KLE Society's
KLE Technological University, Hubballi.

A Minor Project -2 Report

On

Walkthrough of Heritage Sites Using Hand Gesture Controls

submitted in partial fulfillment of the requirement for the degree of

Bachelor of Engineering

In

School of Computer Science and Engineering

Submitted By

Ritesh Hiremath	01FE21BCS151
Shreepad Joshi	01FE21BCS152
Shirisha K M	01FE21BCS167

Under the guidance of
Dr. Uma Mudenagudi, Prof. Ramesh Ashok Tabib

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

HUBBALLI – 580 031

Academic year 2023-24

KLE Society's
KLE Technological University, Hubballi.

2023 - 2024

SCHOOL OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that project entitled “Walkthrough of Heritage Sites Using Hand Gesture Controls” is a bonafied work carried out by the student team of Ritesh Hiremath 01FE21BCS151, Shreepad Joshi 01FE21BCS152, Shirisha K M 01FE21BCS167, in partial fulfillment of the completion of 6th semester B. E. course during the year 2023 – 2024. The project report has been approved as it satisfies the academic requirement with respect to the project work prescribed for the above said course.

Guide	Head, SoCSE
Dr. Uma Mudenagudi Prof. Ramesh Ashok Tabib	Dr. Vijayalakshmi.M

External Viva -Voce:

Name of the Examiners	Signature with date
1.	
2.	

Acknowledgement

We would like to thank our faculty and management for their professional guidance towards the completion of the project work. We take this opportunity to thank Dr. Ashok Shettar, Vice-Chancellor, Dr. B.S.Anami, Registrar, and Dr. P.G Tewari, Dean Academics, KLE Technological University, Hubballi, for their vision and support.

We also take this opportunity to thank Dr. Meena S. M, Professor and Dean of Faculty, SoCSE and Dr. Vijayalakshmi M, Professor and Head, SoCSE for having provided us with direction and facilitated for enhancement of skills and academic growth.

We thank our guides Uma Mudenagudi, Dean R & D and Ramesh Ashok Tabib, Asst. Professor at SoECE for the constant guidance during interaction and reviews. We also extend our gratitude to Research Scholars at the Centre of Excellence in Visual Intelligence and SEED Lab for their suggestions and inputs.

We extend our acknowledgment to the reviewers for critical suggestions and inputs. We also thank Project Co-ordinator Dr. Uday Kulkarni, and reviewers for their suggestions during the course of completion.

We express gratitude to our beloved parents for their constant encouragement and support.

Ritesh Hiremeth - 01FE21BCS151
Shreepad Joshi - 01FE21BCS152
Shirisha K M - 01FE21BCS167

ABSTRACT

Cultural heritage sites are invaluable to our historical and cultural understanding, yet many are vulnerable to natural disasters, urbanization, and other threats. To address these challenges, this project leverages augmented reality (AR), virtual reality (VR), and gesture-based controls to create immersive virtual experiences of these sites. We employ advanced 3D reconstruction techniques to transform high-quality multiview images into detailed and realistic 3D models. These models are further refined and textured in Blender to enhance their visual fidelity and accuracy. The final models are integrated into Unity, a real-time development platform, enabling users to navigate and interact with the virtual environments through intuitive hand gesture controls supported by AR/VR devices.

The project results demonstrate the effectiveness of using cutting-edge technologies to create highly detailed and interactive virtual representations of heritage sites. Users can explore these sites in an immersive manner, interacting with the environment naturally through hand gestures. This approach not only enhances user experience but also significantly expands access to cultural heritage, allowing people worldwide to experience and learn about these sites without the need for physical travel. Furthermore, the digital preservation of these sites ensures their historical and cultural significance is maintained for future generations.

In conclusion, this project highlights the potential of integrating AR, VR, and gesture-based controls for the digital preservation and presentation of cultural heritage sites. The developed platform preserves the intricate details of heritage sites while providing an engaging and educational tool for users. By making these sites accessible in a virtual format, we contribute to cultural preservation, education, and tourism, offering new opportunities for global audiences to connect with historical landmarks. This innovative approach underscores the importance of technology in preserving and promoting cultural heritage in the digital age.

Keywords : *3D Reconstruction, Photogrammetry, Gesture-Based Controls, Digital Heritage, AR (Augumented Reality), VR (Virtual Reality).*

CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

Augmented reality (AR) and virtual reality (VR) technologies, combined with gesture-based controls, have transformed the way we interact with digital environments. These technologies provide an immersive experience that allows users to engage with virtual spaces through natural and intuitive movements. By using hand gestures, users can effortlessly explore and interact with digital spaces, creating a seamless and engaging experience that feels more real and interactive. This integration of gesture controls not only enhances the user experience but also makes these technologies more accessible and user-friendly.

Figure 1.1: Heritage Site Exploration Using VR Headset.

In recent years, the development of 3D reconstruction techniques has revolutionized the way we experience historical landmarks and cultural heritage sites. By employing advanced 3D modeling software such as Blender and Unity, these technologies enable the creation of detailed and accurate digital reconstructions. This allows individuals to go beyond the limits of time and space to explore and engage with these locations in a virtual environment. The use of 3D reconstruction provides a captivating and informative source for many people, helping to preserve and spread historical knowledge.

Furthermore, 3D reconstruction can capture precise details and textures using photogrammetry and other advanced techniques, offering viewers realistic and immersive experiences that closely resemble visiting the actual site. This experience is further enhanced by the us-age of gesture-based controls, making interactions with the virtual world more intuitive and natural. Digital reconstructions allow users to walk through and explore sites conveniently, providing a fun and informative walkthrough experience. This innovative approach not only makes cultural heritage more accessible but also creates new opportunities for tourism and education by presenting the past in an engaging, digital way.

