A Framework for Creating Customized Virtual Environments

by

Riccardo Caprile

DIBRIS, Università di Genova

Via Opera Pia, 13 16145 Genova, Italy   
http://www.dibris.unige.it/

A drawing of a person

Description automatically generated

Laurea Magistrale in Computer Science

MSc in Computer Science

Data Science and Engineering Curriculum

**A Framework for Creating Customized Virtual Environments**

**Riccardo Caprile**

Advisor: Manuela Chessa Examiner: Nicoletta Noceti

January, 2023

**Abstract**

This thesis aims to exploit generative language models to create customized virtual reality environments.  
In many fields of application, such as exergames for cognitive or physical training, one of the main issues is to create a wide range of simulated situations, thus avoiding the onset of boredom and habituation phenomena. However, creating many simulated situations may be a long and tedious task.   
This thesis aims to cope with this issue, creating a framework for generating different virtual reality scenarios described in natural language. The framework will allow unskilled people to add virtual scenarios and simulations to existing sw, such as cognitive exergames.  
The student will analyse the use of generative language models (e.g. chatgpt) and their use to create virtual environments (e.g., in Unity).  
A further development will be the combination with real-world elements, detected in real-time, to create dynamic environments where virtual and real objects coexist.

Table of Contents

[Chapter 1 Introduction 6](#_Toc158407493)

[1.1 Motivations and Goals 6](#_Toc158407494)

[1.2 Virtual Reality (VR) 7](#_Toc158407495)

[1.2.1 Applications 8](#_Toc158407496)

[Chapter 2 Large Language Model 9](#_Toc158407497)

[2.1 Definition 9](#_Toc158407498)

[2.2 Applications 9](#_Toc158407499)

[2.2.1 ChatGPT 9](#_Toc158407500)

[2.3 LLM and VR 9](#_Toc158407501)

[Chapter 3 Methods 10](#_Toc158407502)

[3.1 Software 10](#_Toc158407503)

[3.1.1 Unity 10](#_Toc158407504)

[3.1.2 API OpenAI 10](#_Toc158407505)

[3.1.3 Roslyn C# Runtime Compiler 10](#_Toc158407506)

[3.1.4 Vocal Commands 10](#_Toc158407507)

[3.2 Hardware 10](#_Toc158407508)

[3.2.1 Meta Quest 2 10](#_Toc158407509)

[Chapter 4 Design of the System 11](#_Toc158407510)

[Chapter 5 Implementation 12](#_Toc158407511)

[Chapter 6 Results 13](#_Toc158407512)

[6.1 System Testing 13](#_Toc158407513)

[6.1.1 Experiments 13](#_Toc158407514)

[6.1.2 Script Generation: Execution Time 13](#_Toc158407515)

[6.1.3 Script Generation: Error Rate 13](#_Toc158407516)

[6.2 User Testing 13](#_Toc158407517)

[6.2.1 Testers Description 13](#_Toc158407518)

[6.2.2 System Usability Scale (SUS) 13](#_Toc158407519)

[6.2.3 User Experience Questionnaire (UEQ) 13](#_Toc158407520)

[Chapter 7 Conclusions 14](#_Toc158407521)

[7.1 Future Works 14](#_Toc158407522)

[Bibliography 15](#_Toc158407523)

# Introduction

The first chapter briefly gives a summary of the topics covered inside the other chapters and the principal objectives of this thesis and an introduction to Virtual Reality (VR).

## Motivations and Goals

Large language models trained for code generation can be applied to speaking virtual world into existence, the main objective is using them for creating custom virtual environments suitable with user requests. In the recent years, AI models have reached a very high level of reliability when a question has asked them. So, in this thesis we wanted to exploit those models and to give to the user a framework that can overcome the problem of having different environments for different situations with a small effort. We want to give the possibility to build environments with the support of the natural language in order to have a framework accessible to unskilled people not related to computer science world. For instance, a physiotherapist would ask for a virtual environment that is tuned the needs of a patient with an unusual injury. Without the support of a framework like this, the building of the environment would a boring a tedious task, that must be repeated hundreds of times. Students in a laboratory could inspect, analyse, and become familiar with materials or machines that they could never handle, because of the school’s budget or because they are too dangerous to be used inside such structures. There are thousands of possible examples of custom virtual environments.

