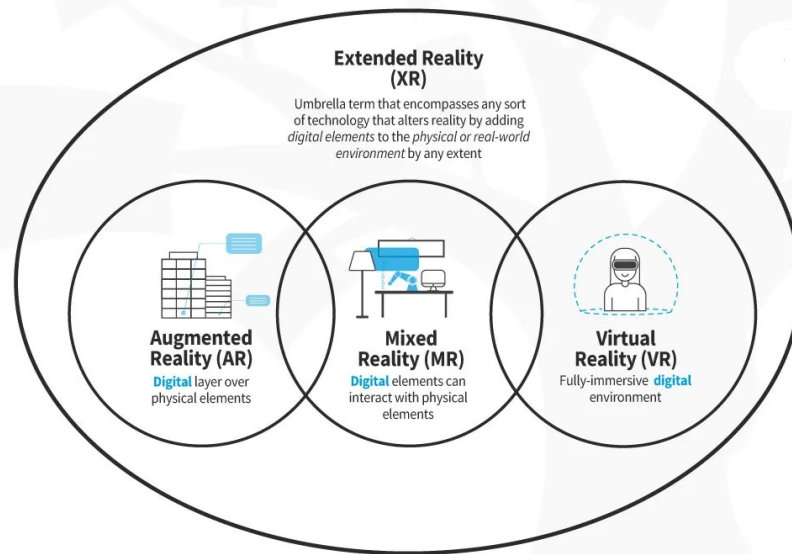


Figure 1.2: Extended reality (XR) [5]

1.2. Augmented Reality (AR):

Augmented Reality (AR) is a technology that blends the real-world with the virtual world with a certain degree of interaction between them. AR can be delivered or felt in a variety of formats, including smartphones, tablets, and AR glasses.



Figure 1.3: Augmented Reality (AR) [3]

1.3. Virtual Reality (VR):

Virtual Reality (VR) is a technology, which immerses itself in a specifically designed/Stimulated fully digital environment generated by the computer. It is a self-controlled environment where the user can control the simulated environment via a system. The VR experiences are felt only in the extreme virtual continuum.



Figure 1.4: Oculus Quest 2 [12]

By combining elements of both Virtual Reality (VR) and Augmented Reality (AR), Mixed Reality creates a unique user experience that merges the real-world with digital content. Unlike VR, which completely immerses users in a virtual environment, Mixed Reality overlays virtual objects onto the real world, allowing users to see and interact with both simultaneously. This integration of virtual and real-world elements provides a more seamless and intuitive interaction with digital content.

The ability to interact with digital content in a natural and immersive manner is one of the key advantages of Mixed Reality. Traditional methods of interacting with computers, such as using a keyboard and mouse or touchscreen, can often feel disconnected from the physical world. Mixed Reality bridges this gap by enabling users to use their hands and body movements to directly manipulate virtual objects. This hands-on approach not only enhances user engagement but also facilitates a deeper understanding of the digital content being interacted with.

Furthermore, Mixed Reality offers a level of realism that is unparalleled by other technologies. By overlaying virtual objects onto the real-world, Mixed Reality creates a sense of presence and depth that is not possible with traditional 2D displays. Users can perceive virtual objects as if they were physically present in their environment, allowing for more realistic interactions and simulations. This realism opens up new possibilities for various applications, including training simulations, design prototyping, entertainment experiences, and educational tools.

In terms of input methods, Mixed Reality offers multiple options for interacting with virtual objects. Hand-tracking technology allows users to use their hands and fingers to directly manipulate virtual objects. This intuitive method of interaction closely mimics real-world actions and gestures, making it easy for users to grasp and manipulate virtual objects. Additionally, motion controllers provide a more precise and versatile input method, allowing users to perform complex interactions and gestures with ease. These input methods, combined with the immersive nature of Mixed Reality, create a highly interactive and engaging user experience. In short, Mixed Reality offers a level of realism and immersion that is unparalleled by other technologies, opening up new possibilities for various applications.

2 | HOLOLENS 2

Hololens2 is a Mixed Reality(MR) headset developed by Microsoft Corporation. It is the successor to the original Hololens and offers significant improvements in terms of comfort, field of view, and interaction capabilities. This advanced device combines Virtual Reality (VR) and Augmented Reality (AR) technologies to create a unique Mixed Reality experience. One of the key features of this device is its ergonomic redesign, which enables users to wear it comfortably for extended periods. It features a lightweight carbon fiber frame with an adjustable headband. Additionally, it provides a significantly larger field of view, enabling users to see more virtual objects in their surroundings. This wider field of view enhances immersion and makes the Mixed Reality experience more realistic. Hololens2 also introduces improved interaction capabilities, enabling hand-tracking technology that allows users to interact with virtual objects using natural hand gestures, such as pinch, grab, zoom-in, zoom-out, etc. This eliminates the need for external controllers or input devices, making the user experience more intuitive and seamless. Furthermore, it supports eye tracking and head tracking, enabling users to navigate menus and interact with holographic content simply by looking at and interacting with them. The device also incorporates spatial sound technology, which enhances the audio experience by providing accurate sound positioning based on the user's location and orientation. This creates a more immersive environment where sounds seem to come from specific directions about virtual objects. It is powered by a custom-built Microsoft Holographic Processing Unit (HPU), which enables real-time processing of sensor data and advanced spatial mapping. The HPU combines depth sensing, spatial mapping, and artificial intelligence algorithms to accurately track the user's movements and position virtual objects in the real-world. This ensures seamless integration of virtual content with the user's physical environment. In terms of platform, it operates on Windows Mixed Reality and Open XR, enabling developers to create a diverse range of Mixed-Reality applications. With this platform, developers can build immersive experiences that seamlessly combine elements from both virtual and Augmented Reality. Hololens2 has found applications in various industries, including manufacturing, healthcare, education, and entertainment.

In manufacturing, it can be used for remote assistance, allowing expert teams from the machine's manufacturer to guide the local maintenance group in real time using mixed reality annotations.

In education, it can provide interactive learning experiences by bringing virtual objects into the classroom.

In entertainment, Hololens2 can be used for immersive gaming experiences that blend virtual and real-world elements. Hololens2 paves the way for advanced Mixed Reality technology, which has the competence to seamlessly blend virtual and real-world environments. Furthermore, it opens up new possibilities for an immersive experience and revolutionizes digital information in the forthcoming future.

Hololens2 is built in with a few advanced features that provide us the immersive experiences which are:

1. Spatial mapping: It is a detailed representation of the real-world environment in Hololens2, It helps the user to gain a better understanding of the real-world when using Hololens2 and also it provides a more natural view for the user while interacting.

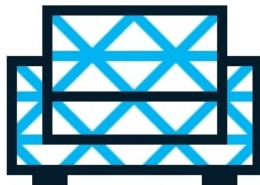


Figure 2.1: Spatial Mapping [8]

2. Natural interaction: Hololens2 helps the user to experience Natural interaction with holograms projected from it via hand gestures.



Figure 2.2: Natural Interaction [8]

3. Spatial sound: It connects holograms more intensively with the Mixed Reality world and gives more detailed information updates about the environment and objects around the user. It helps the user to recognize the environmental sound directions and reciprocate according to them. This gives confidence to the user in gesture and voice interactions.



Figure 2.3: Spatial Sound [8]

4. Voice Recognition: It helps the user to interact with the Hololens2 through voice commands. It gives access to hands-free modalities, where we can command the Hololens2 to close/Open an application.

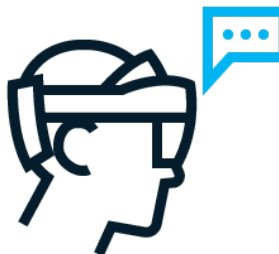


Figure 2.4: Voice Recognition [8]

5. Eye Tracking: It helps to measure the user's eye movement by a set of sensors. It records the user's eye gaze and gives input to the system more precisely which hologram is to be interacted with. Provided it helps in the calibration of holograms according to the user's height, pupil movements, etc.

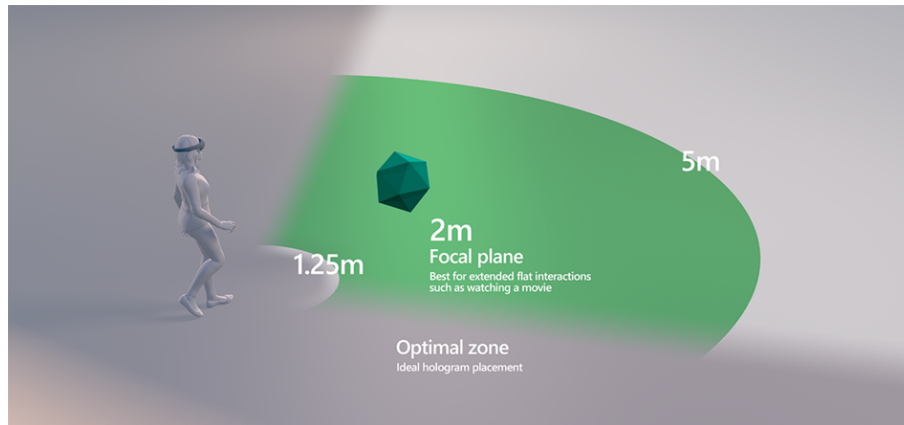


Figure 2.5: Eye Tracking [6]

6. Six degrees of freedom (6DOF) tracking: Six degrees of freedom (6DOF) tracking systems give HoloLens2 the ability to track and respond to the user's movement in all 6 degrees of freedom. This feature allows the user to move and rotate their body freely in physical space, and the movements are reflected accurately in the virtual environment. This gives a realistic and intuitive interaction experience by using special hardware like sensors, IMU, HPU, etc...

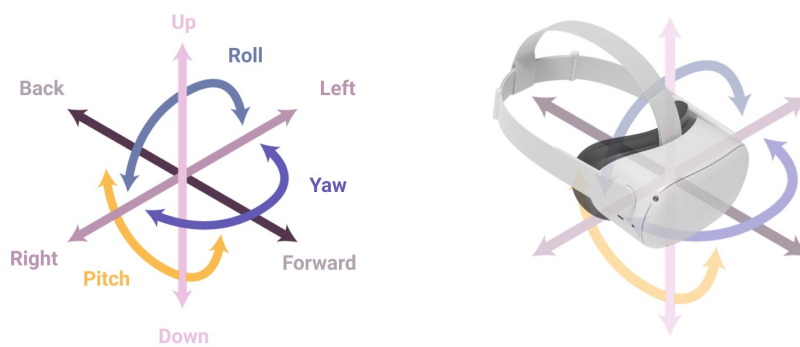


Figure 2.6: 6 DoF in HoloLens2 [1]

HoloLens2 Exploded View: The Exploded view depicts the numerous modules which are responsible for delivering the immersive and realistic Mixed Reality experience to the user.

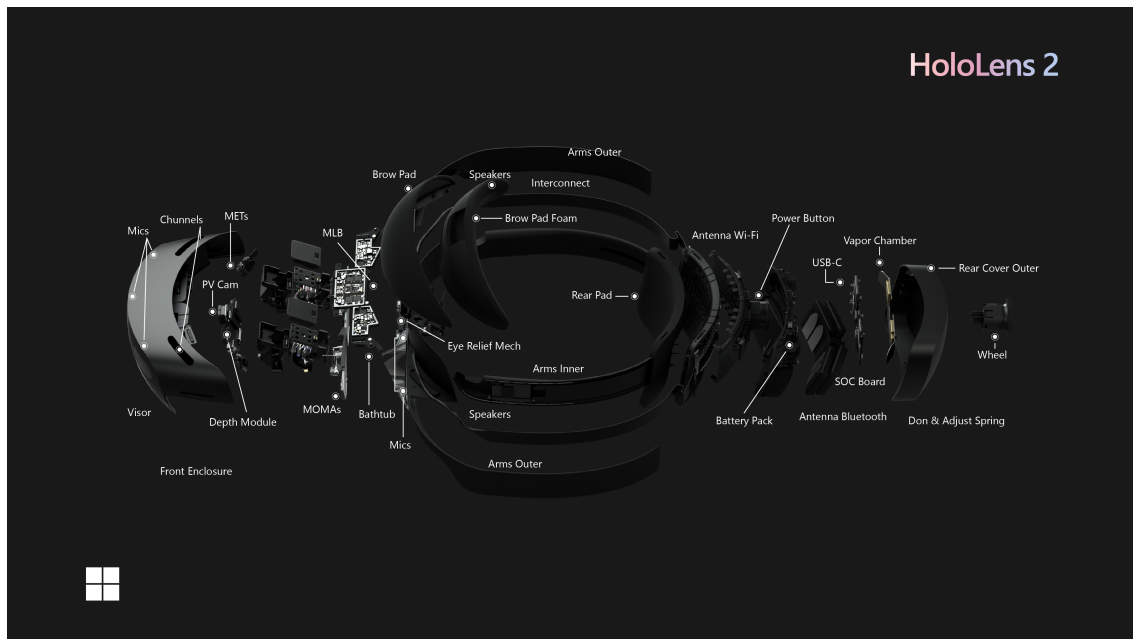


Figure 2.7: Hololens2 Exploded View [7]

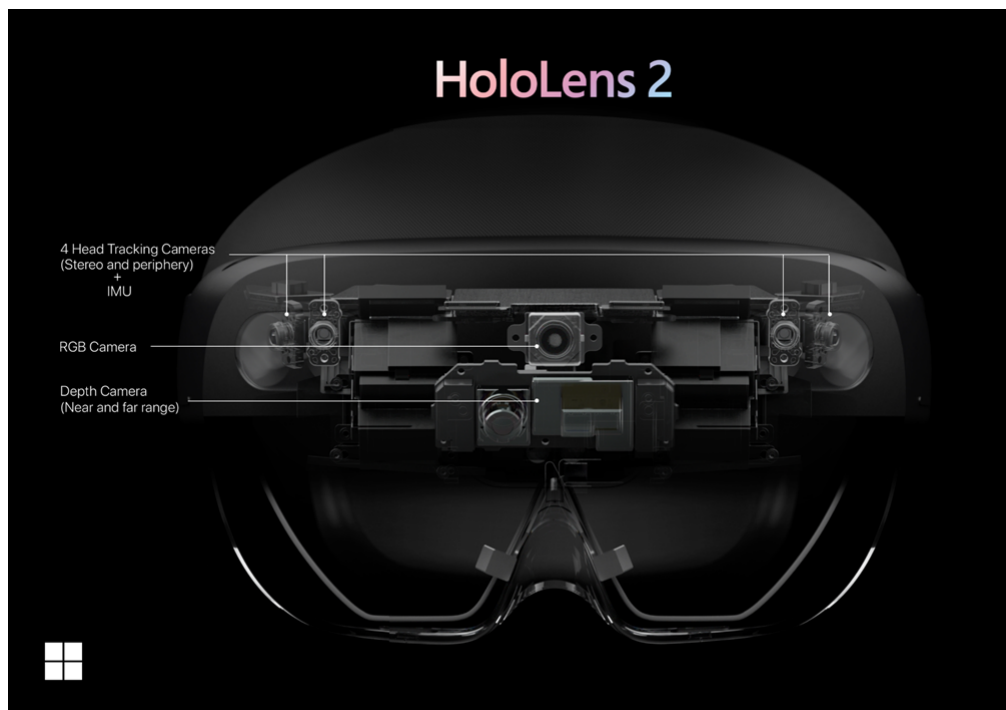


Figure 2.8: Hololens2 Front View [7]

Technical Data	Hololens2
Resolution	2048 x 1080 (per eye)
Weight	566 g
Operating system	Windows Holographic
Chipset	Qualcomm Snapdragon 850
Memory	4 GB RAM
Storage	64-GB UFS 2.1
Camera	8MP, 1080p video
Microphone array	5 channels
Speakers	Built-in spatial sound
Head tracking	4 visible light cameras
Eye tracking	2 infrared (IR) cameras
Depth	1-MP Time-of-flight depth sensor
Inertial measurement unit (IMU)	Accelerometer, gyroscope, magnetometer
Eye-based rendering	Display optimization for 3D eye position
Holographic Processing Unit (HPU)	Second-generation custom-built holographic processing unit
Hand Interaction Gestures	Yes
Voice command	Yes
Biometric identification	Yes
Battery life	2-3 hours of active use
Release date	November 2019
Price	3.849,00 €

Table 2.1: Technical Data of HoloLens2

3 | Developing Remote Assistance Application for FIMI-Machinery

Machinery maintenance constitutes an integral aspect of every industry, ensuring the continuous operation of machines. This maintenance is typically carried out at regular intervals and can be part of a routine check before issues such as lubrication leaks or drive train malfunctions occur. At a higher level, maintenance encompasses a strategy that involves inspecting, monitoring, and addressing problems that may arise during machinery operation. It may also involve assessing damage and replacing parts when machines break down. Industrial machines are often large and complex, requiring regular maintenance to ensure the smooth operation of industrial processes. Consequently, complex machinery necessitates specialized maintenance solutions. Advancements in technology have enhanced the assessment of machinery conditions by incorporating wireless portable data-collecting sensors and condition-monitoring software. These innovations allow for the development of sophisticated maintenance plans tailored to various machines used in different processes. Performing maintenance on these intricate machines demands a high level of knowledge and expertise in assembling and dismantling them. Typically, industries procure heavy machinery from manufacturers based on their specific requirements. While routine maintenance can be managed by in-house maintenance departments, more complex maintenance tasks, especially those within less accessible parts of the machinery, may require assistance from the manufacturers. In such cases, manufacturers often dispatch experts to perform maintenance tasks. However, when the machinery is purchased from a different country, the manufacturer's representative must travel to the location to carry out repairs. This can result in production downtime, incurring both financial costs and time delays.

To mitigate both cost and time constraints, a potential solution involves remotely assisting the local maintenance team with specialized experts from the machine's manufacturing company. As previously mentioned, the advancement of new technologies allows us to

monitor the machines' proper functioning by deploying various sensors and collecting data. Once a problem is identified through the collected data, experts can offer remote training and guidance to the local maintenance team, eliminating the need for travel and minimizing plant downtime.

As a solution to the aforementioned problem, the following context describes the development of a Mixed Reality application for HoloLens2. This application facilitates global connectivity through internet-based video conferencing, enabling seamless communication. The main feature of this application is that the local maintenance group can wear these Mixed Reality headsets and connect with the expert teams from the machine's manufacturer and the expert can see what the local team is seeing from a First-person perspective by accessing the camera feed from the HoloLens2. Simultaneously, the expert team can use holographic animations to guide the local team in performing necessary tasks. Developing this application involves two main components: creating backend operations to establish web communication between the HoloLens2 and other systems (including the development of signaling servers and security protocols) [31], and generating the required holographic animations for task guidance. Ultimately, these distinct functions need to be seamlessly integrated into a single application, ensuring they can run concurrently without issues. The subsequent chapters will provide a detailed overview of this development process.

3.1. Backend Network Development

Backend Development is also known as the Server-side Development for the application. It helps in saving and organizing important data on the client side of the application. It involves the development of the right security protocols to ensure secure data transmission. For example, To contact someone through an application the process of transmitting data to the server, identifying the correct client ID to establish a connection, and handling all behind-the-scenes operations within the application falls under the domain of Backend Development. On the other hand, the development of user interfaces falls under front-end development. There are two architecture types used in video conferencing technology, which are the following,

- Client-Server Architecture
- WebRTC

3.1.1. Client-Server Architecture

Client Server model is a process of dividing and sharing the workload between the provider of data called **server** and the one which requests the services called **client**. The server runs multiple programs, which it tends to share with one or more clients. In contrast, clients do not share their resources but instead request services from the server and await a response. The server, in turn, provides services to one or many clients in response to their requests. The servers are classified based on the service they provide like web server, file server, etc... All the requests and services in this model are delivered over a network called **client server network** [32].

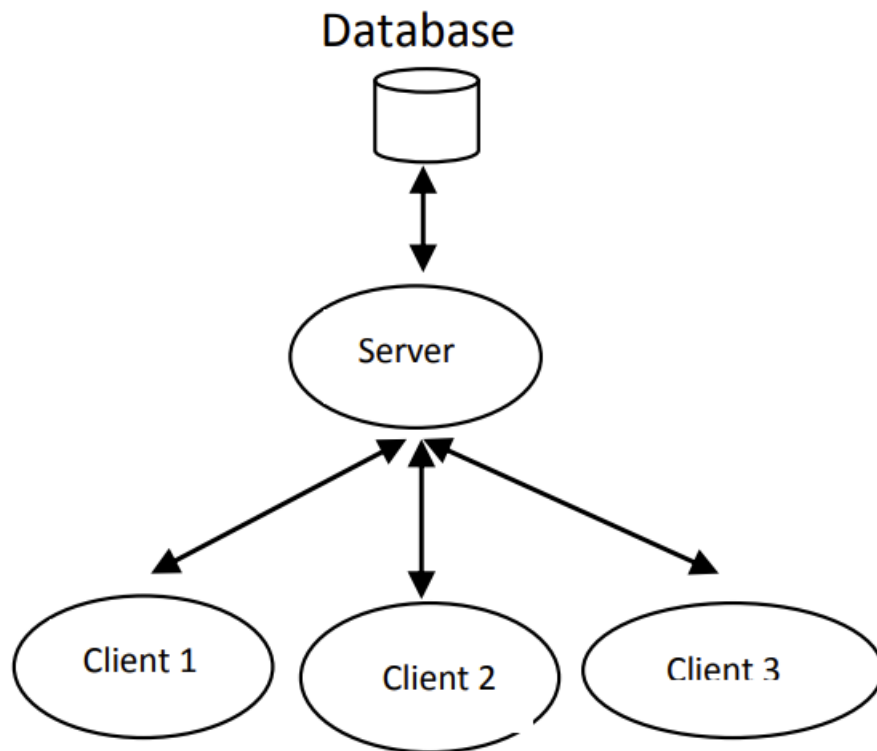


Figure 3.1: Client-server Architecture [32]

Operating Systems Client/Server Communication

The different methods used by client-server communications are as follows,

SOCKETS: Sockets facilitate the data transfer between the clients and the server. Sockets act as one of the two endpoints on the communication link between two programs running on the network. Each endpoint is a combination of an IP (Internet Protocol) address and port number so that every TCP (Transmission Control Protocol) connection

can be uniquely identified by its endpoints to have multiple connections between the host and server. To establish a connection, clients need to know the name of the machine where the server is running, and the port number the server is listening on, and the client also needs to identify itself to the server to establish the connection. The client tries to communicate with the server running on the machine through its port. After the client identifies itself, it binds to a local port assigned by the system. Once the server accepts the client's request, it obtains a new socket bound to the local port assigned by the system and another remote endpoint set to the IP address and port of the client. Simultaneously, the server creates a new socket dedicated to listening for additional connection requests while servicing the needs of the already connected client. Following this process, the client can exchange messages with the server via the socket.