1.1 Motivation

The preservation of cultural heritage is essential for maintaining the historical and cultural fabric of our societies, ensuring that endangered or deteriorating heritage sites are accessible to future generations.An effective way to do this is with digital reconstruction and immersive AR/VR technologies, which enable diverse interactive experiences that perfectly capture the essence of these locations. By attracting tourists with captivating virtual tours, this approach not only supports conservation initiatives but also strengthens the local economy.Furthermore, these technologies offer educational potential by enabling in-depth exploration of historical locations by students as well as educators. They also improve accessibility, making cultural heritage places familiar to individuals all across the world—even those who are unable to visit in person.

1.2 Literature Survey

1. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis (Ben Mildenhall et al., ECCV 2020) [1] :

This paper introduces NeRF (Neural Radiance Fields), a novel method for synthesizing new views of complex 3D scenes using neural networks. NeRF leverages a fully connected deep neural network to encode a volumetric scene representation, treating scenes as continuous 5D functions. This representation maps 3D spatial locations and 2D viewing directions to color and volume density, enabling the rendering of photorealistic images with fine details, soft shadows, and reflections that traditional methods struggle to achieve. Training involves 2D images from various viewpoints and corresponding camera poses, allowing the network to predict radiance and density across scenes. NeRF has shown exceptional capability in generating high-fidelity images with intricate details and realistic lighting effects, marking a significant advancement in computer vision and graphics.

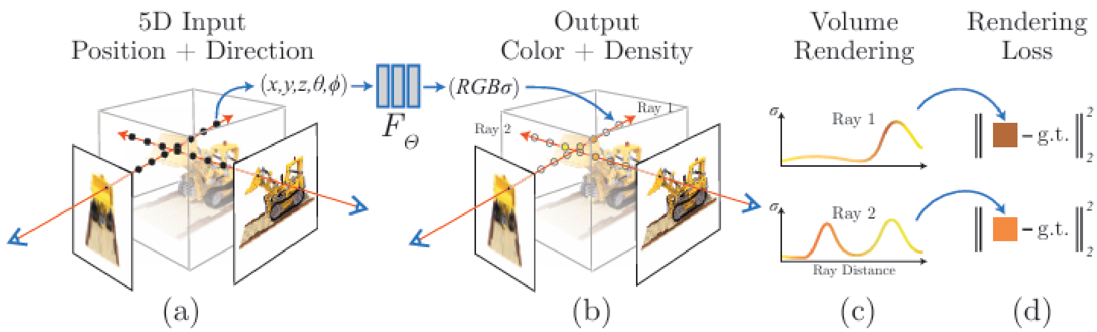

Figure 1.2: NeRF process by Ben M et al., Sample 5D coordinates along camera rays (a), input to MLP for color and density (b), composite via volume rendering (c), and optimize by minimizing differences from ground truth (d).

2. NeRF for heritage 3D Reconstruction(G. Mazzacca et al., CIPA 2023) [2] :

The paper explores the use of Neural Radiance Fields (NeRF) for digitally preserving and reconstructing cultural heritage sites. NeRF is a cutting-edge technique that synthesizes novel views of 3D scenes from a sparse set of 2D images, creating highly detailed and accurate 3D models of historical structures. This approach enhances visual quality and offers a scalable solution for documenting and preserving cultural heritage digitally.

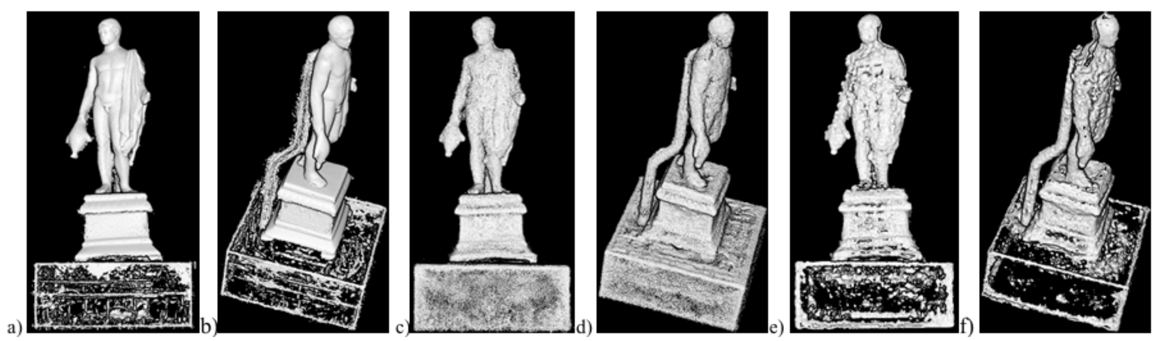

Figure 1.3: Experimental Reults of G Mazzacca et al. Photogrammetric MVS (a-b), Nerfacto (c-d) and Instant-NGP (e-f) point clouds of the bronze statue with a transparent basement and support.

The study includes several case studies demonstrating NeRF's ability to handle intricate architectural details and varying lighting conditions. The results show NeRF's potential to revolutionize heritage documentation by addressing limitations of traditional methods like photogrammetry and laser scanning. The paper concludes that NeRF significantly advances heritage 3D reconstruction, enhancing the accessibility and preservation of global cultural landmarks.

3. 3D reconstruction using Structure From Motion (SFM) algorithm and Multi View Stereo (MVS) based on computer vision(M Kholil et al., IOP 2021) [3] :

The paper discusses techniques for creating 3D models from 2D images. Structure From Motion (SFM) estimates camera positions and reconstructs the 3D structure of a scene from multiple overlapping photographs, generating an initial sparse 3D point cloud by triangulating common points.

