UNIVERSITÀ DEGLI STUDI DI CATANIA

DIPARTIMENTO DI INGEGNERIA ELETTRICA ELETTRONICA E INFORMATICA

PhD in ENERGY SYSTEMS, COMPUTER SCIENCE AND TELECOMMUNICATIONS ENGINEERING

Efficient Photogrammetric Reconstruction from 360 Images for Enhanced Scene Exploration and Immersive Observation

AUTHOR

Giovanni Lo Monaco

ADVISOR

Prof. Sergio Palazzo

COMPANY TUTOR

Ing. Valeria D'Amico

XXXVII Cycle

ACADEMIC YEAR 2023 - 2024

Contents

Introduction **5**

1 General Overview of IoT Technologies **7**

1.1 Introduction to the Internet of Things (IoT) 7

1.1.1 Definition of IoT 7

1.1.2 Brief History and Development of IoT 8

1.1.3 Importance and Impact of IoT in Modern Society 10

1.2 Architecture and Components of an IoT System 11

1.2.1 Sensors and Actuators 11

1.2.2 Communication Networks 13

1.2.3 Data Management Platforms 15

1.2.4 Security and Privacy in IoT 19

1.3 IoT Communication Technologies: Focus on LoRa and NB-IoT 20

1.3.1 LoRa (Long Range) 21

1.3.2 NB-IoT (Narrowband IoT) 24

1.3.3 Comparison between LoRa and NB-IoT 28

2 Overview of eXtended Reality (XR) Technologies **30**

2.1 Introduction to XR 30

2.1.1 Definition of XR 30

2.1.2 Components of XR: VR, AR, and MR 31

2.1.3 Sense of Immersion 35

2.2 XR Technologies and Devices 39

2.2.1 HMD (Head-Mounted Displays) 39

2.2.2 Tracking Systems 41

2.2.3 Interfaces and Interaction Tools 42

2.2.4 Difference Between 3DoF and 6DoF 44

2.2.5 Field of View (FoV) 45

2.2.6 Motion Sickness 46

2.3 Applications of XR 47

2.3.1 Gaming 47

2.3.2 Medicine 48

2.3.3 Education . . . 48
2.3.4 Industry . . . 49
2.3.5 Training . . . 49

3 State of the Art in Photogrammetry and Innovative Techniques 51
3.1 Introduction to Photogrammetry . . . 51
3.1.1 Definition of Photogrammetry . . . 51
3.1.2 Traditional Photogrammetry Techniques . . . 52
3.2 Photogrammetry Workflow and Tools . . . 54
3.2.1 Standard Photogrammetry Workflow . . . 54
3.2.2 Tools and Software for Photogrammetry . . . 63
3.3 Applications and Case Studies . . . 66
3.3.1 Archaeology . . . 66
3.3.2 Architecture and Engineering . . . 66
3.3.3 Extended Reality (XR) and Virtual Reality . . . 67
3.3.4 Forensic Science . . . 68
3.3.5 Medical Applications . . . 69
3.3.6 Environmental Monitoring . . . 70
3.4 Innovative Techniques in Photogrammetry . . . 70
3.4.1 Deep Learning in Photogrammetry . . . 70
3.4.2 Real-time Photogrammetry . . . 71
3.4.3 Neural Radiance Fields (NeRF) . . . 72
3.5 Other Techniques for 3D Reconstruction . . . 72
3.5.1 Laser Scanning . . . 73
3.5.2 Computer Vision-based Reconstruction . . . 75

4 Enhancing Quality of Experience (QoE) through IoT, XR, and Photogrammetry 76
4.1 Definition of QoE . . . 76
4.1.1 Techniques for Measuring QoE . . . 77
4.1.2 Distinguishing QoS and QoE . . . 80
4.2 Quality of Experience (QoE) vs. User Experience (UX) . . . 83
4.2.1 User Experience (UX) . . . 83
4.2.2 Distinguishing QoE from UX . . . 84
4.3 Intermediate Position of QoE between QoS and UX . . . 85

5 Doctoral Research Progress and Applied Case Study Analysis 87
5.1 Introduction to the Doctoral Research Path . . . 87
5.2 Research Objectives and Methodology . . . 88
5.2.1 Motivation for Technology Choices . . . 88
5.2.2 Methodological Approach . . . 88
5.3 Research Timeline and Milestones . . . 89

5.4 Overview of Case Studies . . . 91
5.5 Case Study 1: Advancing Proof of Concept Demonstrations through IoT and XR Integration . . . 92
5.5.1 Background . . . 92
5.5.2 State of the Art . . . 94
5.5.3 Motivation and Objectives . . . 95
5.5.4 Tools and Software Utilized . . . 97
5.5.5 Methodology . . . 98
5.5.6 Results Discussion . . . 109
5.5.7 Conclusions . . . 116
5.6 Case Study 2: Efficient 3D Reconstruction of Indoor Environment Using Equirectangular Images in Photogrammetry . . . 117
5.6.1 Background . . . 117
5.6.2 State of the Art . . . 118
5.6.3 Motivation and Objectives . . . 119
5.6.4 Tools and Software Utilized . . . 120
5.6.5 Methodology . . . 125
5.6.6 Results Discussion . . . 136
5.6.7 Conclusions . . . 146
5.7 Case Study 3: Comparative Analysis of Distance Estimation in Spherical Images and 3D Photogrammetry Models . . . 148
5.7.1 Background . . . 148
5.7.2 State of the Art . . . 149
5.7.3 Motivation and Objectives . . . 150
5.7.4 Tools and Software Utilized . . . 151
5.7.5 Methodology . . . 154
5.7.6 Results Discussions . . . 161
5.7.7 Conclusions . . . 162
5.8 Conclusions . . . 163

Conclusion **164**

A Case Study 2 Metashape Reports **165**
A.1 DSLR Image Processing Report . . . 165
A.2 Equirectangular Images Processing Report . . . 168
A.3 Equirectangular Video Processing Report . . . 171

B Case Study 3 Questionnaires **174**
B.1 Igroup Presence Questionnaire (IPQ) [159] . . . 175
B.2 Simulator Sickness Questionnaire (SSQ) [160] . . . 176
B.3 Depth Perception Questionnaire (DPQ) [158] [140] . . . 177

B.4 Image-Lighting Visual Realism Questionnaire (ILVRQ) 178
B.5 Visual Distortion Questionnaire (VDQ) 179

Introduction

The rapid advancements in photogrammetry and immersive technologies have opened up new possibilities for efficiently reconstructing and exploring complex environments in virtual spaces. This research focuses on developing efficient methods for photogrammetric reconstruction using 360-degree (equirectangular) images, with the goal of enhancing scene exploration and creating immersive, high-quality virtual experiences. By leveraging 360-degree imaging, the study aims to provide faster, cost-effective solutions that allow for high-fidelity virtual observation of real-world environments, particularly those that are difficult to capture with conventional photogrammetric methods.

The approach explored in this thesis centers on the application of equirectangular images for photogrammetric modeling, which offers a unique advantage in rapid acquisition and comprehensive scene coverage. Photogrammetry, a technique that generates 3D models from 2D images, has traditionally required extensive image sets and significant processing time. However, the use of 360-degree images can streamline this process, offering a more efficient method for capturing spatial data, reducing both acquisition and processing times, and maintaining the level of detail necessary for immersive observation and analysis.

A substantial portion of this research investigates and evaluates the outcomes of this approach through three distinct case studies, presented in Chapter 5:

- **Case Study 1: Advancing Proof of Concept Demonstrations through IoT and XR Integration** - This study focuses on creating an effective method for presenting proof of concept (PoC) demonstrations within a company setting, using foundational photogrammetry techniques to model and present digital scenes. This initial study provided essential insights into basic photogrammetric methods, laying the groundwork for further advancements.
- **Case Study 2: Rapid 3D Reconstruction of Indoor Environment Using Equirectangular Images in Photogrammetry** - This study presents a significant scientific contribution by establishing a standardized operational procedure for using 360-degree images in photogrammetry to reconstruct indoor spaces. To the best of current knowledge, no prior research has documented this specific application. The results highlight the potential of 360-degree imaging for reducing the time and resources required for 3D reconstruction, making it suitable for applications such as virtual tours, crime scene

documentation, and architectural surveys.

- **Comparative Analysis of Distance Estimation in Spherical Images and 3D Photogrammetry Models** - Building on the results from Case Study 2, this final study assesses the effectiveness of spherical images for photogrammetric reconstructions by comparing user perception and depth estimation in spherical images and traditional photogrammetric 3D models. This comparative analysis aims to validate the applicability of equirectangular images for creating realistic, immersive environments, with implications for enhancing the accuracy and usability of virtual reconstructions.

By examining efficient photogrammetric methods through the innovative use of 360-degree imaging, this thesis aims to contribute a streamlined and effective approach to immersive observation. These findings establish a foundation for future studies in reality virtualization, offering insights into alternative photogrammetric methods that prioritize efficiency, realism, and accessibility in virtual scene exploration.