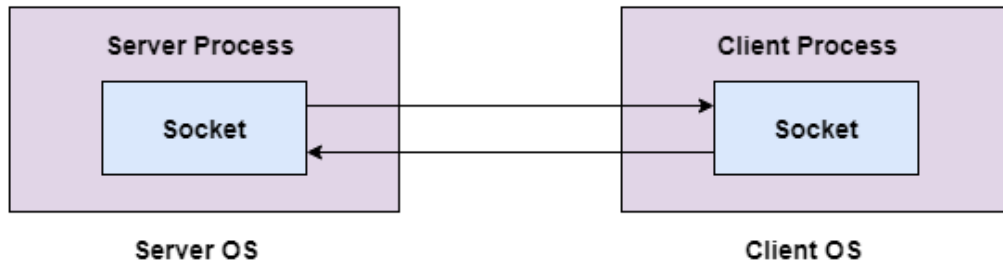


Figure 3.2: Socket Architecture [15]

REMOTE PROCEDURE CALLS: Remote Procedure Call (RPC), also known as a function or subroutine call, is a communication protocol used by the client-server model to request a service from another program on a networked computer without needing to understand the network's intricacies. It is a synchronous program that suspends the requesting program until the results of the remote procedure are returned. In the RPC model, the calling procedure remains essentially the same whether it is a local or remote calling, but they are not identical, so local calls can be easily distinguished from remote calls. The primary purpose of an RPC program is to allow the client to start and execute a procedure on the remote server [23].

Working of RPC: When the client invokes the *client stub* (i.e. it is a piece of code that helps to convert the data passed between the client and the server during a Remote Procedure Call) it converts the parameters from the client into a standard format and copy each parameter into the message which is also called **Client stub marshalls** (Pack). Now, the client stub passes the message to a transport layer which passes the message to another transport layer in the server, and then the message through the server stub

which *demarshalls* (Unpack) the received parameters and executes the procedures in the server once the necessary procedures are correctly executed the server stub again *marshalls* (Pack) the results and send it to the transport layer in the server side which passes the results to the transport layer in the client side. Then, the client stud *demarshalls* (Unpack) the results and executes the results to the client [23].

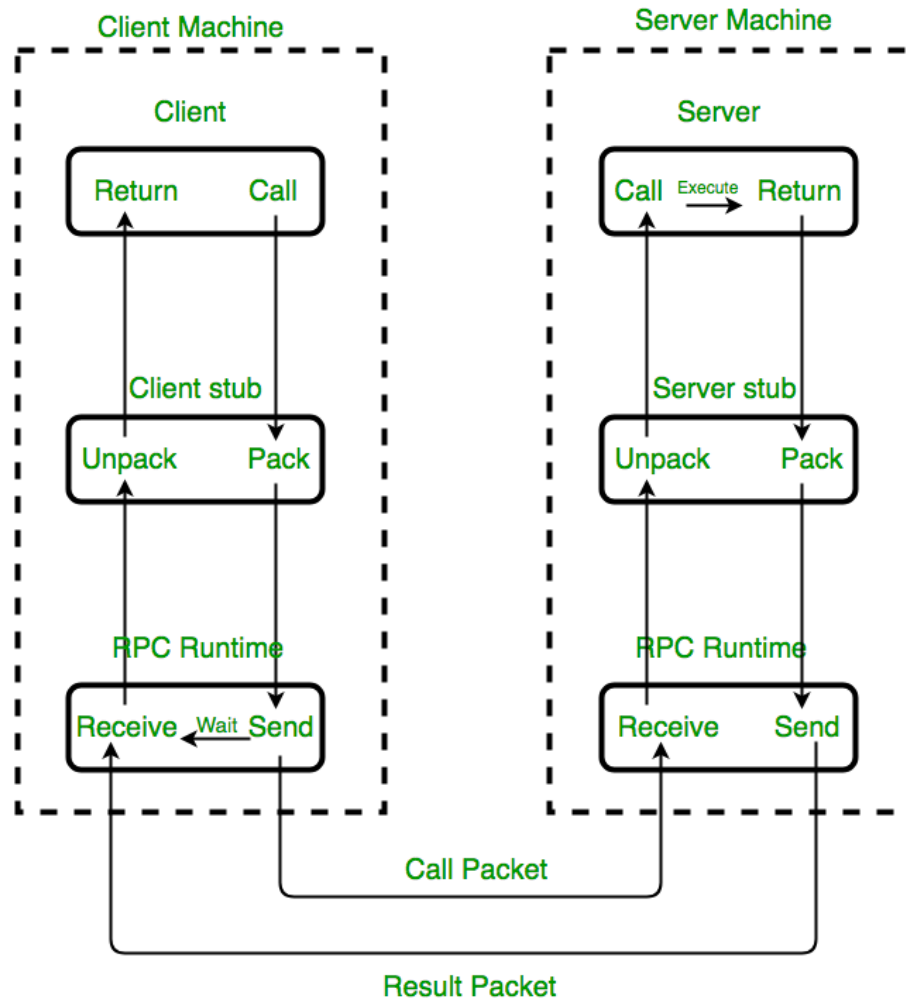


Figure 3.3: Structure of RPC Mechanism [10]

PIPES: Pipes represent one of the simplest and oldest Inter-Process Communication (IPC) methods. They facilitate data exchange between two or more processes by creating either a unidirectional or bidirectional channel between them. A pipe consists of two ends: one for input and one for output. When one process writes data to the input end, the other process reads it from the output end. Ordinary pipes are unidirectional, enabling communication in one direction. For bidirectional communication, two pipes must be implemented.

Ordinary pipes are unidirectional, and they only allow communication in one direction, and for multi-directional we need to implement two pipes. If two process wants to communicate with each other, then the process that writes the data should keep the read end closed and vice versa. A pipe cannot be able to send messages to multiple data at the same time. Pipes are a simple and efficient method in inter-process communication, but they have limitations that have to be considered when choosing client-server communication.

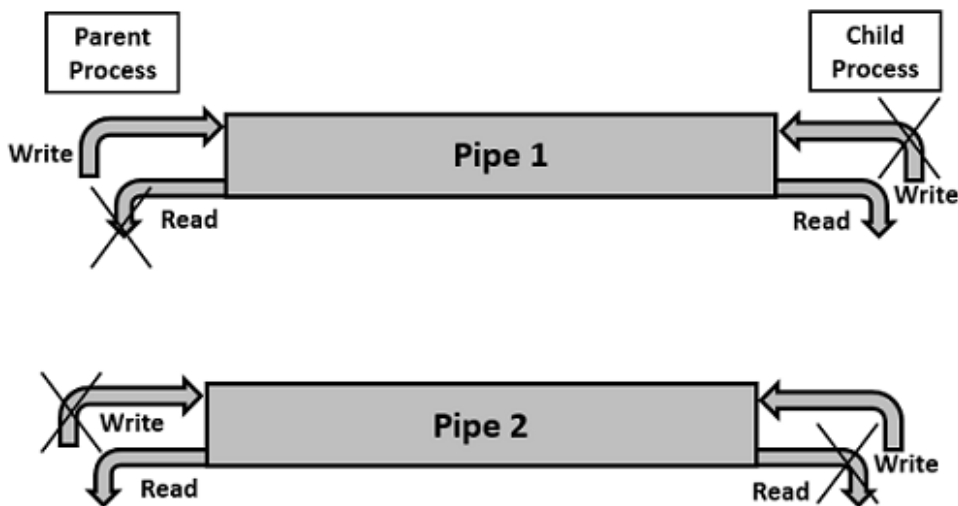


Figure 3.4: Mechanism of Multi-directional communication using Pipes [15]

Characteristics of Client Server Computing

- The client-server computing works on the basis of requests and responses between systems. When a client sends a request to the server and the server responds with the desired data that the client needs.
- Client-Server computing uses a common communication protocol available at the application layer to easily communicate with each other.
- The server uses priority-based response because the server can only able to accommodate a limited number of client requests at a time.
- An example of a client-server model is the web server, it returns the web pages to the clients that requested them.

Upsides of Client-Server Model:

- It is easy to protect the data and provide authorization and authentication because all the data is concentrated in a single place.
- The server and the client don't need to be located in the same place, yet the data can be transferred efficiently.
- It allows a high level of scalability and organization with high efficiency.
- Since all the nodes are independent and request data only from the server, it is easy to replace, update, and relocate the nodes.

Downsides of Client-Server Model:

- The main shortcoming of the client-server model is that if all the clients simultaneously request data from the server, then the server may overload.
- If there is a virus or Trojan in the server then the client will also likely get affected since both are connected to the same network and also it is vulnerable to the Denial of Service (DOS) attacks.
- The expenditure of initial setting and maintaining a client-server model is quite high.

3.1.2. WebRTC

WebRTC (Web Real-Time Communications) provides application programming interfaces (APIs) for developers to create peer-to-peer (P2P) communications between the web browser and the application without worrying about the support for audio, video, or text-based content. WebRTC enables communications without the need for extra plugins, special software for web integration, or a custom interface because it establishes data communication, like opening a standard web page. There are several key functions done by WebRTC which include initiating the recording of video, audio, and text data, transferring those data to the client and then stopping the communications, and so on. WebRTC usually establishes connections between two peers when they are connected through the same Internet Protocol (IP) but in some situations when these peers are operating on different Internet Protocol (IP) the Network Address Translation (NAT) firewalls will prevent the RTC. In this case, WebRTC can be used in conjunction with the Session Traversal Utilities for NAT (STUN) server, which changes the IP address to a common public IP address to establish peer-to-peer connections.

The increase in adoption is responsible for three aspects of WebRTC [26]. First, the lack of external plugins for the architecture of peer-to-peer (P2P) communications. Second is that this whole architecture gives the possibility of scaling into large size for different types of solutions because of its decentralized architecture, WebRTC doesn't need trunk nodes to handle all the traffics generated by simultaneous connections like classic client-server architecture. The third major point is that the WebRTC leans on using mature protocols such as RTP, RTSP, and SDP [26] which gives any vendors the opportunity to support it in its code-base for free and also the exchange of information between any given platform built by using the technology at close to zero development time overhead [33].

WebRTC Video Conferencing architecture types:

WebRTC Video conferencing is a technology that helps us to connect different users at different places by which they can communicate with integrated video, and audio and can collaborate on shared files and documents. For every video conferencing application, the architecture of sending and receiving the audio/video streams is necessary. This can be implemented in many ways, but only three are widely used in the development of video conferencing technology,

I Peer-Peer

II Multi Point Conferencing (MCU)

III Selective Forwarding Unit (SFU)

PEER-PEER COMMUNICATION

Peer-to-peer communication is a distributed architecture used by the application to divide the tasks and workloads between peers. It is sometimes also called as **mesh-architecture**. This is the most basic type of architecture used by WebRTC. Each participant will share their video and audio with all the peers by establishing a peer connection with each of them. Peers make a portion of the resources available for the other peers to use without the need for a central server to coordinate. In a P2P system, each node or peer has the same capability and holds part of the resources, and some peers hold most resources to contribute more to the network while others only consume from the network. These peers are visible and free to join or leave any other peer simultaneously to transmit and receive the data. In peer-to-peer communication, tasks are not handled by a single server rather it was distributed among all the peers. Therefore it is more resilient to failure because there is no single point of failure. In most cases, a centralized server is used to help peers identify each other through the network, but it doesn't have any control over the network itself [27].

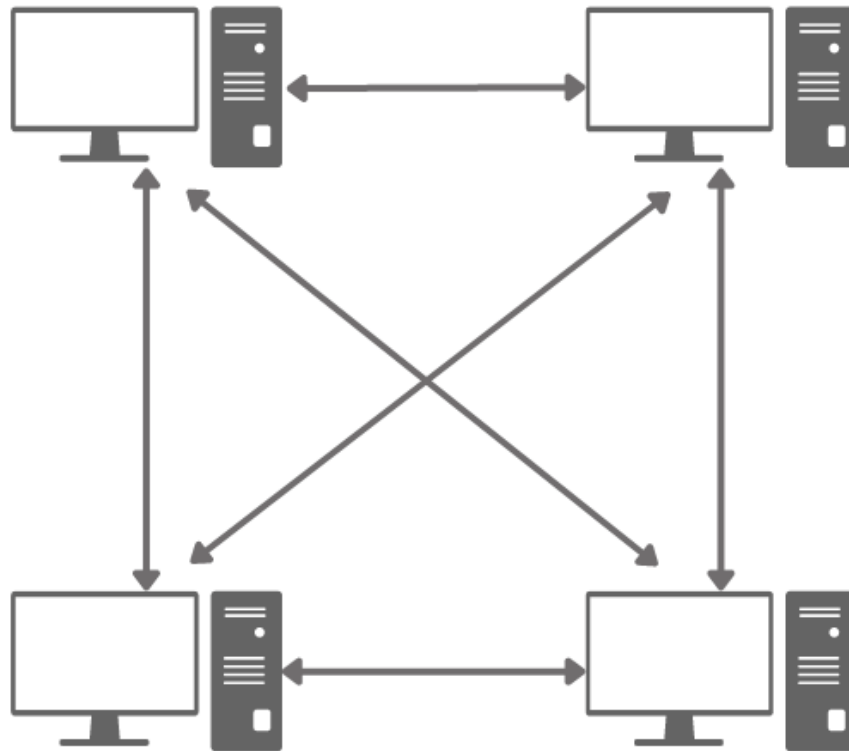


Figure 3.5: Peer-Peer Communication Architecture [13]

Types of P2P networks:

Unstructured P2P networks: These networks are formed when the overlays are established arbitrarily. In this type of network, the query has to be flooded through the network in order to find as many peers as possible that share the data that we are looking for in the network. The main shortcoming of this type is that the queries may not always be resolved. Popular content will be available to one or more peers, making it easily discoverable. However, if a peer is searching for rare data not shared by many peers, it's less likely to find the specific data it seeks. When the network is flooded with data, it will cause signal traffic which decreases the efficiency. Since, in the unstructured P2P network the nodes are arranged randomly, the communication between node to node is also random. It is suited for high-activity cases like social platforms. However, it has one big limitation, they require a significant amount of CPU and memory to function properly.

Structured P2P networks: It overcomes the limitations of unstructured P2P networks by allowing each peer to be responsible for a specific part of the content in the network, which is called a Distributed Hash Table (DHT). After assigning values to a specific part of the content, it will use a global protocol to determine which peer is responsible for which content. So, whenever a peer wants to find data, it uses the global protocol to find the peer responsible for the data and then directs the search towards the responsible peers. Since in the structured P2P network, the nodes are organized they have some centralization happen which means they are more expensive to maintain and setup compared to the Unstructured P2P network structure.

Hybrid P2P networks: Hybrid Peer-Peer combines the architecture of real P2P and the client/server model to overcome the limitations of the other two WebRTC architectures. The architecture consists of two main parts super peers and regular peers. Super peers are responsible for tracking resources within the network and maintaining a list of regular peers who can access these resources. This centralized information management system helps alleviate the burden on regular peers. In Pure P2P architecture, the number of devices that can be connected is limited by the capacity of the network but in Hybrid P2P architecture more devices can be connected which gives large scalability opportunities which is the biggest advantage of this type of architecture. In client-server architecture, if one server fails then the entire network will be down, but in Hybrid P2P architecture even when the super peer fails it will not take down the entire network. So, the hybrid P2P is more robust in nature. One of the limitations is that when a single super peer fails then all the regular peers connected to the super peer fail, this causes a single-point failure in the architecture, and also the networks are vulnerable when multiple devices try to overload the network with data traffic.

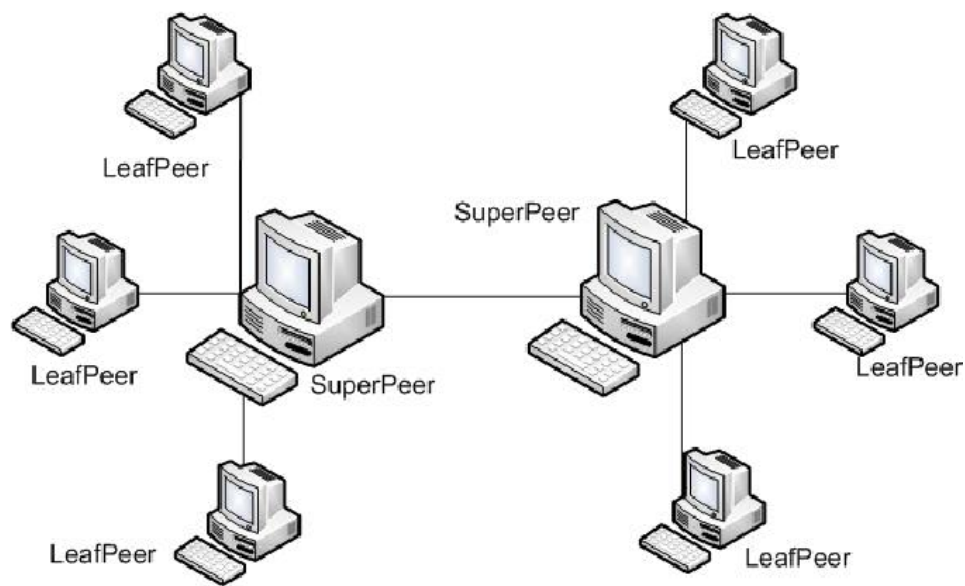


Figure 3.6: Hybrid Peer-Peer Communication Architecture [24]

Advantages Of P2P Communication Model:

- Cost-effective because of the absence of the centralized server.
- Each node is independent of one another, so it is easy to maintain.
- In the client/server model if a server goes down all our data will be lost but in P2P architecture the data is distributed among multiple peers.
- It is more scalable because it is decentralized, and more nodes can be added to the system to access the data.
- Data is more secure because of the absence of the third-party involvement.

Disadvantages Of P2P Communication Model:

- If one peer is affected by a virus and shares the data with the network, it will then affect all the other peers in the network.
- It is difficult to assign network access permission to peers when a large number of peers are connected to the network.
- When the nodes in the network go offline, the data becomes unavailable.

MULTI POINT CONFERENCING (MCU)

MCU is another type of WebRTC video conferencing architecture that acts like a mixture that collects all the inbound video streams and combines them into a single outbound video stream. The MCU encodes the videos into a quality that meets the bandwidth requested from clients on the endpoints. When a streaming platform wants to display a large group of participants on a screen greater than a 3x3 grid, often MCU is used to achieve these requirements and display the larger group of participants in a clustered layout. A larger group of participants can be displayed at the endpoints, only the specification of the equipment and the communication channel can limit the number of participants that can be displayed [34].

Architecture Of MCU:

In P2P communication architecture, each client has to send and receive media directly to and from the other clients that are connected in the network by establishing a peer connection with each of them. Therefore, in P2P, each client has to share a portion of their resources with other clients that are in contact. So, it can quickly become unmanageable when the number of clients increases. In MCU architecture, the server collects media streams from various clients, decodes them, reduces their resolution, and creates a new video with the required quality and resolution for each client. The server then re-encodes the stream and sends it to the respective clients. This process demands substantial computational power and can degrade video quality due to re-compression. As a result, the scalability of this architecture is limited, and the cost of implementing it is exceedingly high, often without sufficient justification, despite its ability to display a large group of participants at the endpoints [34].

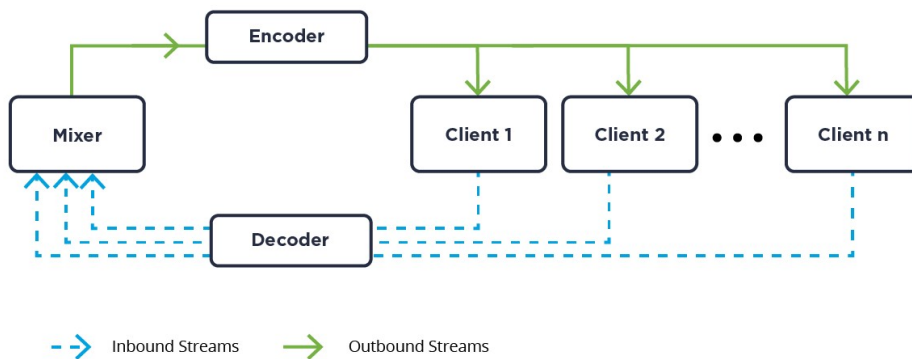


Figure 3.7: MCU Communication Architecture [11]

Advantages of the MCU architecture:

- There is no limitation on how many groups of participants can be displayed on the endpoints, it was only limited by the structural and processing power required.
- MCU can generate and share videos with different quality depending on the specific downlink requirements of the users at the endpoint.
- MCUs' are very reliable for low-capacity works.
- Since the client only consumes one stream, front-end development becomes straightforward as most of the complexity is shifted to the backend.

Disadvantages of the MCU architecture:

- The central server determines the layout that all the participants see, and the participants cannot be able to alter the video quality that the other participants can see. Thus, the layout is not flexible.
- The processing power and the complexity of the server sides are higher if the number of participants participating increases, which in turn increases the cost of setup and maintenance.
- Mixing of numerous streams into a single stream creates a slightly larger delay.

SELECTIVE FORWARDING UNIT (SFU)

SFU is another type of WebRTC architecture that acts as a media server component capable of receiving multiple media streams from different clients and decides which one of these media streams should be sent to which participants. In the peer-to-peer connection, the client creates an offer describing the desired connection and sends it to the remote peer. Now the remote peer can either reject or accept the offer by creating an answer and sending it back to the client. This type of communication architecture is suitable for a small call process, but when the number of participants increases we need something more like SFU to handle the loads. When compared to MCU there is no compressing, encoding, or decoding happening to the video streams therefore there is no loss of video stream quality [29].

Architecture Of SFU:

SFU works by forwarding the media streams between the sender and one or more participants at the receiver end, also it analyzes the video streams that are coming back from the receivers to use that information intelligently to manage the flow of media streams in the network. Let's consider a three-peer connection shown in Figure 3.8, in this all the peers send the video streams to the SFU server and then the server analyses the video streams and sends the video streams of the remaining two peers that are connected to each other peers in the network without any quality loss. In most cases, all peers will send several versions of their video streams each with different quality and then the SFU server decides to send the low-quality video streams to peers that have a very poor internet connection and vice versa [29].

