Bachelor of Science Thesis

DEVELOPMENTS AND LIMITATIONS OF THE METAVERSE FOR ENGAGING AUGMENTED REALITY VIDEO CONFERENCING

Georgios Aristodemou

UNIVERSITY OF CYPRUS

COMPUTER SCIENCE DEPARTMENT

May 2023

UNIVERSITY OF CYPRUS

COMPUTER SCIENCE DEPARTMENT

DEVELOPMENTS AND LIMITATIONS OF THE METAVERSE FOR ENGAGING AUGMENTED REALITY VIDEO CONFERENCING

Georgios Aristodemou

Research Supervisor

Dr. Andreas Aristidou

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science at the University of Cyprus

May 2023

Acknowledgments

First, I would like to thank my advisor and supervisor Professor Andreas Aristeidou for the guidance he provided while conducting this research. His knowledge and support were essential during this study, and he was always available when I needed his help.

Additionally, I would like to thank Marios Charalambous of CYENS, for his help in the creation process of the Metahuman Avatar.

Lastly, I want to thank my family for their support during the conduction of this dissertation and throughout my studies.

Abstract

The Metaverse as a concept is becoming more and more relevant in recent years, entering the public consciousness and with many technologies being developed to support it. Metaverse technologies are recent, still in development and proofs of concept to demonstrate the value and potential the tech can provide to users in the future. In this dissertation we've created a simple proof of concept demo for an Augmented Reality video conferencing experience by creating a Metahuman Digital Sibling Avatar, giving it simple face tracking and inserting it in Augmented Reality. Through the process of creating this demo we have determined some challenges and limitations of creating Mixed Reality technologies for the Metaverse. These challenges and limitations will need to be considered in the development of future technologies to optimize them so they can be as useful and valuable as possible for their users. In this dissertation we discuss these challenges while showcasing the process of developing our proof of concept demo, opening the way for future work in the field to improve and optimize Metaverse and Mixed Reality applications.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

Humanity has been interested in exploring and understanding the virtual world for a long time. In recent years, with the Covid-19 pandemic making online video conference calls and remote work more common, it's becoming increasingly important to make the experience of digital work and communication as real and engaging as possible. Virtual Reality technology has been developing rapidly, with the concept of the Metaverse [7, 23] entering the public consciousness and gaining interest from developers and researchers. Currently, the use of the Metaverse to simulate reality in the digital world is still being tested as a concept, with companies like Meta and Epic Games [6, 30] making a lot of progress in improving the technology. With Unreal Engine's 5 new feature Metahuman Creator, it's possible to create a Metahuman Avatar that looks very realistic, and a Digital Twin of the person creating it. With this accessible tool for creating human digital twin avatars, our aim was to create a simple demo of an Augmented Reality online conferencing experience using real time face tracking on a Metahuman Avatar. The goal is to use this demo as a proof of concept to show to users to evaluate whether this would be a desirable, realistic and engaging experience for users. Advancements in technology must be researched and interrogated to determine the value they bring to the users, to determine how they can be improved and optimized.

1.2 Problem Statement

Metaverse is becoming a more relevant term recently in the tech industry. Businesses are pivoting to developing applications explicitly for the Metaverse, and the concept is entering public consciousness and understanding. The goal of any developing technology is to offer a user-friendly experience and provide value. For this reason, it is important to understand how Metaverse virtual experiences can be engaging for users, as well as the inherent limitations of the Metaverse as a concept.

1.3 Contributions

While conducting this dissertation, we created a simple proof of concept demo of an Augmented Reality video conferencing experience, by creating a Digital Twin Metahuman Avatar and placing it in Augmented Reality with simple Face Tracking applied.

In the process of creating this demo, we've been able to identify the potential problems and limitations the development of this technology might face, as well as any potential value it might have for users. These will be discussed in the Discussion section of this paper.

Chapter 2

Literature

2.1 Metaverse and Metahuman

The term ‘Metaverse’ was coined in Neal Stephenson's 1992 Science Fiction novel “Snow Crash.” [15] There is no universally agreed upon definition of the word yet, but different interpretations of the word retain the same concept. The Metaverse is a virtual world, but that by itself is insufficient to describe it.

Much like the internet, the Metaverse is a virtual universe that allows users to access services, entertainment and socialization. However, the Metaverse goes a step further than the Internet, as it attempts to bridge the gap between our physical reality and the virtual one [2, 20]. In that sense, the Metaverse can be thought of as a collection of virtual “places” that can be accessed from the physical world through technological devices. The Internet can be thought of as part of the Metaverse accessed by a screen. A game world is its own virtual world that can also be included in the Metaverse, accessed by computer or game system. The concept is still broad, and its potential has become a popular topic among researchers and developers, some are joining in the hype [26] while others are skeptical.

Figure 2.1 A Promotional Image showcasing what is believed to be the potential of the Metaverse, a complete and robust virtual world. © Trace Network Labs

Mixed Reality technologies have been advancing at a rapid pace and have been bridging the gap between the virtual and physical worlds. With Virtual Reality, we can explore the virtual world like it's a physical space. With Augmented Reality we can enhance the physical world with virtual elements, blending the digital and the physical. With these technologies we can explore the Metaverse and create various experiences within it, and they have boundless potential and present unique challenges and considerations [9, 11, 17].

When using these technologies to explore the virtual world of the Metaverse, we surpass and transcend the presence of our physical bodies. Metahumans are the digital representations of ourselves we use to explore the Metaverse, and are the latest development in the advancement of Digital Avatars.

2.2 Digital Avatars

Since the creation of Video Games in the 1970s and 1980s, the player controls a player character to do actions inside the game. In the early days of games, the playable characters were simple pixelated icons or cursors. As gaming technology advanced, the playable characters became more complex, and were able to represent human characters more closely and give them more actions to do. The term “Avatar” to refer to a digital representation of a human player was coined by Richard Garriot for his 1985 game ‘Ultima IV: Quest of the Avatar’. Ultima was a series of Role-Playing Games, where the player took part in a story by making choices and affecting the narrative. For Ultima IV, Garriot wanted the player character to be a representation of the player’s real-world self in the digital world so that the player felt more responsible for the ethical and moral choices they made in the game. The word “Avatar” is of Hindu origin and refers to the manifestation of a deity on earth as a physical form, so he applied it for the manifestation of the player inside the digital world [22]. The term was further popularized in Neal Stephenson’s 1992 cyberpunk novel Snow Crash, in which it refers to a human’s virtual representation in the Metaverse [12]. As seen in Figure 2.2, the Avatar in Ultima IV is still crude and pixelated, yet it remains one of the first attempts to have the pixelated player character be a one-to-one representation of the human player controlling it, making it an important development in Digital Avatars and gaming.

Figure 2.2 A screenshot of Richard Garriot's 1985 game "Ultima IV: Quest of the Avatar" the game that coined the term Avatar as a digital representation of the player. The pixelated avatar can be seen on the left holding a sword and shield, with information about the avatar's location, actions, and status on the right side of the screen. © Origin Systems

Since the origins of the term, 'Avatar' refers to a playable character in a game that is meant to be a representation of a human player inside the virtual world of a game. Not every playable character in a game is an Avatar. Sometimes the playable characters are more defined, like when they are specific fictional characters like Super Mario, or more abstract playable objects like a cursor or spaceship. Avatars are defined by the fact that they represent the player controlling them in more ways than just being controlled by the player, and they vary in the degrees in which they represent the player. For example, games that are in a First-Person Perspective allow the player to view the world through the eyes of their avatar, which increases immersion even if the player doesn't know what the avatar looks like. The playable character can even have a specific appearance but still act as an avatar for the player. For example, in the Half Life games, the main character the player controls is Gordon Freeman, who has a specific name, appearance, and backstory. However, because the game is in First Person Perspective and Gordon Freeman never speaks, he still acts as a blank avatar for the player to insert themselves into and control. As shown in Figure 2.3 the main character of the game is never visible, so games that are in

First Person Perspective make it easy for the player to project themselves in them.

Figure 2.3 A screenshot of Valve's 1998 Computer Game "Half Life." The player character of this game is Gordon Freeman, a character that has a specific name, look, and background, however because of the game's First-Person perspective and lack of dialogue, the player can project themselves onto Gordon Freeman easily, therefore he plays the role of an Avatar. © Valve

Similarly, in 2007's Mass Effect, the player character is Commander Shepherd. However, unlike Half Life, the player gets to choose Shepherd's look, Shepherd's gender, and even their personality and how they respond in dialogue. Although they are similar in concept, Commander Shepherd in Mass Effect is an example of a video game Digital Avatar that that gives the player a lot more options and freedom in their expression, and thus is a much closer representation of them.

Figure 2.4 A screenshot of the 2007 BioWare game “Mass Effect” featuring the main character Commander Shepherd who can be customized by the player to an extent. Despite having a set name and role, their look, gender, and dialogue are chosen by the player. © BioWare

As video games advanced and more types of digital avatars were created to suit their needs, developers were slowly bridging the gap between the players and their digital avatars. Certain games did not need realistic avatars, others did not need avatars at all, but for virtual experiences meant to be more realistic and truer to life, Avatars that were closer and more realistic representations of the human players were necessary. Avatars that are meant to be as close in appearance and behavior to their human equivalents are called Digital Siblings.