Chapter 1

General Overview of IoT Technologies

1.1 Introduction to the Internet of Things (IoT)

1.1.1 Definition of IoT

The Internet of Things (IoT) refers to the vast network of physical objects—devices, vehicles, appliances, and other items—embedded with sensors, software, and other technologies that enable them to connect and exchange data with other devices and systems over the internet. These interconnected devices are often referred to as ”smart” devices due to their ability to gather data, communicate with other devices, and make autonomous decisions based on pre-defined rules or real-time data analysis.

IoT devices can range from everyday household items like smart thermostats, refrigerators, and lighting systems to complex industrial machinery and systems used in manufacturing, transportation, and healthcare. The key components of IoT include:

- **Sensors and Actuators:** Sensors collect data from the physical environment (for example, temperature, humidity, motion), while actuators perform actions based on this data (for example, adjusting a thermostat, turning on lights).
- **Connectivity:** IoT devices communicate with each other and with centralized systems through various communication technologies such as Wi-Fi, Bluetooth, Zigbee, cellular networks (for example, 4G, 5G), and low-power wide-area networks (LPWAN) like LoRa and NB-IoT.
- **Data Processing:** Collected data is processed either locally (on the device) or in centralized data centers (cloud computing), where it can be analyzed to derive insights and inform decision-making.
- **User Interface:** Interaction between users and IoT devices is facilitated through various interfaces, including mobile applications, web dashboards, and voice commands.

The evolution of IoT is driven by advancements in technology such as miniaturization of

sensors, increased computing power, improved connectivity, and the proliferation of cloud and edge computing. IoT powers diverse applications, ranging from smart homes to industrial automation.

The transformative potential of IoT lies in its ability to create an interconnected ecosystem where physical and digital worlds merge, leading to improved efficiency, productivity, and quality of life. However, the widespread adoption of IoT also raises significant challenges, particularly in the areas of security, privacy, and data management. Addressing these challenges is critical to ensuring the sustainable and secure growth of IoT technologies.

1.1.2 Brief History and Development of IoT

The Internet of Things (IoT) has evolved significantly since its inception, marked by various milestones and technological advancements. This section provides a chronological overview of the key developments in the history of IoT.

Early Beginnings

The concept of connecting physical objects to the internet dates back to the early days of the internet itself. In 1989, the idea of connecting ”things” to the internet began to take shape. One of the earliest examples is the Trojan Room coffee pot, possibly the first application of its kind, which was created to monitor the status of a coffee pot remotely [1].

In 1990, John Romkey created the first internet-connected device, a toaster that could be turned on and off over the internet [2]. This demonstration highlighted the potential of connecting everyday objects to the internet for remote control and automation.

Pioneering Innovations

In 1994, Steve Mann invented WearCam [3], a wearable camera system with near-real-time performance using a 64-processor system. This innovation showcased the possibilities of wearable technology and real-time data processing.

Paul Saffo, in 1997, provided an early description of sensors and their future applications, foreshadowing the integral role that sensors would play in the IoT ecosystem [4].

Coining the Term ”Internet of Things”

The term ”Internet of Things” was coined by Kevin Ashton in 1999 while working as the executive director of the Auto-ID Center at MIT. In the same year, the center developed a global RFID-based item identification system, laying the groundwork for the widespread use of RFID technology in IoT applications [5].

Commercialization and Early Adoption

In 2000, electronics giant LG announced plans to release a smart refrigerator capable of determining whether the food items stored within needed replenishment [6]. This marked a significant step towards the commercialization of IoT technology in consumer appliances.

The year 2003 saw the deployment of RFID technology on a massive scale by the US Army in their Savi program. Retail giant Walmart also began using RFID technology extensively in its stores worldwide, demonstrating the technology's potential for supply chain management and inventory tracking.

Mainstream Recognition and Standardization Efforts

By 2005, mainstream publications such as The Guardian, Scientific American, and the Boston Globe were featuring articles about IoT and its future potential, indicating growing public and academic interest in the field.

In 2008, a group of companies launched the IPSO Alliance to promote the use of Internet Protocol (IP) in networks of smart objects, further advancing the standardization and interoperability of IoT devices. Additionally, the FCC approved the usage of the "white space spectrum," enabling new possibilities for IoT connectivity.

Technological Milestones and Industry Adoption

The launch of IPv6 in 2011 was a critical milestone that triggered massive growth and interest in IoT. IPv6 provided a virtually limitless number of IP addresses, essential for the large-scale deployment of IoT devices.

Following this, major IT companies like Cisco, IBM, and Ericsson took significant educational and commercial initiatives to promote IoT. These companies invested heavily in IoT research, development, and deployment, contributing to the rapid advancement of IoT technologies.

Current Trends and Future Directions

In recent years, IoT has continued to evolve with advancements in artificial intelligence (AI), machine learning (ML), and edge computing. These technologies enable IoT devices to process data locally, reducing latency and enhancing real-time decision-making capabilities [7].

Key trends in IoT include the integration of AI and ML for advanced analytics, the adoption of edge computing for efficient data processing, and the deployment of 5G networks to provide ultra-low latency and high-bandwidth connectivity. Additionally, blockchain technology is being explored to enhance IoT security and data integrity [8].

The journey of IoT from its early conceptualization to its current state has been marked by continuous innovation and the convergence of multiple technological domains. As IoT contin-

ues to mature, it promises to transform industries, improve quality of life, and drive economic growth, while also presenting challenges related to security, privacy, and standardization.

1.1.3 Importance and Impact of IoT in Modern Society

The Internet of Things (IoT) is revolutionizing various aspects of modern society by integrating everyday objects with the internet, enabling seamless communication and interaction. This technological advancement has profound implications across multiple domains, enhancing efficiency, productivity, and quality of life.

The study conducted an analysis on the impact of the Internet of Things (IoT) on social change in Japan, Germany, and Australia. The findings revealed that the IoT positively influences social change in all three countries, albeit with some variations in the specific impacts observed [9]. The impact of IoT technologies is largely positive, promoting job creation, enhancing research and development, and improving business operations. The changes are incremental, reflecting a gradual adaptation and integration of IoT technologies within these societies.

Economic Impact

IoT technologies have a significant economic impact, driving innovation and growth in various industries. By facilitating better data collection and analysis, IoT helps businesses optimize operations, reduce costs, and improve decision-making processes. For instance, smart manufacturing systems use IoT to monitor machinery in real-time, predict maintenance needs, and minimize downtime, resulting in more efficient production lines and higher output. The retail sector also benefits from IoT through inventory management systems that track stock levels and customer preferences, enabling more personalized and timely services [10].

Social and Cultural Impact

On the social front, IoT technologies are transforming how individuals interact with their environments and each other. Smart homes equipped with IoT devices, such as smart thermostats, security systems, and home assistants, provide enhanced comfort, security, and energy efficiency. These devices learn user preferences and behaviors, making homes more responsive and adaptive to the needs of their inhabitants. Moreover, smart city initiatives leverage IoT to improve urban living conditions by managing traffic flow, reducing energy consumption, and enhancing public safety through connected infrastructure [10].

Health and Well-being

In the healthcare sector, IoT plays a crucial role in enhancing patient care and operational efficiency. Wearable devices monitor vital signs and track fitness activities, providing valuable data for preventive healthcare and personalized treatment plans. Hospitals use IoT to manage

medical equipment, monitor patient conditions remotely, and streamline workflows, which improves the quality of care and reduces the burden on healthcare providers [10].

Environmental Sustainability

IoT also contributes to environmental sustainability by enabling more efficient resource management. Smart grids, for example, optimize energy distribution and consumption, reducing waste and lowering carbon footprints. IoT-enabled agricultural practices, such as precision farming, use sensors and data analytics to monitor soil conditions, weather patterns, and crop health, leading to more sustainable farming practices and higher yields [10].

Challenges and Considerations

Despite its benefits, the widespread adoption of IoT poses several challenges. Privacy and security concerns are paramount, as the vast amount of data generated by IoT devices can be susceptible to breaches and misuse. Ensuring robust cybersecurity measures and protecting user data is crucial to maintaining trust and preventing potential risks associated with IoT technologies. Additionally, the digital divide must be addressed to ensure that the benefits of IoT are accessible to all segments of society, avoiding exacerbation of existing inequalities [10].

In conclusion, the IoT holds transformative potential for modern society, offering numerous economic, social, and environmental benefits. However, it is essential to address the associated challenges to harness its full potential responsibly and equitably. As IoT technologies continue to evolve, they will undoubtedly play an increasingly central role in shaping the future of our interconnected world.

1.2 Architecture and Components of an IoT System

1.2.1 Sensors and Actuators

Sensors and actuators are fundamental components of any IoT system. Sensors collect data from the environment, while actuators perform actions based on the data and instructions received. Together, they enable the seamless interaction between the physical and digital worlds, forming the backbone of IoT applications.

Sensors are devices that detect and measure physical properties such as temperature, humidity, light, motion, pressure, and more. They convert these physical parameters into electrical signals that can be processed and analyzed. There are various types of sensors, including:

- **Temperature Sensors:** Measure the temperature of the environment. Common types include thermocouples, resistance temperature detectors (RTDs), and thermistors.