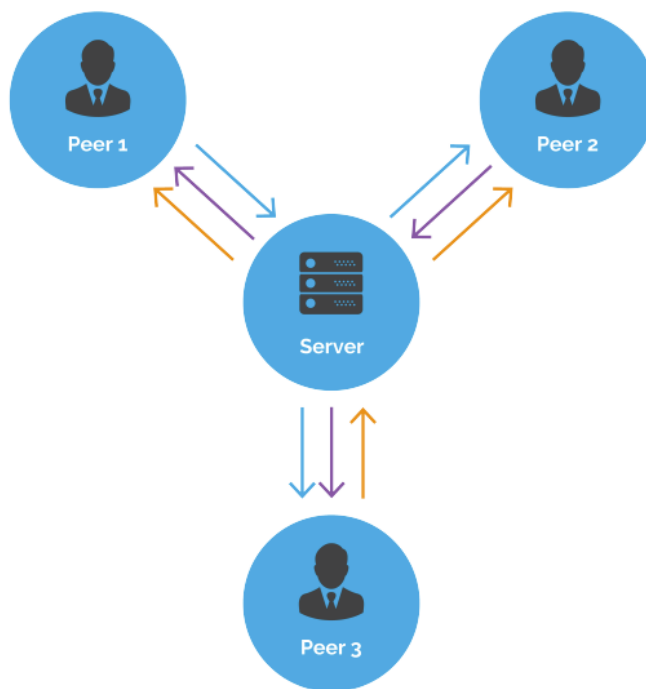


Figure 3.8: SFU Communication Architecture [14]

Advantages of SFU:

- SFU is an upload bandwidth-friendly architecture because the participants have to publish only one outgoing video stream.
- SFU is more scalable than P2P architecture and the clients don't need to share part of their resources like P2P which puts less strain on the client side.
- Participants can decide which the other participant's video streams they want and in which quality.
- Compared to other configurations, SFU demands less from the centralized server.

Disadvantages of SFU:

- The presence of an intermediate server increases the complexity and cost of the process.
- The server demands higher bandwidth when more copies of the same video stream are created to send it to different participants.
- The client has to decode numerous incoming video streams, so it will require a lot of processing power compared to MCU which sends only one video stream.
- There is a weak link problem. Since the server cannot be able to adjust the outgoing stream if one participant fails to connect to the network, then the quality of the others decreases.
- Live-streaming is not possible because all the video streams are not mixed into a single stream.
- Full privacy is not achieved by default because the intermediate server can access the raw data of the video stream when forwarding, but it can be solved by adding encryption to the video stream before sending it to the SFU server.

Signaling server

As discussed in Chapter 3.1.2 in Peer-to-peer communication, each connection shares its video and audio streams directly with other peers by establishing direct communication between each other. So, the small problem in P2P is that it needs to have some kind of tool to help the peers find each other who wants to exchange information. This is where the signaling server comes into play, It helps the peers to identify one another and

initiate communication between them, but once it initiates the communication its job is done, and it can also be used for additional events. The other main function for which the signaling server is used in Peer-to-peer communication is to keep track of the peers who join and leave the communication and also to create or dispose of the connections. Whenever a new peer wants to join a conversation it sends the request to the server and then the server makes sure that the request from that peer is visible to all the peers who are already in the conversation, If some peers accept the invite then the server includes the other peers in the conversation [26] [33].

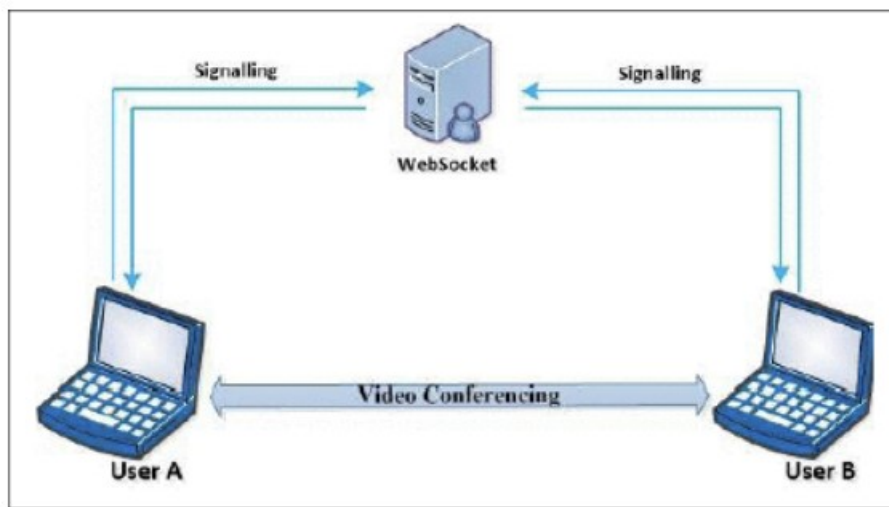


Figure 3.9: Signalling Server Mechanism [26]

Node-Dss Signaler:

In standard WebRTC Peer-Peer communication it uses Session Description Protocol (SDP) to establish communication between the peers but in order to perform this task there is a need for some kind of tool for the peers to discover and connect with other peers this is where the Node Dss signaler comes in to play. It acts as a message brokerage service between the peers and allows them to pass messages between predetermined receivers because it saves the peers from more complex tracking of clients who join or leave the network, then The two peers can send and receive the SDP messages in order to establish the connection. This NodeDss signaler is a built-in solution offered by the Mixed Reality-WebRTC tool kit, which will be used in this application for the front-end development in Unity3D.

Note: This NodeDss Signaler is not a production-ready tool because it offers no security to the established connection between peers. All the communication takes place using HTTP, not HTTPS without any proper authentication, all it requires is to know the identifiers of the clients to establish a connection. The weakest link in this whole back-end process of the application which is at risk is definitely at the NodeDss signaler part [17].

Installing and Configuring Node Dss Signaler: For the application, the server is installed in a remote desktop environment created in Amazon Web Services, in which the server runs 24x7. As discussed in Chapter 3.1.2 whenever the local peer wants to establish a connection to the known remote peer both peers will establish their interests by sending an offer request to the server running in the AWS instance and the server checks the data of both peers and initiate the handshake, once the initial connection is established the peers will start sending the data directly to one another. In order to leverage the service that comes with the Mixed Reality-WebRTC Toolkit, there comes a need to install the *node.js* server, and then the necessary files are downloaded to install node-Dss signaler from the GitHub [19]. After all the necessary files are downloaded, proceed to install the server through the command prompt using the following commands.

```
set DEBUG=dss*  
npm install
```

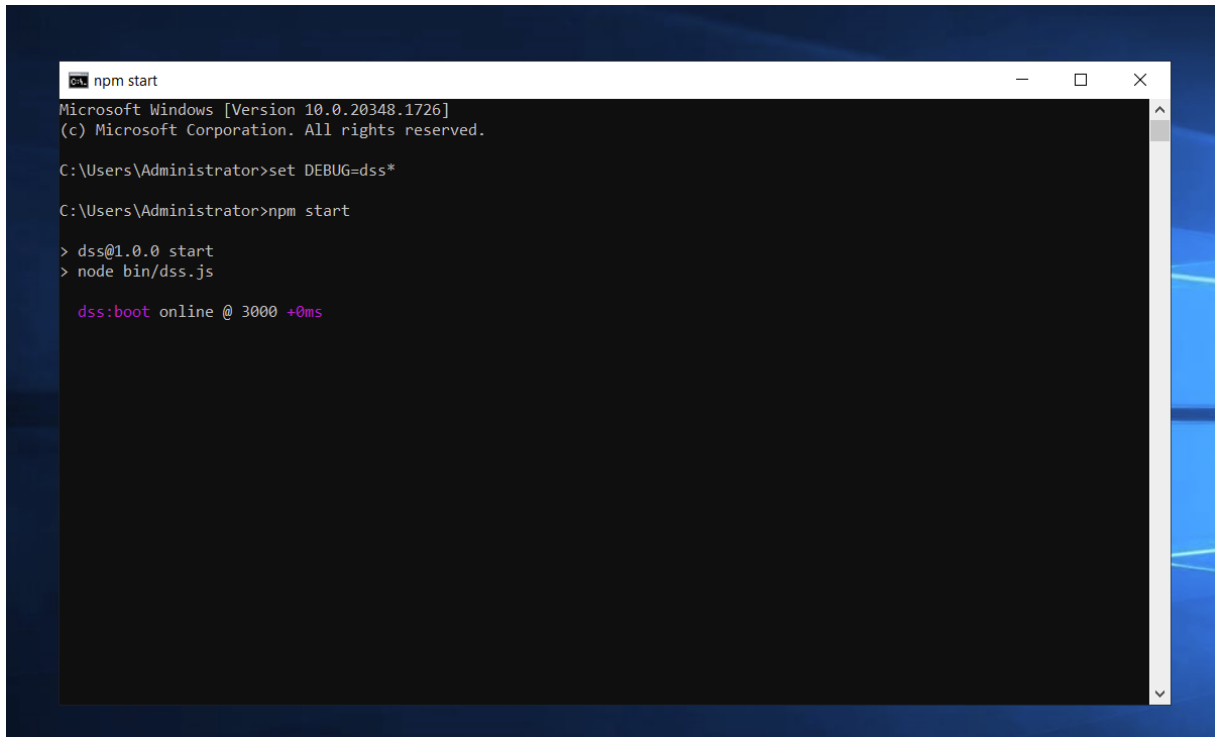
Figure 3.10: Command to install the server [19]

Once all the dependencies for the server are installed, the server can be activated by using the following command,

```
npm start
```

Figure 3.11: Command to start the server [19]

Now, the server is listening through the predefined port for the possible offers from the peers across the network, once it receives the offers from the client and remote peer it establishes the connection between them. Figure 3.12 shows the server which has been created for this Remote Maintenance Application is running on an AWS instance and constantly listening for offers from the clients.



```
npm start
Microsoft Windows [Version 10.0.20348.1726]
(c) Microsoft Corporation. All rights reserved.

C:\Users\Administrator>set DEBUG=dss*

C:\Users\Administrator>npm start

> dss@1.0.0 start
> node bin/dss.js

dss:boot online @ 3000 +0ms
```

Figure 3.12: Running NodeDss server

SDP (Session Description Protocol):

SDP is the first part of peer-to-peer communication where the peers have to exchange data through the signaling server in order to establish a connection. When one peer sends an offer and another peer is interested in joining the network, they need to establish a connection without prior knowledge of each other's systems. This is where the Session Description Protocol (SDP) comes into play. SDP is a string-based profile that contains all the necessary system information, including hardware specifications, desired media content for exchange, IP addresses, audio and video encryption settings, and other relevant details [27]. The SDP configuration can either represent an answer or an offer. Both Offer and Answer are bidirectional, which means that it doesn't matter from which side it initiates the conversation, but the outcome will be the same. In This context, when the client (i.e., Laptop) creates an offer it sends an SDP message through the signaling server which contains the information of the system like the IP address of the client, audio and video format information and what hardware it supports in return the Remote Peer (i.e., Hololens2) answers with another SDP message which contains all the hardware and media stream data's just like the client. After receiving the answer, a connection will be established between the two peers and data will be exchanged.

The following Figure 3.13 and Figure 3.14 shows the SDP message exchange happened between both the laptop and the Hololens2 which established a remote connection to exchange media streams [33] [26].

```

dss:body ("MessageType":3,"Data":{"candidate:1959295338 1 udp 2122197247 2001:b07:646a:b845:7f88:dc46:8e07:88c5 58051 typ host generation 0 ufrag 5c0a network-k-id 3 network-cost 10|1|","IceDataSeparator":""}) +ms
dss POST /data/HOOLENS 200 - 0.367 ms +8ms
dss:body ("MessageType":3,"Data":{"candidate:3275632839 1 udp 2122129151 192.168.1.58 58049 typ host generation 0 ufrag 5c0a network-id 1 network-cost 10|0|","IceDataSeparator":""}) +ms
dss POST /data/HOOLENS 200 - 0.383 ms +1ms
dss:body ("MessageType":3,"Data":{"candidate:1959295338 1 udp 2122197247 2001:b07:646a:b845:7f88:dc46:8e07:88c5 58048 typ host generation 0 ufrag 5c0a network-k-id 3 network-cost 10|0|","IceDataSeparator":""}) +ms
dss POST /data/HOOLENS 200 - 0.714 ms +5ms
dss:body ("MessageType":3,"Data":{"candidate:3275632839 1 udp 2122129151 192.168.1.58 58052 typ host generation 0 ufrag 5c0a network-id 1 network-cost 10|1|","IceDataSeparator":""}) +ms
dss POST /data/HOOLENS 200 - 0.327 ms +1ms
dss:body ("MessageType":1,"Data":{"v=0\r\n= - 545451740391002945 2 IN IP4 127.0.0.1\r\ns=-\r\nrt=0 0\r\na=group:BUNDLE 0\r\na=msid-semantic: WWS/\r\nm=audio 9 UDP/TLS/RTP/SAVPF 111 103 104 9 102 0 8 106 105 13 110 112 113 126\r\nlnc=IP IN IP4 0.0.0.0\r\na=ice-ufrag:5c0a\r\na=ice-pwd:tqtK0BoA T8w0mZsb5FXKK\r\na=ice-options:trickle\r\na=fingerprint:sha-256 3655:cf7c:bc81:8f:8d:b3:09:ac:26:e8:37:d1:a8:c2:f0:ae:99:ac:25:a2:94:49:b2:6f:7 E:6B\r\na=transport:actpass\r\na=mid:0\r\na=extmap:1 urn:ietf:params:rtp-hdrext:ssrc-audio-level\r\na=extmap:9 urn:ietf:params:rtp-hdrext:sdes:mid\r\na=sendrecv\r\na=msid:- 69758ff6-f3c7-4975-88f2-af8480589749\r\na=rtcp-mux:\r\na=rtcp:111 opus/48000/2\r\na=rtcp-fb:111 transport-cc\r\na=rtcp:111 mintime:10;useinbandf c=1\r\na=rtcp:map:86 ISAC/16000\r\na=rtcp:map:104 ISAC/32000\r\na=rtcp:map:9 6722/8000\r\na=rtcp:map:102 ILBC/8000\r\na=rtcp:map:0 PCM/8000\r\na=rtcp:map:8 PCMA/8000\r\na=rtcp:map:106 CN/32000\r\na=rtcp:map:105 CN/16000\r\na=rtcp:map:13 CN/8000\r\na=rtcp:110 telephone-event/48000\r\na=rtcp:112 telephone-event/32000\r\na=rtcp:map-113 telephone-event/16000\r\na=rtcp:map:126 telephone-event/8000\r\na=ssrc:339283228 cname:IDaLPCNcYhQEDUBU\r\na=ssrc:339283228 mid:- 69758ff6-f3c7-4975-88f2-a 0b586794\r\na=ssrc:339283228 mslabel:86c80aca-61a7-4966-85b3-72205dfdddc\r\na=srtp:127 90000\r\na=srtp:124 90000\r\na=srtp:126 90000\r\na=srtp:128 90000\r\na=srtp:129 90000\r\na=srtp:130 90000\r\na=srtp:131 90000\r\na=srtp:132 90000\r\na=srtp:133 90000\r\na=srtp:134 90000\r\na=srtp:135 90000\r\na=srtp:136 90000\r\na=srtp:137 90000\r\na=srtp:138 90000\r\na=srtp:139 90000\r\na=srtp:140 90000\r\na=srtp:141 90000\r\na=srtp:142 90000\r\na=srtp:143 90000\r\na=srtp:144 90000\r\na=srtp:145 90000\r\na=srtp:146 90000\r\na=srtp:147 90000\r\na=srtp:148 90000\r\na=srtp:149 90000\r\na=srtp:150 90000\r\na=srtp:151 90000\r\na=srtp:152 90000\r\na=srtp:153 90000\r\na=srtp:154 90000\r\na=srtp:155 90000\r\na=srtp:156 90000\r\na=srtp:157 90000\r\na=srtp:158 90000\r\na=srtp:159 90000\r\na=srtp:160 90000\r\na=srtp:161 90000\r\na=srtp:162 90000\r\na=srtp:163 90000\r\na=srtp:164 90000\r\na=srtp:165 90000\r\na=srtp:166 90000\r\na=srtp:167 90000\r\na=srtp:168 90000\r\na=srtp:169 90000\r\na=srtp:170 90000\r\na=srtp:171 90000\r\na=srtp:172 90000\r\na=srtp:173 90000\r\na=srtp:174 90000\r\na=srtp:175 90000\r\na=srtp:176 90000\r\na=srtp:177 90000\r\na=srtp:178 90000\r\na=srtp:179 90000\r\na=srtp:180 90000\r\na=srtp:181 90000\r\na=srtp:182 90000\r\na=srtp:183 90000\r\na=srtp:184 90000\r\na=srtp:185 90000\r\na=srtp:186 90000\r\na=srtp:187 90000\r\na=srtp:188 90000\r\na=srtp:189 90000\r\na=srtp:190 90000\r\na=srtp:191 90000\r\na=srtp:192 90000\r\na=srtp:193 90000\r\na=srtp:194 90000\r\na=srtp:195 90000\r\na=srtp:196 90000\r\na=srtp:197 90000\r\na=srtp:198 90000\r\na=srtp:199 90000\r\na=srtp:200 90000\r\na=srtp:201 90000\r\na=srtp:202 90000\r\na=srtp:203 90000\r\na=srtp:204 90000\r\na=srtp:205 90000\r\na=srtp:206 90000\r\na=srtp:207 90000\r\na=srtp:208 90000\r\na=srtp:209 90000\r\na=srtp:210 90000\r\na=srtp:211 90000\r\na=srtp:212 90000\r\na=srtp:213 90000\r\na=srtp:214 90000\r\na=srtp:215 90000\r\na=srtp:216 90000\r\na=srtp:217 90000\r\na=srtp:218 90000\r\na=srtp:219 90000\r\na=srtp:220 90000\r\na=srtp:221 90000\r\na=srtp:222 90000\r\na=srtp:223 90000\r\na=srtp:224 90000\r\na=srtp:225 90000\r\na=srtp:226 90000\r\na=srtp:227 90000\r\na=srtp:228 90000\r\na=srtp:229 90000\r\na=srtp:230 90000\r\na=srtp:231 90000\r\na=srtp:232 90000\r\na=srtp:233 90000\r\na=srtp:234 90000\r\na=srtp:235 90000\r\na=srtp:236 90000\r\na=srtp:237 90000\r\na=srtp:238 90000\r\na=srtp:239 90000\r\na=srtp:240 90000\r\na=srtp:241 90000\r\na=srtp:242 90000\r\na=srtp:243 90000\r\na=srtp:244 90000\r\na=srtp:245 90000\r\na=srtp:246 90000\r\na=srtp:247 90000\r\na=srtp:248 90000\r\na=srtp:249 90000\r\na=srtp:250 90000\r\na=srtp:251 90000\r\na=srtp:252 90000\r\na=srtp:253 90000\r\na=srtp:254 90000\r\na=srtp:255 90000\r\na=srtp:256 90000\r\na=srtp:257 90000\r\na=srtp:258 90000\r\na=srtp:259 90000\r\na=srtp:260 90000\r\na=srtp:261 90000\r\na=srtp:262 90000\r\na=srtp:263 90000\r\na=srtp:264 90000\r\na=srtp:265 90000\r\na=srtp:266 90000\r\na=srtp:267 90000\r\na=srtp:268 90000\r\na=srtp:269 90000\r\na=srtp:270 90000\r\na=srtp:271 90000\r\na=srtp:272 90000\r\na=srtp:273 90000\r\na=srtp:274 90000\r\na=srtp:275 90000\r\na=srtp:276 90000\r\na=srtp:277 90000\r\na=srtp:278 90000\r\na=srtp:279 90000\r\na=srtp:280 90000\r\na=srtp:281 90000\r\na=srtp:282 90000\r\na=srtp:283 90000\r\na=srtp:284 90000\r\na=srtp:285 90000\r\na=srtp:286 90000\r\na=srtp:287 90000\r\na=srtp:288 90000\r\na=srtp:289 90000\r\na=srtp:290 90000\r\na=srtp:291 90000\r\na=srtp:292 90000\r\na=srtp:293 90000\r\na=srtp:294 90000\r\na=srtp:295 90000\r\na=srtp:296 90000
```