2.3 Digital Twin and Digital Sibling

A Digital Twin is a one-to-one true to life representation of a human, object, or system in the virtual world [21]. A Digital Twin is meant to mirror its physical equivalent in every way possible, including appearance, physics, and behavior. An early version of the concept of the Digital Twin was introduced by NASA as early as the 1960s for the Apollo 13 mission. NASA had simulators of the systems that were in the Apollo 13 spacecraft, which allowed them to run simulations and tests before their actual launch mission. Digital Twins can be simulations and virtual versions of simple structures, but to be a true Digital Twin, it must also replicate its behavior.

Figure 2.5 A Digital Twin model of a kitchen presented side by side to its physical counterpart as presented in [21]. Digital Twins can replicate their physical equivalents with a lot of detail and fidelity in their models, even more so for inanimate objects and structures.

A Digital Twin can be a virtual simulation of a system. For example, the digital twin of a factory would replicate its look and machinery, as well as their function, schedules and processes. Such virtual simulations of factories are very useful for understanding their function and optimization, as it is easier to alter a virtual factory than a physical one. Plans, schedules, and alterations to the physical factory or workplace can be tested in advance, making them incredibly useful as simulations.

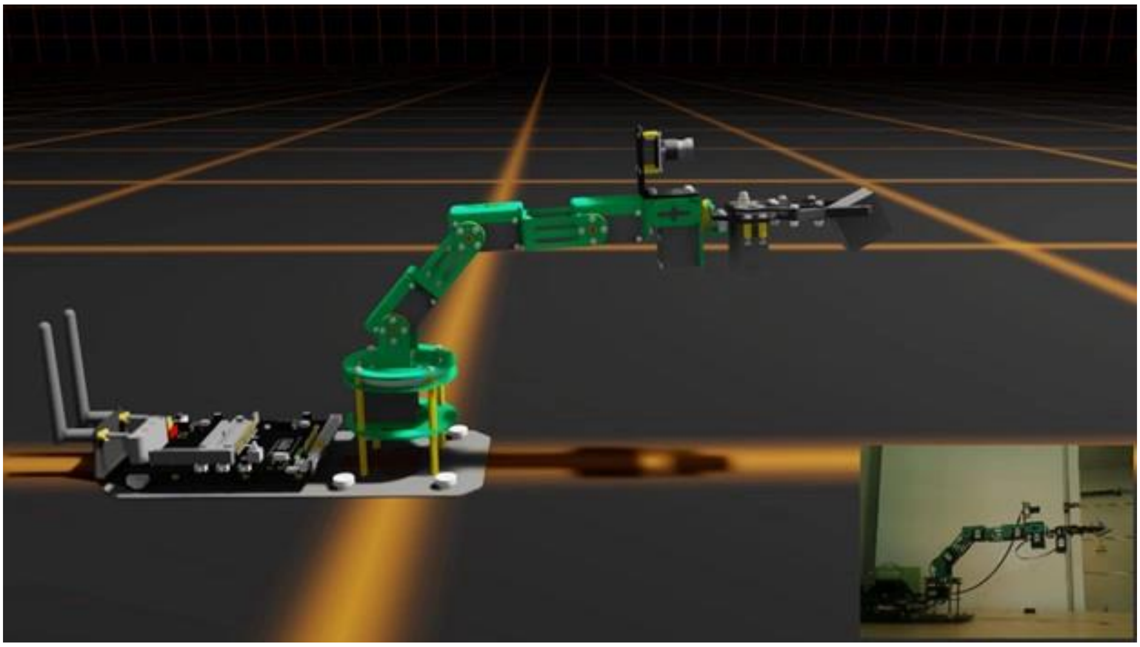


Figure 2.6 A piece of robotic factory equipment with its Digital Twin as presented in [21]. Sharing the same design and functionality allow the machines to be tested and optimized digitally.

Digital Twins can apply to networks and cities for simulation purposes. Complex systems like cities can be recreated digitally to simulate their behavior. A Digital Twin of a human is called a Digital Sibling, an avatar that is as true to life as possible of the human it is meant to represent in the digital world. For the purposes of this dissertation, we have created a Digital Sibling avatar using Blender and Unreal Engine.

Chapter 3

Creating a Metahuman Digital Sibling

3.1 Overview

We will begin with an overview of the process required to make a Metahuman Digital Sibling avatar. To create the Metahuman avatar, we used Blender and the Metahuman Creator Plugin that was introduced by Epic Games in Unreal Engine 5 [30]. For Blender we used the Facebuilder plugin which can build a face model and texture with pictures as input. The texture and model are then imported to Unreal Engine 5 where the head model has its key animation points mapped so that it may be converted and animated as a Metahuman in the Metahuman Creator. Afterwards the facial features and texture are added to the Metahuman model. Hairstyles, facial hair and clothing must be added in the Metahuman Creator from a set selection, so there are some limitations for how one-to-one the Digital Sibling representation is.

3.2 Facebuilder

Facebuilder is a plugin created by Keentools for the 3D Modeling software Blender that takes pictures of a face from different perspectives as an input and digitizes it by creating a 3D Model that attempts to match the given face. For this to work it requires multiple photos, each from different perspectives so that it can recreate every side of the face, as well as convert the 2D images to a 3D model and give it depth. The images used to generate the model used to create the Digital Sibling used in this dissertation are shown in Figure 3.1.

Figure 3.1 The set of photos used as input for Facebuilder for this dissertation. The plugin needs many perspectives to properly model the face in 3D. The first three are simple front and side perspectives. The 4th, 5th, and 6th photos are from various angles so it can accurately portray depth in 3D.

The plugin can take many pictures to increase its fidelity and accuracy, more perspectives and angles can give better and more reliable results, but they can also act as noise. No matter how many photos it takes as input, the result it gives is only an estimate of what the model of the face given looks like in 3D, and it needs to be manually corrected with the photos as reference to produce a truly accurate result. The plugin supports facial expressions, but it works best with a neutral expression. The shape points generated from each photo can either overlap or contradict each other, so they need to be manually adjusted and corrected for the head model to be correct. The head model generated by the process after being manually corrected is shown in Figure 3.2

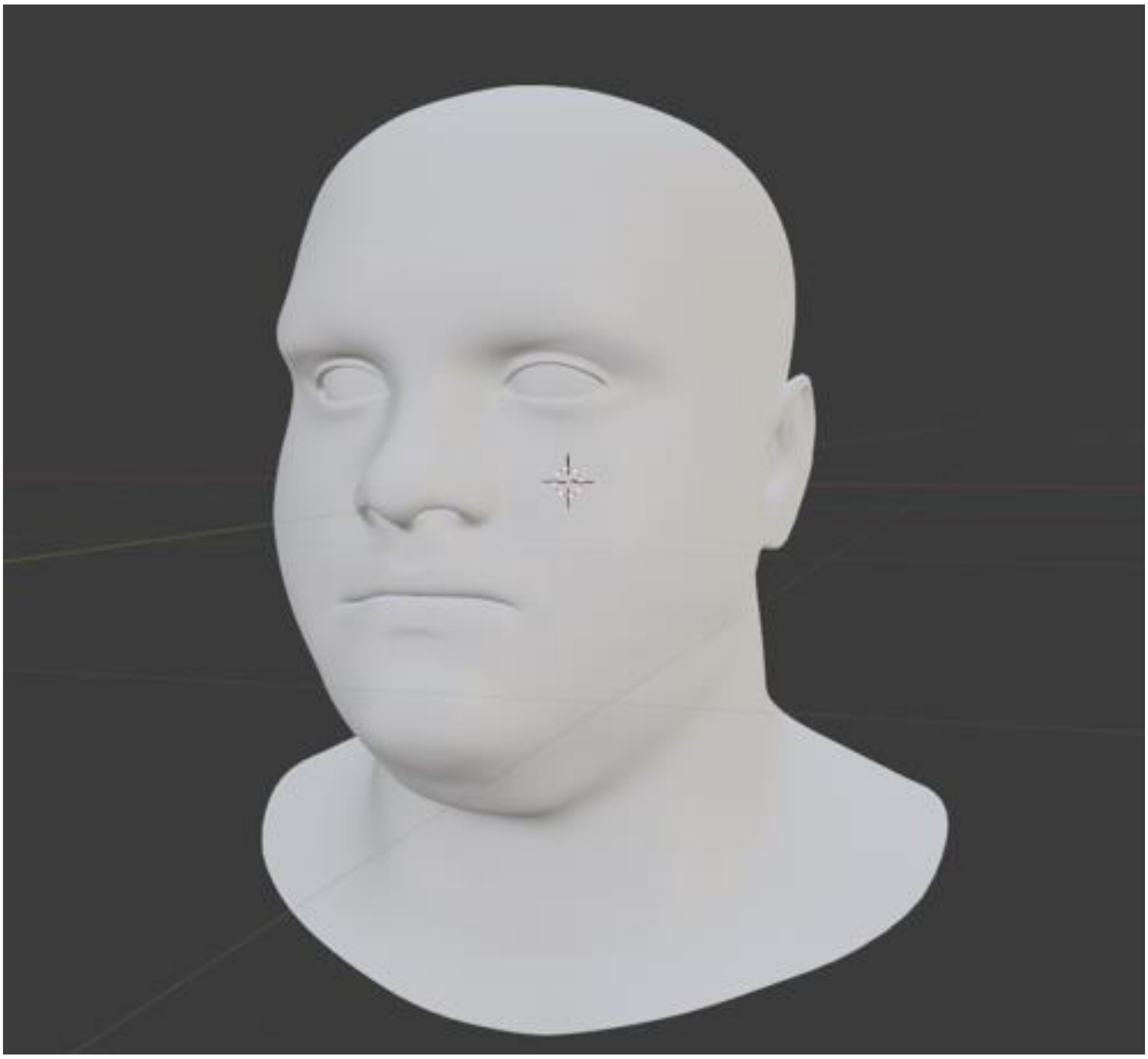


Figure 3.2 The head model generated by Facebuilder without any lighting or texture added.