- **Humidity Sensors:** Detect the moisture level in the air. Capacitive and resistive humidity sensors are widely used.
- **Light Sensors:** Measure the intensity of light. Photodiodes and light-dependent resistors (LDRs) are common examples.
- **Motion Sensors:** Detect movement or acceleration. These include accelerometers, gyroscopes, and passive infrared (PIR) sensors.
- **Pressure Sensors:** Measure the pressure of gases or liquids. Common types include piezoelectric and capacitive pressure sensors.
- **Proximity Sensors:** Detect the presence or absence of an object near the sensor. Examples include capacitive, inductive, and ultrasonic proximity sensors.

Sensors play a critical role in gathering real-time data, which is essential for making informed decisions and automating processes in IoT systems.

Actuators are devices that perform actions or initiate responses based on the signals received from the IoT system. They convert electrical signals into physical movement or action. Common types of actuators include:

- **Electric Motors:** Convert electrical energy into mechanical motion. They are widely used in robotics, industrial machines, and home appliances.
- **Solenoids:** Electromechanical devices that convert electrical energy into linear motion. Commonly used in valves, switches, and relays.
- **Relays:** Electrically operated switches that control a circuit by opening or closing contacts. Used in automation and control systems.
- **Pneumatic Actuators:** Use compressed air to create mechanical motion. Often used in industrial applications for controlling valves and other mechanical systems.
- **Hydraulic Actuators:** Use hydraulic fluid to produce mechanical motion. Commonly used in heavy machinery and industrial equipment.

The rapid advancements in sensor and actuator technologies have proven crucial in the development of smart cities. These technologies seamlessly integrate into IoT frameworks, forming the backbone of data-driven urban environments. Sensors, designed for various IoT applications, ensure precision and durability, essential for reliable data collection in urban management. Actuators complement sensors by enabling real-time adjustments, enhancing responsive urban management practices. Data security is paramount in these systems, necessitating robust information-sharing mechanisms to uphold data integrity and privacy. Recent developments include the incorporation of blockchain technology and AI-driven solutions, promising further improvements in urban efficiency and sustainability [11].

Figure 1.1: Graphical abstract illustrating advancements in sensor and actuator technologies.

1.2.2 Communication Networks

The effectiveness of IoT in smart cities heavily relies on robust communication networks that facilitate seamless data exchange. Various technologies such as Wi-Fi, Bluetooth, Zigbee, LoRa, NB-IoT, NFC, and Cellular Networks play pivotal roles in enabling connectivity across urban landscapes.

Wi-Fi

Wi-Fi remains a cornerstone technology in urban IoT deployments due to its high bandwidth and widespread availability. It supports high-speed data transmission, making it suitable for applications requiring real-time data processing and analysis.

Bluetooth

Bluetooth technology, known for its low energy consumption and short-range communication capabilities, finds applications in localized IoT networks within smart city infrastructures. It is often used for connecting devices in close proximity, such as in smart homes and personal area networks.

Zigbee

Zigbee operates on low-power, low-data-rate wireless mesh networks, making it ideal for smart city applications that prioritize energy efficiency and scalability. Its mesh networking capability allows for extensive coverage and resilience against network disruptions.

LoRa (Long Range)

LoRa technology excels in providing long-range communication with low power consumption, making it suitable for wide-area IoT deployments in smart cities. It supports applications like environmental monitoring, smart agriculture, and asset tracking.

NB-IoT (Narrowband IoT)

NB-IoT is designed for low-power, wide-area networks, offering enhanced coverage and penetration compared to traditional cellular networks. It is well-suited for applications requiring reliable connectivity over large urban areas, such as smart metering and smart parking systems.

NFC (Near Field Communication)

NFC enables short-range wireless communication between devices, typically within a few centimeters. It is commonly used for contactless payments, access control, and identification in smart city applications like public transportation and secure entry systems.

Cellular Networks

Traditional cellular networks (for example, 4G LTE, 5G) continue to play a crucial role in urban IoT deployments, offering wide coverage and high-speed data transmission capabilities. They support various IoT applications requiring ubiquitous connectivity, such as real-time traffic management, remote monitoring, and emergency response systems.

Comparison and Integration

Each communication technology in IoT networks offers unique advantages and trade-offs. Integration of multiple technologies enables smart cities to leverage diverse capabilities, optimizing network performance based on specific application requirements. The table below summarizes key aspects of different communication technologies used in smart city deployments:

Technology	Coverage	Data Rate	Power Consumption	Applications
Wi-Fi	High	High	Moderate	Real-time data processing
Bluetooth	Low	Low	Low	Personal area networks
Zigbee	Medium	Low	Low	Energy-efficient mesh networks
LoRa	Long-range	Low	Very low	Wide-area sensor networks
NB-IoT	Wide-area	Low	Low	Smart metering, parking systems
NFC	Short-range	Low	Very low	Contactless payments, access control
Cellular Networks	Very wide	High	High	Ubiquitous connectivity

Table 1.1: Comparison of communication technologies in smart city IoT deployments.

The choice of communication network depends on factors like coverage area, data rate requirements, power consumption constraints, and scalability needs. This ensures efficient

and sustainable urban IoT deployments tailored to specific application scenarios.

The scientific literature continues to explore a wide array of IoT communication technologies, from established protocols like Ethernet and IP to innovative solutions such as LoRa and 5G cellular networks. These technologies cater to diverse IoT applications, ensuring connectivity, security, and efficiency across varied environments and scales. Ongoing research focuses on optimizing these technologies to meet evolving demands in IoT deployments worldwide [12].

1.2.3 Data Management Platforms

Limitations of Traditional Approaches

Effective management of data is crucial for maximizing the potential of IoT deployments in smart cities. Data management platforms play a pivotal role in collecting, storing, processing, and analyzing vast amounts of data generated by IoT devices.

Traditionally, data management encompasses the architectures, practices, and procedures governing the lifecycle of data within a system. In the IoT context, data management serves as an essential intermediary layer between data-generating objects and the applications that analyze and utilize this data. Given the diverse, heterogeneous, and voluminous nature of data generated by IoT devices, traditional Relational Database Management Systems (RDBMS) are often inadequate for several reasons.

Firstly, RDBMS are optimized for structured, predefined data sources and employ strict normalization rules for data storage. In contrast, IoT environments involve a multitude of data sources such as sensors, RFIDs, and embedded systems, each generating continuous streams of data that do not fit neatly into predefined schemas.

Secondly, traditional RDBMS primarily handle static data storage, retrieval, and basic transactional operations. In contrast, IoT data management systems must manage real-time data ingestion, storage, and processing with capabilities for dynamic data updates, logging, and audit trails to support real-time analytics and decision-making.

Unlike traditional databases where queries and updates are periodic and relatively simple, IoT data management systems must handle continuous and complex data flows from myriad sources. This necessitates flexible data models that can accommodate various data types and adapt to evolving application requirements without the rigid constraints of traditional relational schema and normalization practices.

Nevertheless, principles from traditional RDBMS remain pertinent in IoT data management. These include techniques for integrating remote storage at the object level, supporting diverse and unstructured data formats, and relaxing ACID (Atomicity, Consistency, Isolation, Durability) properties to ensure data availability, consistency, and integration across IoT platforms. Despite deviating from the transparency norms of traditional DBMS, IoT data management adheres to principles that dictate logical and physical data management structures tailored to

the dynamic and varied demands of IoT applications.

The proliferation of devices such as ubiquitous sensors and RFID readers in IoT systems leads to rapid data generation. This exponential growth in data volume necessitates efficient data storage solutions that handle large-scale data and ensure scalability. Moreover, IoT data originates from diverse sources and includes both structured and unstructured formats, posing a significant challenge for storage mechanisms that must manage heterogeneous data resources effectively.

Traditional approaches where all data is stored permanently in centralized systems are inadequate for IoT applications due to their inability to cope with the dynamic nature and sheer volume of data generated continuously. To address these challenges, researchers are exploring innovative alternatives to traditional data storage methods. Streaming and real-time databases have emerged as crucial solutions in IoT environments, where data is processed and analyzed in real-time as it flows from sensors and devices.

However, the design of these database systems traditionally revolves around centralized storage architectures, which may not be optimal for the decentralized and scalable requirements of IoT data. As a result, the development of effective data management systems for IoT necessitates robust mechanisms for updating and accessing data efficiently across distributed networks. Ongoing research underscores the critical need for adaptable and scalable data storage solutions tailored specifically to the unique demands of IoT ecosystems [13]–[15].

Data Collection and Storage

Once collected, IoT data undergoes processing and analysis to extract actionable insights. Data management platforms employ advanced analytics techniques including machine learning algorithms and statistical models. This enables smart cities to derive valuable information for decision-making, predictive maintenance, and optimization of urban services.