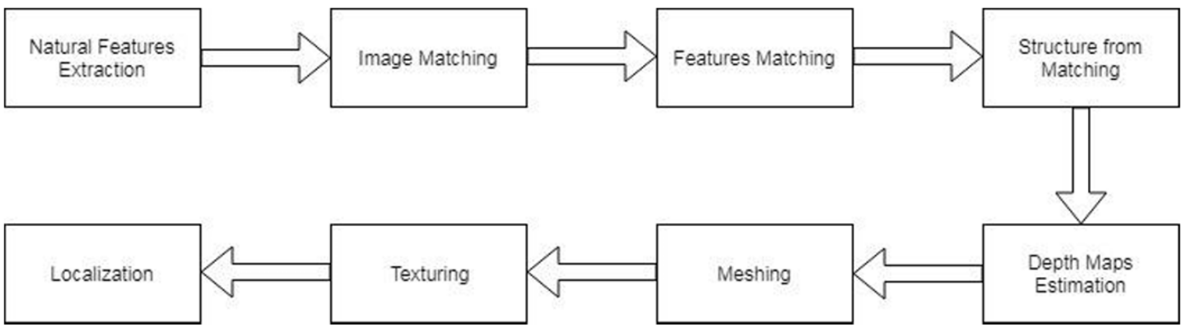

Figure 1.4: Photogrammetric pipeline

The paper then employs Multi View Stereo (MVS) to refine this model, enhancing its density and accuracy by processing the images to create a dense point cloud. This combination of SFM and MVS leverages computer vision to produce detailed 3D models, useful for applications such as cultural heritage preservation and virtual reality. The study demonstrates the effectiveness of integrating these algorithms for robust 3D reconstructions.

4. A Glance into Virtual Reality Development Using Unity(Alexandru D. Filip, Infotmatica Economica 2018) [4] :

The paper provides an overview of developing virtual reality (VR) applications using the Unity game engine. The author discusses the advantages of using Unity for VR development, highlighting its user-friendly interface, extensive documentation, and strong community support. The paper outlines the essential components and tools within Unity that facilitate VR development, such as the integration of VR SDKs, asset management, and real-time rendering capabilities.

Filip also delves into practical aspects of VR development, including best practices for optimizing performance and ensuring a smooth user experience. The paper includes case studies demonstrating various VR applications built with Unity, emphasizing its versatility and potential across different industries, from gaming to education and healthcare. By providing

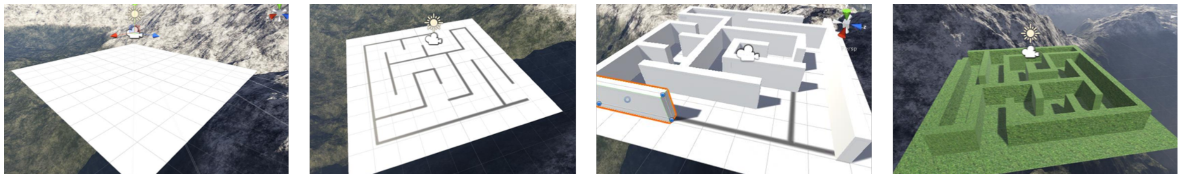

Figure 1.5: Unity maze development.

insights into the development process and practical tips, the paper serves as a valuable resource for developers looking to explore VR technology using Unity.

5. Examination of fire scene reconstructions using virtual reality to enhance forensic decision-making. A case study in Scotland.(Vincenzo Rinaldi et al.,Virtual Reality 2024) [5] :

The paper investigates the use of VR technology in forensic fire scene analysis. By creating a virtual reconstruction of a fire incident in Scotland, the study demonstrates how VR provides a more immersive and detailed examination, aiding investigators in understanding fire progression and identifying key evidence.

The case study highlights VR's potential to improve forensic science by enhancing visualization, facilitating collaboration, and presenting findings more effectively. The authors conclude that VR integration in forensic practices can significantly enhance the accuracy and efficiency of fire scene investigations, leading to more reliable forensic outcomes.

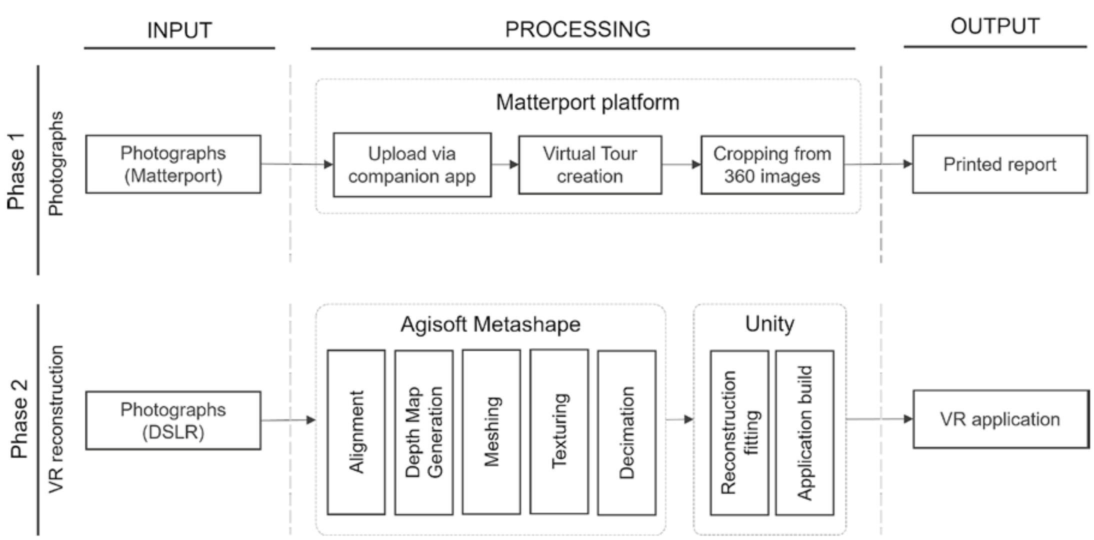

Figure 1.6: Data workflow from acquisition to deployment.

1.3 Problem Statement

Developing a virtual experience of heritage sites through 3D reconstruction and integrating gesture-based controls.

1.4 Applications

- Heritage sites must be digitally reconstructed and maintained to preserve its historical significance for future generations.
- To attract tourists, promote cultural heritage, and boost local economies, create immersive virtual tours.
- To improve learning opportunities, give students and learners an interactive platform to thoroughly explore and assess historical sites.
- Enable global access to cultural heritage sites for those unable to visit in person, including people with physical limitations.
- Provides realistic digital records of historical places to support efforts to restore and recover from natural disasters and other damages.