Figure 3.13: SDP message offered by Hololens2

```
dss GET /data/HOLELNS 200 - - 0.073 ms +66ms
dss GET /data/HOLELNS 200 - - 0.073 ms +182ms
dss:body {"MessageType":2,"Data":{"candidate:6154969959956389818 2 IN IP4 127.0.0.1\r\nns=group-BUNDLE 0 \r\nna=mid-semantic; WSH\r\nnameid=9 \r\nna=option:trickle\r\nfingerprnt=sha-256 05:F8:D8:A4:F9:0E:7A:45:8D:C6:E9:72:8A:E5:12:32:92:8E;7F:B5:18:36:54:6C:9B:2D:8C:70:64:DA:9C\r\nna=setup:active\r\nna=mid:0\r\nna=extmap:1 urn:iETF:params:rxtx-hdr-ext:offset\r\nna=extmap:9 urn:iETF:params:rxtx-hdr-ext:sds:mid\r\nna=sen-drecv\r\nna=sen-dcblf\r\nna=sen-dcblf-5342-4546-8546-9438dBabb21c\r\nna=rtpcm-mux\r\nna=rtpmap:111 opus/48000/2\r\nna=rtpct-fb:111 transport-cc\r\nna=fmtp:111 minptime=10;usebandfec\r\nna=rtpm-ap:103 ISAC/16000\r\nna=rtpmap:9 G722/8000\r\nna=rtpmap:102 TLBC/8000\r\nna=rtpmap:8 PCMU/8000\r\nna=rtpmap:8 PCMVA/8000\r\nna=rtpmap:105 CN/16000\r\nna=rtpmap:13 CN/8000\r\nna=rtpmap:110 telephone-event/48000\r\nna=rtpmap:111 telephone-event/16000\r\nna=rtpmap:126 telephone-event/8000\r\nna=srcsr:1151897181 cname:7WlUoSjARHxCofR\r\nna=video:9 UDP/TLS/RTP/SAVPF 96 97 98 99 127 124 125\r\nna=IN IP4 0.0.0.0\r\nna=rtpct-fb:126 0.0.0.0\r\nna=event-fragg:GLhk\r\nna=ice-pwd:irs74VshSMnTKGwvm949T0X\r\nna=ice-options:trickle\r\nna=fingerprnt=sha-256 05:F8:D8:A4:F9:0E:7A:45:8D:C6:E9:72:8A:E5:12:32:92:8E;7F:B5:18:36:54:6C:9B:2D:8C:70:64:DA:9C\r\nna=setup:active\r\nna=mid:1\r\nna=extmap:2 urn:iETF:params:rxtx-hdr-ext:offset\r\nna=extmap:3 http://www.webrtc.org/experiments/rtp-hdrext/abs-send-time\r\nna=extmap:4 urn:3gpp:video-orientation\r\nna=extmap:5 http://tools.ietf.org/id/draft-holmer-rmcat-transport-wide-cc-extensions-01\r\nna=extmap:6 http://www.webrtc.org/experiments/rtp-hdrext/playout-delay\r\nna=extmap:7 http://www.webrtc.org/experiments/rtp-hdrext/video-content-type\r\nna=extmap:8 http://www.webrtc.org/experiments/rtp-hdrext/video-id\r\nna=extmap:10 http://tools.ietf.org/html/draft-ietf-avtext-framemarking-07\r\nna=extmap:9 urn:iETF:params:rxtx-hdr-ext:sds:mid\r\nna=sen-drecv\r\nna=mid-timing-a29a2aed-46e2-46d5-ac24-736bbfa0697\r\nna=rtpcm-mux\r\nna=rtpct-fb:rszle\r\nna=rtpmap:96 VP8/90000\r\nna=rtpct-fb:96 goog-remb\r\nna=rtpct-fb:96 transport-cc\r\nna=rtpct-fb:96 nack\r\nna=rtpct-fb:96 nack pli\r\nna=rtpmap:97 rtcp/90000\r\nna=fmtp:97 apt=96\r\nna=rtpmap:98 VP9/90000\r\nna=rtpct-fb:98 goog-remb\r\nna=rtpct-fb:98 transport-cc\r\nna=rtpct-fb:98 nack\r\nna=rtpct-fb:98 nack pli\r\nna=rtpmap:98 x-google-profile-id=0\r\nna=rtpm-ap:99 rtcp/90000\r\nna=fmtp:99 apt=98\r\nna=rtpmap:127 rtx/90000\r\nna=rtpmap:124 rtx/90000\r\nna=fmtp:124 apt=127\r\nna=rtpmap:125 ulpfec/90000\r\nna=srcsr-group:FID 3455294876 724334291\r\nna=srcsr:candidate:748905WARHXrCoFr\r\nna=src:candidate:724334291 cname:7WlUoSjARHxCofR\r\nna=ICEDataSeparator:""}' +1s
dss POST /data/PC 200 - - 12.526 ms +140ms
dss GET /data/PC 200 - - 0.073 ms +25ms
dss:body {"MessageType":3,"Data":{"candidate:2489864918 1 udp 2122197247 2001:b07:646a:b845:81ee:66d7:1b9a:45f3 52479 typ host generation 0 ufrag GLhk network-id 3 network-cost 10|0|","IceDataSeparator:"|"'} +40ms
dss POST /data/PC 200 - - 0.074 ms +48ms
dss:body {"MessageType":3,"Data":{"candidate:1341159495 1 udp 2122262783 2001:b07:646a:b845:6a16:e7ad:d2ef:4cf8 52478 typ host generation 0 ufrag GLhk network-k id 2 network-cost 10|0|","IceDataSeparator:"|"'} +23ms
dss POST /data/PC 200 - - 2.113 ms +24ms
dss GET /data/HOLELNS 200 - - 0.085 ms +1ms
dss:body {"MessageType":3,"Data":{"candidate:339728134 1 udp 2122129151 192.168.1.53 52480 typ host generation 0 ufrag GLhk network-id 1 network-cost 10|0|","IceDataSeparator:"|"'} +4ms
```

Figure 3.14: SDP message offered by Laptop

Advantage and Disadvantage of WebRTC:

- WebRTC shares the media streams between the peers at a high speed directly, which offers low latency in communication.
- WebRTC can be integrated with a range of applications without the need for separate infrastructure because it supports most of the available platforms and browsers.
- Offers a more reliable connection as it enables the user to generate videos in different quality and bit rates.
- Enables smooth multi-party video sharing by organizations globally.
- Provides secure connection by encrypting video and audio streams using Secure RTP protocol (SRTP) for encryption and authentication.
- WebRTC doesn't offer any standard protocol, so different developers implement this in different methods.
- Peer's internet connection can influence the quality of the video/audio streams.
- WebRTC is relatively a new technology, it is still under development.

3.2. Front end Application Development

The primary focus of front-end development is the user interface and user experience when they open the application. Everything that the user sees upon opening the application, such as buttons, fonts, and the overall layout, falls under front-end development. In this application, the front end is developed using Unity3D, which involves creating a canvas to display incoming video streams, implementing buttons for initiating back-end server communications, and simulating specific animations using CAD models.

Function of Mixed Reality and MixedReality-WebRTC Toolkit

Mixed Reality-WebRTC Toolkit: The toolkit provides the necessary Unity3D components to help integrate video, audio, and real-time communication into the Mixed Reality application. The toolkit includes components like Peer Connection, NodeDss Signaler, Microphone Source, and Video Source. This set of Unity3D components will help to achieve rapid prototyping and integration.

MixedReality Toolkit: On the other hand, the general Mixed Reality toolkit provides a collection of components and features that can enhance the Mixed Reality experience for users and also able to achieve rapid prototyping by using different components by adding and removing depending on the features that are desired as an outcome.

Two different toolkits are used because the former only gives the necessary components for integrating real-time communication to the application but later is the one that provides all the components needed for implementing human interaction features like hand tracking, eye tracking, and object manipulation into the application, and also it helps with developing the application for different environments.

3.2.1. Developing process of Mixed Reality application with Real-time Communication

The application comprises various components from the MixedReality-WebRTC toolkit integrated into different GameObjects to facilitate real-time audio and video communication. This part of the Chapter 3.2.1 is built under the scene name *Video Sharing*.

PeerConnection: A new GameObject **PeerConnection** is created in the project to add the following MixedReality-WebRTC Component *Peer Connection*. This component is responsible for sending and receiving the video audio streams from local

to remote peer (i.e., Laptop and Hololens2) and vice versa once the connection has been established. This component can't work on its own just by simply adding it to the project, it needs to be configured for the incoming audio and video streams. In the Inspector window under the *Peer Connection* component, the incoming video and audio streams that have been collected from the later steps need to be added by selecting **+Audio** (i.e., Which adds local and remote audio streams) and **+Video** (i.e., which adds local and remote video streams).

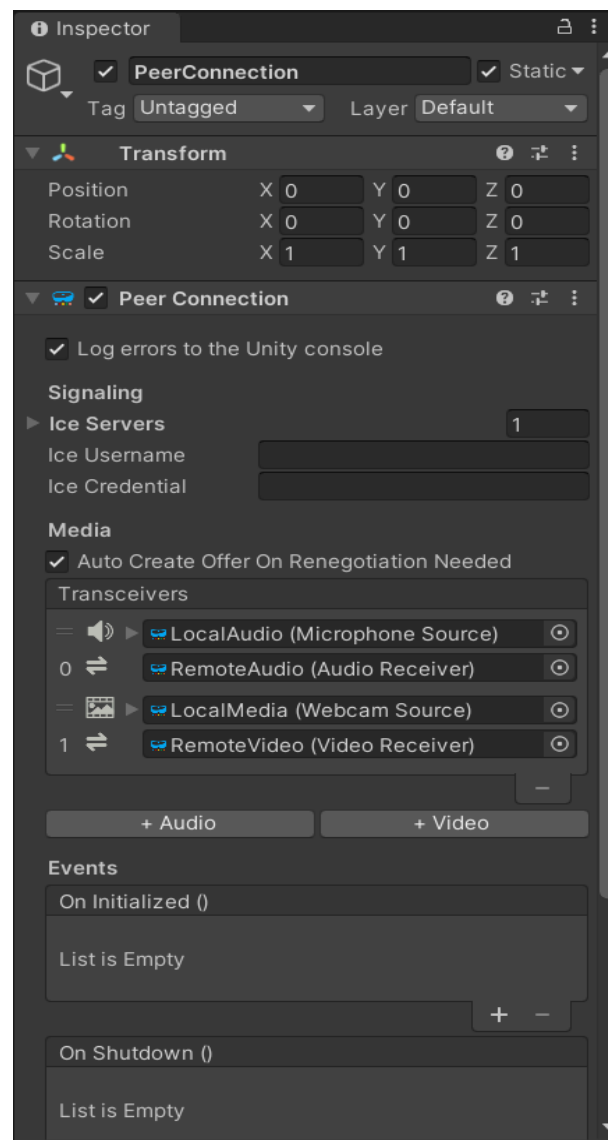


Figure 3.15: Peer Connection component

NodeDssSignaler: The importance of the *NodeDss Signaler* component has been discussed already in Chapter 3.1.2 which states that the Peer-Peer communication protocol uses SDP message to exchange information about their system architecture, audio, and video formats supported, but before that the peers has to be able to identify each other through the network. To achieve this, pre-defined IDs (**i.e., PC and Hololens2**) are assigned to both peers, eliminating the need for them to search for specific clients on the network. The *NodeDss Signaler* component is employed to define these Peer IDs during the application's development stage. To add the *NodeDss Signaler* 3D Unity component to the project, create a new GameObject called **NodeDssSignaler**. In the Inspector window of this project, there will be three input fields: the first two for acquiring the Local and Remote peer IDs and the third one for obtaining the IP (Internet Protocol) address of the system where the signaling server is running. A detailed explanation of this signaling server can be found in Chapter 3.1.2. An Important consideration is that, Once the application is deployed on two different devices, the Local and Remote Peer IDs have to be inverted.

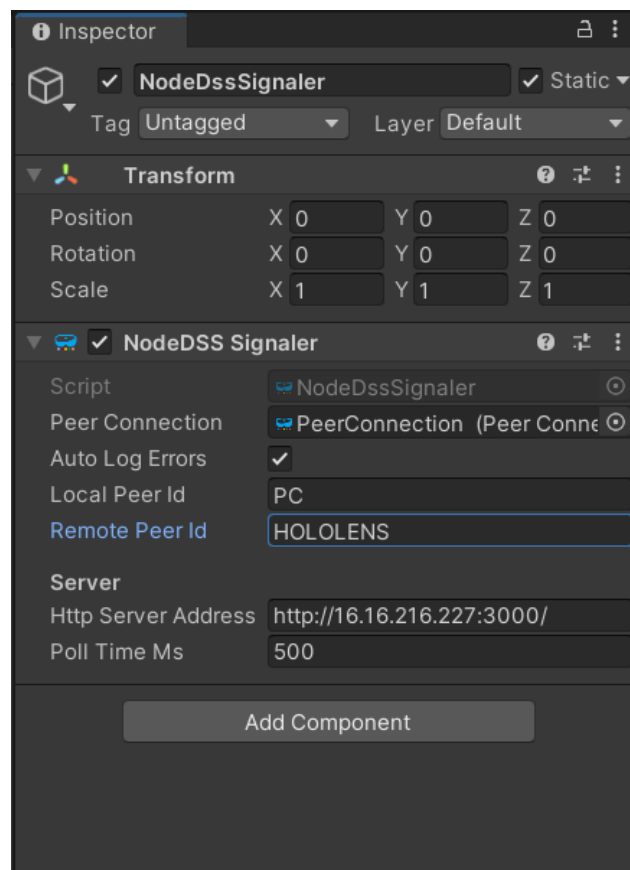


Figure 3.16: NodeDss Signaler component

LocalAudio: The GameObject **LocalAudio** is created in order to add the following WebRTC component *Microphone Source* to the system, which helps the application access the audio frames from the microphone of the local machine and decompose it into an audio track which can be later sent to the Remote peer parallelly with the video track. However, for this application, there is no local audio callback on audio track sources because here it has very few use cases.

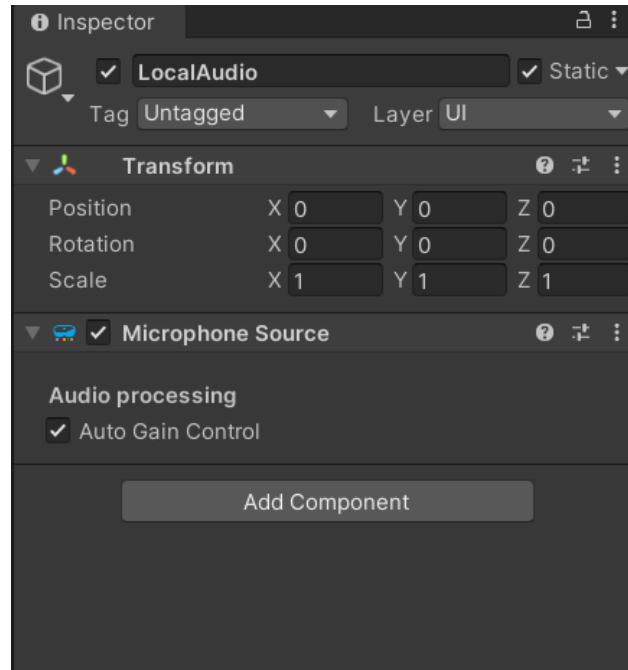


Figure 3.17: Microphone component

LocalMedia: The main reason to create the GameObject **LocalMedia** is to integrate the WebRTC component *Webcam Source* into the application that will help in capturing the videos through the local capturing devices available in the peers, whether it is a laptop or a Hololens 2. This component accesses the local cameras to gather the video frames and then composes it into a video track to share with other peers, once the connection is established between the peers. The local video track of the Local peer will also be rendered locally via a **LocalPlayer** GameObject in parallel to being streamed to the remote peer. After adding the *Webcam Source* component, start and stop accessing video frames are defined from the Local webcam. This can be achieved by adding two commands *VideoRenderer.StartRendering* and *VideoRenderer.StopRendering* in the inspector window under the component *Webcam Source* like shown in the Figure 3.18

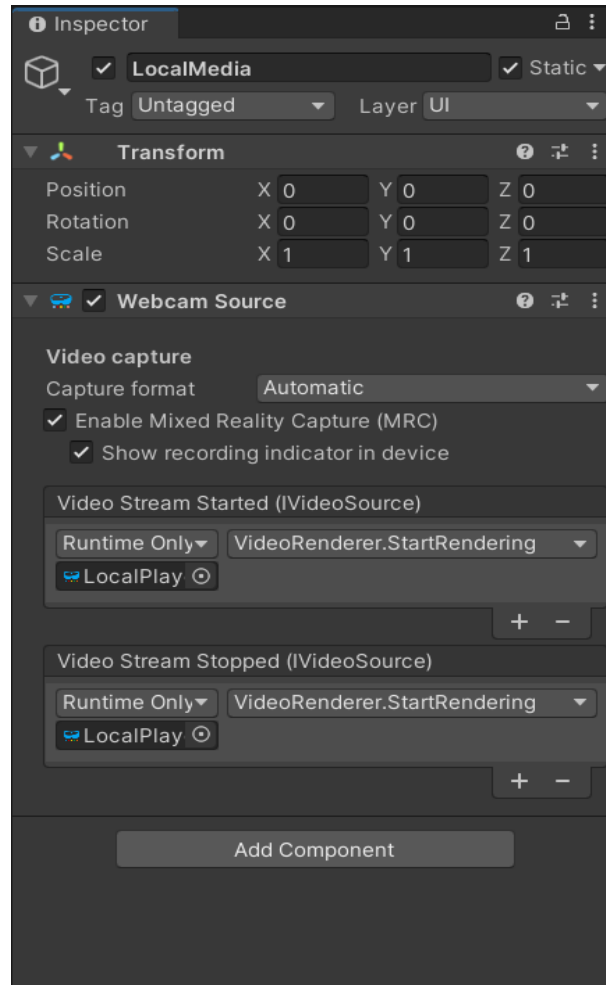


Figure 3.18: Webcam Source component

RemoteAudio: Similar to the GameObject **LocalAudio** this GameObject is created for the same purpose to add the following WebRTC component *Audio Source* into the application which receives and renders the audio track shared by the Remote peer to the local peer, once the connection has been established this audio stream will be rendered simultaneously along with the video rendered from the remote peer. Additional commands on when to start and stop rendering the remote audio stream (i.e., *AudioRenderer.StartRendering* and *AudioRenderer.StopRendering*) and the component *Audio Renderer* to render the incoming audio should be added under the *Audio Receiver* component like shown in the Figure 3.19

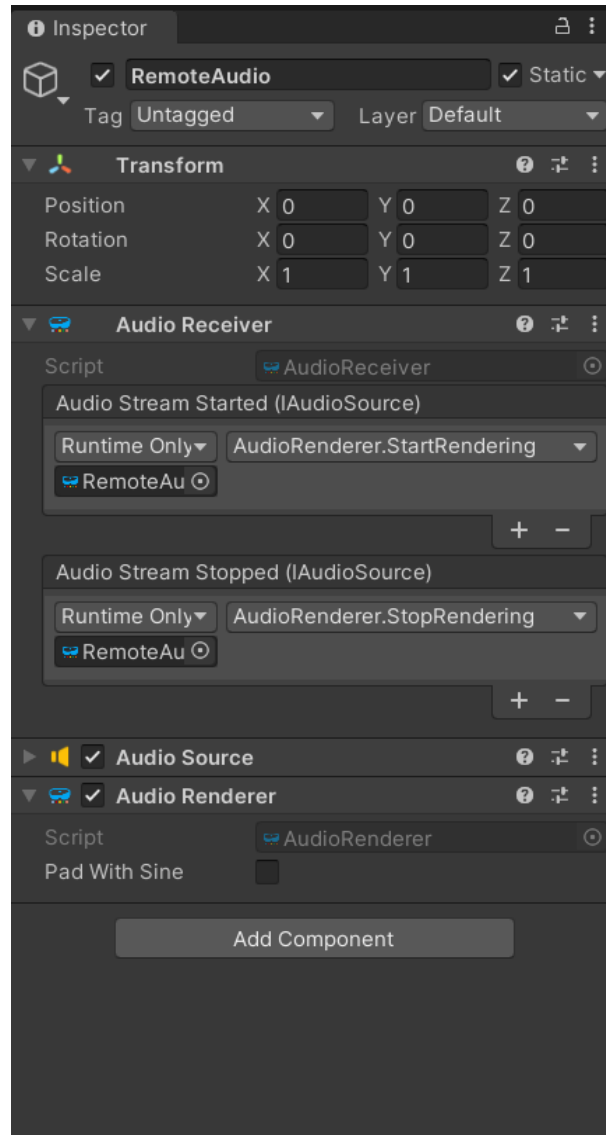


Figure 3.19: Audio Renderer Component

RemoteVideo: The WebRTC component *Video Receiver* function is similar to that of the *Audio Receiver* component, which is to render the videos received from the remote peer, which will then be displayed on the **Remote Player** created for this purpose. This WebRTC component is integrated into the application by creating a new GameObject **RemoteVideo**. Under the *Video Receiver* component on the inspector window, the command to start and stop the rendering of the incoming video stream (i.e., *VideoRenderer.StartRendering* and *VideoRenderer.StopRendering*) should be added like in the Figure 3.20 to control the video stream when the application is either opened or closed accordingly.

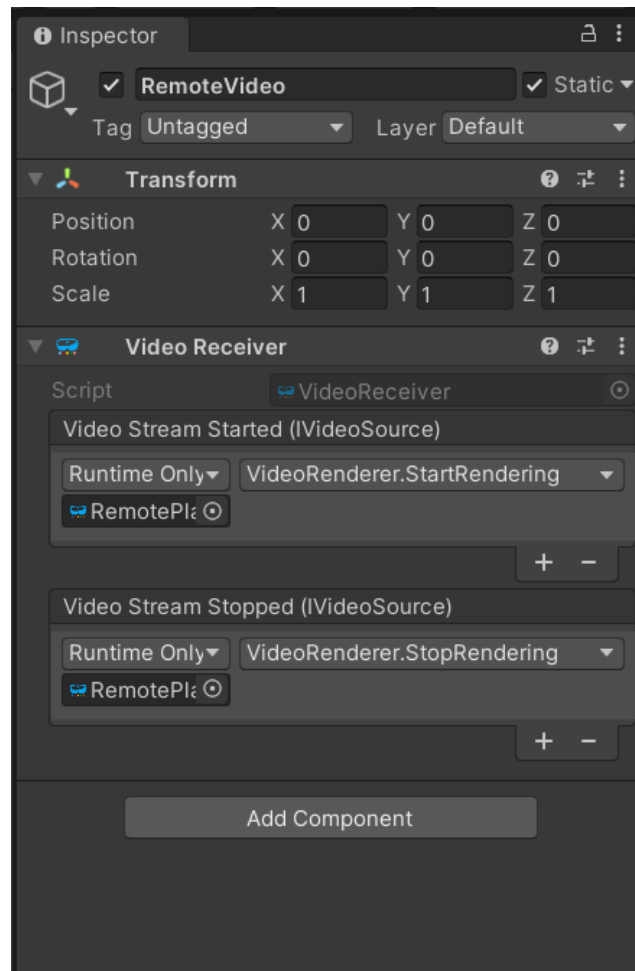


Figure 3.20: Video Receiver Component

LocalPlayer and RemotePlayer: These two GameObjects are created for the purpose of displaying the captured and received video frames of both the local and remote peers. In order to display the local video stream, a GameObject **LocalPlayer** is created and the following unity3D components (*Plane (Mesh Filter)*, *Mesh Renderer* and *Video Renderer*) are added under the Inspector window of GameObject **LocalPlayer** like in the Figure 3.21. Similarly, the video stream received from the remote peer is also displayed to the user through the GameObject **RemotePlayer**, which can be created by following the same steps used for creating the GameObject **LocalPlayer**. The only difference is that the **LocalPlayer** GameObject is a lot smaller than the **RemotePlayer** like in the Figure 3.22, this can be achieved by changing the scale and orientation values under the transform section on both GameObjects to the values shown in Figure 3.21.