1. **Decision-Making:** By analyzing real-time and historical data, city administrators can make informed decisions promptly. This includes optimizing resource allocation, improving public safety measures, and enhancing overall city operations.
2. **Predictive Maintenance:** IoT data analytics allows for predictive maintenance of critical infrastructure and public utilities. By analyzing sensor data, cities can forecast equipment failures or performance degradation, enabling proactive maintenance actions to prevent service disruptions and reduce operational costs.
3. **Urban Service Optimization:** Smart cities utilize data-driven insights to optimize urban services such as transportation systems, waste management, energy distribution, and environmental monitoring. By understanding usage patterns and demand fluctuations, cities can deploy resources more efficiently and improve service delivery.

The role of Big Data in this context is paramount. IoT generates massive volumes of data from

diverse sources. Big Data technologies enable the collection, storage, and analysis of this data at scale, providing valuable insights into urban dynamics, infrastructure performance, and citizen behaviors. These insights drive evidence-based decision-making and facilitate proactive management of city resources and services [16].

Data Processing and Analysis

Data processing and analysis are essential for realizing the full potential of IoT deployments, where vast amounts of data are generated from diverse sources such as sensors, devices, and networks. These processes involve sophisticated techniques to extract actionable insights and facilitate decision-making across various domains.

1. **Real-Time Processing:** IoT data arrives in real-time streams from sensors and devices embedded in smart city infrastructures. Real-time processing includes data ingestion, filtering, and aggregation to detect immediate events or anomalies. This capability is crucial for applications requiring rapid responses, such as emergency management and traffic control systems.

2. **Batch Processing:** Alongside real-time streams, IoT systems accumulate extensive historical data that necessitates batch processing. This method involves analyzing data in chunks over longer periods, essential for applications like predictive maintenance, which relies on identifying trends and patterns to forecast equipment failures.

3. **Data Integration and Fusion:** IoT environments gather data from heterogeneous sources, such as sensors, RFID tags, and mobile devices. Data integration merges these diverse datasets into unified formats, ensuring consistency and completeness. Data fusion techniques combine multiple streams to enhance data quality and extract precise insights.

4. **Advanced Analytics:** Machine learning algorithms and statistical models are pivotal in analyzing IoT data. These techniques uncover hidden patterns, correlations, and anomalies within data streams. Machine learning models classify data, make predictions, and offer prescriptive recommendations based on historical and real-time analysis.

5. **Visualization and Reporting:** Effective visualization tools and dashboards convert complex IoT data into intuitive graphs, charts, and reports. Visualization aids stakeholders in identifying trends, anomalies, and making informed decisions. Real-time dashboards provide current insights into operational metrics, improving situational awareness and performance monitoring.

Advancements in cloud computing and edge computing technologies bolster IoT data processing and analysis capabilities. Cloud platforms provide scalable resources for storing and processing data, facilitating complex analytics and machine learning tasks. Edge computing decentralizes data processing, reducing latency and enabling faster decision-making in

distributed IoT environments.

AI techniques significantly enhance IoT data processing by extracting meaningful insights from diverse datasets. Machine learning algorithms detect patterns and anomalies in real-time data streams, aiding in the early identification of infrastructure issues or security threats. Predictive analytics powered by AI forecast future trends based on historical data, crucial for preemptive maintenance and operational optimization [17].

Statistical models complement machine learning by providing probabilistic insights into urban phenomena, supporting data-driven decisions for resource allocation and service improvement. Integrating AI into data management platforms enhances operational efficiency and responsiveness, enabling smarter and sustainable urban management strategies.

Integration with IoT Ecosystem

Data management platforms integrate seamlessly with the broader IoT ecosystem, supporting interoperability between diverse devices and systems. They enable data fusion from heterogeneous sources and facilitate data sharing across platforms, enhancing the overall efficiency and effectiveness of smart city operations.

Security and Privacy Considerations

Ensuring robust data security and privacy is critical in the deployment of IoT systems, especially due to their distributed and heterogeneous nature. Data management platforms implement robust security measures such as encryption, access control, and secure data transmission protocols. They comply with regulatory frameworks to protect sensitive information and uphold user privacy in smart city environments.

Scalability and Performance

Scalability is essential to accommodate the growing volume and variety of IoT data. Data management platforms are designed to scale horizontally across distributed systems, ensuring high performance and availability. They support real-time processing capabilities to meet the dynamic demands of smart city applications.

Emerging Trends and Technologies

Future advancements in data management platforms for IoT include edge computing [17], federated learning [18], and blockchain integration [19]. These innovations aim to enhance data processing efficiency, decentralize decision-making, and fortify data integrity and transparency in smart city infrastructures.

The comprehensive integration of data management platforms in IoT architectures underscores their pivotal role in enabling smart cities to harness the full potential of IoT technologies for sustainable urban development.

1.2.4 Security and Privacy in IoT

Security and privacy are paramount concerns in Internet of Things (IoT) ecosystems due to their distributed and heterogeneous nature.

Threat Landscape

Security threats in IoT can arise from various sources, including:

1. **Network Attacks:** Threat actors exploit vulnerabilities in wireless protocols (for example, Wi-Fi, Bluetooth Low Energy) and cellular networks for intercepting, manipulating, or disrupting data flows.
2. **Device Exploitation:** IoT devices often have weak authentication mechanisms, default credentials, and susceptibility to remote exploits, facilitating unauthorized access and control.
3. **Data Breaches:** Sensitive data collected by IoT sensors (for example, personal, location-based) is at risk of interception during transmission or unauthorized access if inadequately protected.

Security Measures

To address these challenges, the following strategies and practices are essential:

1. **End-to-End Encryption:** Employ strong cryptographic protocols to safeguard data integrity and confidentiality during transmission between IoT devices and backend servers.
2. **Authentication Mechanisms:** Implement robust authentication mechanisms, such as multifactor authentication (MFA) and digital certificates, to verify device identities and authorize access to resources.
3. **Patch Management:** Regularly update IoT device firmware and software to mitigate known vulnerabilities and apply security patches promptly.
4. **Intrusion Detection Systems (IDS):** Deploy IDS solutions to monitor network traffic and detect anomalous activities indicative of potential security breaches.
5. **Privacy Enhancements:** Adopt ”Privacy by Design” principles to minimize data collection, anonymize sensitive information, and provide transparent data handling practices.

Challenges and Future Directions

As technologies evolve and the volume of IoT deployments grows, new security and privacy challenges continue to emerge. Addressing these challenges requires ongoing research, innovation, and collaboration across industry, academia, and regulatory bodies. The future

of IoT security and privacy hinges on proactive measures, robust frameworks, and scalable solutions that safeguard data and protect user privacy in an increasingly connected world.

Ensuring trustworthiness in IoT devices and services entails robust security measures. IoT systems must prevent unauthorized access, data breaches, and malicious activities that can compromise user privacy and system integrity. Furthermore, ensuring IoT safety is critical to mitigate risks of physical harm or damage caused by IoT components.

In parallel, ethical considerations are crucial to govern the responsible use of IoT technologies. Ethical frameworks must guide the development and deployment of IoT systems to uphold societal norms, privacy expectations, and ethical standards. By embedding security, safety, and ethical considerations into IoT design and implementation processes, stakeholders can foster an environment conducive to trustworthy and sustainable IoT deployments [20].

1.3 IoT Communication Technologies: Focus on LoRa and NB-IoT

The proliferation of the Internet of Things (IoT) has necessitated the development of robust communication technologies capable of supporting the vast number of connected devices. Among the diverse array of IoT communication protocols, Long Range (LoRa) and Narrowband IoT (NB-IoT) have emerged as prominent technologies due to their distinct advantages and suitability for various applications.

LoRa, a member of the Low Power Wide Area Network (LPWAN) family, is designed to provide long-range communication with low power consumption. Its capacity to penetrate obstacles and operate in diverse environments positions it as a preferred choice for rural and remote deployments, such as agriculture monitoring, environmental sensing, and smart metering.

Conversely, NB-IoT, a cellular-based LPWAN technology standardized by the 3rd Generation Partnership Project (3GPP), leverages existing cellular infrastructure to deliver reliable and secure communication. Its emphasis on in-building penetration and high device density makes it well-suited for urban environments and applications such as smart cities, industrial IoT, and healthcare monitoring.

Focusing on LoRa and NB-IoT facilitates the exploration of the complementary nature of these technologies in addressing the varied requirements of IoT applications. While LoRa excels in scenarios demanding long-range and low-power operation, NB-IoT's strengths lie in leveraging cellular networks for robust and scalable connectivity [21].

1.3.1 LoRa (Long Range)

LoRa (Long Range) is a robust RF modulation technology designed specifically for low-power, wide-area networks (LPWANs). LoRa is notable for its ability to provide extensive communication ranges while maintaining minimal power consumption, making it ideal for battery-operated devices with extended lifespans. The technology's ability to cover extensive distances, approximately 15 kilometers in rural settings and up to 5 kilometers in urban environments, positions it as a leading solution for various IoT applications.

LoRa operates in a star network topology where multiple end devices communicate with a central gateway. This architecture simplifies network management and enhances scalability. The modulation technique used by LoRa is a proprietary spread-spectrum method based on Chirp Spread Spectrum (CSS). This approach enables robust and long-range communication, resistant to interference and environmental challenges. The use of orthogonal spreading factors allows multiple signals to be transmitted simultaneously without interference, significantly increasing network capacity [22].