1.5 Objectives and Scope of the project

The aim of the project is to use the latest technologies to improve cultural heritage sites accessibility, engagement, and preservation. By focusing on gesture-based interaction, AR/VR deployment, and 3D reconstruction, this project aims to develop interactive and immersive experiences that bring historical landmarks to life in a modern, dynamic way.The following objectives outline the specific goals and intended scope of the project, providing a comprehensive framework for its development and implementation.

1.5.1 Objectives

- Generating 3D models of heritage sites using multiview images.
- Deploying scenes of heritage sites in AR/VR space.
- Creating a human Walkthrough experience with hand gesture-based movements in virtual environment.

1.5.2 Scope of the project

The scope of this project includes the 3D reconstruction of historical landmarks making use of multi view images.This process involves employing advanced photogrammetry and 3D modeling techniques to create highly detailed and accurate digital replicas of these sites.The project aims to generate 3D models that retain the historical and cultural significance of the heritage sites by capturing its features and textures. The historical significance of the sites will be preserved for future generations with the help of these digital replicas.

In addition to creating detailed 3D models, the project involves deploying these reconstructed scenes in augmented reality (AR) and virtual reality (VR) spaces.Also attracting tourists and highlighting cultural heritage, AR and VR deployments will offer a realistic and captivating experience that can stimulate interest in these locations and develop local economies.

The initiative also intends to improve the user experience by developing a human walkthrough experience that uses movements based on hand gestures. This will require creating a user interface that is natural and intuitive so that people can use hand gestures to engage and traverse the virtual environment. The project aims to increase educational opportunities for researchers and students, provide global access to heritage sites for those unable to visit in person, and provide realistic digital records to aid in restoration and recovery efforts following natural disasters and other damages by making the virtual tours more interactive and accessible.

1.6 Introduction to Software Tools

This section briefly describes the various open-source software tools used to develop our project.

1.6.1 COLMAP

COLMAP, an acronym for COmputer Vision and MAchine Perception, is a widely used open-source photogrammetry software. It was developed by Johannes L. Schönberger during his PhD at ETH Zurich and has since become a popular tool for 3D reconstruction due to its robust and accurate algorithms. In our project, COLMAP is utilized for the initial stages of 3D reconstruction, processing multiview images of heritage sites. COLMAP's advanced photogrammetry capabilities, including Structure-from-Motion (SfM) and Multi-View Stereo (MVS) techniques, enable us to accurately reconstruct the geometry of the sites. By converting a series of 2D images into a coherent 3D model, COLMAP provides a detailed and accurate foundation for the subsequent refinement and texturing processes. This step is crucial for ensuring the digital preservation and virtual representation of the heritage sites.

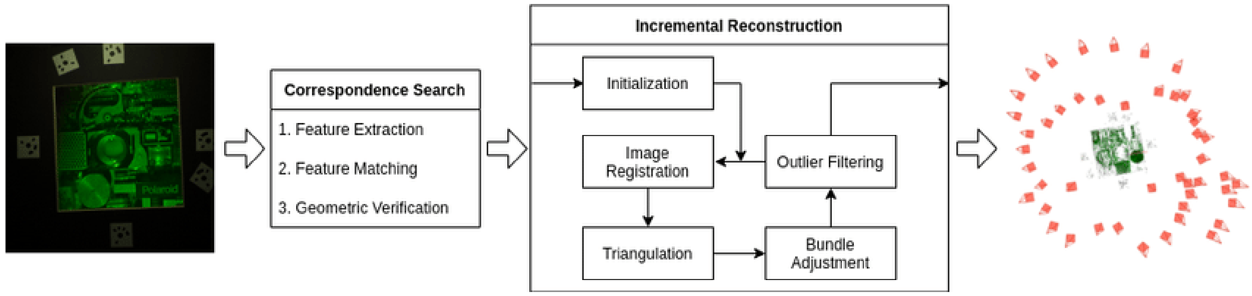

Figure 1.7: COLMAP incremental SfM pipeline for 3D reconstruction.

1.6.2 BLENDER

Once the 3D models are generated using COLMAP, we refine and enhance them in Blender, an open-source 3D creation suite renowned for its comprehensive set of tools. Blender, originally developed by Ton Roosendaal in 1995 and released as open-source software in 2002, has grown to become a leading tool in the 3D modeling community. In our project, Blender allows us to perform detailed mesh editing, ensuring that the reconstructed models are not only accurate but also visually appealing. Additionally, Blender's robust texture mapping capabilities enable us to apply realistic textures to the models, enhancing their visual fidelity. This refinement process in Blender is essential for preparing the models for integration into an interactive virtual environment, ensuring they meet high standards of detail and realism.

1.6.3 UNITY

The final step of our project involves integrating the refined and textured 3D models into Unity, a leading real-time development platform created by Unity Technologies and released in 2005. Unity allows us to create a fully interactive virtual environment where users can navigate and interact with the heritage sites using gesture-based controls. This immersive experience is further enhanced by Unity's robust support for AR/VR technologies, enabling us to provide a seamless and engaging user experience. By leveraging Unity's capabilities, we ensure that the virtual tours of the heritage sites are not only informative but also captivating, making cultural heritage more accessible and engaging for a global audience.

Chapter 2

REQUIREMENT ANALYSIS

The purpose of this project's requirement analysis is to determine the essential functional and non-functional needs required to accomplish the goals of using gesture-based controls and 3D reconstruction to create a virtual experience of heritage sites.

2.1 Functional Requirements

The functional requirements focus on the core capabilities and features that the project must implement to achieve its objectives.These consist of the procedures for 3D reconstruction, setting up virtual scenarios in AR/VR settings, including gesture-based controls.