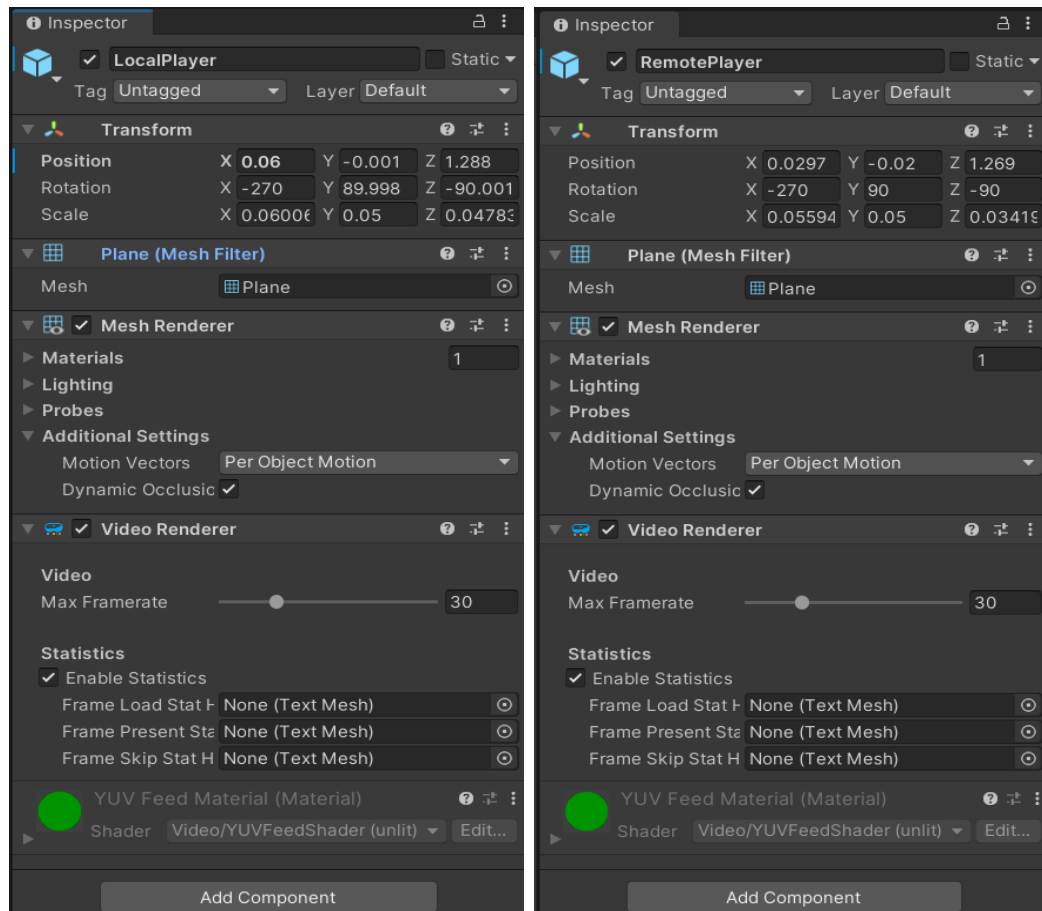


Figure 3.21: LocalPlayer and RemotePlayer



Figure 3.22: Local and Remote Video Stream display

CreateOffer Button: The Peer-to-peer communication between the application is established on an offer and answer basis. This has been discussed in detail in Chapter 3.1.2 from that it is clear that once both the peers are listening through the server, one peer has to send an offer through the network in the form of an SDP message so that the other peer can respond. This offer is initiated by the Client, by using **CreateOfferButton** which is available in MixedReality-toolkit. This button has to be configured for our purpose of initiating an offer by adding the necessary components under the GameObject **CreateOfferButton** in the inspector window as shown in Figure 3.23.

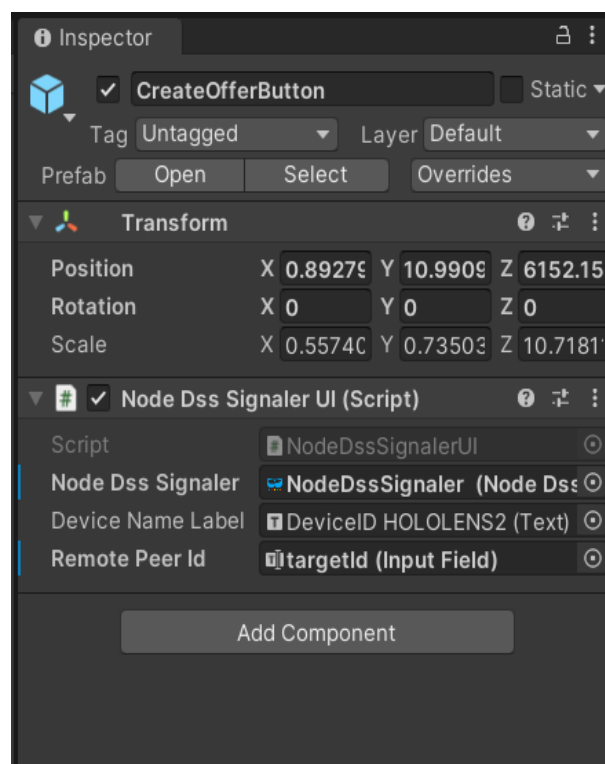


Figure 3.23: Configuring CreateOfferButton

LocalRemote Display: The video streams of both the local and remote peer will be displayed using the GameObjects **VideoPlayer** and **RemotePlayer** separately like shown in the Figure 3.22, but in order for this application to work well on Hololens2 these GameObjects have to be equipped with *HandTracking* functions so that the user can interact and change the position of these VideoPlayers when necessary. To achieve this result, the following GameObjects **LocalPlayer**, **RemotePlayer**, and **CreateOfferButton** have to be developed under a common GameObject named **LocalRemoteVideoDisplay**. Now the *HandTracking* function (i.e., touch, grab, and resize) can be added to the GameObject by attaching the following scripts

Constraint Manager, *Object Manipulator*, *NearInteractionGrabbale* like shown in Figure 3.24 which helps the user to interact with all the GameObjects as one.

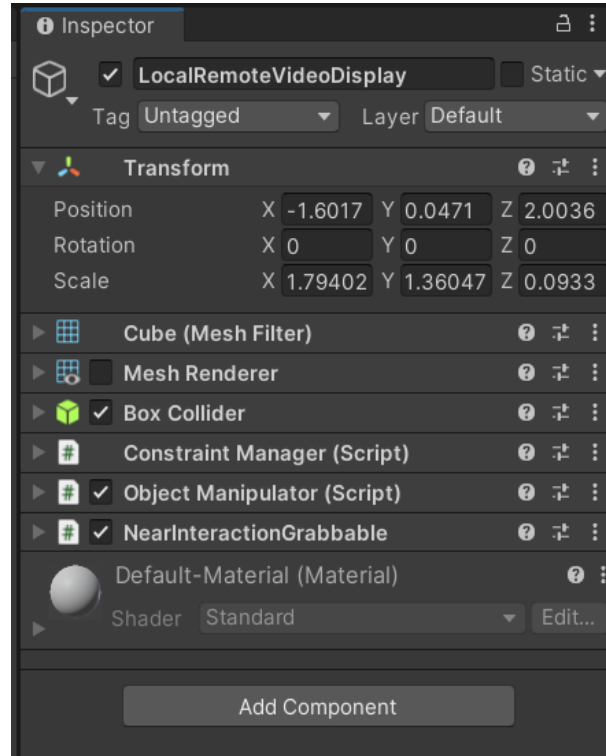


Figure 3.24: Configuring GameObjects for HandTracking Capabilities

3.2.2. Developing Mixed Reality environment for FIMI Simulations of Machinery

This chapter depicts the parallel working of the interactive animations developed by the Mixed Reality toolkit, which acts as a Virtual demo for the user's better understanding added with the Audio and video sharing as seen in Chapter 3.2.1. In order to facilitate parallel working in a main scene of **Video Sharing**, it is crucial to introduce a concept known as ADDITIVE SCENE. This concept entails the utilization of multiple individual scenes that are required to make use of the ADDITIVE SCENE.

Individual FIMI Scene

In this, Individual FIMI Scene an FBX file with drive constraints from FIMI machinery is loaded in and then imported to Unity Project to integrate with the Mixed Reality toolkit. Initially, the Mixed Reality Toolkit is added to the current scene to enable human interaction features.

Animation 01: The FBX file with drive constraint is imported to the Hierarchy window by enabling the animation type as LEGACY so that the imported drive constraint is converted as an animation file, which then finally is controlled by *Playable Director* Component once inside the Unity project. The Animation file **Drive B1** is added as Child under a 3DObject **Animation01** which is composed of primary components *Mesh renderer*, *Box collider*, *Constraint Manager(script)*, *Object Manipulator(script)* and *NearinteractionGrabbable* responsible for enabling the interaction features as shown in Figure 3.25

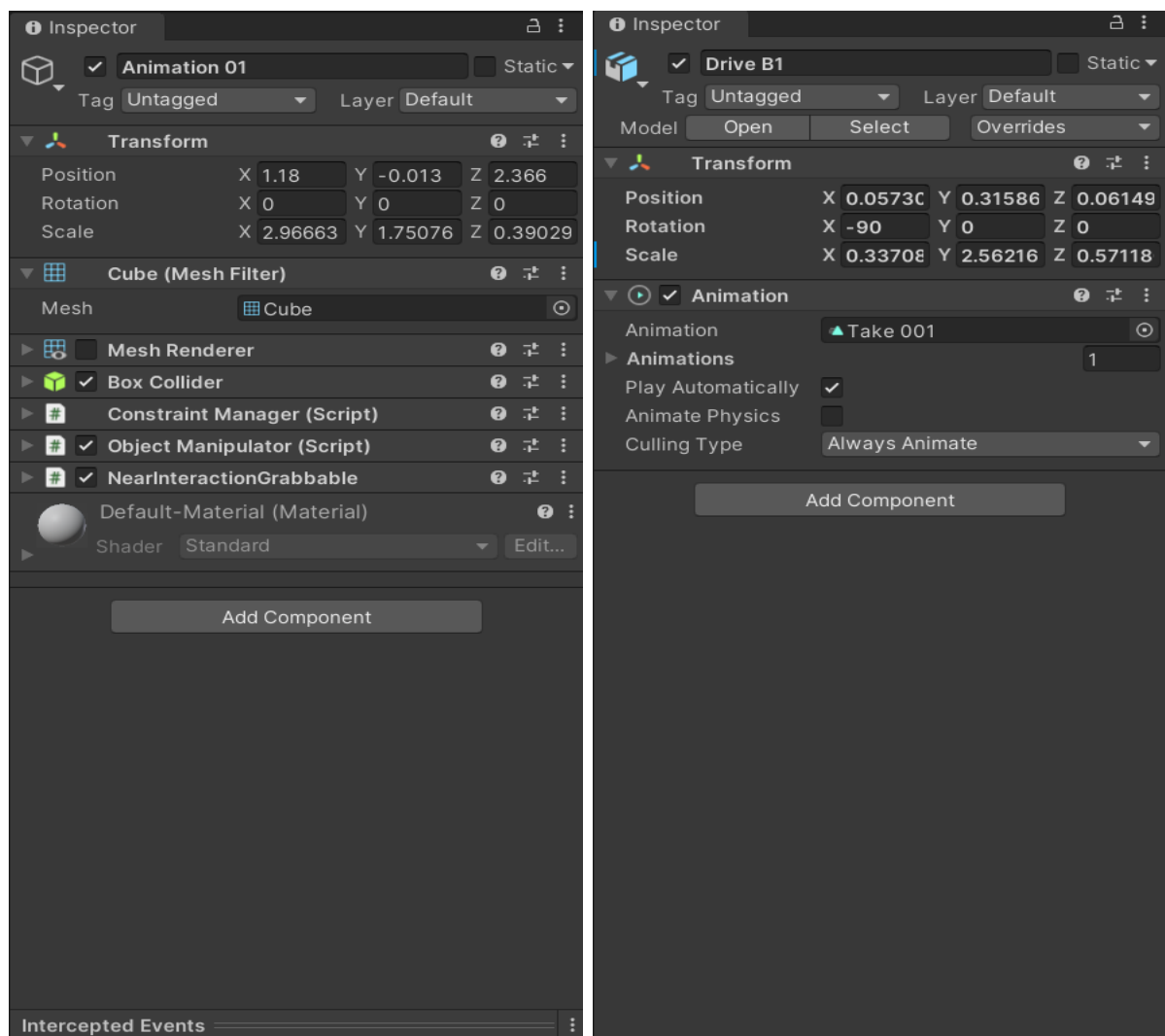


Figure 3.25: Animation 01 and Drive B1

The Same importation steps are applied to two more 3D Objects **Animation02** and **Animation03**

FIMI Scene01: In order to control the imported animations a Timeline asset with a control track is used, which is then later added to a GameObject **FIMI Scene01** under a component *Playable Director* assigning the wrap mode in Loop.

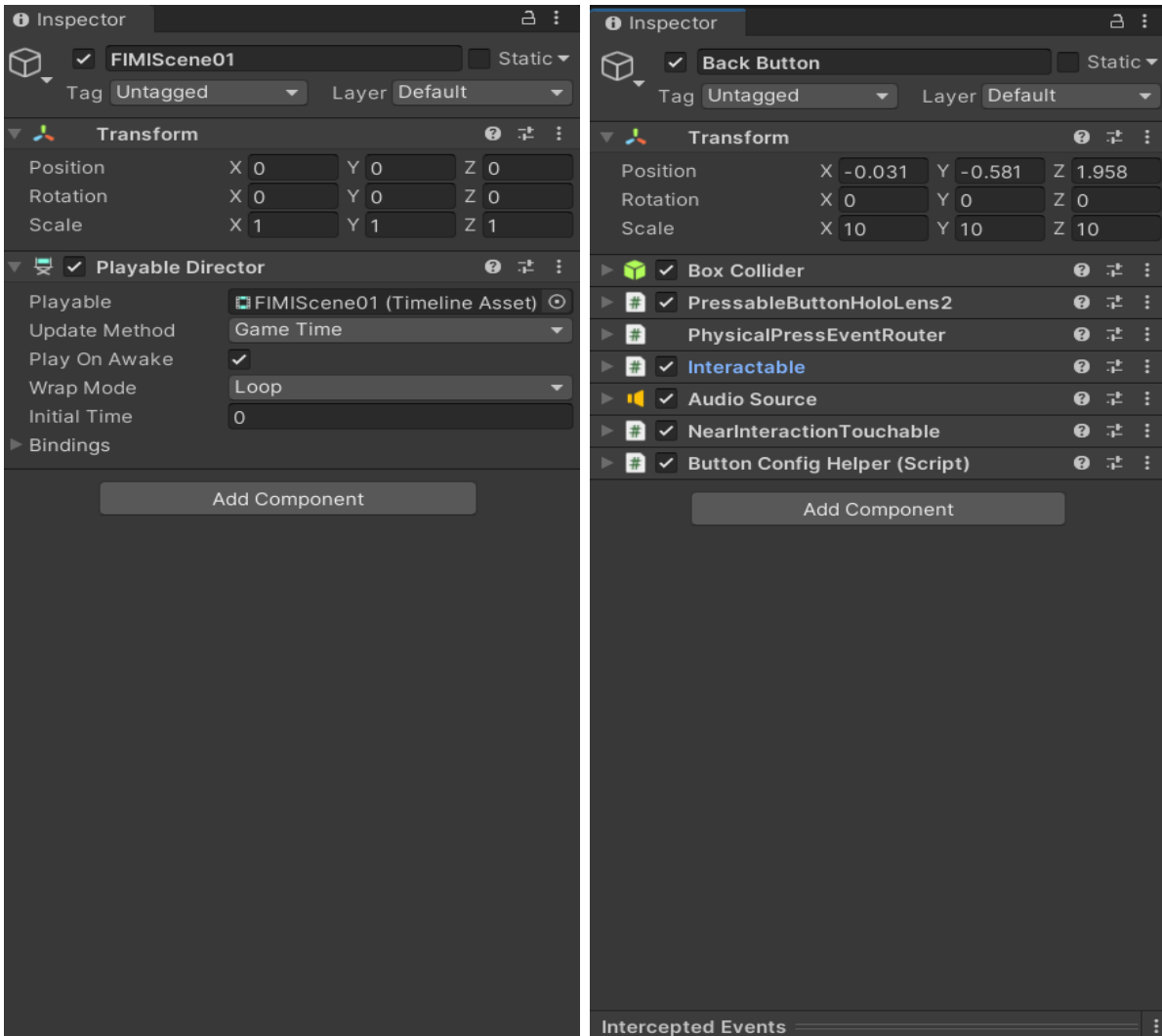


Figure 3.26: FIMI Scene01 and Back Button

Back Button: To Switch back to the main scene *Video Sharing* this UI button is equipped with few primary components *Box collider*, *PressableButtonHoloLens2*, *Interactable*, *NearInteractionTouchable*, and *Button Config Helper(scripts)*.

Close Scene01: The GameObject **Close Scene01** entails the component *Additive Script BACK button Scene01* which is responsible for the parallel working of Both scenes together. This Component is enabled with a C-sharp Script that handles the loading and unloading scene. Interacting with this Button, the ADDITIVE SCENE function is initiated. A Couple of Individual scenes are built as seen in the above Chapter 3.2.2

Additive Scene

Integration of all the individual scenes together in a main scene is enabled, with the help of the Additive scripts which are briefed in this Chapter.

FIMI Scene01 Button: An UI button is embedded in the main Scene *Video Sharing* which is linked with a GameObject **Nav.toFIMIScene01** through a component *Interactable* under the tab Events.

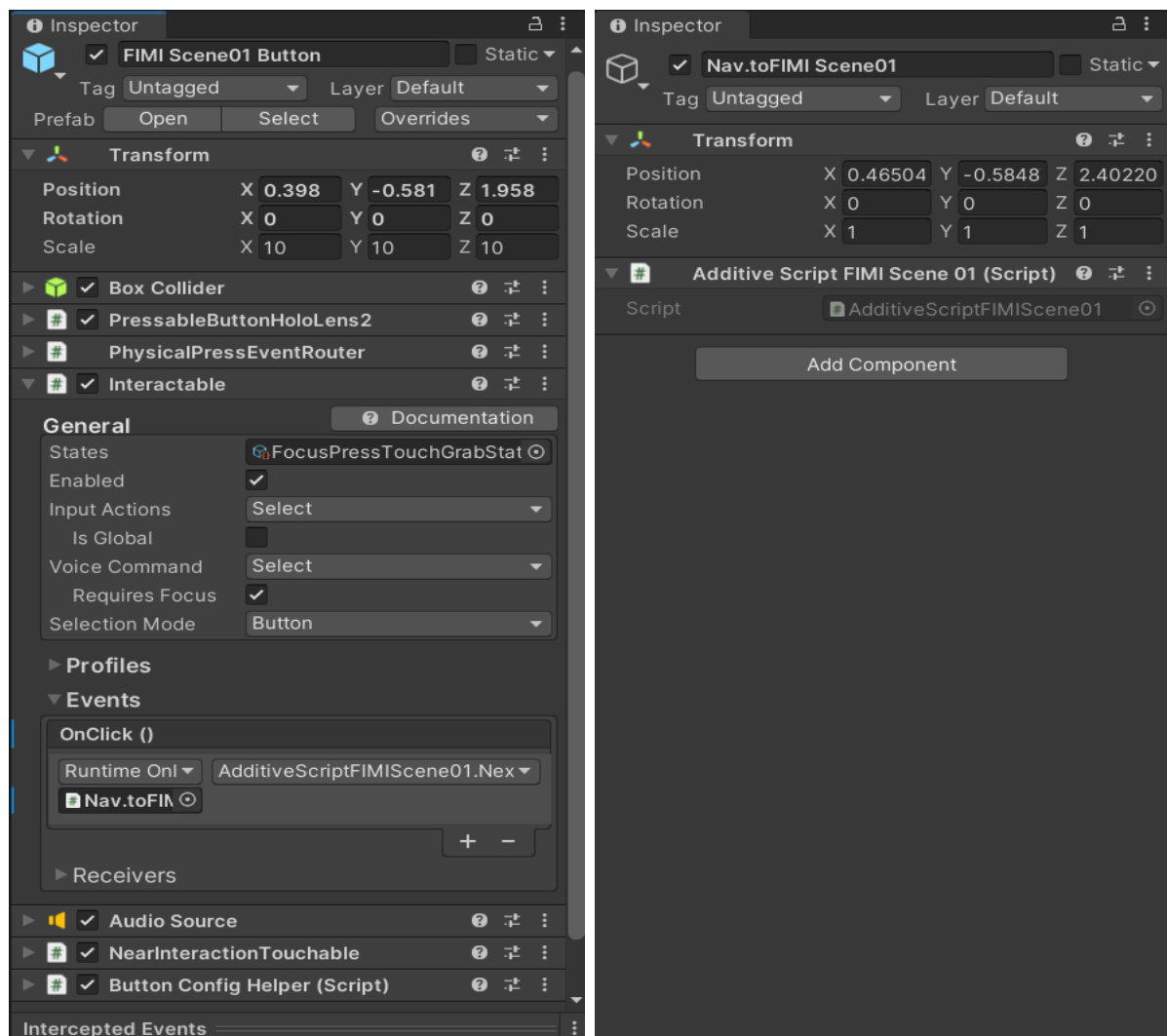


Figure 3.27: FIMI Scene01 Button and Nav.toFIMI Scene01

Nav.toFIMI Scene01: This Empty GameObject is built to hold the C-sharp script *AdditiveScriptFIMIScene01* which is later called out in the above GameObject **FIMI Scene01 Button**

Remote Assistance Application Final Output

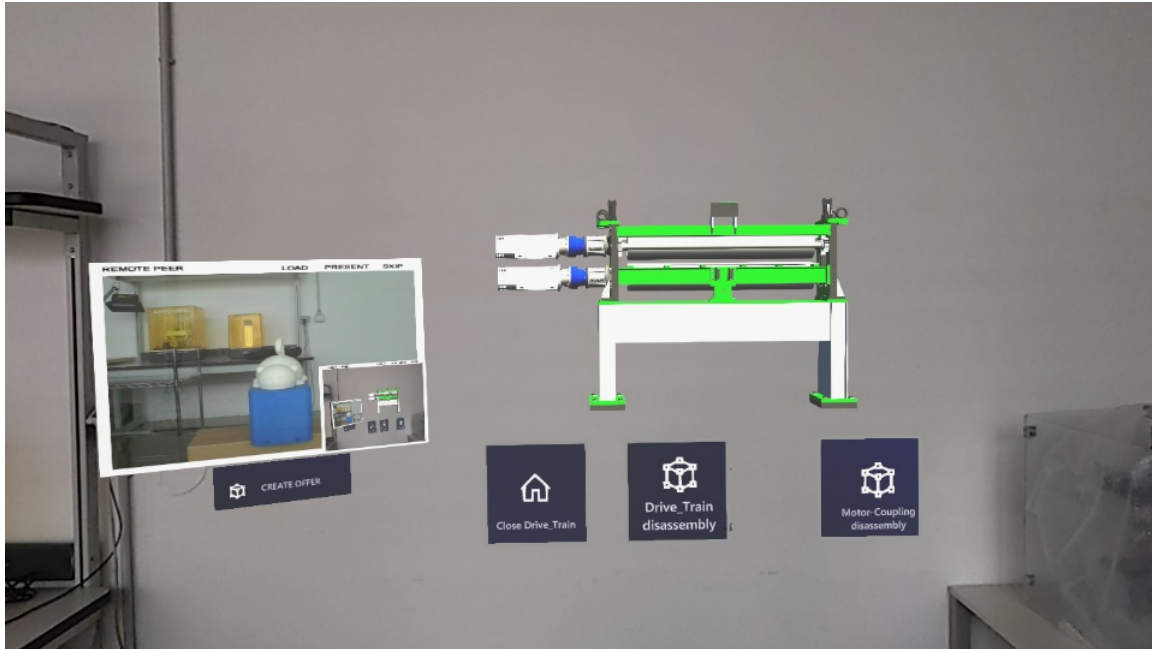


Figure 3.28: Final Output

3.3. Summary

The primary objective behind the development of this Application was to simplify and streamline maintenance work, particularly in industrial fields. Traditionally, the local maintenance group would receive important instructions from expert teams from the machine's manufacturer, often located far away from the actual site. This application aims to eliminate the need for hands-on assistance by leveraging advanced technology such as Hololens2, enabling the local maintenance group to effectively collaborate with designers in real time, seek clarifications, and communicate with them remotely. With the help of Hololens2, the local maintenance group can visualize complex instructions in a virtual format, making it easier for them to understand and execute tasks accurately. This seamless collaboration not only improves efficiency but also reduces errors that may occur due to miscommunication or misunderstanding. This application has a positive impact on overall productivity across various industrial sectors. By leveraging advanced technology like Hololens2, it enables remote collaboration, eliminates the need for printed manuals, reduces errors, and boosts productivity in the industrial sector. Figure 3.28 provides a concise overview of the ultimate result or outcome generated by the Application on the Hololens2. It encapsulates the key information or data that can be derived from using the Application, presenting it in a visual format for easy understanding and interpretation.