After the head model is created by Facebuilder, it uses the pictures given to it to create a texture of the face and wrap it around the head model. The texture generated by the tool and the head model with the generated texture wrapped around it is shown in Figures 3.3 and 3.4 respectively.

Figure 3.3 The texture generated by Facebuilder using the pictures given as reference and input.

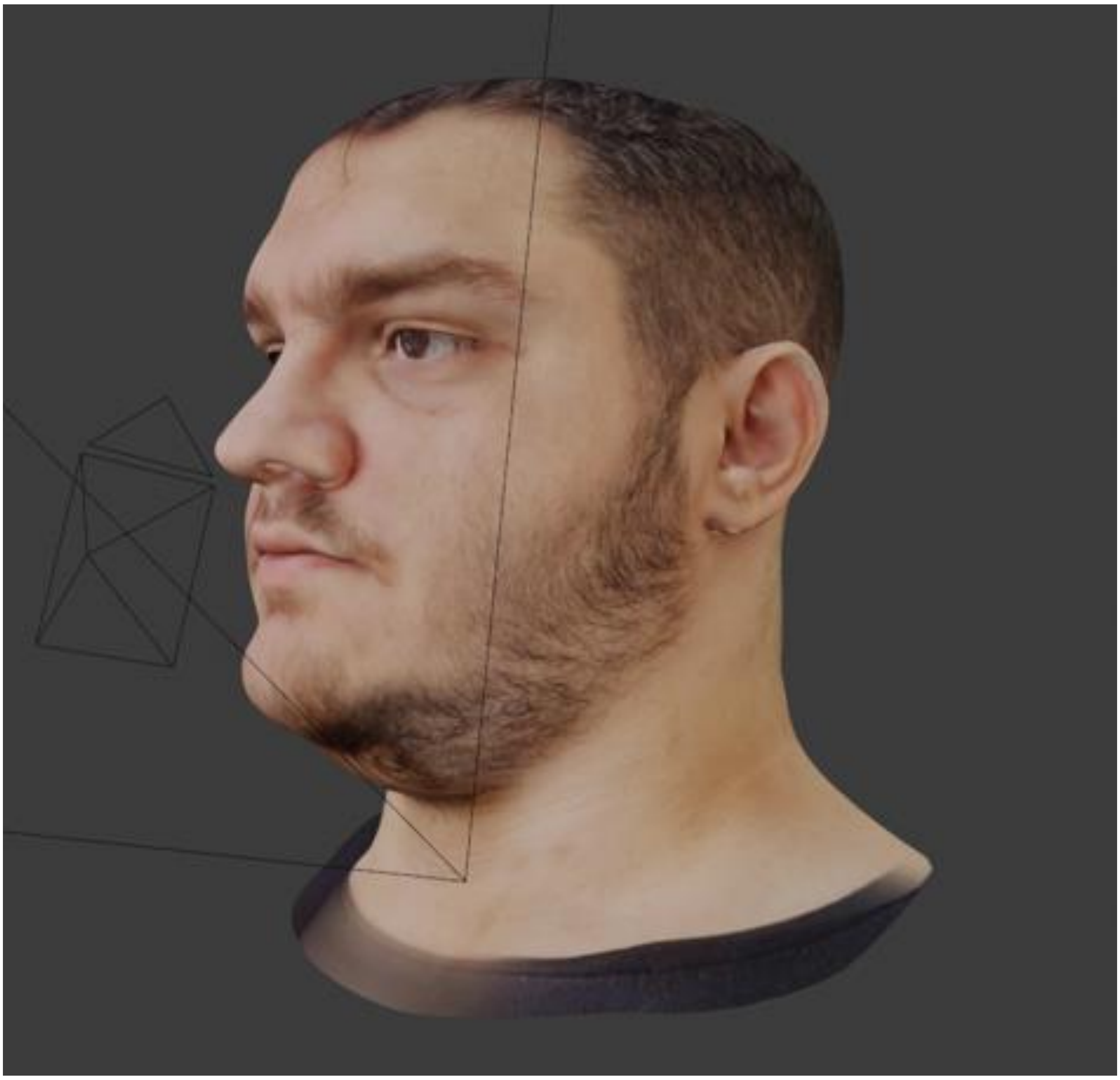


Figure 3.4 The texture generated by Facebuilder using the input pictures as reference wrapped around the head model. The texture matches the head as shown by the lips, ears and eyes matching their respective placings in the model, so this means that the 3D head model generated is reasonably accurate to the pictures.

This is the first step in the process of creating the Digital Sibling Metahuman Avatar. With this completed, Facebuilder exports both the head model and texture generated from the pictures so they can be used and converted to a Metahuman avatar in Unreal Engine 5.

3.3 Metahuman Creator

Unreal Engine 5 introduced Metahumans, incredibly realistic and high fidelity human digital avatars that can be customized and animated in a variety of ways [30]. Alongside Metahumans, the Metahuman Creator was also introduced as a tool to easily create and customize them. The Creator tool is still in development, only providing a specific number of options for hairstyles, facial hair and clothing.

Figure 3.5 A promotional image for the Metahuman Creator tool showcasing the diversity, fidelity, and realism of the avatars it can create. The facial features of these models can be customized heavily, but the bodies do not offer the same variety.

The 3D Head model and texture created with the Facebuilder plugin must be imported in Unreal Engine 5 to be converted into a Metahuman head model, so that the Metahuman Creator tool can create a Metahuman Avatar with the correct head shape and facial features. Similarly to the Facebuilder plugin, this involves importing the mesh and manually adjusting the shapepoints generated by the software until we get an acceptable result. This process is shown in the following images in Figure 3.6.

Figure 3.6 The Shape Points needed to create and animate the Metahuman created by the Metahuman Creator tool. Like Facebuilder, different perspectives are needed to create an accurate result, and the points must be corrected manually to match the model and texture.

After generating and adjusting the shapepoints for the main facial features, a model for the Metahuman Creator to use will be created. The results of that process, both with and without the texture wrapped around it are shown in Figure 3.7.