- Gather high-quality multiview images, use photogrammetry and advanced techniques to create detailed 3D models, and apply realistic textures for visual fidelity.
- Deploy 3D models into AR/VR environments using platforms like Unity or Unreal Engine, develop a user-friendly navigation system, and ensure cross-platform compatibility.
- Implement gesture recognition technology to detect and interpret hand movements accurately, map gestures to actions, and ensure accurate calibration.
- Provide robust storage solutions for 3D models and user data, enable easy updates, and ensure secure access for different user roles.

2.2 Non Functional Requirements

The quality characteristics, performance benchmarks, and limitations that the project must follow are specified by non-functional criteria. These specifications guarantee the system's effectiveness, scalability, security, and ease of use.

- Maintain low latency for smooth interaction, achieve a high frame rate (minimum 60 FPS), optimize load times, and manage system resources efficiently.
- Support a large number of concurrent users, allow easy scaling of new heritage sites, design with a modular architecture, and utilize cloud services for increased demand.

- Ensure robust security measures for user data and content, implement secure authentication, use encryption for data transmission and storage, and adhere to privacy regulations.
- Guarantee high availability and minimal downtime, provide comprehensive error handling and recovery, implement regular backups, and use redundant systems for continuous operation.

2.3 Hardware Requirements

To ensure the successful development and deployment of the virtual experience of heritage sites through 3D reconstruction and gesture-based controls, the following hardware components are required:

High-Performance Workstations:

- CPU: Multi-core processors (e.g., Intel Core i7 or AMD Ryzen 7) to handle complex computations during 3D reconstruction and real-time rendering.
- GPU: High-end graphics cards (e.g., NVIDIA GeForce RTX 3080 or AMD Radeon RX 6800) to support photogrammetry processes and deliver smooth AR/VR experiences.
- RAM: Minimum 16GB to facilitate multitasking and processing large datasets.
- Storage: Solid State Drives (SSD) with at least 1TB capacity for fast data retrieval and storage of high-resolution images and 3D models.

3D Scanning and Imaging Equipment:

- Cameras: High-resolution DSLR or mirrorless cameras for capturing detailed multiview images of heritage sites.
- Drones: Equipped with high-resolution cameras for aerial imaging to capture comprehensive views of larger heritage sites.
- 3D Scanners: LiDAR or structured light scanners for precise 3D data collection of intricate details and textures.

AR/VR Headsets:

- VR Headsets: Devices such as the Oculus Rift, HTC Vive, or Valve Index for an immersive virtual reality experience.
- AR Headsets: Devices like Microsoft HoloLens or Magic Leap for augmented reality applications, allowing users to see digital reconstructions overlaid on the real world.

Gesture Control Devices:

- Hand Tracking Sensors: Devices such as Leap Motion Controller or built-in sensors in VR headsets for detecting and interpreting hand gestures.
- Wearable Controllers: Devices like the Oculus Touch Controllers or HTC Vive Controllers to facilitate precise and intuitive interaction within the virtual environment.

2.4 Software Requirements

To successfully develop and deploy the virtual experience of heritage sites through 3D reconstruction and gesture-based controls, the following software components are required:

3D Modeling and Reconstruction:

- Photogrammetry Software: Applications such as Agisoft Metashape, RealityCapture, or Autodesk ReCap for creating 3D models from multiview images.
- 3D Modeling Software: Tools like Blender, Autodesk Maya, or 3ds Max for refining and detailing 3D models.
- Texture Mapping Tools: Software such as Substance Painter for applying realistic textures to the 3D models.

AR/VR Development Platforms:

- Game Engines: Unity or Unreal Engine for developing and deploying AR/VR environments. These engines provide robust support for 3D graphics, physics, and interaction.
- AR Development Kits: ARCore (Google) or ARKit (Apple) for creating augmented reality applications on mobile devices.
- VR SDKs: OpenXR, Oculus SDK, or SteamVR SDK for integrating VR hardware and ensuring compatibility with various VR headsets.

Gesture Recognition Software:

- Hand Tracking SDKs: Leap Motion SDK or Ultraleap for developing hand gesture recognition and integrating it with VR/AR environments.
- Machine Learning Libraries: TensorFlow or PyTorch for implementing advanced gesture recognition algorithms.

Chapter 3

SYSTEM DESIGN

This chapter outlines the architectural framework, detailed design components, and interaction mechanisms of the project. The primary objective of this system is to achieve 3D reconstruction from 2D images of heritage sites, deploy the reconstructed models into a virtual environment, and enable intuitive walkthroughs using hand gestures.

3.1 Architecture Design

The architecture design for creating a virtual 3D environment of heritage sites involves several key steps. Initially, multiview images of the heritage sites are collected for 3D reconstruction. The reconstructed models are refined in Blender with mesh and texture mapping. Subsequently, the terrain, 3D objects, and hand gesture control modules are integrated within Unity to create the virtual environment. Finally, the project is compiled into an executable file for deployment, ensuring compatibility with AR/VR devices through the use of appropriate plugins. This process ensures a detailed and interactive virtual representation of heritage sites, enhancing user engagement and experience.

