Virtual Reality Interfaces for 3D model reconstruction and sharing

**Student: Andrei-Alexandru Bunea**

**Supervisor: Prof. Anthony Steed**

MSc Computer Graphics, Vision & Imaging

September 2023

This report is submitted as part requirement for the MSc Degree in 'Computer Graphics, Vision & Imaging', at University College London. It is substantially the result of my own work except where explicitly indicated in the text. The report may be freely copied and distributed provided the source is explicitly acknowledged.

# Acknowledgements

I would like to express my profound gratitude to Professor Anthony Steed, my supervisor, for his invaluable support throughout the whole collaboration. His guidance and mentorship have been instrumental in successfully finishing this project.

I would also like to thank Nels Numan and Ziwen Lu, teaching assistants of the Virtual Environments course, who guided me in the early stages to find and design the idea that was implemented in this final project.

Finally, I would also like to thank my family for their unconditional and unwavering support. They have been the driving force behind every step of my journey, shaping me into the person I am today.

# Abstract

In recent years, 3D reconstruction has revolutionised various industries, proving to be a transformative practice in scientific research, e-commerce, and beyond. Studying a 3D model has proven more useful and intuitive than 2D images in analysing objects, helping professionals improve their expertise in medicine, archaeology, art, and many other fields.

This immersive experience not only enhances the assessment of virtual objects but also bridges the gap between the virtual and physical worlds. The integration of this technology has revolutionized the way professionals approach their work, enabling them to gain unparalleled insights into complex structures and artefacts. However, the reconstruction is done mostly by professionals that are passionate about photogrammetry and possess skills and software for 3D reconstruction, and can later choose to publish their work on specialized platforms for 3D portfolios such as “Unity Asset Store” or “Sketchfab”.

This project comes as a solution for users who would like to create their own 3D models without having any prior knowledge of photogrammetry, visualize the results using a virtual reality platform, store their work and easily share it with others.

The project encompasses a web-based user interface acting as a social media platform where users publish 3D models based on uploaded images or videos, a structure from motion 3D reconstruction pipeline using “Meshroom”, and a Unity-based application for VR visualization. The VR application allows users to immerse themselves alongside the reconstructed 3D objects, enabling them to inspect the models from various angles and distances.

The project's methodology combines image processing, computer vision, and VR development, resulting in a robust and user-friendly system. Throughout the development process, a thorough review of related literature was conducted to ensure the project's alignment with current advancements in VR applications and photogrammetry.

The evaluation of the system involved user feedback and usability tests, providing valuable insights into the overall effectiveness and user experience. The results demonstrate the system's ability to generate accurate 3D models and deliver an engaging VR experience to users.

Contents

[Acknowledgements 2](#_Toc142418382)

[Abstract 3](#_Toc142418383)

[1. Introduction 5](#_Toc142418384)

[1.1. Motivation 5](#_Toc142418385)

[1.2 Photogrammetry 6](#_Toc142418386)

[1.3 Structure from Motion 9](#_Toc142418387)

[1.4 Visualisation tools 10](#_Toc142418388)

[1.5 Project structure 10](#_Toc142418389)

[1.6 Structure of the report 10](#_Toc142418390)

[References 11](#_Toc142418391)

# Introduction

## Motivation

In today’s rapid technological advancement, scientists and researchers have vastly improved their analysis by replacing the ordinary image interpretation with a more comprehensive concept: 3D models.

Using a virtual entity of a physical object and a VR visualisation tool simulates the sensation of looking at real-world objects. Immersing the user in such way enables in-depth examination and investigation, as it confers a holistic understanding of the object, rather than the analysis of a batch of scans or images taken from different points of view. Moreover, the hands-on experience provides the researchers with a boost of confidence in their expertise, as they are more assured when working with real-world models.

Three-dimensional (3D) modelling has become a transformational force in today's quickly changing technology landscape, having a significant influence on a variety of industries from education to cultural preservation.

The goal of this project is to reduce the gap between the common users who do not possess advanced skills in photogrammetry, and the 3D model creation realm, providing a reliable and easy-to-use platform for 3D reconstruction and model sharing. Moreover, this project is investigating the complex interactions between user-generated 3D models, immersive visualisation and their many applications.

While there are a significant number of tools for either structure from motion reconstruction and VR visualisation of 3D models, there is no such platform that combines both. Moreover, each one of these activities are performed my professionals who own 3D reconstruction software and have advanced expertise in how to perform a reconstruction algorithm or how to create a VR scene where entities can be observed.