Technical Characteristics

The following characteristics define the technical prowess of LoRa:

- **Communication Range:** LoRa technology excels in providing long-range communication capabilities:
  - **Rural Areas:** Up to 15 kilometers.
  - **Urban Areas:** Up to 5 kilometers.
- **Power Consumption:** One of LoRa's key features is its ultra-low power consumption, enabling long battery life:
  - **Battery Life:** Up to 10 years for battery-operated devices.
- **Modulation Technique:** LoRa uses Chirp Spread Spectrum (CSS) modulation, which provides robust communication:
  - **Channel Bandwidths:** 125 kHz or 500 kHz for uplink; 500 kHz for downlink.
- **Orthogonal Spreading Factors:** The use of orthogonal spreading factors is a significant advantage, allowing multiple signals to be transmitted simultaneously without interference:
  - **Network Capacity:** A single gateway can handle up to 1.5 million messages per day.
- **Robustness:** LoRa signals are highly robust against interference and environmental challenges:
  - **Interference Resistance:** Robust against in-band and out-of-band interference.

- **Multipath and Fading:** Immune to multipath and fading effects.
- **Doppler Shift:** Tolerant to Doppler shifts, ideal for mobile applications.

Specification	Value
Communication Range (Rural)	Up to 15 km
Communication Range (Urban)	Up to 5 km
Battery Life	Up to 10 years
Channel Bandwidth (Uplink / Downlink)	125 kHz (uplink) 500 kHz (downlink)
Network Capacity	1.5 million messages/day
Modulation Type	Chirp Spread Spectrum (CSS)

Table 1.2: LoRa Network Specifications

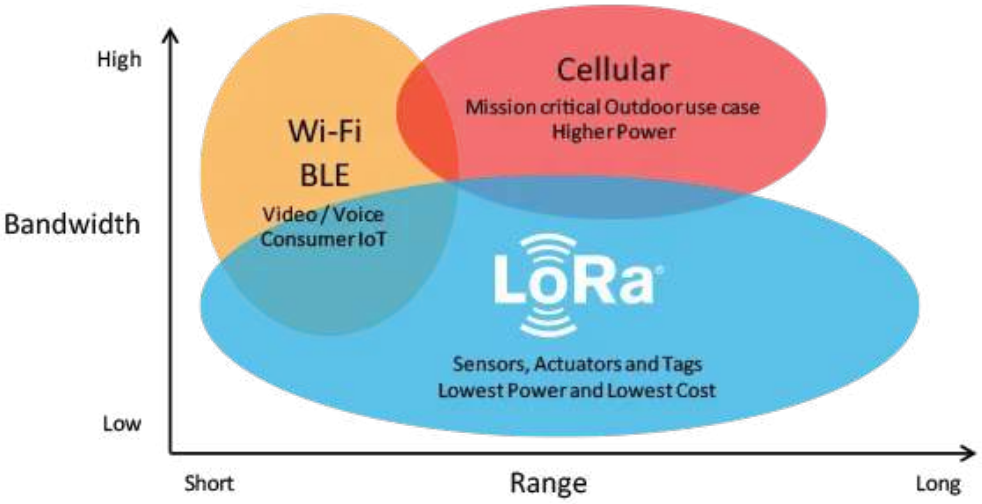


Figure 1.2: Bandwidth vs Range: comparison between LoRa and other RF technologies

Main Applications

LoRa technology finds widespread application across various domains due to its long-range communication capabilities and low power consumption. In smart city environments, LoRa facilitates large-scale sensor deployments for applications such as smart parking, waste management, and environmental monitoring. Its ability to cover extensive urban areas efficiently makes it ideal for these applications. In agriculture, LoRa enables precision farming through soil moisture monitoring, crop health management, and livestock tracking. The technology's low-power characteristics ensure longevity in remote agricultural settings.

Industrial IoT leverages LoRa for asset tracking, predictive maintenance, and monitoring of equipment and environmental conditions. LoRa's robustness against interference and its suitability for long-term deployment in industrial environments make it a preferred choice. In

healthcare, LoRa supports remote patient monitoring, asset tracking within medical facilities, and drug temperature monitoring during transport. Its reliability and scalability contribute significantly to enhancing healthcare service delivery.

In logistics and supply chain management, LoRa enables real-time tracking of goods, warehouse management, and fleet monitoring. The technology's cost-effectiveness and scalability make it feasible for widespread adoption across various logistical operations. Lastly, in environmental monitoring and disaster management, LoRa provides critical data for early warning systems, pollution monitoring, and disaster response coordination [23].

Application	**Use Case**
Smart Cities	Smart parking Waste management Environmental monitoring
Agriculture	Soil moisture monitoring Crop health management Livestock tracking
Industrial IoT	Asset tracking Predictive maintenance Environmental monitoring
Healthcare	Remote patient monitoring Asset tracking Drug temperature monitoring
Logistics and Supply Chain Management	Real-time tracking of goods Warehouse management Fleet monitoring
Environmental Monitoring and Disaster Management	Early warning systems Pollution monitoring Disaster response coordination

Table 1.3: Main Applications of LoRa Technology

Advantages and Disadvantages

LoRa technology offers several advantages that make it suitable for various IoT applications. Its long transmission distance, with a high sensitivity down to -148 dBm and a link budget of up to 175 dB, enables communication over extensive areas, up to 15 kilometers in ideal conditions. Moreover, LoRa devices operate at low power consumption, drawing 200 nA when idle and 10 mA when actively receiving signals, resulting in battery lifespans of up to 5 years. This extended battery life and robust signal transmission architecture make LoRa preferable for applications prioritizing reliability and longevity. Furthermore, LoRa technology supports powerful multi-channel information processing capabilities, facilitating parallel and asynchronous processing of data across nodes, with gateways capable of handling up to 5,000,000 data packets daily. Additionally, LoRa utilizes advanced modulation techniques

for superior interference resistance, ensuring reliable data transmission even in challenging environments.

However, LoRa has its limitations. It offers a low data transfer rate, which is suitable for periodic and brief exchanges but may not meet the requirements of applications needing fast data transmission. Moreover, the small payload size restricts the amount of data that can be transmitted in a single communication. LoRa's suitability for real-time applications is limited due to high latency and jitter, and its deployment requires building new networks, which can be a barrier in some scenarios. As LoRa networks expand, there is a risk of increased spectrum interference, affecting overall network performance. Additionally, while LoRa modules are cost-effective, their initial setup costs and ongoing maintenance may be higher compared to other wireless technologies [24].

Advantages	**Disadvantages**
Transmission Distance up to 15 km in ideal conditions	Low data transfer rate
Low power consumption (200 nA idle, 10 mA active)	Small payload size
Long battery life up to 5 years	High latency and jitter
Powerful multi-channel information processing	Requires building new networks
Strong interference resistance	Potential spectrum interference
	Higher initial setup and maintenance costs
	Limited bandwidth due to duty cycle

Table 1.4: Advantages and Disadvantages of LoRa Technology

1.3.2 NB-IoT (Narrowband IoT)

Technical Characteristics

NB-IoT (Narrowband Internet of Things) offers a range of technical features that cater specifically to the needs of IoT applications, particularly those requiring low power consumption, extended coverage, and efficient data transmission. One of the standout features of NB-IoT is its capability to operate in power-saving modes, facilitated by technologies like Power Saving Mode (PSM) and extended Discontinuous Reception (eDRX). PSM allows devices to remain registered while conserving power during deep sleep periods, ensuring longer battery life. In comparison, eDRX enhances this further by extending the sleep cycle of terminals in idle mode, thus reducing unnecessary power consumption during periods of inactivity. These mechanisms not only optimize power management but also support the longevity of battery-powered IoT devices [25].

Moreover, NB-IoT excels in coverage and latency sensitivity, crucial for IoT deployments requiring reliable connectivity over extensive areas. Leveraging mechanisms such as retransmission and low-frequency modulation, NB-IoT achieves significant coverage improvements, reaching up to 164 dB in independent deployment scenarios. Despite higher latency due to data retransmission in high coupling loss environments, NB-IoT supports latency as low as 6 seconds under optimal conditions, meeting the stringent latency requirements of various IoT applications.

In terms of transmission mode, NB-IoT builds upon LTE technology, modifying it to suit IoT-specific requirements. Operating on a narrow RF bandwidth of 200 kHz, NB-IoT utilizes different modulation schemes like QPSK and BPSK, alongside OFDMA and SC-FDMA technologies in downlink and uplink transmissions, respectively. This flexibility enables NB-IoT to adapt to diverse IoT terminal capabilities and channel conditions efficiently.

The spectrum resource allocation for NB-IoT is backed by major telecom operators, ensuring robust network support across various deployment scenarios including independent, guard-band, and in-band modes. These deployment modes leverage LTE spectrum bands, providing seamless integration with existing cellular networks while accommodating the specific needs of IoT applications.