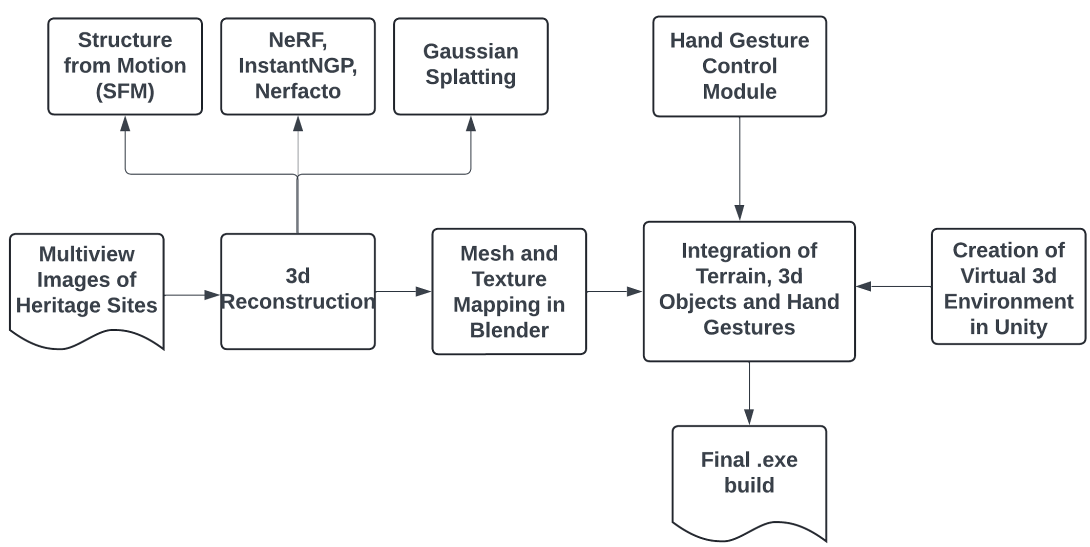

Figure 3.1: Functional block diagram for our proposed work.

3.2 Data Design

The project utilizes a comprehensive dataset of multi-view images of the Hampi temples to create detailed and accurate digital reconstructions. Each temple in the dataset is documented with a minimum of 4000 high-resolution images, captured from various angles and perspectives to ensure every intricate architectural detail and texture is preserved. These images were taken under different lighting conditions to enhance the realism of the 3D models. Additionally, Unity assets were used for terrain building, including high-quality terrain textures, vegetation models, and environmental effects. These assets facilitated the creation of realistic and immersive environments.

Chapter 4

IMPLEMENTATION

This chapter briefly describes the system's implementation details by describing each component with its code skeleton in terms of algorithm.

4.1 3D Reconstruction

In the initial phase of our project, we focused on 3D reconstruction to create accurate and detailed models from 2D images of heritage sites. We explored several advanced methods, including Structure from Motion (SfM), Neural Radiance Fields (NeRF), NeRF-Acto, and Instant Neural Graphics Primitives (Instant NGP), each with unique strengths. After a thorough evaluation, we selected SfM for its proven reliability and mature toolset, utilizing the COLMAP pipeline to reconstruct heritage models. COLMAP's robust features, such as automatic camera calibration, feature extraction, matching, and dense reconstruction, enabled us to achieve high-precision 3D models, meticulously capturing and preserving the intricate details of heritage artifacts. Once the 3D models were generated, we used Blender for further refinement, including mesh and texture mapping, before exporting the models into Unity. Blender's powerful editing tools ensured that the models were optimized and ready for integration into the virtual environment. Algorithm 1 shows the implementation details for the COLMAP pipeline, detailing each step involved in the 3D reconstruction process from image acquisition to the final model export.

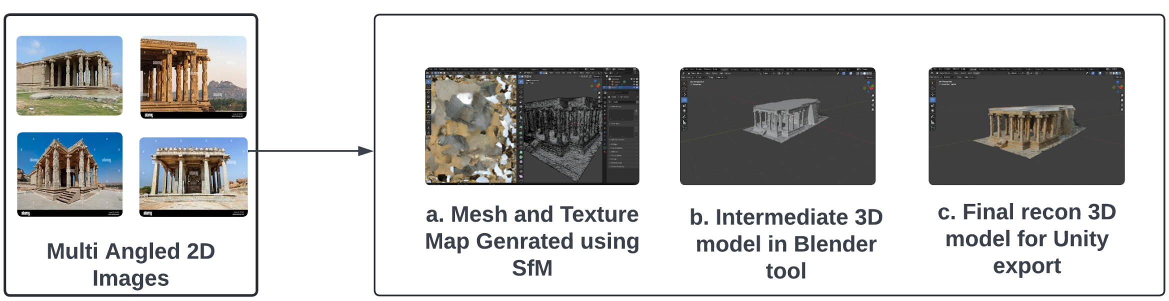

Figure 4.1: 3D reconstruction and Unity export model of Kadalekalu Ganesha temple in Hampi.

Algorithm 1 COLMAP Pipeline for 3D Reconstruction

Require: Set of 2D images `Images`
Ensure: Reconstructed 3D model `Model`
Load Images: Import all images from `Images`
for each image $I \in$ `Images` **do**
  Detect features F_I using SIFT
  Extract descriptors D_I for each feature
end for
for each pair of images (I_1, I_2) **do**
  Match features between I_1 and I_2 to find correspondences C_{I_1,I_2}
  Verify matches using RANSAC to filter out outliers
end for
Initialize reconstruction with the best image pair (I_b1, I_b2)
Estimate relative pose between I_b1 and I_b2
Triangulate initial 3D points P_{init} from C_{I_b1,I_b2}
while there are unprocessed images **do**
  Select the next image I_n
  Match features and triangulate new 3D points P_{new}
  Estimate camera pose for I_n
  Perform bundle adjustment to refine camera poses and 3D points
  Remove outliers from P_{init} and P_{new}
end while
Perform dense reconstruction:
for each pair of images (I_1, I_2) **do**
  Compute depth maps D_{I_1,I_2} using Semi-Global Matching (SGM)
end for
Fuse depth maps D_{I_1,I_2} into a dense point cloud P_{dense}
Generate surface mesh $M_{surface}$ from P_{dense}
if texture mapping is required **then**
  Apply textures from `Images` to $M_{surface}$
end if
Export Model: Save the final 3D model `Model` in a suitable format (e.g., .obj, .ply)
return `Model`

4.2 Virtual Environment in Unity

The second module involves building a virtual environment in Unity with a forest theme, featuring rocks, trees, and greenlands. We imported the 3D reconstructed heritage site models into Unity and integrated them into a sculpted landscape using Unity's terrain tools. The environment was populated with assets from the Unity store and custom models, seamlessly blending with the heritage site models. Using Unity's vegetation system, we created dynamic greenlands and applied realistic lighting for depth and immersion. Performance optimizations, including level of detail (LOD), efficient texture mapping, and occlusion culling, ensured a smooth user experience. This virtual environment serves as a realistic backdrop for the interactive walkthrough of the heritage sites.