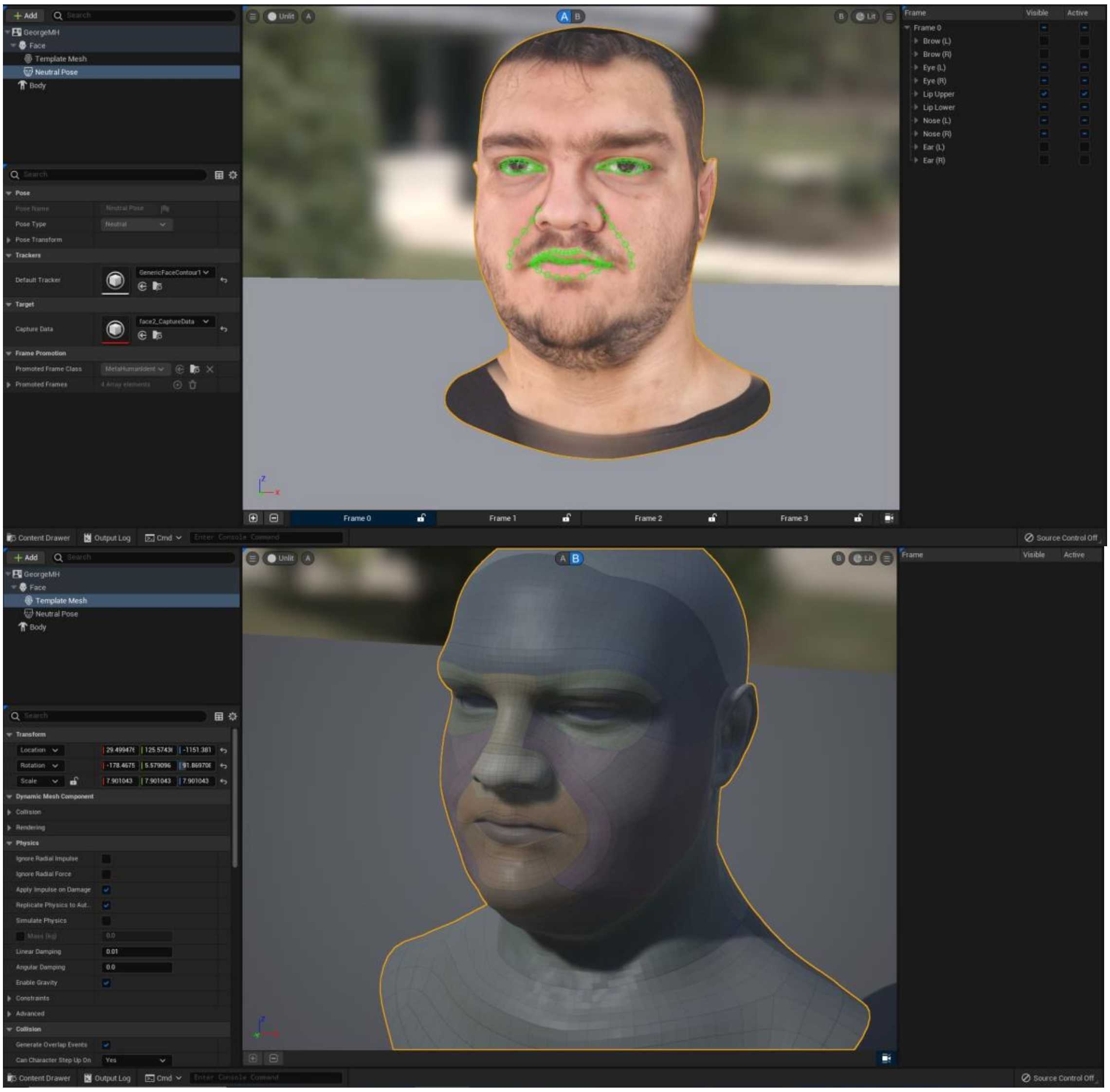
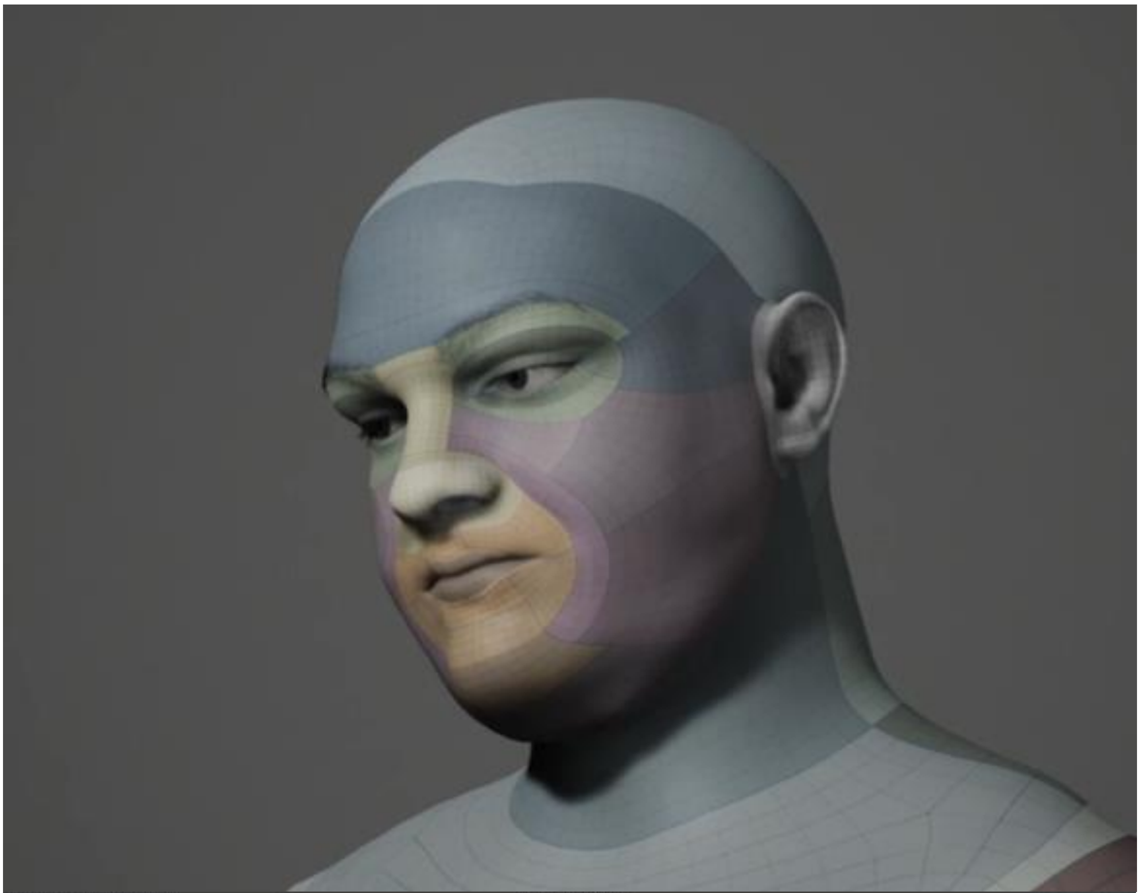


Figure 3.7 The imported texture and mesh being turned into a Metahuman head model with the generated shape points. With the above result, a Metahuman with the correct features can be created by the Metahuman Creator to be a Digital Sibling.

The Metahuman Creator converts the above result into a Metahuman avatar, giving it the correct head shape and features to be animated. The process and results of the Metahuman Creator is shown below in Figure 3.8




My MetaHumans / GeorgeMH
METAHUMAN
Studio
Auto
Medium
LOD 0
FACE
Custom Mesh
Blend
Skin
Eyes
Teeth
Makeup
HAIR
Head
Eyebrows
Eyelashes
Mustache
Beard
BODY
Proportions
Tops
Bottoms
Shoes
HOTKEY REFERENCE
This MetaHuman uses grooms currently in development, only displayed at LODs 0 & 1
Idle
Face ROM
Body ROM
Happy A
Surprise

Figure 3.8 The results of the Metahuman Creator with the model created from the Facebuilder mesh. The first image shows the Metahuman without the skin and facial features added, just the Unreal Engine model added on the avatar and animated. The second and third images are screenshots from the animation demos with the final Metahuman showing its movement. These facial features are the closest options to the pictures given, which means that with this model, head shape and body type this is the closest representation it can create, making this model a Digital Sibling.

With the Metahuman created, it can be imported to Unreal Engine 5 to be properly animated. To finish creating the Digital Sibling for the purposes of our demo, we used Faceware Studio to rig it with some simple face tracking.

3.4 Face Tracking

For the demo to work the Metahuman avatar must be animated with the movements of the person controlling it so it can play the role of a Digital Sibling. For this purpose we have used Faceware Studio to add face tracking to the avatar in Unreal Engine 5. We will outline the process of the face tracking and the result.

Faceware Studio is a motion tracking software that can track the main facial features and animate them accordingly. The result of its face tracking capabilities is shown in Figure 3.9.