Furthermore, NB-IoT employs semi-static link adaptation, which categorizes coverage classes based on minimal coupling loss levels (MCL) to optimize communication performance. This approach allows terminals to select appropriate modulation and coding schemes dynamically, ensuring reliable communication under varying signal conditions.

Additionally, NB-IoT emphasizes ultra-low power technologies to minimize energy consumption, crucial for battery-operated IoT devices. By introducing advanced power-saving mechanisms and optimizing energy efficiency during data transmission, NB-IoT achieves a balance between power consumption and performance, extending operational lifespans of IoT devices.

Finally, NB-IoT's network architecture integrates seamlessly with existing LTE infrastructures, comprising terminals, base stations, core networks, cloud platforms, and vertical business centers. This holistic approach supports scalable deployment and efficient management of IoT services across diverse industry verticals [26].

Main Applications

Narrowband Internet of Things (NB-IoT) is a standard defined by the 3GPP collaborative standards organization to address the need for low-power, wide-area communications technology. This standard supports a wide range of applications that require low-cost devices with minimal power consumption and the capability to be deployed in massive numbers. The main applications of NB-IoT include smart grids, smart cities, and smart industries.

Feature	Description
Low Power Consumption	PSM and eDRX modes extend battery life by enabling deep sleep cycles and reducing idle mode power consumption.
Enhanced Coverage	Supports up to 164 dB of coverage, utilizing retransmission and low-frequency modulation for reliable communication.
Transmission Mode	Based on LTE, uses narrow RF bandwidth (200 kHz) with QPSK and BPSK modulation in downlink and SC-FDMA in uplink.
Spectrum Resource	Utilizes LTE bands for deployment in independent, guard-band, and in-band modes, ensuring network integration and scalability.
Semi-Static Link Adaptation	Adapts modulation and coding schemes based on coverage class (MCL levels), optimizing communication under varying conditions.
Ultra-Low Power Technology	Implements advanced power-saving mechanisms to minimize energy consumption and extend device operational lifespans.
Network Integration	Integrated with LTE infrastructure, comprising terminals, base stations, core networks, cloud platforms, and vertical business centers.

Table 1.5: Technical Features of NB-IoT

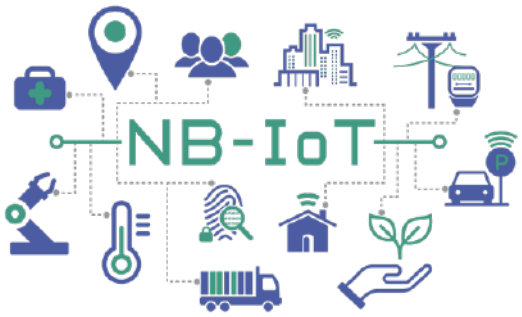


Figure 1.3: NB-IoT main applications

In smart grids, NB-IoT offers secure communications across various network types, such as home area networks and neighborhood area networks. Its features, including excellent coverage and low power consumption, facilitate enhanced deployment in smart metering, real-time monitoring, and electric vehicle (EV) charging. For smart metering, NB-IoT reduces complexity and costs by enabling direct communication via the cloud, eliminating the need for gateways. This technology supports automated meter reading services, real-time power metrics collection, and comprehensive advanced metering infrastructure (AMI) solutions, significantly optimizing energy management and consumption. Furthermore, NB-IoT supports intelligent EV charging systems by ensuring efficient interaction between vehicles, batteries, and charging stations under the control of a power supply grid management

system.

The deployment of NB-IoT in smart cities enables intelligent management in areas such as transportation, tracking systems, environmental monitoring, healthcare, and smart buildings. NB-IoT's capability to connect various devices, such as monitors and weather sensors, enhances wireless performance without compromising power efficiency. In healthcare, NB-IoT supports secure, long-range communication and high energy efficiency, enabling the integration of wearable devices with hospital management systems, thus promoting personalized healthcare solutions under the Health 4.0 paradigm. In transportation, NB-IoT facilitates real-time vehicle tracking, traffic analysis, and navigation, contributing to congestion reduction and improved public transport management. Additionally, NB-IoT-enabled street lighting systems improve energy utilization and management efficiency by providing real-time status updates and performance analysis.

In smart industries, NB-IoT connects numerous sensors within factories to the cloud, allowing for optimal planning, flexible manufacturing, and proactive maintenance. The technology enables predictive maintenance by monitoring machine conditions and alerting facility managers when maintenance is required, thus minimizing downtime. NB-IoT also supports industrial automation by integrating with time-sensitive networking (TSN) and industrial control systems, enhancing the efficiency of critical industrial operations. The synergy between NB-IoT and advanced technologies such as AI, cloud computing, and big data processing fosters a highly efficient and intelligent industrial environment, facilitating seamless communication and real-time insights into production and operational states.

Application	**Use Case**
Smart Grid	Secure communication, low power consumption Real-time monitoring, smart metering EV charging
Smart Cities	Intelligent transport, tracking systems Environmental monitoring, healthcare integration Smart buildings, street lighting
Smart Industry	Predictive maintenance, industrial automation Real-time monitoring, cloud integration AI and big data processing

Table 1.6: Summary of Main Applications of NB-IoT

Advantages and Disadvantages

NB-IoT presents both advantages and disadvantages in its deployment and application. One of the primary advantages of NB-IoT is its excellent coverage, which allows devices to operate in remote or difficult-to-reach areas, such as underground or inside buildings. Additionally, NB-IoT boasts low power consumption, which is crucial for extending the battery life of devices,

enabling them to operate for years without needing a replacement. The technology also supports a high device density, allowing for the massive deployment of sensors and devices in various applications such as smart cities, smart grids, and industrial IoT. Furthermore, NB-IoT offers secure communication channels, ensuring the integrity and privacy of the data transmitted, even with AI improvements [27].

However, NB-IoT also has its disadvantages. The low data rate, which is a design feature to conserve power, can limit the types of applications it supports, particularly those requiring high bandwidth or real-time data transmission. Additionally, while NB-IoT's extended coverage is beneficial, it may result in higher latency, which can be a drawback for applications needing immediate responses. Another disadvantage is the initial cost of deploying NB-IoT infrastructure, which can be substantial, particularly in regions without existing compatible networks. Finally, the interoperability with other IoT technologies and standards can pose challenges, potentially complicating integration efforts in mixed-technology environments.

Advantages	**Disadvantages**
Excellent coverage in remote areas	Low data rate limits high-bandwidth applications
Low power consumption for extended battery life	Higher latency can be an issue for real-time applications
Supports high device density for massive deployments	Initial deployment costs can be high
Secure communication channels	Interoperability with other IoT technologies can be challenging

Table 1.7: Advantages and Disadvantages of NB-IoT

1.3.3 Comparison between LoRa and NB-IoT

LoRa and NB-IoT are two prominent LPWA (Low Power Wide Area) technologies, each with unique characteristics and advantages suited to different IoT applications. LoRa operates in the unlicensed spectrum below 1 GHz, employing a proprietary spread spectrum modulation (CSS) which trades data rate for sensitivity. This technology is designed for long-range communication with low power consumption, making it ideal for applications requiring extensive coverage and long battery life. In contrast, NB-IoT, a standard defined by 3GPP, operates within the licensed spectrum and utilizes a simplified LTE protocol to minimize costs and battery usage. It employs OFDM for downlink and SC-FDMA for uplink, offering high data rates and low latency but at the expense of higher power consumption and more complex deployment.

In terms of quality of service (QoS), NB-IoT has an advantage due to its use of licensed spectrum and time-slotted synchronous protocol, providing better reliability and less interference compared to LoRa's asynchronous protocol in the unlicensed spectrum. For battery life and

latency, LoRa's asynchronous nature allows for flexible sleep cycles, leading to longer battery life but higher latency. NB-IoT, with its regular synchronization, consumes more power but provides lower latency and higher data rates. Coverage and range also differ; LoRa can cover wide areas with fewer gateways, making it suitable for rural or suburban deployments, while NB-IoT's coverage is typically limited to areas with existing cellular infrastructure, although it can achieve greater range within those areas.

Cost considerations are significant in choosing between the two. LoRa benefits from lower deployment and operational costs since it operates in the free spectrum and requires fewer gateways. Conversely, NB-IoT incurs higher costs due to spectrum licensing and the need for more base stations. The deployment model also varies, with LoRa offering more flexible and quicker setup due to its maturity and ready components, while NB-IoT's deployment is dependent on existing cellular networks and is still in the rollout phase.

Factor	**LoRa**	**NB-IoT**
Quality of Service (QoS)	Lower QoS Unlicensed spectrum Asynchronous protocol	Higher QoS Licensed spectrum Synchronous protocol
Battery Life	Longer battery life Flexible sleep cycles	Shorter battery life Regular synchronization
Latency	Higher latency	Lower latency
Coverage and Range	Wide coverage with fewer gateways Suitable for rural areas	Limited to cellular coverage Greater range in covered areas
Cost	Lower cost Free spectrum Fewer gateways	Higher cost Spectrum licensing More base stations
Deployment Model	Flexible Quick setup Mature components	Dependent on cellular networks Still in rollout phase

Table 1.8: Comparison between LoRa and NB-IoT

This comparison highlights that the choice between LoRa and NB-IoT depends on specific application requirements, such as the need for QoS, battery life, latency, coverage, and cost. Moreover, the scientific community persistently investigates the comparative analysis of these two LPWAN technologies [28], even within specialized sectors [25].