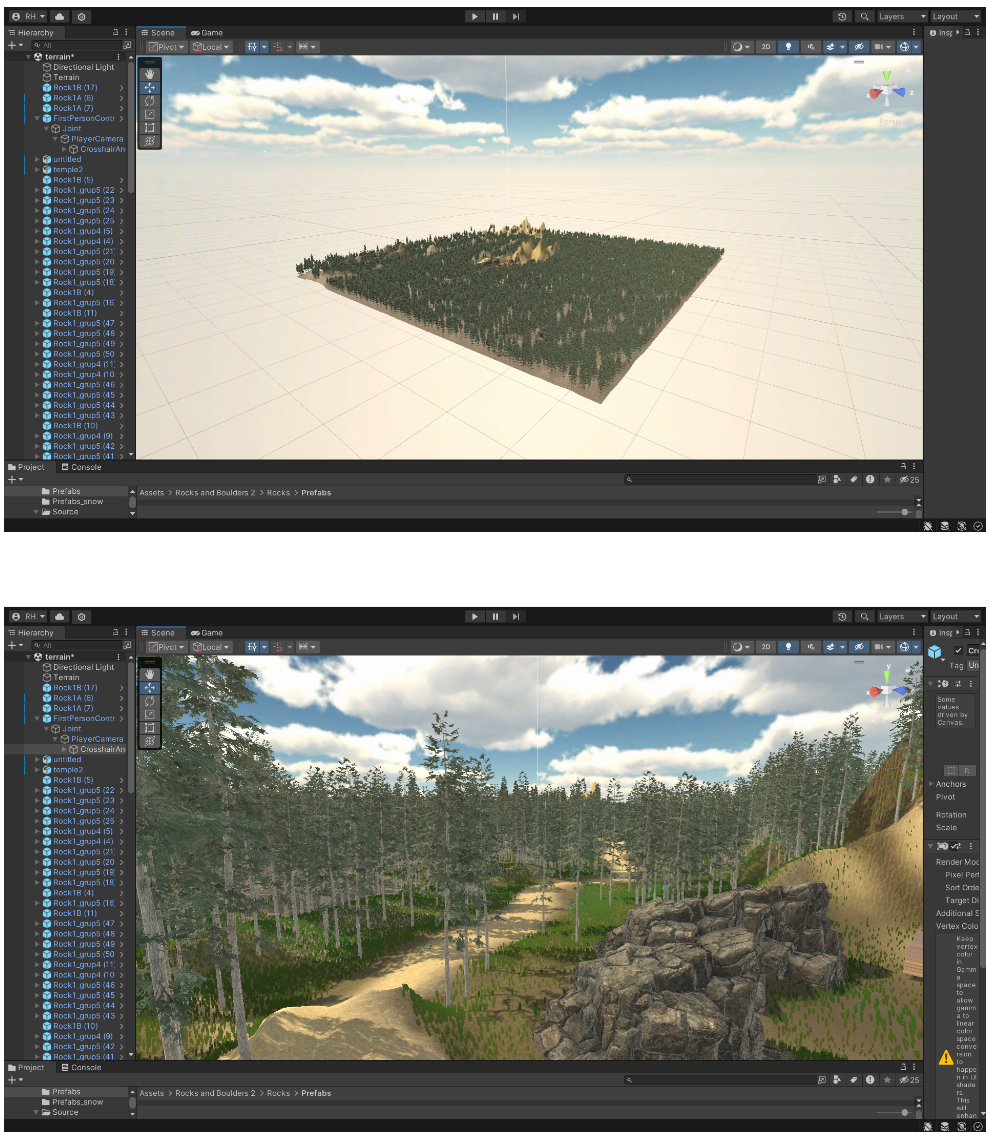

Figure 4.2: Forest theme-based terrain for the virtual environment on Unity platform

4.3 Hand Gesture Control

In the next phase, we implemented a hand gesture control system within a Unity-based environment, focusing on enhancing user interaction, particularly in virtual walkthroughs of heritage sites. The implementation commenced with integrating MediaPipe's hand tracking module, enabling real-time capture and processing of video frames to detect hand landmarks accurately. We developed a custom gesture recognition pipeline using machine learning classifiers to leverage these landmarks. This allowed us to interpret gestures such as swipes, pinches, and thumbs-ups, translating them into meaningful commands within our Unity application. The primary application scenario involved creating immersive virtual walkthroughs of heritage sites, where users could explore and interact with historical locations using intuitive hand gestures. Users could navigate virtual environments, inspect architectural details, and trigger contextual information about historical landmarks by performing specific gestures. Visual feedback mechanisms were incorporated to provide immediate confirmation of gesture recognition, ensuring a seamless and engaging user experience. Algorithm 2 shows the implementation of the hand gesture-based control.

Algorithm 2 Hand Gesture-Based Movements Control

```
Require: Video frames with hand landmarks
Ensure: Detected gesture command
Initialize MediaPipe hand tracking module
Load pre-trained neural network model for gesture classification
while video frames available do
    Capture and preprocess video frame
    Detect hand landmarks using MediaPipe
    Extract features from detected landmarks
    Classify gesture using the neural network model
    for each detected gesture do
        Calculate confidence score
        if Confidence score above threshold then
            Interpret gesture command
            Execute the corresponding action in Unity environment
        end if
    end for
end while
```

4.4 AR/VR Integration

To ensure our Unity 3D scenes are seamlessly compatible with AR/VR devices, we employ essential plugins and tools such as Unity MARS, Vuforia Engine, and AR Foundation. These platforms facilitate the transformation of our digital environments into immersive AR/VR spaces, enabling us to overlay virtual elements in real-world settings. This approach enhances spatial interaction and supports various devices like Google ARCore and Apple ARKit, thereby enriching our exploration and presentation of heritage sites through augmented reality experiences. Additionally, frameworks like ARToolkit and EasyAR provide advanced functionalities such as image recognition and 3D tracking, further enhancing our ability to create compelling and interactive AR applications for heritage site preservation and education.