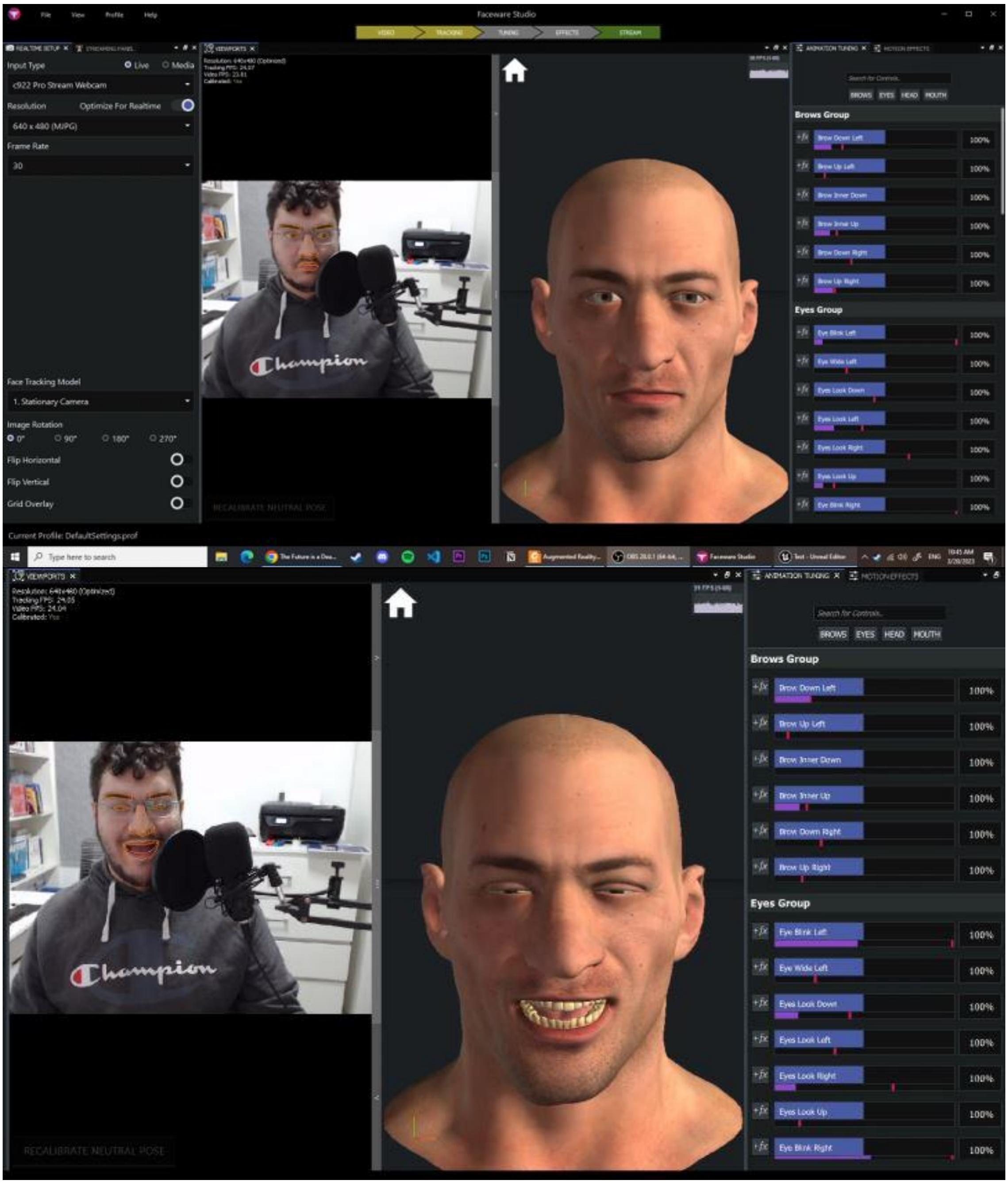


Figure 3.9 Face tracking result from Faceware Studio. The first image is a neutral pose while the second is with mouth movement to show the motion tracking.

The motion tracking must be linked with Unreal Engine 5 to be used on the Metahuman avatar we created. For this the Unreal Livelink plugin is used, which allows the movements recorded by Faceware Studio to be streamed and transferred to Unreal Engine, allowing them to be applied and used on any model that can support facial animation such as the Metahuman avatar. The Blueprint used to rig the model so that it's animated live by Faceware Studio is shown below in Figure 3.10.

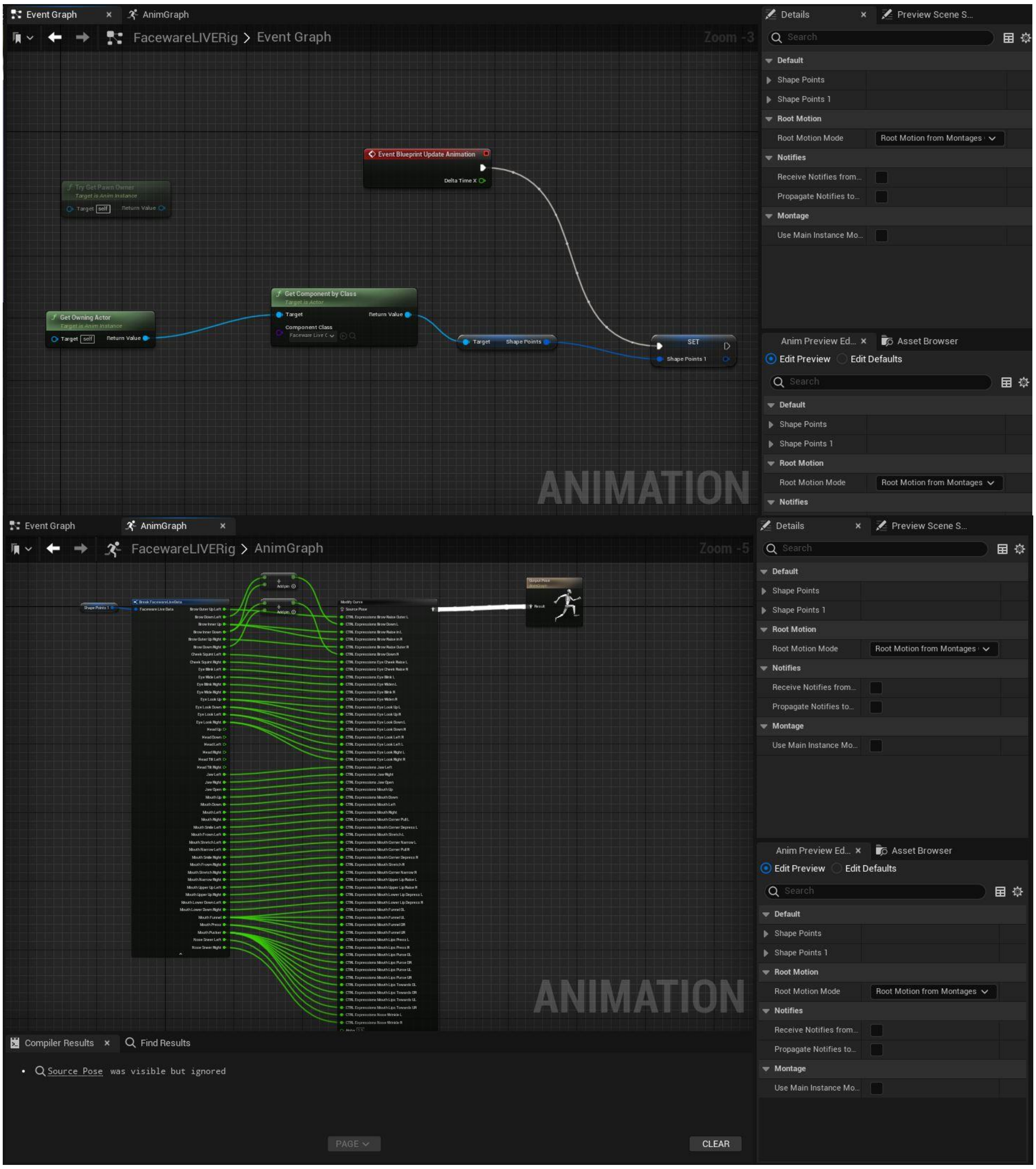


Figure 3.10 The Animation Blueprint used to rig the Metahuman avatar. The first image is the Event Graph which activates the animation. The second image is the Animation Graph which maps the shape points to the motion tracking to be animated.

With the Blueprint attached on the Metahuman avatar it can be animated with motion capture from Faceware Studio, which allows for real time lip syncing. The result of the process and the link between Faceware Studio and Unreal Engine 5 is shown in Figure 3.11.