Chapter 2

Overview of eXtended Reality (XR) Technologies

2.1 Introduction to XR

2.1.1 Definition of XR

eXtended Reality (XR) is a comprehensive term encompassing various immersive technologies that merge the physical and virtual worlds to create enhanced or entirely virtual experiences. XR encompasses technologies like Virtual Reality (VR), which creates fully immersive environments, Augmented Reality (AR), which overlays digital content on real-world settings, and Mixed Reality (MR), which integrates digital and physical elements in real-time. VR involves creating a fully immersive digital environment separate from the real world, usually experienced through headsets or Head-Mounted Displays (HMDs), allowing users to interact within a simulated space. AR overlays digital information, such as images, videos, or data, onto the real world, enhancing the user's perception and interaction with their physical surroundings without replacing them. MR blends elements of both VR and AR, enabling physical and digital objects to coexist and interact in real-time, providing a seamless integration of the real and virtual worlds through advanced sensors and tracking techniques.

Figure 2.1: Key Components of XR Systems, including VR, AR, and MR Applications.

A significant application of XR technologies is the interaction with digital twins [29]—virtual replicas of physical entities—allowing users to monitor, analyze, and manipulate these digital counterparts in real-time, thus enhancing decision-making processes, predictive maintenance, and overall system optimization . XR technologies aim to provide a high degree of immersion, realism, and user engagement across various applications.

2.1.2 Components of XR: VR, AR, and MR

Virtual Reality (VR)

Virtual Reality (VR) refers to the creation of a computer-generated environment that simulates a realistic experience, fully immersive and separate from the real world. This is achieved through the use of headsets or Head-Mounted Displays (HMDs) that cover the user's field of vision and often include auditory components. In a VR environment, users can interact with 3D objects and navigate through the virtual space as if they were physically present. The sense of immersion in VR is enhanced by high-resolution graphics, spatial audio, and haptic feedback, which together create a convincing simulation of reality.

The concept of VR dates back to the 1960s [30] when the first rudimentary VR systems were developed. One of the pioneering efforts was the Sensorama [31], created by Morton Heilig in 1962, which provided a multi-sensory experience using stereoscopic 3D images, sound, vibrations, and even smells. Despite the early enthusiasm, the technology of the time was not advanced enough to support widespread adoption, leading to a decline in interest.

However, advancements in computing power, graphics processing, and sensor technology have reignited interest in VR (2.2). Modern VR systems are now capable of delivering highly realistic and interactive experiences, making the technology viable for a range of practical applications. Contemporary VR solutions are being employed in fields such as educational gaming [32], training simulations [33], medical applications [34], and virtual tours [35], demonstrating tangible benefits and use cases that were previously unattainable.

VR is widely used in various fields, providing users with experiences that would otherwise be impractical or impossible in the real world. The renewed focus on VR technology highlights its potential to revolutionize how we interact with digital content and perform complex tasks in simulated environments.

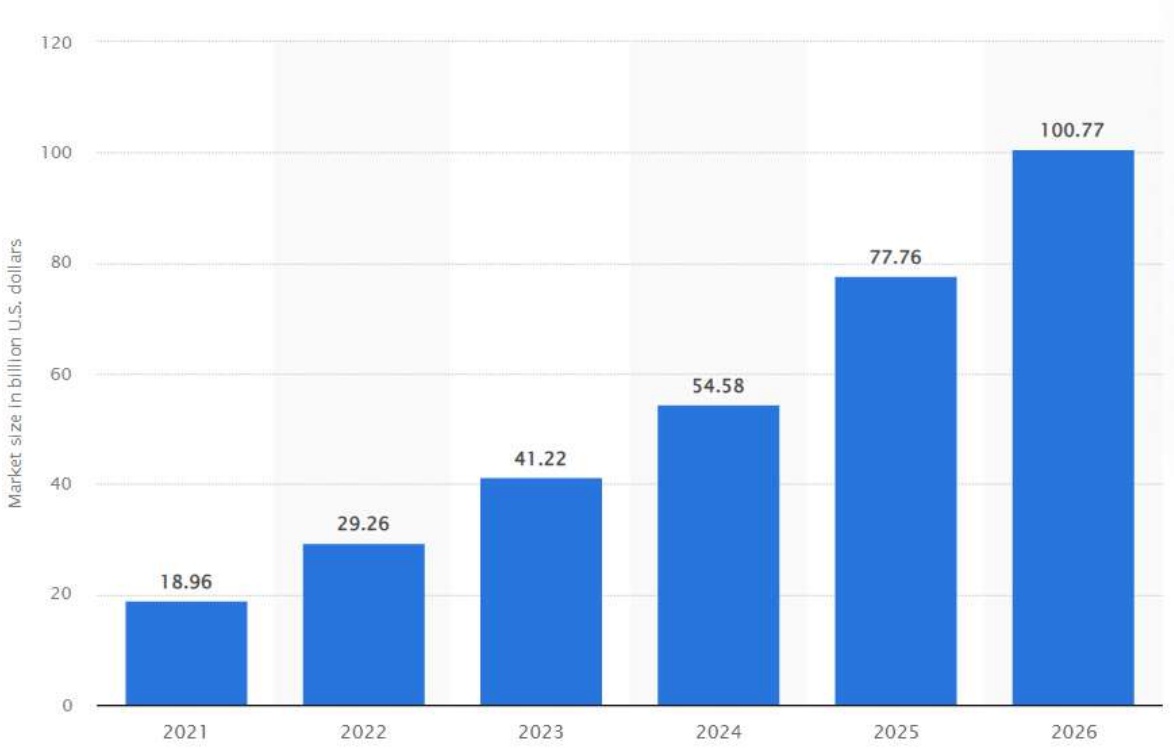


Figure 2.2: Extended reality (XR) market size worldwide from 2021 to 2026 [36]

Augmented Reality (AR)

Augmented Reality (AR) is a technology that enriches the user's perception of the real world by overlaying digital content, such as images, videos, or data, onto their view of the physical environment. Unlike Virtual Reality (VR), which immerses users in a completely simulated environment, AR integrates virtual elements seamlessly into real-world surroundings in real-time. This integration is made possible by sophisticated technologies including computer vision, which enables devices to interpret and understand the surrounding environment, image recognition algorithms that identify objects and patterns in real-time, and spatial mapping techniques that map physical spaces and surfaces for accurate placement of virtual objects.

AR applications span a wide range of industries and use cases. In gaming and entertainment, AR enhances gameplay by superimposing interactive elements onto real-world settings, creating immersive experiences that blend virtual challenges with physical surroundings. In education [37], AR transforms learning experiences by providing interactive visualizations and simulations, allowing students to explore complex subjects in a more engaging and intuitive manner. In healthcare [38], AR assists in surgery, diagnostics, and training, enhancing precision and efficiency. Industrial applications of AR include remote assistance [39], where experts can provide real-time guidance and instructions to field technicians using AR-enabled

devices, reducing downtime and improving maintenance efficiency.

Overall, AR technologies continue to evolve, offering innovative ways to enhance user experiences, improve operational efficiency across various industries, and unlock new possibilities for interaction between digital and physical worlds.

Mixed Reality (MR)

Mixed Reality (MR) is an immersive technology that blends elements of Virtual Reality (VR) and Augmented Reality (AR) to create environments where physical and digital objects coexist and interact in real-time. MR environments are generated by devices like specialized headsets or smart glasses that integrate virtual elements into the user's view of the real world, while also allowing users to interact with and manipulate these digital objects. Unlike AR, which overlays digital content onto the physical environment, MR enhances this interaction by anchoring virtual objects to specific locations and allowing them to interact with physical surfaces and objects.

Key technologies enabling MR experiences include spatial mapping for accurate environment understanding, depth sensing to perceive physical space, and advanced tracking systems that ensure precise alignment and interaction between virtual and physical elements. These technologies facilitate realistic simulations and interactive scenarios that can be applied across various sectors.

In practical terms, MR finds applications in fields such as design and engineering, where architects can visualize and manipulate virtual prototypes within real-world contexts, improving design precision and collaboration [40]. Healthcare utilizes MR for surgical planning [41], allowing surgeons to visualize patient anatomy overlaid with diagnostic data in real-time, enhancing surgical accuracy and patient safety.

Overall, Mixed Reality continues to evolve as a powerful tool for merging physical and virtual worlds, offering enhanced user experiences, improved productivity, and innovative solutions across industries.