Chapter 5

RESULTS AND DISCUSSIONS

In this section, we present the final build of our project, showcasing the integration of refined 3D models of heritage sites and the implementation of hand gesture-based walkthroughs in the virtual environment.

5.1 Results

The final build of our project showcases an immersive virtual environment where users can explore reconstructed heritage sites, including ancient temples. Leveraging the capabilities of Unity, we have created an interactive walkthrough experience that supports hand gesture controls, enhancing the user's engagement and interaction with the virtual space. Users can navigate through the temple sites, examine intricate details of the architecture, and access additional information about the heritage using simple hand gestures. This intuitive interaction model, combined with high-quality 3D models and realistic textures, provides a rich and educational experience, making it possible for users to virtually visit and appreciate these cultural landmarks from anywhere in the world.

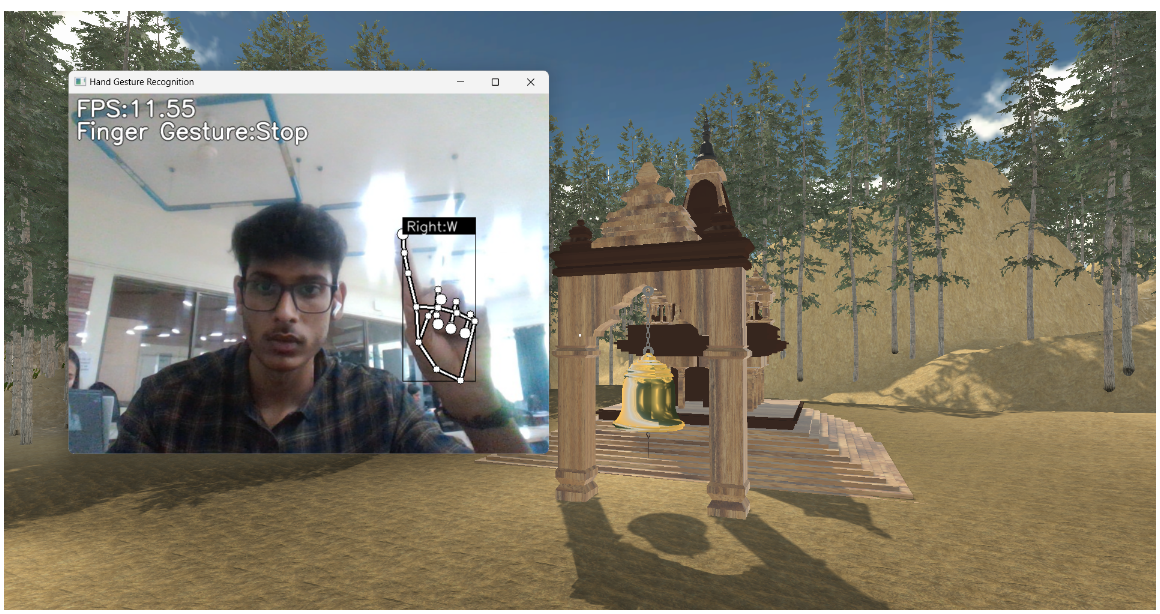

Figure 5.1: Hand gesture-based walkthrough of heritage site.

Figure 5.2: Deployed 3D reconstructed Kadalekalu Ganesh Temple in virtual environment of Unity platform.

Figure 5.3: Ayodhya Ram Mandir deployed in 3D virtual environment.

Chapter 6

CONCLUSION AND FUTURE SCOPE

In this chapter, we summarize the key findings of our project and outline potential future developments to enhance and expand its impact.

6.1 Conclusion

This project highlights the potential of integrating AR, VR, and gesture-based controls for the digital preservation and presentation of cultural heritage sites. The developed platform preserves the intricate details of heritage sites while providing an engaging and educational tool for users. By making these sites accessible in a virtual format, we contribute to cultural preservation, education, and tourism, offering new opportunities for global audiences to connect with historical landmarks. This innovative approach underscores the importance of technology in preserving and promoting cultural heritage in the digital age.

6.2 Future Scope

- **Enhanced User Interaction:**
  - Develop more sophisticated gesture recognition algorithms to improve the accuracy and responsiveness of user interactions.
  - Integrate voice commands and other sensory inputs to create a more immersive and interactive user experience.
- **Expanded Heritage Site Library:**
  - Extend the project's scope to include a broader range of heritage sites from different cultures and regions.
  - Collaborate with cultural institutions and historians to ensure the authenticity and accuracy of the digital reconstructions.
- **Educational Applications:**
  - Create educational modules and interactive lessons based on the virtual heritage sites to support history and cultural studies in schools and universities.

  - Develop guided virtual tours led by experts, providing in-depth knowledge and insights about the heritage sites.

- **Accessibility Enhancements:**
  - Optimize the platform for various AR/VR devices, including mobile devices, to ensure broader accessibility.
  - Implement features to support users with disabilities, ensuring the virtual tours are inclusive and accessible to everyone.

REFERENCES

[1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis, 2020.

[2] G. Mazzacca, A. Karami, S. Rigon, E. M. Farella, P. Trybala, and F. Remondino. Nerf for heritage 3d reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023:1051–1058, 2023.

[3] Moch Kholil, I Ismanto, and M Fu'ad. 3d reconstruction using structure from motion (sfm) algorithm and multi view stereo (mvs) based on computer vision. *IOP Conference Series: Materials Science and Engineering*, 1073:012066, 02 2021.

[4] Cosmina Isar. A glance into virtual reality development using unity. *Informatica Economica*, 22(3):14–22, 2018.

[5] Vincenzo Rinaldi, Karen Ann Robertson, Graham George Strong, and Niamh Nic Daéid. Examination of fire scene reconstructions using virtual reality to enhance forensic decision-making. a case study in scotland. *Virtual Reality*, 28(1):57, 2024.