4 | Developing Educational Applications for Polimi and ASPOC Students

The Major Goal of this application is to perform a dual serving for students with and without disabilities, including those who have congenital retardation. The development of the special education application is intended to especially meet the needs of students who have Congenital Retardation, but it will also be helpful for students without impairments to learn in a secure environment. It provides specialized features and tools to enhance their academic path while taking into account their particular learning problems and aptitudes. A more inclusive and powerful educational experience is promoted by redesigning the application, which also improves accessibility, engagement, and overall learning outcomes for these students. Developing applications for students with Congenital Retardation requires a comprehensive understanding of their unique needs and abilities. These individuals often have cognitive impairments that affect their learning and communication skills. So, they require special attention and care. Customization, simplicity, accessibility, and inclusivity are key factors that need to be considered during the development process. This chapter depicts the challenges and needs of Students with Cognitive Retardation so that, a user-friendly application can be created that effectively support the learning and development of students with Congenital Retardation.

1 Students with Congenital Retardation

2 Special Education Application development

4.1. Students with Congenital Retardation:

Congenital Retardation, also known as intellectual disability, refers to a condition in which individuals have limitations in intellectual functioning and adaptive behavior. It is present from birth or early childhood and affects various aspects of a person's life, including their ability to learn, communicate, and perform everyday tasks.

An association named ASPOC [2] supports students with Congenital Retardation and aims to enhance their overall development and quality of life. They collaborated with Politecnico di Milano-Lecco in providing Education to students with Cognitive Retardation, which is believed that nurturing cognitive potential should be an integral part of formal education to prepare individuals for the challenges of the modern world [16].

Individualized Education Plan (IEP): Students with Congenital Retardation often benefit from an IEP, which is a personalized educational program tailored to their specific needs. This plan outlines the goals, accommodations, and support services required to facilitate their learning and development. It may include specialized instruction, assistive technology, speech therapy, occupational therapy, and behavioral interventions.

Supportive Environment: Creating a supportive environment is essential for pupils with Congenital Retardation. This involves ensuring inclusive classrooms that promote acceptance, understanding, and equal opportunities for all students. Teachers and school staff should receive training on how to effectively support pupils with intellectual disabilities. These students are provided with a better platform to explore, learn, and experience social and other skills.

Students with Cognitive Retardation, face a range of challenges in their academic pursuits. This has an impact on student's ability to learn, understand, and apply information. While the specific difficulties experienced by students with Cognitive Retardation may vary depending on the individual, there are several common problems that they often encounter. It is important to note that each student is unique, and their needs should be addressed on an individual basis.

4.1.1. Common Challenges faced by students with Cognitive Retardation in education programs:

1. Learning Difficulties: One of the primary challenges faced by students with Cognitive Retardation is learning difficulties. These individuals may have difficulty acquiring new knowledge and skills at the same pace as their peers without disabilities. They often require assistance when acquiring new skills in the initial phase, and individuals may take longer to excel in acquiring these skills. To overcome these challenges, teachers should employ effective instructional techniques, utilize visual aids to enhance comprehension, provide ample practice opportunities, and offer frequent feedback and reinforcement.

2. Communication Challenges: Communication difficulties are another significant problem faced by students with Cognitive Retardation. They may struggle with expres-

sive language skills (verbal or written) and receptive language skills (understanding spoken or written language). Limited vocabulary, poor grammar, difficulty organizing thoughts coherently, and challenges in understanding abstract or complex language are common issues. To address these challenges, they must be given their own environment, time, and practice for better communication skills. The educator can use Augmentative and alternative communication (AAC) systems may also be employed to enhance communication abilities.

3. Social and Emotional Difficulties: Students with Cognitive Retardation often face social and emotional challenges that can impact their overall well-being and academic performance. They may struggle with social interactions, making friends, understanding social cues, and regulating their emotions. To address social and emotional needs, educators can create inclusive and supportive classroom environments that foster positive relationships among students. Teaching social skills explicitly through role-playing, modeling, and guided practice can help students develop appropriate social behaviors.

By taking into account the aforementioned challenges [4] experienced by students with Cognitive Retardation, the utilization of advanced technology can be employed to address these issues and create a more conducive environment for these students to enhance their academic and individual skills.

The Findings of the present research [22] have indicated that the application of mixed reality by creating experiential learning, enjoyment, and interactivity, has a positive and significant impact on students' novel learning experiences. One of the key outcomes of having learning experiences with mixed reality technology is the satisfaction of students with the content, technology, and learning environments. Mixed Reality offers several advantages in the educational sector. Firstly, it provides a more immersive and interactive learning experience for students, allowing them to engage with the subject matter in a hands-on way. This can enhance their understanding and retention of complex concepts. Additionally, Mixed Reality can simulate real-world scenarios, enabling students to practice and develop practical skills in a safe and controlled environment. These benefits make Mixed Reality a valuable tool for educators looking to enhance their teaching methods and improve student outcomes. It is important to note that while these technologies can be beneficial for individuals with cognitive ability, it is essential to consider individual needs and preferences when selecting specific applications or experiences. Using Mixed Reality technology for cognitive training has low cost, ease of accessibility (it can be set up at home), ease of use, and high user tolerance.

4.1.2. Challenges faced by students with Cognitive Retardation in using Mixed Reality Technology:

The Mixed Reality technology is deployed in Hololens2, where the students can feel the immersive learning experiences but using this MR headset even causes discomfort to normal students without disabilities. Thinking of using these advanced technologies for students with Cognitive Retardation is a nightmare. For each and every problem there is a solution, First, the problems faced or experienced by the disabled students are listed and then the solution is found.

Students with Cognitive disabilities may **face several challenges** when using Mixed Reality. These challenges can vary depending on the specific cognitive ability and the severity of the disability. However, there are some common issues that individuals with Cognitive disabilities may encounter:

1. Complexity of Interaction: Mixed Reality (MR) experiences often require users to navigate complex virtual environments and interact with various elements using gestures. This complexity can be overwhelming for individuals with Cognitive disabilities, making it difficult for them to understand and engage with the technology effectively.

2. Cognitive Overload: MR and VR experiences can provide a vast amount of sensory information simultaneously, including visual, auditory, and sometimes haptic feedback. This sensory overload can overwhelm individuals with cognitive disabilities, leading to difficulties in processing and interpreting the information presented to them.

3. Attention and Focus: Students with Cognitive disabilities often struggle with attention and focus-related issues. The immersive nature of MR and VR experiences can make it challenging for them to maintain attention on specific tasks or instructions, leading to reduced engagement and comprehension.

4. Learning Curve: Using tablets and mobile devices may require learning new skills or understanding complex user interfaces. Individuals with Cognitive disabilities may find it challenging to grasp these concepts or remember the steps involved in operating these devices.

5. Sensory Sensitivities: Some individuals with cognitive disabilities may have sensory sensitivities, such as sensitivity to bright lights, loud sounds, or certain textures. The immersive nature of MR/VR experiences can exacerbate these sensitivities, making it uncomfortable or overwhelming for them to use such technologies.

These challenges can be overcome through effective training, guided instruction, repeated practice, and hands-on experience. Moreover, fostering a supportive and encouraging environment can empower individuals to conquer these obstacles and boost their confidence when facing new tasks.

4.1.3. Design strategies implemented for individuals with Cognitive disabilities:

To address these challenges, **certain design strategies** are implemented to make MR experiences, more accessible for individuals with Cognitive disabilities. These strategies include:

- 1. Simplified Interfaces:** Designing intuitive and simplified user interfaces can help individuals with Cognitive disabilities navigate through MR/VR experiences, tablets, and mobile devices more easily. This can involve reducing clutter, using clear and concise instructions, and providing visual cues to aid comprehension.
- 2. Customizable Settings:** Allowing users to customize the settings of MR/VR applications, tablets, and mobile devices can cater to individual needs and preferences. This includes options to adjust the level of sensory input, font size, and color contrast, or enable features like text-to-speech or voice commands.
- 3. Clear Instructions and Feedback:** Providing clear instructions and feedback within the MR/VR experience or on tablets/mobile devices can help individuals with cognitive disabilities understand tasks and progress through activities more effectively. This can involve using visual cues, audio prompts, or haptic feedback to reinforce actions or guide users.

It is important to note that while these strategies can enhance accessibility for individuals with Cognitive disabilities, there is no one-size-fits-all solution.

4.2. Special Education Application development:

The students with Cognitive impairment are trained in the Polimi lab using a Milling machine called Charly 4U. It is extremely risky to operate a milling machine without any prior knowledge; hence, a virtual demonstration through Hololens2 or virtual educator on how to operate the milling machine is produced.

4.2.1. CHARLY 4U

The CHARLY 4U milling machine is a state-of-the-art piece of equipment that offers high precision, efficiency, and versatility in machining operations. Its advanced features, such as CNC control, automation capabilities, and a wide range of tooling options, make it an indispensable tool for modern manufacturing industries.

It can be operated in 2 different axes, they are:

- **3 axis** – x y z axes movement
- **4 axis** – x y z and rotation around x-axis

Each axis has nearly 9 basic functions to be educated on before getting hands-on experience in the real machine. Mastering the basic steps for each axis is crucial for operating the machine effectively and ensuring accurate machining results. Additionally, understanding the machine's capabilities and workings allows the students to optimize their learning. By understanding the machine's capabilities and workings, students can identify its limitations and explore ways to optimize their learning experience. This knowledge also enables them to understand the basic settings required for efficient and smooth operations.

In actuality, the Charly 4U is controlled via a dedicated computer where the instructions are fed into it. When used with the Hololens2, the computer tabs are shown as a canvas for training purposes.

Autodesk Inventor 2022:

Using Autodesk Inventor 2022, the Charly 4U machine architecture is transformed into a CAD-designed model with the necessary drive limitations for animation.

Autodesk 3ds Max:

A file from Autodesk Inventor 2022 that has drive restrictions is fed to 3Dmax, where the specific file is converted into a .FBX file, which will then be accepted in Unity.

Unity:

The Unity version used for this application is Unity 2020.3.22f1, and Visual Studio Community 2019 was used for coding the scripts. Lots of scripts are re-coded as per the requirements, which would help the students with cognitive retardation.

4.2.2. UNITY

To experience the real-world and virtual world environment together in an immersive way, we use the Mixed Reality Toolkit as seen in the previous Chapter 3.2. In this section, the basic features required to build an interactive application where real-world and virtual-world environment blends together are detailed.

1. CANVAS FOR TEXT:

To present the text elements in scene names and instructions a **Unity Canvas** is used, it is a crucial component for displaying texts, scene names, instructions, and other

visual elements within a Unity project. It consists of scripts like *Canvas Utility* and *Near Interaction Touchable Unity UI*, as well as components like *Canvas*, *Canvas Renderer*, *Rect Transform*, *TextMesh-Pro UI*. These components and scripts work together to create an interactive and visually appealing user interface.

- *Canvas*: It represents the abstract space in which the UI is laid out and rendered. All UI elements must be children of a GameObject that has a Canvas component attached.
- *Canvas Renderer*: It is used to render a graphical UI Object contained within a Canvas.
- *Rect Transform*: It is the 2D layout counterpart of the Transform component. Where Transform represents a single point, Rect Transform represents a rectangle that a UI element can be placed inside. If the parent of a Rect Transform is also a Rect Transform, the child Rect Transform can also specify how it should be positioned and sized relative to the parent rectangle.
- *TextMesh-Pro (UI)*: It is an easy-to-use system for high-quality text. It has many text appearance and formatting options and is an easy way to add a professional touch to any project's user interface.

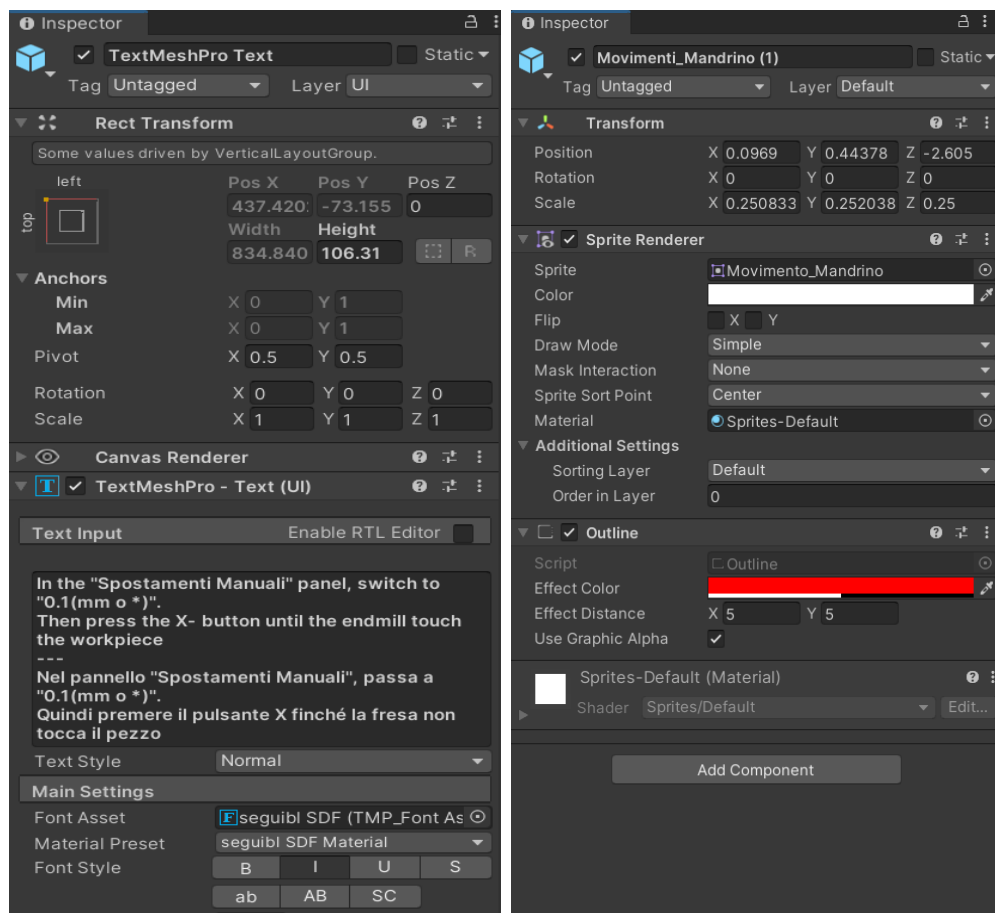


Figure 4.1: Inspector Tab for TextMesh-Pro and Image

2. CANVAS FOR IMAGES:

To present the images, a *sprite renderer* component in Unity allows for the rendering of 2D images onto GameObjects. By assigning a UI sprite as the input, the images or tabs like on the computer screen can be replicated within the scene. This provides a visually appealing and interactive way to present the images to the players.

- *Sprite Renderer:* The Sprite Renderer component renders the Sprite and controls how it visually appears in a Scene for both 2D and 3D projects.

3. ANIMATIONS:

While importing the FBX file into Unity, the **Legacy** option is selected to activate animation constraints. The appropriate animation clips, Take001, are inserted under the *Animation* component, and the **"Play automatically"** check box is checked.

To interact with the animation in an immersive manner, the Animation is set as a Child to a GameObject which consists of a Few components: *Transform*, *Cube (Mesh Filter)*, *Mesh Renderer*, *Box Collider*, *NearInteractiongrabble*, *Constraint Manager (Script)*, *Object Manipulator*, *Boundscontrol*

- *Transform:* The Transform is used to store a GameObject's position, rotation, scale, and parenting state and is thus very important. It consists of ;

Position- It Transforms the GameObject in X, Y, and Z coordinates

Rotation- It Rotations the GameObject around the X, Y, and Z axes.

Scale- It Scales the GameObject along the X, Y, and Z axes.

- *Cube (Mesh Filter):* The Mesh Filter takes a mesh from your assets and passes it to the MeshRenderer for rendering on the screen.

- *Mesh Renderer:* The Mesh Renderer takes the geometry from the Mesh filter and renders it at the position defined by the GameObject's transform Component.

- *Box Collider:* The Box collider is a basic cuboid-shaped collision primitive, which asks as a trigger to interact with virtual hands from hololens2.

- *Bounds Control:* The Boundscontrol.cs script provides basic functionality for transforming objects in mixed reality. A bounds control will show a box around the hologram to indicate that it can be interacted with and also it is equipped with a small node which helps the user to interact with the animation for resizing and rotation.

- *Object Manipulator:* The ObjectManipulator.cs script makes an object movable, scalable, and rotatable using one or two hands. The object manipulator can be configured to control how the object will respond to various inputs. The script should work with most forms of interaction, such as Hololens2 articulated hand, and Hololens2 hand rays' inputs.

- *NearInteractiongrabbable*: The Nearinteractiongrabbable.cs script makes an object that has a collidable on it in order to make that collidable near grabbable. So any Mixedreality near pointer will then dispatch pointer events to the closest near grabbable objects.

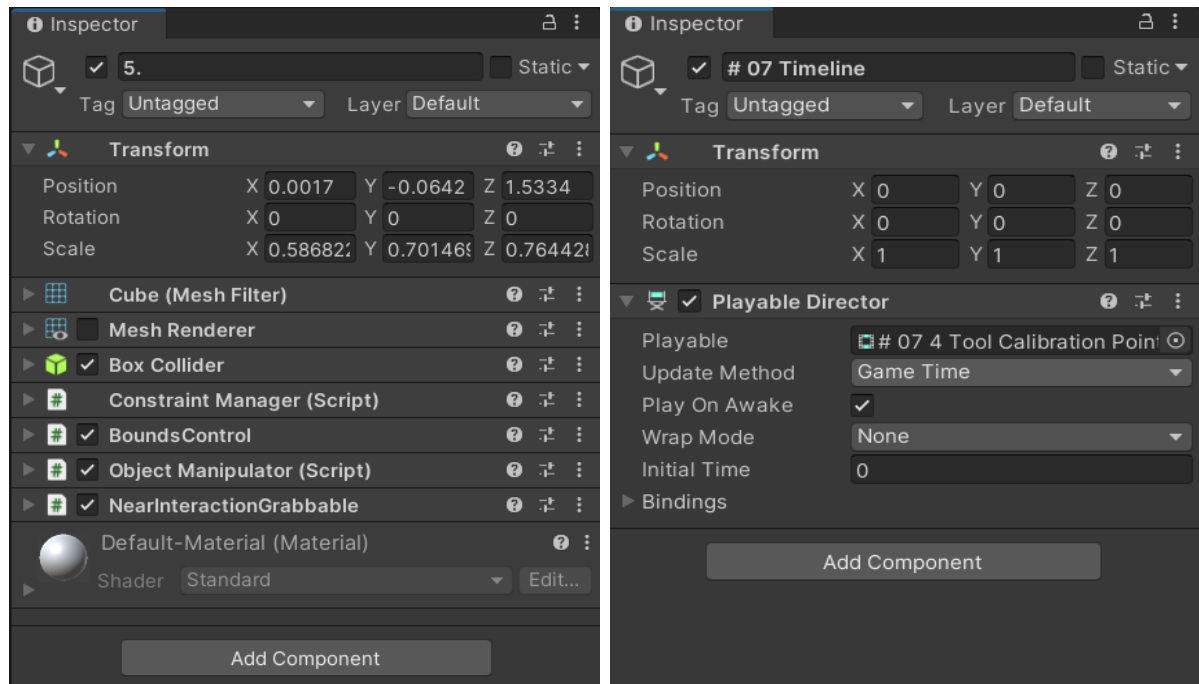


Figure 4.2: Inspector Tab Animation GameObject and Timeline

Importing two or more animations, as explained in the prior phases, in order to show how well the machine functions The use of "**Unity Timeline**" is made possible by the need to produce cinematic content using the animations. By using a control track, each animation can be managed separately. A component known as *Playable Director* is added to a *GameObject* to enable the use of a timeline.

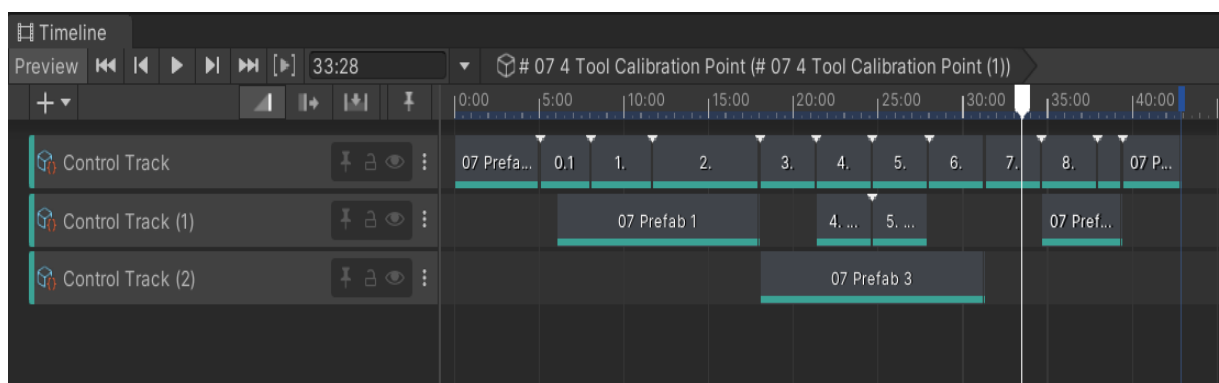


Figure 4.3: Timeline Tab

4. BUTTONS:

A button gives the user a way to trigger an immediate action. It is one of the most foundational components in mixed reality. MRTK provides various types of button prefabs. In this application, three major buttons are used with various scripts are responsible for the aforementioned functions in this application, including switching between scenes, playing and pausing animations, and returning to the previous scene. These buttons can be programmed to trigger specific actions when pressed. For example, the button for toggling scenes can be linked to a script that switches the active scene and updates the display accordingly. Similarly, the play and pause buttons can be connected to scripts that control the animation playback, allowing users to easily control the flow of the animations.