In addition, the industry currently lacks a platform for 3D models sharing that is not oriented towards professionals. Not only there are a couple of websites that provides the opportunity of 3D models browsing, but the most frequently used applications such as “Unity Asset Store” [1] or “Sketchfab” [2] require the designers to perform their own 3D model creation process before uploading the results to their portfolios.

Therefore, the innovation that this project brings to the field is a democratic and accessible process for 3D model production, sharing and exploration by utilising the capabilities of 3D reconstruction complex algorithms, advanced software, and virtual reality. By following this approach, users are relieved from the burden of mastering the state-of-the-art algorithms, allowing them to fully enjoy the benefits of creating and exploring virtual representations of real-world models.

## 1.2 Photogrammetry

Photogrammetry is a dynamic and complex process that combines multiple image processing algorithms to create 3D models from sets of overlapping images. The aim of the process is to determine the three-dimensional geometry of physical objects captured.

In general, a photogrammetric project involves two major steps: acquisition and preparation of the image dataset, and the process of reconstruction. The first step ensures a proper camera setup and scene properties for the desired task while capturing a collection of photos with good overlapping areas, and a full coverage of the area of interest. The preparation of the image dataset may also involve applying special filters or processing techniques to images to potentiate the features of the physical object. Whereas the second step covers the whole process that converts the images into the final 3D model, encapsulating custom algorithms from feature extraction and matching to geometric reconstruction, depth map estimation and texturing. [3]

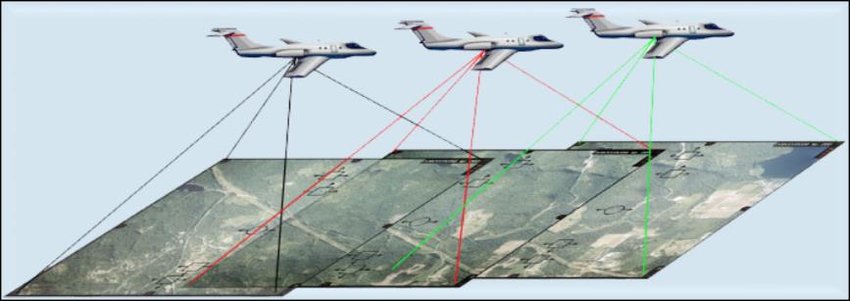


Figure 1 Aerial Photogrammetry illustration

Generally, photogrammetry is divided into aerial and terrestrial photogrammetry. Aerial photogrammetry acquires images taken from lower heights using captures from drones or balloons, medium heights using airplanes or even from astonishing heights using modern satellites. The process of acquiring images for aerial photogrammetry is presented in Figure 1. [4] Terrestrial photogrammetry is performed on images captured from the ground. The images can be acquired from far distances or close range. When photogrammetry is performed on images taken from less then 100 meters, it is considered to be categorised as close-range photogrammetry. [5]

While aerial photogrammetry is a popular practice in cartography, archaeological exploration and environmental monitoring, close-range photogrammetry has been used in detail-oriented practices such as art restoration, industrial design, and medicine.

The close-range photogrammetry takes a more intimate approach, focusing on capturing fine details of objects from a short distance. This technique usually involves the user of high-resolution cameras to photograph subjects from multiple points of view, creating accurate and detailed 3D models. [6]

A statue of a person

Description automatically generated

Figure 2 Close-range Photogrammetry application for digital curation

In Figure 2 is presented a close-range photogrammetry application performed by the Institute of America in Transylvania, Romania in 2023 during an archaeological photogrammetry workshop. [7] In the illustration from above we can observe a multitude of high-quality images taken from different perspectives and ranges, photos that come all together to reconstruct a solid virtual representation of the targeted statue. Using photogrammetry to capture every fine detail from images of an object, not only provides a better understanding of the object but also preserves an invaluable and meticulous digital copy of the entity.

## 1.3 Structure from Motion

Structure from Motion (SfM) is a process derived from photogrammetry and encapsulates all its goals. Both SfM and photogrammetry target the reconstruction of 3D scenes or objects from a collection of 2D images. However, SfM incorporates other additional tasks such as estimating the camera position and parameters, being more concerned with capturing relative positions and orientations of images.