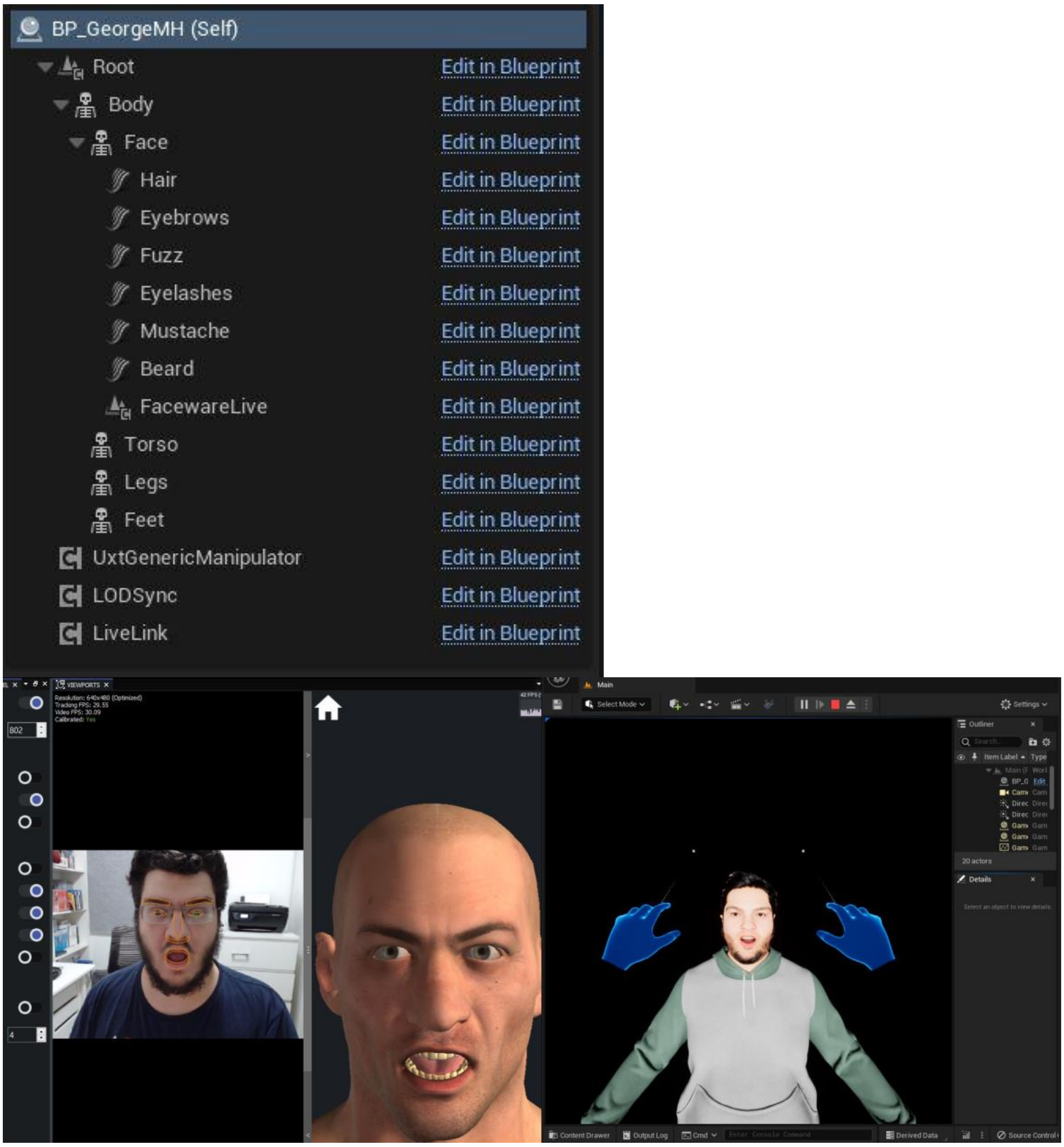


Figure 3.11 The Animation rig Blueprint is attached to the Metahuman model face skeleton as shown in the first image, which allows the Metahuman avatar to link to Faceware Studio and receive motion capture data. As shown in the second image, the Metahuman is being animated by Faceware Studio, as it has the same expression as the Faceware Studio model which tracks movements from the camera.

Chapter 4

Augmented Reality

4.1 Lighting in AR

One of the most complex issues of Augmented Reality is lighting [11,27]. Depending on the purpose of the AR application, the lighting can either prioritize visibility or realism, but the result nonetheless must blend in the scene it's in. This means that the lighting on objects that exist in Augmented Reality must match the lighting of the scenery they are placed in. Objects in Augmented Reality do not have a specific light source they are exposed to, but are instead illuminated by the lighting of the physical world. Having digital objects be lit by the light in the real world is a big challenge. An example of this is shown in Figure 4.1 below.

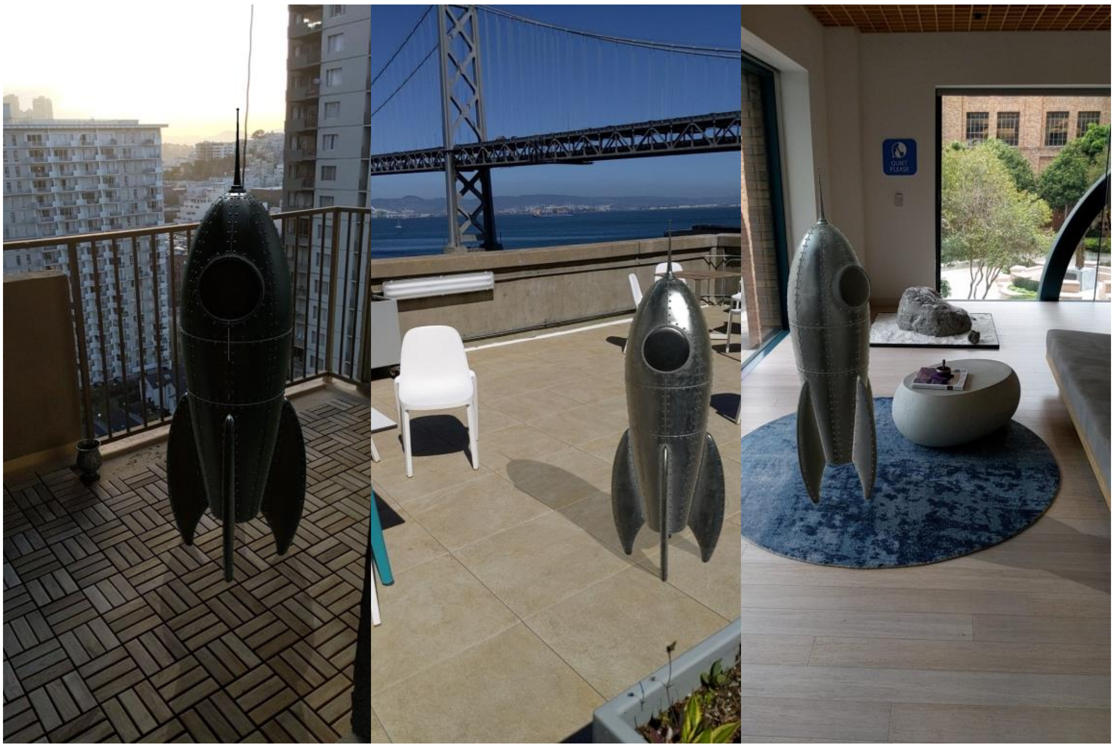


Figure 4.1 The same object placed in Augmented Reality in different lighting gives different results. Even though it is a virtual object, it must be affected by the lighting of the real world to blend in. © Google ARCore

For the purposes of the demo in this dissertation, the Metahuman model has been placed in Augmented Reality scene but will be shown off with screenshots in the Unreal Engine software and in Hololens Emulation. For this reason the environment is just a black background that acts as a placeholder for the physical world.

To prioritize visibility, a simple directional light source was placed in the scene with the Lumens lighting engine introduced in Unreal Engine 5. The Lumens Engine provides realistic environmental lighting, which will be a substitute for the real world ambient lighting that usually affect virtual Augmented Reality objects.

4.2 AR Scene

For the Augmented Reality scene we simply placed the Metahuman Digital Sibling we created, rigged with the face capture, in an AR Scene using Unreal Engine 5's Mixed Reality plugins. The scene as seen in the Unreal Engine 5 viewport is shown in Figure 4.2.

Figure 4.2 The AR scene in UE5. Without taking video of the real world from a camera, the background is shown as simply black.

To provide some simple UX and simulate interaction and the presence of a user in Augmented Reality, we used the Microsoft Mixed Reality Toolkit to add some hands that simulate the user's. Figure 4.3 shows the Blueprint of the AR hands

.

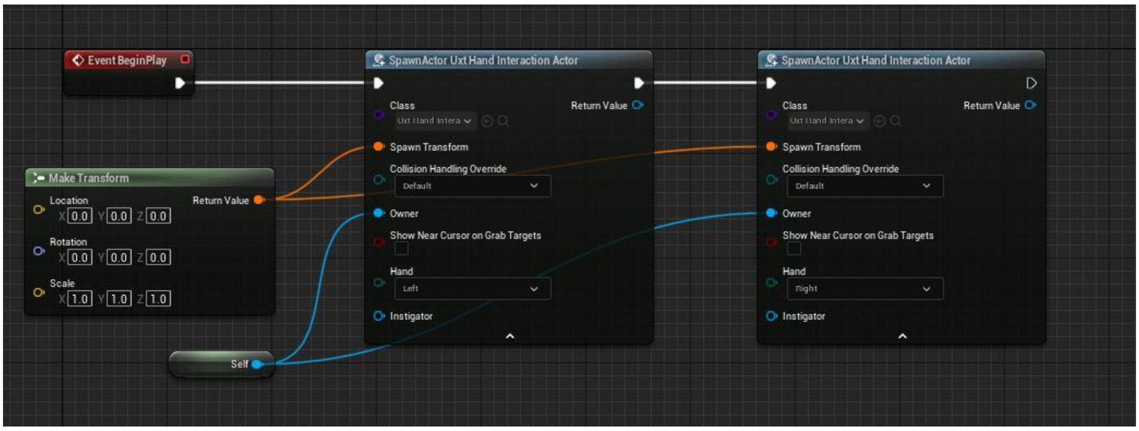


Figure 4.3 The Blueprint of the hand UX controls added with the Microsoft Mixed Reality Toolkit, an open source UI/UX library for mixed reality applications.

4.3 Output

With the AR scene, lighting, and Metahuman Digital Sibling Avatar that we discussed so far, Figure 4.4 shows the output of the scene when “Play” is pressed in Unreal Engine 5.