BACK BUTTON:

A GameObject with a button configuration needs to include some basic Components and scripts that, when combined, explain to the user why the button should be pressed. These typically consist of a Button component that governs the button's visual appearance and interaction and an Event System component that controls user input and initiates the necessary activities. In order to define precise actions to be taken when the button is pressed, scripts like an onClick script can be implemented.

- *Box Collider*: The Box collider is a basic cuboid-shaped collision primitive, which asks as a trigger to interact with virtual hands from Hololens2.
- *Icon and text*: It helps the user to add text and icons for the specific buttons used.
- *Seeitsayitlabel*: It acts as voice assistance, where the user can utter the word displayed on the button to interact with the specific button. This feature enhances the user experience by providing a hands-free interaction option. It enables users to navigate through the interface effortlessly by simply speaking the commands associated with each button.
- *PressableButtonHololens2*: It is a Shell style button that supports the precise movement of the button for the direct hand Tracking input. It combines the Interactable script with the Pressable button script.
- *Audio Source*: It plays back an Audio Clip in the scene. The clip can be played to an audio listener or through an audio mixer. The audio source can play any type of Audio Clip and can be configured to play these as 2D, 3D, or as a mixture (SpatialBlend). The audio can be spread out between speakers.
- *NearInteraction Touchable*: The Nearinteractiongrabbable.cs script makes an object that has a collidable on it in order to make that collidable near grabbable. So any Mixed reality near pointer will then dispatch pointer events to the closest near grabbable objects.
- *Button Config Helper (Script)*: The Button Config Helper (Script) is a component that

can be added to a button in order to assist with the customization of its icon, text, and label. When this script is added to a new empty GameObject, it allows for easy navigation back to the previous scene by defining the necessary functions within the basic event `OnClick()`.

```
using UnityEngine;
using UnityEngine.SceneManagement;

public class TOMAINMENU : MonoBehaviour
{
    public void NextScene()
    {
        SceneManager.LoadScene("# 01 4 Axes Main Menu");
    }
}
```

Figure 4.4: Customized Back Button Script

PAUSE BUTTON:

The only difference for the BACK BUTTON is the customized script to pause and play the animation, which tells the timeline to start and stop depending on the user's needs. By exploiting this unique capability, the user may comprehend how the system functions. Thanks to this unique function, the user has control over the animation, which enables them to halt and resume it whenever they like. By allowing for flexibility and interaction with the working equipment, it improves the user experience.

```
using System.Collections;
using System.Collections.Generic;
using UnityEngine;

public class Pause : MonoBehaviour
{
    bool isPaused = false;

    public void PauseGame()
    {
        if (isPaused)
        {
            Time.timeScale = 1;
            isPaused = false;
        }
        else
        {
            Time.timeScale = 0;
            isPaused = true;
        }
    }
}
```

Figure 4.5: Customized Pause Button Script

REPLAY BUTTON:

The sole difference between the BACK BUTTON's components and functionality is the customized script that instructs the scene manager to replay the current scene. This provides a way for the user to become proficient in a particular scene or machine's operation over and over again in repetition mode. Users who need extra practice or who wish to go over a scene several times might really benefit from this feature. The user can strengthen their comprehension and advance in that particular area of machine operation by having the ability to recreate the present scene with a custom script.

```

using UnityEngine;
using UnityEngine.SceneManagement;

public class ReplayButton : MonoBehaviour
{
    public void ResetScene()
    {
        SceneManager.LoadScene(SceneManager.GetActiveScene().buildIndex);
    }
}

```

Figure 4.6: Customized Replay Button Script

Educational Application Final Output

Figure 4.7: Hololens2 view

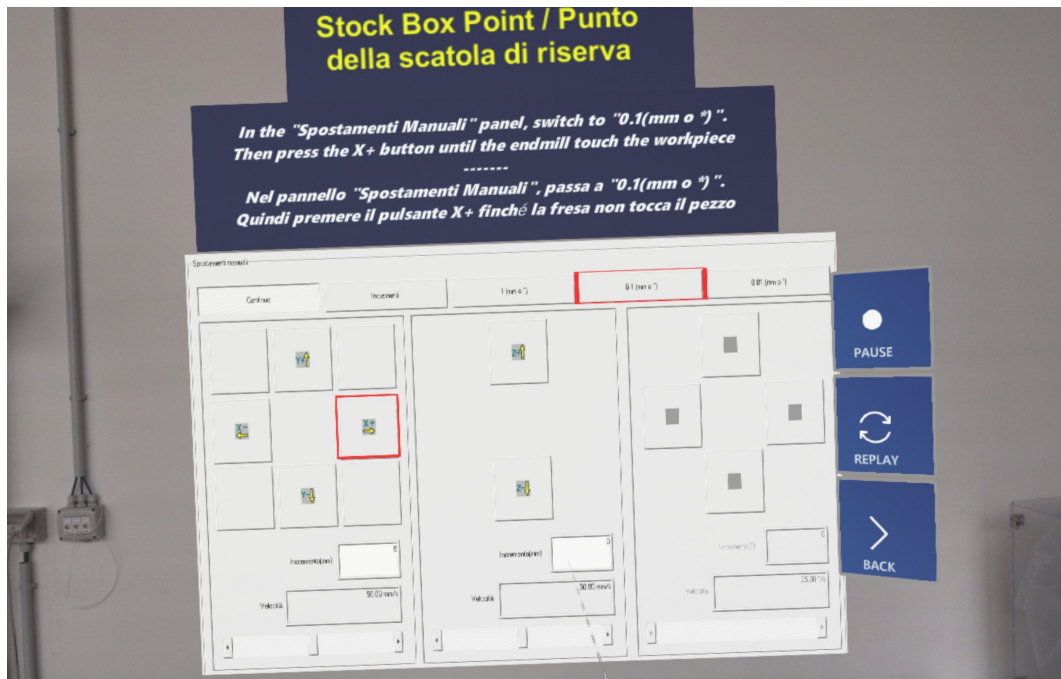


Figure 4.8: Hololens2 view

4.3. Summary

This Educational application caters to both students with and without disabilities, including those with congenital retardation. It incorporates specialized features and tools to enhance their learning experience. The challenges associated with developing applications for students with congenital retardation are addressed to the fullest extent by offering simplified interfaces, customizable settings, and clear instructions with feedback. The application utilizes Unity Canvas for text and images, employs a sprite renderer for images, and integrates a Unity Timeline for cinematic content. Furthermore, it incorporates button prefabs to enable immediate actions such as scene transitions, play/pause animations, and returning to previous scenes. By creating a supportive and encouraging learning environment, the application assists individuals in overcoming challenges and gaining confidence in handling new tasks. Feedback from students played a pivotal role in defining the application's usability, and these insights are detailed in the corresponding Chapter 6.

5 | Developing Special Education Applications for the ILSEME Association

The ILSEME association shares a similar motto with the ASCOP association, as both organizations aim to empower individuals with Congenital Retardation. The ILSEME association specifically focuses on engaging students or pupils with Congenital Retardation in basic and regular activities within a company or office setting. The goal is to provide them with opportunities to take care of their own financial security and live a normal life, just like their peers without retardation.

To accomplish this goal, the ILSEME association [9] acknowledges the significance of offering comprehensive training to these individuals. They prioritize training in fundamental activities such as cleaning, mopping, and maintaining cleanliness within the company or office premises. By entrusting these responsibilities to the students/pupils, the association aims to cultivate a sense of ownership and responsibility in them. This approach not only aids in the development of their inherent skills but also imparts a valuable understanding of the duties associated with both work and daily life.

By engaging students/pupils with Congenital Retardation in such activities, the ILSEME association strives to create an inclusive environment where everyone can contribute and thrive. This approach promotes independence, self-confidence, and a sense of belonging among individuals with Congenital Retardation. It also helps break down societal barriers and stereotypes by showcasing the capabilities and potential of these individuals.

The achievement of a goal may not occur overnight, but with consistent effort and dedication, it can be reached. This particular chapter takes a different approach compared to the previous Chapter 4, where an animated representation of a real machine is redesigned and utilized for training purposes for both students with and without disabilities.

However, there is a significant difference in this chapter as it incorporates the use of the Vuforia engine. This engine possesses the capability to read QR codes and play videos on the Hololens2 device. By leveraging this technology, the students/pupils engaged in cleanliness training activities are able to participate in a hands-free manner and replicate each step simultaneously. This innovative approach to training proves to be particularly beneficial for students with Congenital Retardation, as it provides them with an improved learning method. This technology-driven method proves to be particularly beneficial for students with Congenital Retardation, as it enables them to participate in the training on an equal footing as their peers.

The Common challenges and challenges in using Mixed Reality applications for students/pupils with Cognitive Retardation are already discussed in Chapters 4.1.1, 4.1.2, 4.1.3.

5.1. Special Education Application Development using Vuforia Engine

The students or pupils with Cognitive Retardation are trained to perform the basic cleaning activities using Mixed Reality technology with the Vuforia Engine. This technology allows them to engage in interactive virtual simulations that mimic real-life cleaning tasks via pre-recorded videos. Through this training, students or pupils with cognitive retardation can develop essential life skills and gain a sense of independence and accomplishment. The Vuforia Engine in Hololens2 is used to track the QR code and a pre-recorded video starts playing. The students can then follow along with the Videos, which provide step-by-step instructions and guidance for each cleaning task. This hands-on approach helps them learn and practice the necessary techniques in a safe and controlled environment. Additionally, the use of QR codes ensures that the students can easily access and navigate through different cleaning scenarios, enhancing their learning experience.

VUFORIA ENGINE

Vuforia Engine is a powerful software development kit (SDK) developed by PTC Inc. that enables image tracking and Augmented Reality (AR) experiences on various platforms, including Microsoft Hololens2. This advanced engine provides developers with the tools and capabilities to create immersive AR applications that can recognize and track real-world objects or images, allowing virtual content to be seamlessly integrated into the user's environment.

One of the key features of Vuforia Engine [18] is its robust image recognition and tracking technology. It utilizes computer vision algorithms to analyze and identify specific target images or objects in real time. This allows developers to create AR experiences that can overlay virtual content onto these recognized targets, enhancing the user's perception of the physical world.

Vuforia Engine supports a wide range of image targets, including 2D images, 3D objects, and even environments. Developers can define target images or objects using Vuforia's Target Manager, which provides an intuitive interface for uploading and managing target assets. Once the targets are defined, Vuforia Engine can detect and track them using the device's camera, enabling interactive AR experiences.

In the context of Microsoft HoloLens2, Vuforia Engine offers seamless integration with the device's capabilities. It leverages HoloLens2's advanced sensors, such as depth cameras and eye-tracking technology, to enhance the accuracy and stability of image tracking. This integration allows developers to create more realistic and immersive AR experiences on HoloLens2 [30].

The Vuforia Engine is composed of several managers who play crucial roles in various aspects of its functionality. These managers work together to ensure the smooth integration of Vuforia with Unity and provide a comprehensive augmented-reality experience.

- *License Manager:* The License Manager is a web application on the developer portal that allows you to create and manage license keys. A License Key is an alpha-numeric code that the app must pass to the SDK as a string during SDK initialization. The License Key determines which features are available for the app to use, such as Model Targets or Cloud Recognition.

- *Target Manager:* The Target Manager is a web application to create and manage Device, Cloud, and VuMark databases. The Target Manager is available on the Engine Developer Portal by logging in with a Vuforia Engine Developer account. The Target Manager can generate a Device Database for Images, Cylinders, and Multi Targets. It must be downloaded and stored on the device with the application. The supported image formats are described in the Target Manager article. In short:

- i Images must be in one of the supported formats: PNG or JPEG.
- ii Images must either be grayscale or RGB-8-bit or 24 bit.
- iii Image sizes must not exceed 2 MB.

Finally, the Database is downloaded and opened in the Unity project.

UNITY

The "Vuforia Hololens2 Sample" asset is downloaded from the unity ASSET STORE and added to the particular unity project to allow access to the Vuforia Engine. This asset consists of the Mixed Reality Toolkit, the Hololens2's foundation. The Vuforia contents are set up in the MRTK. In order to guarantee the proper operation of image targeting, the Vuforia content GameObject is made up of numerous components and scripts.

- *Vuforia Content*: It Consists of a Script Frame Rate Settings and acts as a Parent GameObject to hold the Child GameObjects which are responsible for Image targeting and Video playing.
- *Frame Rate Settings*: The Vuforia Engine APK for Unity provides the Frame rate setting (scripts) method to obtain and set a recommended render framerate based on your app's rendering and performance requirements. To set a target framerate, call these methods when initializing the app.
- *Image Target Behaviour (scripts)*: It is responsible for collecting the Specific Targets from the Specific database.
- *Default Observer Event Handler*: It is in charge of monitoring the Specific Target when the Hololens2 camera is tracking it. Here Hololens2 will be able to display the video when the target is found and pause the video when the target is lost thanks to the videoplayer GameObject that is added under Events().
- *Image Content*: To make the Above scripts function they must be added to an empty GameObject [Image Content].
- *Plane (Mesh Filter)*: A Plane Mesh Filter component holds a Reference to a Plane. It renders the mesh with the Plane 3D Object.
- *Mesh Renderer*: The Mesh Renderer takes the geometry from the Mesh filter and renders it at the position defined by the GameObject's transform Component.
- *Mesh Collider*: The mesh collider is a basic Plane collision primitive, which asks as a trigger to interact with virtual hands from Hololens2.
- *Video Player*: It is employed to add video files to game objects and play them on their run-time textures. A source file containing the required video clip is added. When the Hololens2 reads the target, make sure that Play on awake, wait for the first frame, Loop, and Skip on Drop are checked for better video display.

For Better Visualization Import Text, Images, and buttons to Toggle between scenes as seen in the previous Chapter 4.2.2.

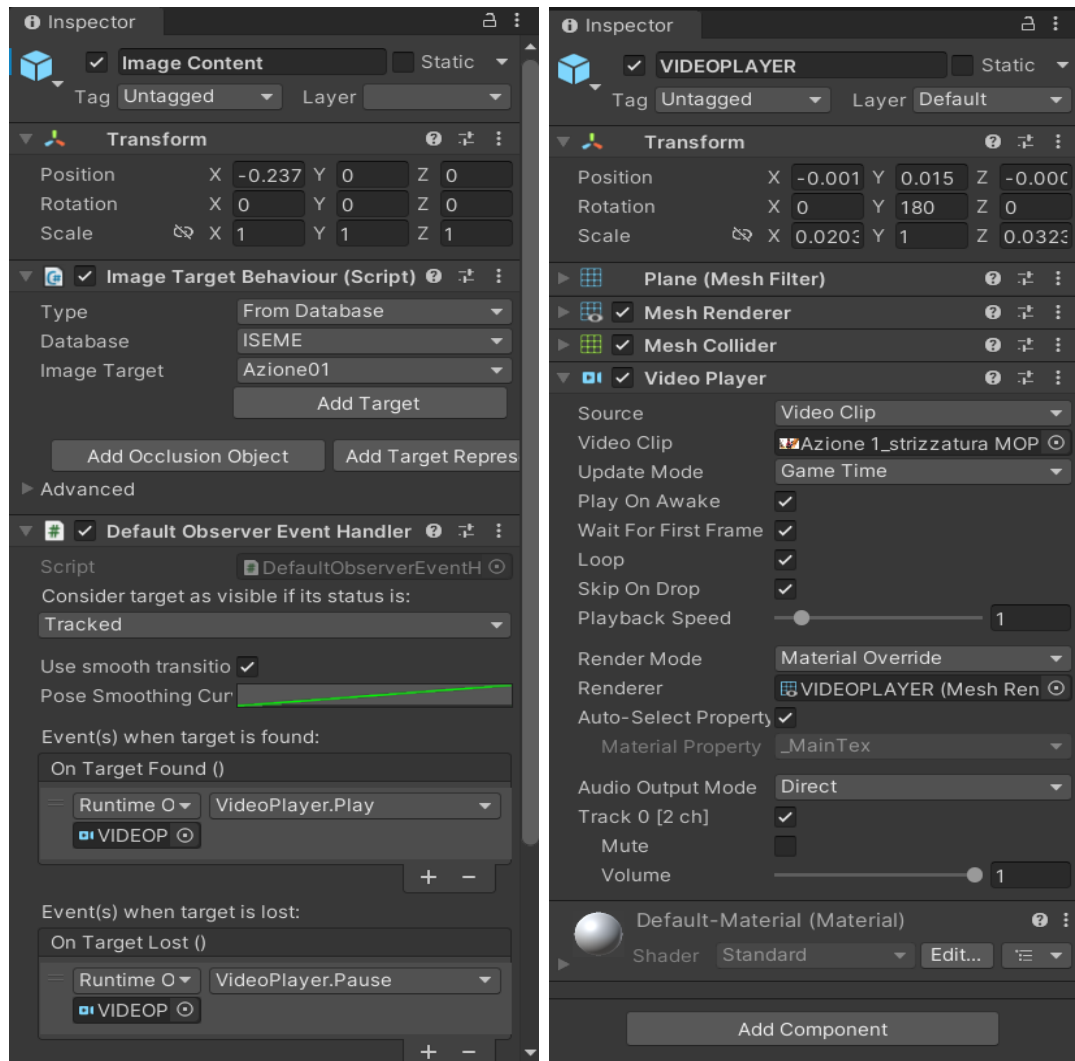


Figure 5.1: Inspector Tab Image Content and VideoPlayer

Special Education Application Final Output

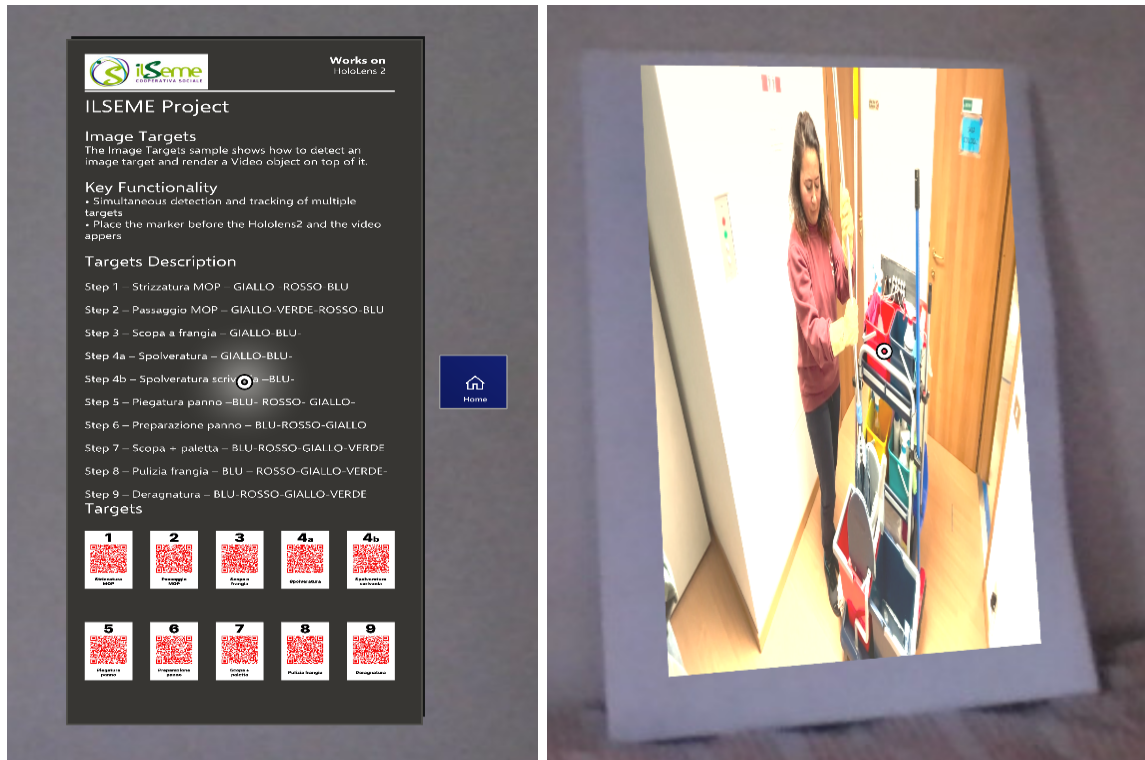


Figure 5.2: Application Menu and QR marker for playing media

5.2. Summary

The development of this application paves the way for students with Cognitive Retardation to actively engage in typical activities, allowing them to maintain their mental and financial stability for personal development. This application uses image targets to create an immersive and interactive experience that plays a video when the corresponding QR code is read by the Hololens2 thanks to the Vuforia engine and Unity. It is very beneficial to teach the fundamental cleaning techniques repeatedly until the student with cognitive disabilities masters that particular task. They are better prepared for the future as it advances both their personal development and technological literacy. These technologies can be combined to create a streamlined, user-friendly interface, ensuring accessibility for all users. Figure 5.2 provides a concise overview of the ultimate result or outcome generated by the application on the Hololens2. It encapsulates the key information or data that can be derived from using the application, presenting it in a visual format for easy understanding and interpretation.

6 | System Usability Scale (SUS)

The System Usability Scale (SUS) has been a reliable evaluation method for nearly three decades, demonstrating its trustworthiness in comparison to industry standards for assessing system usability. SUS assesses the system based on user demographics, usage purposes, and environmental context by soliciting feedback from a group of individuals who answer straightforward questions about their user experience. These questions are typically designed using a *Likert Scale*, where a set of 10 simple questionnaires is given to respondents, each offering five response options, ranging from "Strongly Agree" to "Strongly Disagree." The SUS score is subsequently computed. Typically, SUS is administered after users have interacted with the system being evaluated, without any post-interaction debriefing. It is preferable to provide immediate user feedback after system use rather than later.

The 10 questionnaires listed below are the most often used [25]:

- I think that I would like to use this system frequently
- I found the system unnecessarily complex
- I thought the system was easy to use
- I think that I would need the support of a technical person to be able to use this system
- I found the various functions in this system were well integrated
- I thought there was too much inconsistency in this system
- I would imagine that most people would learn to use this system very quickly
- I found the system very cumbersome to use
- I felt very confident using the system
- I needed to learn a lot of things before I could get going with this system

System Usability Scale (SUS) Score and Calculation

Step 1: Convert the received responses to numbers: Once all the responses are collected, these responses are converted into a numerical point system using the following point scale,

- Strongly Disagree: 1 point
- Disagree: 2 points
- Neutral: 3 points
- Agree: 4 points
- Strongly Agree: 5 points

Step 2: Calculate SUS Score: After converting the responses into numerical points, all the statements are divided into odd-numbers and even-numbers for calculations.