Firstly, we wanted to define and explain the Large Language Models and what is the relationship between them and Virtual Reality, because the framework is based on them, and the combination of these tools will produce the desired result. Without LLM we could not have the outputs requested for creating the customized world.

Following, we wanted to describe hardware and, especially, software needed for the realization of the framework, giving technical aspects of those methods. Then, we would analyse the design of the system and understand how the tools described before are combined with.

At this point, we must go deeper into the implementation of the system, analysing all the C# scripts, and how every step of the design has been implemented.

In the last part of the thesis, we have a section for the results obtained by the framework. The “users analysis” is going to be based on their feedback with a couple of questionnaires, in addition to this there is going to be a technical analysis based on how the system worked in different situations (eg. Error rate and execution time).

To conclude this work, we would like to give how the framework could be improved in future works.

## Virtual Reality (VR)

Virtual Reality (VR) is a technology which allows a user to interact with a computer-simulated environment that can be a representation of the real world or an imaginary world. The main objective is to give an immersive experience of a virtual world. The environment is typically accessed via a display, usually a wearable display (eg. Head mounted display (HMD), gloves or body suits).

The illusion of “being there” is affected by motion sensors that pick up the user’s movements and adjust the view on the screen accordingly in real time.

VR can be divided into five categories:

* **Non-Immersive:** It refers to a virtual experience through a computer where you can control some characters within the software, but the environment is not directly interacting with you. Basically, you are dealing with a virtual world, but you are not at the centre of attention in the game.

All basic forms of gaming devices such as: Playstation, Xbox, Computers, etc, provide us with a non-immersive virtual reality experience.

* **Immersive:** This type of experience is what most people think what they come across with Virtual Reality. It is based on wearable displays to track user’s movements and present the VR information depending on the position of users, which allows them to experience 360 degrees of the virtual environment. It feels like you are within the virtual world physically.
* **Semi-Immersive:** As the name suggests, it is a mix of non-immersive and immersive virtual reality. It is a 3D space or virtual environment where you can move on your own, either through a headset or a computer screen. All the activities are concentrated toward you, but, on the other hand, you have no real physical movements other than your visual experience.

Immagine che contiene schermata, collage, uomo, arte

Descrizione generata automaticamente

1 The three categories of Virtual Reality

In between immersive and non-immersive virtual reality, seen before, there is also:

* **Augmented Reality:** It combines the real world and computer-generated content. It is a technology that enhances the real-world environment with digital information and objects, creating a more immersive and interactive experience. Users can access digital information in a more engaging and interactive way.

An example of augmented reality are interactive displays;for instance, when we scan an object in a shop, we could visualize some extra information about that article.

One of the first appearance of Virtual Reality devices similar to Head Mounted displays was in the first half of 1800s with the stereoscope. It was built by following research that demonstrated that the brain combines two photographs of the same reference object taken from different points to make the image appear to have a sense of depth and immersion.

In 1956, the cinematographer Morton Heilig created Sensorama [[1]](#One) the first VR machine. It combined full colour 3D video, audio, vibrations, smell and atmospheric effects.

His second device created was the so called Telesphere Mask which was the first HMD. It provided stereoscopic 3D images with wide vision and stereo sound.

### Applications

# Large Language Model

Because the framework implemented it is based on a large language model such as ChatGPT , developed by OpenAI, we want to provide , in this chapter , information about these models and their connection with the Virtual Reality development world.

## Definition

## Applications

### ChatGPT

## LLM and VR

# Methods

In this chapter, we want to describe all the software and hardware used during the implementation. For each of them , we are going to give an idea of their working principles plus some technical aspects and specifications.

## Software

### Unity

### API OpenAI

[[2]](#One)

### Roslyn C# Runtime Compiler

### Vocal Commands

## Hardware

### Meta Quest 2

# Design of the System

# Implementation

After the definition of the design of the system, we have to analyse and explain every single step of the implementation deeply. In this chapter, we provide a fully technical description of all the blocks needed to build the final framework.

# Results

## System Testing

### Experiments

### Script Generation: Execution Time

### Script Generation: Error Rate

## User Testing

### Testers Description

### System Usability Scale (SUS)

### User Experience Questionnaire (UEQ)

# Conclusions

## Future Works

Bibliography

[1] Morton Heilig. The cinema of the future. 1955.

[2] https://github.com/RageAgainstThePixel/OpenAI-DotNet