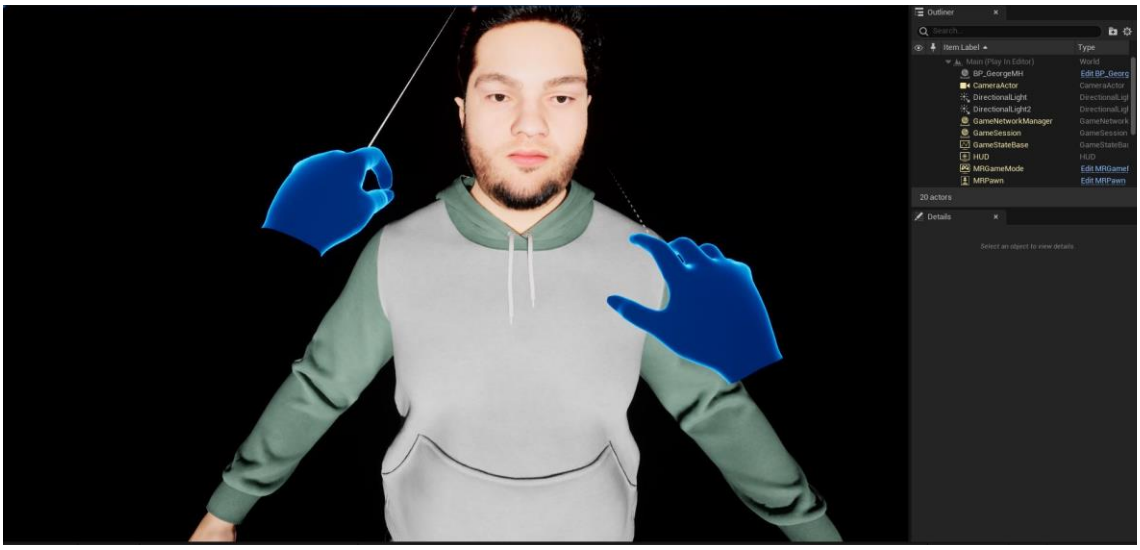


Figure 4.4 The output of the AR scene, the hands are meant to simulate the user of the Augmented Reality app, interacting with the Digital Sibling avatar in Augmented Reality.

As shown in Figure 3.11 above, if Faceware Studio is open and the motion capture data is being streamed to Unreal Engine 5, the Digital Sibling Metahuman avatar is also animated and lip synced from the camera.

Chapter 5

Discussion

5.1 Discussion

As shown in our results above, we have created a Digital Sibling Metahuman Avatar and added face tracking motion capture to it so that its motions mimic reality, thus making it a true Digital Sibling. Placing the Digital Sibling in Augmented Reality allows it to be placed in the physical world, which would allow the Digital Sibling to blend in with reality as it is made to be as realistic as possible. However, in the process of creating the Digital Sibling Avatar and the demo, the limitations of the technology and the Metaverse as a concept have become clear. In this Chapter we will discuss several topics relating to developing technologies for the Metaverse and Mixed Reality applications, as well as the nature of Digital Avatars in this context.

5.2 Relevant Questions

To understand the potential usefulness of this demo and this technology, there are a lot of questions we need to consider regarding the relationship between people and their avatars, or how people perceive the digital world [8].

A key question to ask is when realism is desirable in virtual avatars. The relationship between people and their virtual avatars can be complex [14]. How much a person relates to their avatar can depend on many factors, both physical and virtual. How much experience a person has with gaming and virtual worlds affects how much they can relate to their avatar, even more so for people who have addictive tendencies for gaming [29]. The way people use avatars and the context of the virtual worlds they are in also matter. Studies have found that people tend to use Digital Avatars to either create an idealized version of themselves, creating an Avatar that represents their ideal self, or choose an appearance that is closer to reality in order to inject a version of themselves inside a game or a virtual world [13, 28]. Games with higher levels of customization and games with higher difficulty have been linked with a stronger user-avatar relationship. The difference between a more ideal avatar or a realistic one can be subtle, and concrete data is needed to determine when each is more desirable. For example, it's easy to presume that a more grounded, true to life avatar is better for professional and realistic situations such as working from home or virtual training simulators, since in professional environments realism and accuracy are desirable traits. However, what if data shows that when people are allowed to create whatever avatar they want and operate in the virtual world as their "ideal" self they gain more confidence and their productivity increases? We need more definitive data about the relationship between users and avatars in more virtual worlds, and how that can change depending on the context of the virtual world and the task required to accomplish within it.

Another issue that needs to be considered is not just how users relate to their own avatars, but how they relate to the avatars of others. In both gaming virtual worlds and professional virtual offices or simulators, it’s necessary for a user and their respective avatar to communicate and socialize with the avatars of other users.

Figure 5.1 A virtual office as seen in [1] with many virtual avatars interacting with each other in a professional environment.

There are many questions regarding the way people perceive and relate to other avatars, and by extension other users, since we lack definitive data on the subject. Inside a virtual world, the only reference someone has for the other users is their avatar, so interaction only happens between the avatars [10]. Socialization in virtual worlds has been happening for years, for example in Massive Multiplayer Online games (MMOs) [5] and Virtual Reality social apps such as VRChat and Second Life. These apps and games have been popular since they were created, but it’s unknown how changing the context to a more professional environment like a virtual office, school, or training simulation can change the ways people form interpersonal relationships via their digital avatars [24]. As more schools and companies attempt to use the virtual world to host classes, meetings, and conferences, we will have more data on their effectiveness in creating engaging socialization and networking experiences.

Figure 5.2 A university lecture hosted inside the virtual world Second Life from Kansas State University. © Kansas State University

Second Life so far seems to be quite effective at hosting events and classes, as it is easily accessible and offers a high level of avatar customization, yet it's not concrete yet. It's also possible that an environment with Digital Sibling avatars that are meant to simulate reality as much as possible won't have the same results, more data is needed with this context [19].

Figure 5.3 A training simulation made by Meta Quest for doctors in the Los Angeles Children's Hospital with realistic avatars. © Meta

5.3 Accessibility

One of the most important issues to address with Mixed Reality and Metaverse technologies is accessibility. If a technology is not widely accessible to the public, then it will never truly become normalized and adopted in public and everyday life. In the case of our demo, there are quite a few factors that make it inaccessible. First, creating a Metahuman Digital Sibling is quite a complicated process, so creating one for many users can be a challenge. Furthermore, there is the issue of face and body tracking. For the purposes of this demo, we only implemented face tracking with Faceware Studio, as there are no reliable and accessible methods for full body tracking without significant cost or specific equipment. Tracking the human body in real time is an intensive process and a complex problem [3, 4, 18, 31]. Both face and body tracking require specific technologies to work. Developments in the tech can make it more reliable as well as more accessible, however today it is still quite niche and with very limited use cases. Specific technology such as treadmills and physical space is also needed for the feeling of presence [16, 25] inside a virtual world or while controlling an avatar in Augmented Reality, as shown in Figure 5.4.

Figure 5.4 People using virtual reality with equipment meant to immerse them more in the virtual world within the physical space, increasing their sense of presence in the virtual world. © ITCL

5.4 Limitations of the Digital World

Technology is necessary to be able to access the Metaverse and any type of Digital World. When immersed in a digital world, doing certain actions requires specific technologies that either haven't been developed yet or are still limited in their functionality and use cases. For example, in the earlier stages of this project we attempted to include video call functionality to the demo we were developing. To do this we would use the AgoraSDK library which provides this feature. However, the plugin does is not supported in Mixed Reality applications, and no similar technology exists at present. This means that for interaction to be possible between two Digital Avatars, they still need to exist within the same virtual "space". Thus, we are introducing the restrictions of the physical world to the digital space, instead of using the digital world to surpass the limits of the physical world. Because of the limited amount of technologies available for use in the digital world, the actions that can be performed by avatars are specific, limited, and context-sensitive. For example, an avatar in a videogame can only perform the actions it was programmed to be able to do in the game's world. For Mixed Reality applications and Digital Worlds that are treated like physical spaces in the Metaverse, the actions of the avatars are limited by the bodies of the players controlling them in the physical world and their capabilities, as well as the technologies they have access to. Because of these limitations and the accessibility problems of the technologies we've outlined above, large and interconnected Digital Worlds seem to not be feasible for the time being. For these Digital Worlds to provide as much value as possible to their users and as much functionality to the avatars within them, they need to be designed with the limited and context-specific capabilities of the avatars in mind. For this reason, the most likely version of the Metaverse we will see in the future is a collection of small, separate virtual worlds, each with their own specific use cases, instead of a Digital Universe of interconnected Virtual Spaces that simulate our own reality.

Chapter 6

Conclusions

6.1 Conclusion

Although the purpose of the virtual world is to surpass the limits of the physical world, no matter how advanced and immersive the technology becomes, the limit of existing within the physical world will always exist. To use Digital Siblings within AR/VR, specific spaces and technology is needed, which heavily limits their use cases and accessibility. The concept of the Metaverse, as understood by some, as a singular Virtual Universe that connects every virtual world together to be accessed like a physical place for brands and companies to migrate to, has some big limitations that make it not feasible. Meta, the company that has been chasing this vision of the Metaverse, has since shifted their priorities away from it even after spending billions of dollars to achieve little results. The more feasible interpretation of the Metaverse is a collection of virtual worlds, each game or app being its own "Metaverse" that can be accessed through a device. Mixed Reality experiences and Digital Avatars that allow us to explore these digital worlds are becoming more advanced and more common, so it's becoming more important to create Mixed Reality experiences that are as engaging and accessible as possible, while understanding the way we relate to the digital avatars of ourselves and other users. We have demonstrated the process of creating a Metahuman Digital Sibling, an avatar that transcends the limitations of digital avatars to embody the human it represents in as much detail and fidelity as possible, to be as true-to-life as possible. With that avatar, we have created a proof-of-concept demo for an Augmented Reality video conferencing application, which can be used as a starting point to collect data and feedback from users in order to further improve it and technology similar to it in the future.