$x = (\text{Summation of all the points of Odd numbered questions}) - 5$

$y = (25 - (\text{Summation of all the points of Even numbered questions}))$

Once the x and y values are calculated, the final SUS score can be calculated by the following formula,

$$SUS_{score} = (x + y) * 2.5 \text{ [20]} \quad (6.1)$$

Step 3: Interpreting the SUS Score: The Overall SUS score indicates the system's usability by assessing the responses collected from a single user. Each response results in an SUS score within the range of 0-100; Note that it should not be interpreted as a percentage or percentile. For instance, if the SUS score is 68, it signifies that the response falls at the 50th percentile. The subsequent Table 6.1 provides a general guideline for interpreting SUS scores.

SUS Score	Grade	Adjective Rating
> 80.3	A	Excellent
68 – 80.3	B	Good
68	C	Okay
51 – 68	D	Poor
< 51	F	Awful

Table 6.1: Guidelines for interpretation of SUS score [21]

Advantages and Disadvantages of SUS

- Since it uses only 10 Definitive questions, it is easy for the respondents to respond quickly.
- It is more versatile and applicable to different hardware, software products or web-sites, etc.
- It is inexpensive compared to other solutions and has the ability to evaluate user satisfaction.
- SUS cannot be able to provide accurate information about products in weakness.
- Also, it is impossible to compare two systems systematically using SUS.

6.1. SUS data Collected from FIMI Machinery

The application developed in Chapter 3 is known as the Remote application and is specifically created for FIMI machinery. Its primary purpose is to enable FIMI's clients to receive assistance with small maintenance issues remotely. This is achieved by utilizing the Hololens2 technology. The detailed functionalities of this application and its potential benefits have been extensively discussed in the aforementioned chapters. Additionally, the System Usability Scale (SUS) data, which was collected from FIMI employees who had tested the system, is presented in Figure 6.1.

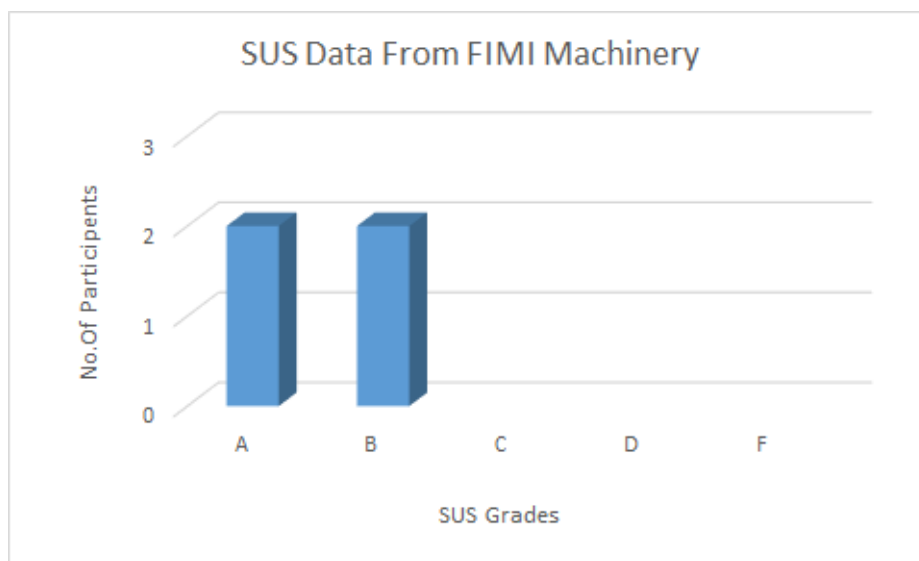


Figure 6.1: SUS data from FIMI Machinery

The application developed underwent testing by the company's employees to evaluate its functionality. Their immediate feedback was collected using the System Usability Scale (SUS). Analysis of the SUS data from Figure 6.1 reveals that all the individuals who tested the system rated it as either **Excellent** or **Good**. None of them found it difficult, as indicated by the SUS grades obtained from their feedback. Therefore, it is evident that this system holds significant potential in the industrial field, especially with further improvements.

6.2. SUS data Collected from students of Politecnico Di Milano

The Hololens2 application developed in Chapter 4 provides an interactive virtual demo instruction on operating the milling machine available in a campus laboratory. A select group of **Politecnico Di Milano** students evaluated the app's usability using the System Usability Scale (SUS), and their feedback is visualized in the accompanying graph shown in Figure 6.2.

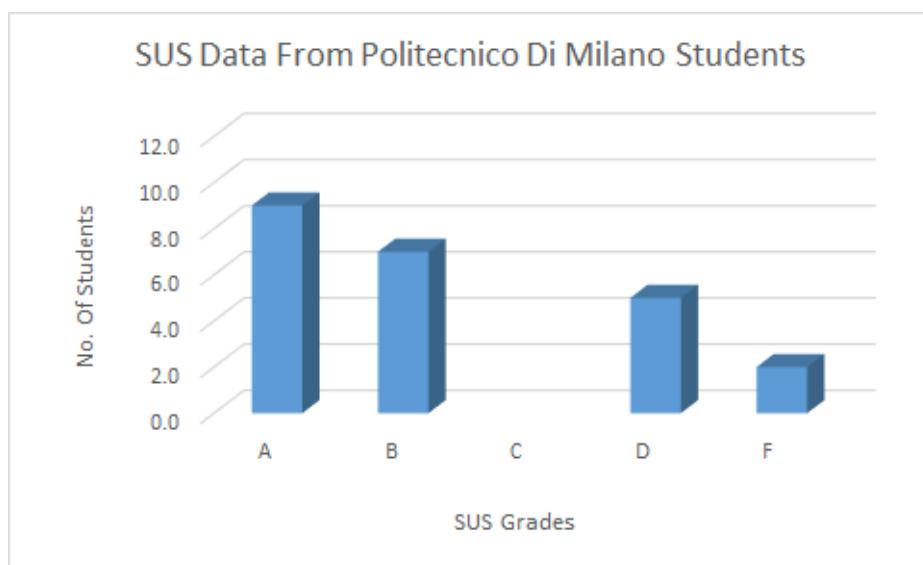


Figure 6.2: SUS data from Politecnico Di Milano students

According to the data presented in Figure 6.2, it can be observed that out of the 24 students who tested the system, only two individuals found it to be very difficult to use. However, a significant majority of approximately seventy to eighty percent of the students expressed that the system is either **Good** or **Excellent**. These findings indicate the high quality and usability of the system, particularly considering that they were collected after the participants' initial attempt at using it. It is important to note that with further practice and familiarity, these results have the potential to improve even more in the future.

6.3. SUS data collected from ASPOC students

The mentioned application in Chapter 4 serves a dual purpose. Firstly, it is developed for the students of Politecnico di Milano, and secondly, it aims to assist the ASPOC association in Lecco. The primary objective of this application is to support the teaching process of the students from ASPOC association and enhance the learning experience for students with Cognitive Retardation. To evaluate the effectiveness of the application, System Usability Scale (SUS) data was collected from the students with Cognitive Retardation after they had an opportunity to try out the Hololens2. The collected data is then analyzed and presented in a graphical format, as depicted in Figure 6.3.

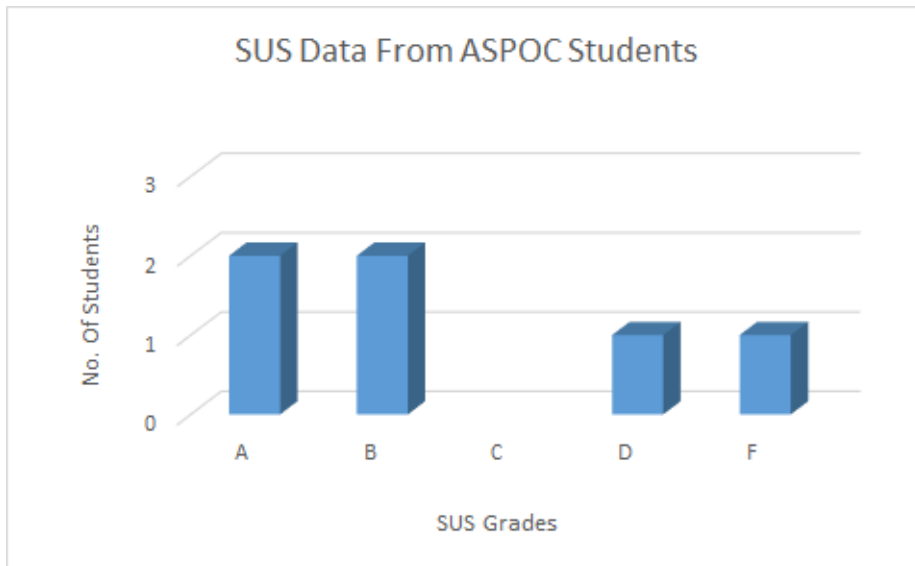


Figure 6.3: SUS data from ASPOC students

According to the data presented in Figure 6.3, it is evident that among the students who had the opportunity to try the Hololens, their opinions are divided. Specifically, 50 percent of the students reported having a positive experience, while the remaining 50 percent expressed a negative perception. This division in opinions can be attributed to the fact that the students who belong to the ASPOC association suffer from varying degrees of congenital retardation. As a result, their ability to comprehend and perform tasks is affected, making it challenging for some students to effectively use the application. However, it is important to note that with consistent practice, their proficiency in using the application can be enhanced.

6.4. SUS data collected from ILSEME Association

The application mentioned in Chapter 5 of the aforementioned source is specifically designed for the students who are part of the ILSEME association. Its main purpose is to provide training and assistance in finding small job opportunities for these individuals. A detailed explanation of this application has already been provided in the preceding chapters. The ILSEME association is located in the Como region of Italy and prioritizes training in fundamental activities such as cleaning, mopping, and maintaining cleanliness within the company or office premises. By entrusting these responsibilities to the students/pupils, the association aims to cultivate a sense of ownership and responsibility in them.

In order to gather feedback on the application, the instructors at the ILSEME association were provided with a Hololens2 device. The System Usability Scale (SUS) was then utilized to collect feedback from these instructors. The collected data is analyzed and the results are presented in a graph, as shown in Figure 6.4.

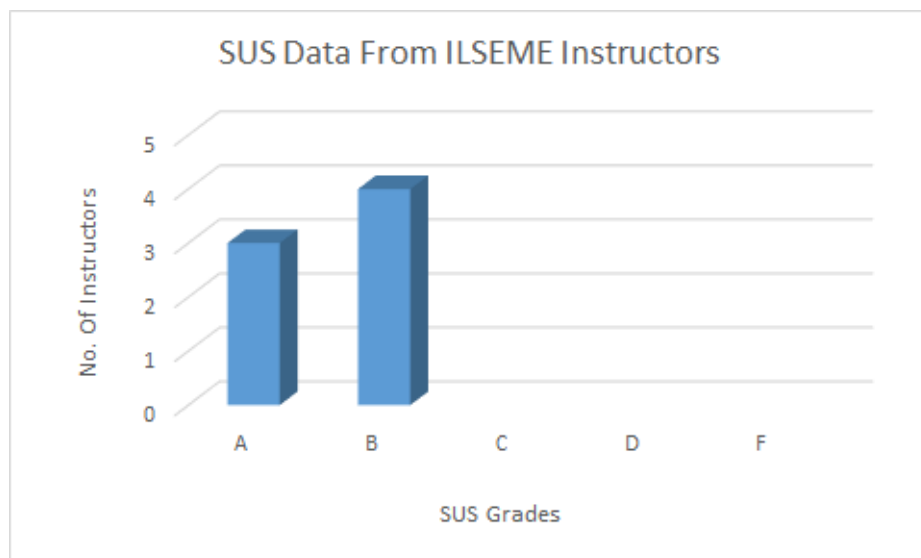


Figure 6.4: SUS data from ILSEME instructors

According to Figure 6.4, it is evident that all the SUS scores fall within the range of Grades A or B, which correspond to the categories of **Excellent** and **Good** in the SUS score system. This outcome strongly indicates that a majority of the instructors affiliated with the ILSEME association perceive this application as beneficial in aiding them with their regular teaching responsibilities.

7 | Conclusions

The Employment of Hololens2 in the industrial sector is crucial in all aspects, from time constraints to budget constraints. With certain modifications in accordance with their needs, the SUS from industry workers provides an insight into how successfully the application will assist them. The utilization of this cutting-edge technology in numerous facets of the industrial sector is made possible by this application. Next in order, Hololens2 also has a significant impact on the educational sector. where it provides both regular students and pupils with impairments or cognitive retardation with an interactive learning experience. Students now have their own environment in which to learn, imitate, and advance both academically and technologically thanks to the utilization of these tools. The SUS from students with cognitive retardation shows that they have some trouble using mixed reality applications, though this difficulty can be improved with time and practice, whereas students without cognitive retardation find learning in an immersive environment simple and helpful. Overall, Hololens2 is a promising tool for enhancing education and technology in various sectors.

8 | Future Improvements

Mixed Reality has gained significant traction in recent years due to the involvement of tech industry leaders like Microsoft, Google, Apple, and others. These companies have invested substantial resources in research and development to meet the growing demand for AR/VR devices across various sectors. These substantial investments in research and development by different companies are expected to yield more advanced AR/VR headsets that are compact, lightweight, and offer enhanced user experiences. With this in consideration, the following are potential enhancements that can be applied to the applications discussed in Chapters 3, 4, and 5.

Remote Maintenance Application:

The application outlined in Chapter 3 focuses on creating remote video conferencing between the HoloLens and other systems. However, one of its primary drawbacks is its limitation to one-to-one communication, and it lacks network security. Future developments could address these issues by implementing secure protocols such as (*https:*) and introducing additional features like multi-conferencing and screen-sharing capabilities. Incorporating screen-sharing functionality would provide an additional advantage for future enhancements.

Education Application:

To enhance the application described in Chapter 4, there is room for improvement by incorporating additional features such as voice commands and tutorials, teacher control, and new interactive functionalities. These additions could enable virtual assembly and disassembly of machines, further enhancing the overall simulation experience.

Special Education Application:

As part of future enhancements of this Chapter 5 regarding the utilization of Image targets from the Vuforia engine, Object tracking can be introduced. This feature would activate when the HoloLens2 cameras successfully track a designated object. It would then provide instant guidance through text and voice commands, along with visualizations, assisting the user through each step until they advance to the next stage.

Bibliography

- [1] Source of the 6dof image used. <https://mixedrealitynow.com/what-is-6dof>.
- [2] Source of aspoc association. <http://www.aspoc.it/>.
- [3] Source of the augmented reality image used. <https://www.ptc.com/en/blogs/cad/try-augmented-reality-with-creo4>.
- [4] Source of challenges faced by students with cognitive retardation in education programs. www.aaidd.org.
- [5] Source of the mixed reality branching. <https://www.interaction-design.org/literature/article/beyond-ar-vs-vr-what-is-the-difference-between-ar-vs-mr-vs-vr-vs-xr>.
- [6] Source of the eye tracking features image used. <https://learn.microsoft.com/en-us/windows/mixed-reality/design/comfort>.
- [7] Source of the hololens2 exploded view image used. <https://learn.microsoft.com/en-us/hololens/hololens2-hardware>.
- [8] Source of the hololens2 interaction features images used. <https://www.plainconcepts.com/microsoft-hololens/>.
- [9] Source of ilseme association. <https://www.ilseme.info/>.
- [10] Source of implementation Of rpc mechanism. <https://www.geeksforgeeks.org/remote-procedure-call-rpc-in-operating-system/>.
- [11] Source of mcu communication architecture image. <https://www.liveswitch.io/blog/scaling-virtual-event-platforms-with-multipoint-control-units>.
- [12] Source of the oculus quest2 image used. <https://www.oculus.com/quest/refurbished/>.
- [13] Source of peer-peer communication architecture image. <https://www.enjoyalgorithms.com/blog/peer-to-peer-networks>.
- [14] Source of sfu communication architecture image. <https://www.frozenmountain.com/hubfs/5%20-%20pdfs/frozenmountain-sfu-selective-forwarding-whitepaper.pdf>.
- [15] Source of socket architecture and pipes mechanism image. <https://www.tutorialspoint.com/operating-systems-client-server-communication>.
- [16] Source of students with congenital retardation. www.nichd.nih.gov.

- [17] source of instructions about node-dss signaler. <https://microsoft.github.io/MixedReality-WebRTC/manual/unity/helloworld-unity-signaler.html>.
- [18] Source of vuforiaengine description used. <https://developer.vuforia.com/>.
- [19] Source of node-dss server installation files and instructions. <https://github.com/bengreenier/node-dss>.
- [20] Source of sus formulae. <https://www.questionpro.com/blog/system-usability-scale/>.
- [21] Source of sus score table. <https://uiuxtrend.com/measuring-system-usability-scale-sus/>.
- [22] A. Almufarreh. Exploring the potential of mixed reality in enhancing student learning experience and academic performance: An empirical study. *Systems*, 11(6), 2023. ISSN 2079-8954. URL <https://www.mdpi.com/2079-8954/11/6/292>.
- [23] A. D. Birrell and B. J. Nelson. Implementing remote procedure calls. *ACM Transactions on Computer Systems (TOCS)*, 2(1):39–59, 1984.
- [24] E. Brands and G. Karagiannis. Taxonomy of p2p applications. In *2009 IEEE Globecom Workshops*, pages 1–8. IEEE, 2009.
- [25] J. Brooke. Sus: A quick and dirty usability scale. *Usability Eval. Ind.*, 189, 11 1995.
- [26] N. M. Edan, A. Al-Sherbaz, and S. Turner. Design and evaluation of browser-to-browser video conferencing in webrtc. In *2017 Global Information Infrastructure and Networking Symposium (GIIS)*, pages 75–78. IEEE, 2017.
- [27] E. A. Emmanuel and B. D. Dirting. A peer-to-peer architecture for real-time communication using webrtc. *J Multidiscipl Eng Sci Stud (JMESS)*, 3(4):1671–1683, 2017.
- [28] P. S. Gasparello, G. Facenza, F. Vanni, A. Nicoletti, F. Piazza, L. Monica, S. Anastasi, A. Cristaudo, and M. Bergamasco. Use of mixed reality for the training of operators of mobile elevating work platforms with the aim of increasing the level of health and safety at work and reducing training costs. *Frontiers in Virtual Reality*, 3, 2022. doi: 10.3389/frvir.2022.1034500.
- [29] B. Grozev, L. Marinov, V. Singh, and E. Iovv. Last n: relevance-based selectivity for forwarding video in multimedia conferences. In *Proceedings of the 25th ACM Workshop on Network and Operating Systems Support for Digital Audio and Video*, pages 19–24, 2015.
- [30] S. Knopp, P. Klimant, R. Schaffrath, E. Voigt, R. Fritzsche, and C. Allmacher. Hololens ar - using vuforia-based marker tracking together with text recognition in an assembly scenario. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pages 63–64, 2019. doi: 10.1109/ISMAR-Adjunct.2019.00030.
- [31] R. Masoni, F. Ferrise, M. Bordegoni, M. Gattullo, A. E. Uva, M. Fiorentino, E. Carrabba, and M. Di Donato. Supporting remote maintenance in industry 4.0 through augmented reality. *Procedia manufacturing*, 11:1296–1302, 2017.
- [32] H. S. Oluwatosin. Client-server model. *IOSR Journal of Computer Engineering*, 16(1):67–71, 2014.

- [33] W. Wang and L. Mei. A design of multimedia conferencing system based on webrtc technology. In *2017 8th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 148–153. IEEE, 2017.
- [34] M. Willebeek-LeMair, D. D. Kandlur, and Z.-Y. Shae. On multipoint control units for videoconferencing. In *Proceedings of 19th conference on local computer networks*, pages 356–364. IEEE, 1994.

List of Figures

1.1	Mixed Reality branching [5]	4
1.2	Extended reality (XR) [5]	5
1.3	Augmented Reality (AR) [3]	5
1.4	Oculus Quest 2 [12]	6
2.1	Spatial Mapping [8]	10
2.2	Natural Interaction [8]	10
2.3	Spatial Sound [8]	11
2.4	Voice Recognition [8]	11
2.5	Eye Tracking [6]	12
2.6	6 DoF in Hololens2 [1]	12
2.7	Hololens2 Exploded View [7]	13
2.8	Hololens2 Front View [7]	13
3.1	Client-server Architecture [32]	17
3.2	Socket Architecture [15]	18
3.3	Structure of RPC Mechanism [10]	19
3.4	Mechanism of Multi-directional communication using Pipes [15]	20
3.5	Peer-Peer Communication Architecture [13]	23
3.6	Hybrid Peer-Peer Communication Architecture [24]	25
3.7	MCU Communication Architecture [11]	26
3.8	SFU Communication Architecture [14]	28
3.9	Signalling Server Mechanism [26]	30
3.10	Command to install the server [19]	31
3.11	Command to start the server [19]	31
3.12	Running NodeDss server	32
3.13	SDP message offered by Hololens2	33
3.14	SDP message offered by Laptop	33
3.15	Peer Connection component	36
3.16	NodeDss Signaler component	37

3.17	Microphone component	38
3.18	Webcam Source component	39
3.19	Audio Renderer Component	40
3.20	Video Receiver Component	41
3.21	LocalPlayer and RemotePlayer	42
3.22	Local and Remote Video Stream display	42
3.23	Configuring CreateOfferButton	43
3.24	Configuring GameObjects for HandTracking Capabilities	44
3.25	Animation 01 and Drive B1	45
3.26	FIMI Scene01 and Back Button	46
3.27	FIMI Scene01 Button and Nav.toFIMI Scene01	47
3.28	Final Output	48
4.1	Inspector Tab for TextMesh-Pro and Image	55
4.2	Inspector Tab Animation GameObject and Timeline	57
4.3	Timeline Tab	57
4.4	Customized Back Button Script	59
4.5	Customized Pause Button Script	59
4.6	Customized Replay Button Script	60
4.7	Hololens2 view	60
4.8	Hololens2 view	61
5.1	Inspector Tab Image Content and VideoPlayer	67
5.2	Application Menu and QR marker for playing media	68
6.1	SUS data from FIMI Machinery	72
6.2	SUS data from Politecnico Di Milano students	73
6.3	SUS data from ASPOC students	74
6.4	SUS data from ILSEME instructors	75

List of Tables

2.1	Technical Data of HoloLens2	14
6.1	Guidelines for interpretation of SUS score [21]	71