In photogrammetry, camera calibration and image orientation are crucial for accurate measurements and are established from the start of the experiment. Ground control points may be used for georeferencing. In contrast, structure from motion is more flexible and automated, as it does not rely on camera calibration and geolocation. Instead, it uses feature tracking across the images to determine the camera parameters. [8]

The key stages of the structure from motion pipeline include:

1. **Feature Detection**

This first step covers the detection of distinctive features in each image such as corners, edges or other key points. Classic algorithms for feature detection are Harris Corner Detector and ORB.

1. **Feature Matching**

The second step is responsible for establishing corelations between features extracted in different images. Classic algorithms for this step are RANSAC and SIFT

1. **Camera Pose Estimation**

Using the feature corelation between images, the fundamental matrix for camera calibration is estimated. From this matrix, the extrinsic and intrinsic parameters are extracted, measuring the camera position and orientation.

1. **Sparse Point Cloud Reconstruction**

Based on the information provided by the essential matrices estimated at the previous step, information about the local 3D coordinates of the mesh vertices is extracted, thus constructing a sparse point cloud reconstruction of the model.

1. **Bundle Adjustment**

This fifth step is an optimization process that refines the estimated camera poses and vertex positions, minimizing the reprojection error. This step ensures that the reconstructed 3D points align as closely as possible with the characteristics of the physical object.

1. **Mesh Generation**

Additional 3D points are reconstructed by triangulating pixels from all images, completing the point cloud of the scene or model’s surface.

1. **Texture Mapping**

This final step involves mapping the texture to every face of the mesh, using the original images as reference to obtain similar looking textures.

The SfM pipeline is a reliable and flexible approach in creating 3D models, recommended for projects that do not have the facilities of setting up the camera parameters. This technique has the same applications as close-range photogrammetry. Advances in computer vision enrich the technical capabilities of the SfM pipeline to produce more accurate structures, to render corresponding textures and to achieve a faster rendering with smaller running time.

## 1.4 Virtual Reality

Virtual Reality (VR) represents a groundbreaking idea that involves crafting an environment where individuals can detach from their actual surroundings and immerse themselves in diverse sensory experiences. The aim of virtual reality is to provide a highly realistic immersive display that stimulates the senses of users in order to induce the feeling of belonging to a virtual scenario. Therefore, this distinctive idea enables users to engage their senses through realistic simulations of real-world experiences, all from the convenience of their home or office, and having access to multiple virtual worlds in a small device such as a pair of VR headset.

Immersion is the main goal of Virtual Reality. To achieve full immersion, users need to experience a complete detachment from reality, engaging all their senses solely within the virtual environment. To elicit immersion, a virtual environment should include four main components: sensors, simulators, rendering and display. Examples of sensors are trackers, microphones or bio signal devices that turn physical effects into signals. VR headsets such as the Oculus Quest used in the developing process of this project rely on controller’s sensors and the headset’s cameras to provide critical information about the user position and gestures. The output of the sensors is passed to the simulators that simulate the virtual environment and its interaction with the user based on the data provided. Further, the renderer creates the Virtual Environment visually and produces signals, which are then transformed into physical images by the display.

Unity is one of the most popular tools for creating immersive Virtual Reality applications. It is an optimal solution for developers as its flexibility, feature-rich environment, and wide range of compatibility foster the creation of interactive and captivating VR experiences. Even though Unity was conceived initially for game development, its applications cover a larger usability spectrum, making it a cornerstone development tool for VR applications.



Figure 3 Oculus Quest 2 (the headset used in this project)

One of Unity’s strongest features is its integration with VR hardware, like the Oculus Quest headset used in the project. It provides special support for specialised VR technology such as the XR controllers’ tracking mechanism and actionable buttons. It also offers a specialised collection of 3D models, enhancing the performance and creativity of the projects. In Figure 3, the headset used in this project for both development and testing is displayed, as advertised on the Meta official website. [9]

Furthermore, the implementation of realistic interactions within the virtual worlds is facilitated by Unity’s physics engine that enables users to authentically interact with objects from the virtual setup. Assessing this project’s needs, Unity was chosen as the optimal platform for the development of the VR visualisation application. It provided fast protocols for the data acquired from the server at runtime, a detail-oriented interaction with the environment and an aesthetic touch on the design.