6.2 Limitations

Because of the nature of Mixed Reality development, we have run into a lot of limitations while conducting this dissertation and developing our proof-of-concept demo. For face tracking, we used Faceware Studio which is a motion capture software and not a live capture software, it works by streaming information it captures to Unreal Engine 5 while the software is open, so it cannot be integrated and exported in a live application. Furthermore, the demo only uses face tracking, as we did not have access to a reliable and accessible method for upper body motion tracking that did not require high cost or specific equipment that would severely limit its use cases. As discussed earlier, for the purposes of the demo and the screenshots shown in this paper, the background in the engine and HoloLens emulator is a black grid, which is a placeholder for the footage captured by a camera to depict the physical world to place objects inside. The lighting used is also a simple directional light with the Lumen Engine of Unreal Engine 5 so that the Avatar is visible within the scene, which is meant to be a placeholder for the realistic lighting usually seen in Augmented Reality. Additionally, attempts have been made in earlier stages of this project to add video call functionality to the demo by using the AgoraSDK library. However, the library is not yet compatible with Mixed Reality applications so for the purposes of the demo we have created a simulation that approximates the experience of an Augmented Reality Video Call with a Metahuman Avatar.

6.3 Future Work

For improvement of the AR demo app and the Meahuman Digital Sibling, user data is required. A User Evaluation Study must be conducted so that users can be exposed to the demo we created and see if they feel that it is truly more engaging than the simpler alternatives that already exist today, such as video conferencing via Zoom. With this data and user feedback we can make improvements where they are deemed necessary, whether it’s on the face tracking, the avatar itself, or the lighting. More useful feedback we can collect from users is how they relate to the Digital Sibling avatar, to determine if a realistic and true to life avatar truly gives a more engaging experience while collaborating in a professional setting.

References

[1] Arthur - Enterprise Collaboration in Virtual Reality

[2] Artlabs. (2021). Metaverse 101: What It Is & Why You Should Care. Access Date: 24/04/2022. https://artlabs.ai/blog/metaverse-101-what-it-iswhy-you-should-care/

[3] Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, and Matthias Grundmann. 2020. Blazepose: On-device real-time body pose tracking. arXiv preprint arXiv:2006.10204 (2020).

[4] Ricardo R Barioni, Lucas Figueiredo, Kelvin Cunha, and Veronica Teichrieb. 2018.Human pose tracking from rgb inputs. In 2018 20th Symposium on Virtual and Augmented Reality (SVR). IEEE, 176–182.

[5] Blinka, L. (2008). The relationship of players to their avatars in MMORPGs: Differences between adolescents, emerging adults and adults. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 2*(1), 1–7. https://cyberpsychology.eu/article/view/4211.

[6] Chopra, D. (2019). Metahuman: Unleashing your infinite potential. Harmony.

[7] Cortes M.S. (2021). What is the metaverse? A (kind of) simple explainer. Access Date: 24/04/2022. https://mashable.com/article/what-is-the-metaverseexplainer

[8] Alanah Davis, John Murphy, Dawn Owens, Deepak Khazanchi, and Ilze Zigurs. 2009. Avatars, people, and virtual worlds: Foundations for research in metaverses. Journal of the Association for Information Systems 10, 2 (2009), 1.

[9] Dick, E. (2021). Public Policy for the Metaverse: Key Takeaways from the 2021 AR/VR Policy Conference. Information Technology and Innovation Foundation.

[10] Dooley Murphy. 2017. Building a hybrid virtual agent for testing user empathy and arousal in response to avatar (micro-) expressions. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology. 1–2.

[11] Yogesh K Dwivedi, Laurie Hughes, Abdullah M Baabdullah, Samuel RibeiroNavarrete, Mihalis Giannakis, Mutaz M Al-Debei, Denis Dennehy, Bhimaraya Metri, Dimitrios Buhalis, Christy MK Cheung, et al. 2022. Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management 66 (2022), 102542

[12] Rosalynd Fraser, Jan Slattery, Igor Yakovenko. Escaping through video games: Using your avatar to find meaning in life. Computers in Human Behavior 2022

[13] Ganesh, S., Van Schie, H. T., De Lange, F. P., Thompson, E., & Wigboldus, D. H. (2012). How the human brain goes virtual: Distinct cortical regions of the person-processing network are involved in self-identification with virtual agents. *Cerebral Cortex, 22*(7), 1577–1585. https://doi.org/10.1093/cercor/bhr227

[14] Guo Freeman and Divine Maloney. 2021. Body, avatar, and me: The presentation and perception of self in social virtual reality. Proceedings of the ACM on Human Computer Interaction 4, CSCW3 (2021), 1–27

[15] Joshua, J. (2017). Information Bodies: Computational Anxiety in Neal Stephenson's Snow Crash. Interdisciplinary Literary Studies, 19(1), 17-47.

[16] Shunichi Kasahara, Keina Konno, Richi Owaki, Tsubasa Nishi, Akiko Takeshita, Takayuki Ito, Shoko Kasuga, and Junichi Ushiba. 2017. Malleable embodiment: changing sense of embodiment by spatial-temporal deformation of virtual human body. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 6438–6448

[17] Klastrup, L. (2003). A Poetics of Virtual Worlds. Proceedings of Melbourne DAC2003. Melbourne. Available at http://hypertext.rmit.edu.au/dac/papers. Accessed in 10.08.2008.

[18] Kit Yung Lam, Liang Yang, Ahmad Alhilal, Lik-Hang Lee, Gareth Tyson, Pan Hui. 2022. Human-Avatar Interaction in Metaverse: Framework for Full-body Interaction

[19] Marc Erich Latoschik, Daniel Roth, Dominik Gall, Jascha Achenbach, Thomas Waltemate, and Mario Botsch. 2017. The effect of avatar realism in immersive social virtual realities. In Proceedings of the 23rd ACM symposium on virtual reality software and technology. 1–10

[20] Lik-Hang Lee, Tristan Braud, Pengyuan Zhou, Lin Wang, Dianlei Xu, Zijun Lin, Abhishek Kumar, Carlos Bermejo, and Pan Hui. 2021. All one needs to know about metaverse: A complete survey on technological singularity, virtual ecosystem, and research agenda. arXiv preprint arXiv:2110.05352 (2021)

[21] S. Martin. “What is a Digital Twin?” - NVidia, 2019

[22] Nalbant, K.G., Uyanık, Ş., “Your Digital Twin: Metahuman”, 2nd International Conference on Applied Engineering and Natural Sciences, 10th13th March, Konya, Turkey, 2022

[23] Kemal Gokhan Nalbant, Sevval Uyanik. 2022. A Look At The New Humanity: Metaverse and Metahuman, International Journal of Computers

[24] Rabindra Ratan and Béatrice S Hasler. 2014. Playing well with virtual classmates: relating avatar design to group satisfaction. In Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing. 564–573.

[25] Eliane Schlemmer, Daiana Trein, Cristoffer Oliveira. 2009. The Metaverse: Telepresence in 3D Avatar-Driven Digital-Virtual Worlds

[26] Shipman, K. (2021). MetaHumans - a look past the Hype. AccessDate:10/02/2022.https://www.pugetsys tems.com/labs/articles/MetaHumans---a-look-past-the-Hype-2139/.

[27] Silberling, A. (2021). Niantic reveals its vision for a ‘real-world metaverse,’ releases Lightship AR Developer Kit. Access Date: 26/04/2022. https://techcrunch.com/2021/11/08/niantic-revealsits-vision-for-a-real-world-metaverse-releaseslightship-ar-developer-kit/

[28] Kim Szolin, Daria J. Kuss, Filip M. Nuyens, Mark D. Griffiths. “’I am the character, the character is me’: A thematic analysis of the user-avatar relationship in videogames. Computers in Human Behavior 2022

[29] Kim Szolin, Daria J. Kuss, Filip M. Nuyens, Mark D. Griffiths. Gaming Disorder: A systematic review exploring the user-avatar relationship in videogames. Computers in Human Behavior 2022

[30] Unreal Engine. (2021). Access Date:26/04/2022. https://www.unrealengine.com/en-US/metahumancreator.

[31] Xiaozhou Wei, Lijun Yin, Zhiwei Zhu, and Qiang Ji. 2004. Avatar-mediated face tracking and lip reading for human computer interaction. In Proceedings of the 12th annual ACM international conference on Multimedia. 500–503.