## 1.5 Project structure

The 3D model reconstruction and visualisation project is methodically organised into various components that contribute harmoniously to a coherent and efficient workflow. The process guarantees that 2D photos are successfully transformed into 3D models and transposed in an immersive VR environment. The architecture of the project includes four main components: an Angular web application, a 3D reconstruction pipeline and a python, flask backend, and a Unity VR application, all components working in tandem to offer a smooth and engaging user experience.

The interaction with the project begins by accessing the web-based user interface. The frontend component is designed to facilitate friendly designed and intuitive features. To use the web application, the users need to sign up for an account or log in with an existing one. The login credentials are verified by interrogating the project’s database. After the successful validation, users are allowed to perform all the activities that the platform offers. First, there is a Dashboard component where the users upload their images or videos in order to create 3D models. Another feature is the personal portfolio, where users can preview their work, download the models, edit their names, or delete them. In addition, users can search for other users’ profiles in the explore section, follow their work and see their 3D models. Every 3D model can be seen both in VR but also in the web application, where a scene is rendered including the targeted reconstructed object.

The second component is providing the 3D models using a structure from motion automatization on the images or video provided by the user. Using the free, and open-source Meshroom software from AliceVision, the project introduces a script that runs 13 reconstruction stages. At each stage, the pipeline’s script configures the parameters that are passed to Meshroom, having multiple scripts, one for each reconstruction step. After feature extraction and matching, camera estimation, dense scene, depth map filtering, mesh decimation and texturing, the 3D models are created and ready to be rendered in the visualisation tools.

The third component is the one that organises all the processes and data flows – the python, flask backend. Inside the backend, the endpoints for the applications API are defined, providing routes that are used by the web user interface and the VR visualisation tool. Moreover, it processes the upload of images and videos, while preparing the input dataset for the reconstruction pipeline. In order to store details about the users, the application stores their email, name and IDs of the models in a MySQL database, that is queried by the backend when needed.

Last, the fourth component is the Unity VR application which serves as an immersive visualisation tool. The VR application provides support for login, after which the user can preview and spawn their own model and search for the portfolios of the people they follow. In other words, the application has the same features as the web application except for the signup and image upload for reconstruction. In addition, the VR application provides a unique and immersive tool that facilitates a faster and more genuine exploration of 3D models. Inside the virtual scene, the users can move around the object, rotate, scale, or translate the object in order to achieve a better understanding of the model.

## 1.6 Structure of the report

The goal of this report is the understand the mechanics behind the implementation of the implemented tool and it’s use cases. In order to achieve this, the report will further on describe previous implementations of structure from motion application, the implementation of this project, it’s evaluation and the conclusions.

In section 2, we will analyse the literature review and examine the implementation of similar tools, as well as their usage and achievements.

In section 3, the implementation of the project will be discussed, examining step by step the processes behind the four main components: the web user interface, the reconstruction pipeline, the backend, and the Unity VR application.

Section 4 will be designated to the evaluation of the project, studying the performance of the reconstruction process and the feedback from the users that were asked to complete a form after participating in an experiment where they tested both the web user interface and the VR visualisation tool.

Last, section 5 will assess the conclusions regarding this project, its strong features and weaknesses and will consider future work for its improvement.

# References

|  |  |
| --- | --- |
| [1] | “Unity Asset Store,” [Online]. Available: https://assetstore.unity.com/. |
| [2] | “Sketchfab,” [Online]. Available: https://sketchfab.com. |
| [3] | T. Schenk, “Introduction to Photogrammetry,” *The Ohio State University, Columbus,* 2005. |
| [4] | A. &. A. F. M. &. M. M. Jebur, “3D City Modelling by Photogrammetric Techniques,” 2017. |
| [5] | D. V. J. K. R. W. Ruinian Jiang, “Close-range photogrammetry applications in bridge measurement: Literature review,” *Measurement,* pp. 823-834, 2008. |
| [6] | E. M. B. J. S. &. M. J. C. Mikhail, Introduction to modern photogrammetry, John Wiley & Sons, 2001. |
| [7] | “Archeological Institute of America,” [Online]. Available: https://www.archaeological.org/fieldwork/archaeological-photogrammetry/. |
| [8] | J. B. N. G. M. H. J. R. M.J. Westoby, “‘Structure-from-Motion’ photogrammetry: A low-cost, effective tool for geoscience applications,” *Geomorphology,* vol. 179, no. 2012, pp. 300-314. |
| [9] | Meta, [Online]. Available: https://about.fb.com/news/2020/09/introducing-oculus-quest-2-the-next-generation-of-all-in-one-vr/. |