Table 2.1: Comparison of VR, AR, and MR

Aspect	Virtual Reality (VR)	Augmented Reality (AR)	Mixed Reality (MR)
Environment	Fully digital environment, immersive, often isolated from real world	Real-world environment overlaid with digital content, enhancing real-world perception	Blends digital content with real-world environment, maintains interaction with physical surroundings
User Interaction	Immersive interactions with virtual objects and environments	Interacts with virtual objects overlaid onto real-world surfaces	Interaction with both digital and physical objects, enables spatial understanding and manipulation
Applications	Gaming, simulations, training, virtual environments	Education, navigation, marketing, maintenance, remote assistance	Gaming, interactive storytelling, education, training, remote collaboration
Hardware	Head-mounted displays (HMDs), motion controllers, haptic devices	Smartphones, tablets, AR glasses	HMDs, AR glasses, spatial computing devices
Examples	Oculus Rift, HTC Vive, PlayStation VR	Pokémon GO, Snapchat AR lenses, Microsoft HoloLens	Magic Leap, HoloLens 2, Oculus Quest 2
Immersive Experience	Fully immersive, isolating users from real-world distractions	Partially immersive, overlays digital content onto real-world scenes	Blends real and virtual worlds, maintains awareness of physical environment
Challenges	Motion sickness, isolation, hardware complexity	Real-world integration, tracking accuracy, context awareness	Seamless blending, spatial mapping, consistent user experience across varying environments
Future Directions	Enhanced realism, social VR experiences, accessibility improvements	Improved AR glasses, integration with IoT, spatial computing advancements	Enhanced spatial awareness, expanded use cases in industry, improved user interfaces

2.1.3 Sense of Immersion

Definition of Immersion

Immersion in the context of virtual reality (VR) refers to the extent or degree to which a user feels completely engaged and integrated into a virtual environment. It is characterized by a heightened sense of presence and realism, where the user perceives and interacts with the digital environment as if it were real. An immersive experience aims to suspend disbelief, creating a compelling illusion that transports the user away from their physical surroundings into the virtual world.

Key aspects that define immersion include:

- **Realism:** The degree to which the virtual environment convincingly replicates aspects of the real world, such as visual fidelity, accurate physics, and natural interactions.
- **Sensory Engagement:** The involvement of multiple senses (visual, auditory, tactile) in the virtual experience, enhancing the user's perception and interaction with the digital environment.
- **Cognitive Absorption:** The level of cognitive involvement and attention devoted to the virtual experience, leading to a deep engagement and suspension of disbelief.

Achieving a high level of immersion is crucial for creating compelling VR applications across various domains, including gaming, simulations, training, and virtual collaborations. By leveraging advanced technologies and design principles, developers strive to enhance immersion and deliver immersive experiences that captivate and engage users.

Factors Influencing Immersion

Several key factors significantly influence the level of immersion experienced in virtual reality (VR):

- **Graphical Quality:** High-quality graphics and visual fidelity are essential for creating a realistic virtual environment. This includes detailed textures, accurate lighting, realistic shadows, and smooth animations, which collectively enhance the visual immersion and believability.
- **Spatial Sound:** Realistic and spatially accurate audio enhances immersion by providing directional and positional cues that match the user's movements within the virtual space. Immersive sound design contributes to the overall sense of presence and realism.
- **Haptic Feedback:** Tactile feedback through haptic devices or controllers enhances immersion by simulating physical interactions with virtual objects. This feedback can include vibrations, textures, or resistance, adding a sense of touch and physical presence to the virtual experience.

- **Precise Motion Tracking:** Accurate motion tracking of head movements and gestures is crucial for maintaining immersion and responsiveness in VR applications. Seamless tracking ensures that virtual objects and environments react realistically to the user's actions, enhancing the feeling of presence.
- **Natural Interactions:** Intuitive and responsive user interactions with virtual objects and environments contribute to immersion. This includes hand gestures, motion controllers, and other input methods that mimic real-world actions and facilitate a seamless interaction with the virtual environment.

These factors collectively contribute to creating a compelling and immersive VR experience, where users can feel fully engaged and immersed in the digital environment. Developers and designers continually strive to optimize these elements to enhance immersion and deliver more immersive VR applications across various industries.

Presence

Presence in virtual reality (VR) refers to the subjective sensation of "being there" within the virtual environment. It measures the degree to which users perceive and feel present in the digital space compared to the physical world. Achieving a strong sense of presence is essential for creating an immersive and realistic VR experience.

Key factors influencing presence include:

- **Visual Realism:** High-quality graphics, accurate object scale, and detailed environments contribute to visual realism, enhancing the believability of the virtual environment and reinforcing the sense of presence.
- **Audio Immersion:** Spatial audio techniques and realistic sound propagation enhance immersion by providing accurate auditory cues that align with the user's position and movements within the virtual space. This contributes to a more immersive and engaging experience.
- **Interaction and Feedback:** Responsive interactions with virtual objects and environments help users establish a sense of physical presence and connection within the virtual world. Real-time haptic feedback and precise motion tracking further enhance the feeling of being present.

The concept of presence is crucial for designing effective VR applications that engage users and provide realistic experiences. By optimizing visual, auditory, and interaction elements, developers aim to maximize presence and deliver compelling VR environments across various domains.

Sense of Immersion: XR vs VR

Sense of immersion in Extended Reality (XR) and Virtual Reality (VR) refers to the extent to which users feel engaged and connected within a digital environment, but there are distinctions between the two.

In Virtual Reality (VR), immersion aims to completely transport users into a fully digital environment, typically through a head-mounted display (HMD) that covers the user's vision. VR relies heavily on high-quality visuals, spatial audio, and responsive interactions to create a convincing illusion of presence and realism. Users are often isolated from the physical world, focusing solely on the virtual environment.

Extended Reality (XR), encompassing VR, Augmented Reality (AR), and Mixed Reality (MR), integrates digital content with the real world. This integration poses unique challenges and opportunities for immersion. XR leverages spatial computing and real-world context to overlay digital information onto physical environments, enhancing the user's perception of reality without completely replacing it.

Factors influencing immersion in both XR and VR include visual realism, spatial sound, haptic feedback, precise motion tracking, and natural interactions. In XR, the challenge lies in seamlessly blending digital content with real-world surroundings to maintain immersion and believability. Interaction with physical objects and environments adds a layer of depth to immersion in XR, compared to the virtual isolation typically experienced in VR.

Achieving a high sense of immersion in both XR and VR involves optimizing these factors to create compelling user experiences across various applications, from gaming and entertainment to education, training, and industrial use. Advances in technology continue to push the boundaries of immersion in both XR and VR, offering new possibilities for interactive and engaging digital experiences.

Table 2.2: Comparison of Sense of Immersion: VR vs. XR

Aspect	Virtual Reality (VR)	Extended Reality (XR)
Definition of Immersion	Fully immersive, users feel completely engaged and part of a digital environment.	Partially immersive, blending digital and real environments to engage users in both worlds.
Graphical Quality	High-quality, realistic graphics necessary for full immersion.	High-quality graphics that seamlessly blend with real-world visuals.
Spatial Sound	Critical for creating a sense of presence, must match the virtual environment.	Spatial sound that enhances the blending of digital and real environments.
Haptic Feedback	Provides tactile sensations and physical interactions within the virtual environment.	Integrates tactile feedback for interactions with both digital and physical objects.
Motion Tracking	Precise tracking of head and body movements for seamless interaction.	Precise tracking that supports interaction with both digital overlays and real-world objects.
User Interaction	Interaction through motion controllers, hand-tracking, and gesture recognition.	Interaction through hand-tracking, gesture recognition, and manipulation of physical objects.
Presence	Strong sense of ”being there” within the fully digital environment.	Sense of ”being there” in both digital and real environments, maintaining awareness of physical surroundings.
Applications	Gaming, simulations, virtual training environments.	Education, training, industrial applications, enhanced real-world navigation and assistance.
Challenges	Motion sickness, isolation from real world, hardware complexity.	Seamless blending of digital and real, consistent experience across varying environments, tracking accuracy.
Future Directions	Enhanced realism, social VR experiences, accessibility improvements.	Improved AR glasses, spatial computing advancements, expanded use cases in industry.

2.2 XR Technologies and Devices

2.2.1 HMD (Head-Mounted Displays)

Head-Mounted Displays (HMDs) are wearable devices that cover the user's eyes and often ears, creating an immersive visual and auditory experience. These devices are fundamental to both Virtual Reality (VR) and Mixed Reality (MR) applications, offering a gateway to fully or partially digital environments.

Virtual Reality HMDs

- **Oculus Rift (Meta Quest):** Originally known as Oculus Rift, now branded as Meta Quest. Renowned for high-resolution displays and advanced tracking systems, providing a robust VR experience for gaming, simulations, and interactive applications.
- **HTC Vive:** Offers precise motion tracking and high-resolution visuals, enabling an immersive VR experience with extensive room-scale capabilities.
- **PlayStation VR:** Designed for use with PlayStation consoles, providing accessible VR experiences with quality visuals and ease of use.

Mixed Reality HMDs

- **Microsoft HoloLens:** Overlays digital information onto the physical world through see-through lenses, supporting applications in design, education, and healthcare.
- **Magic Leap:** Utilizes light-field display technology to blend digital objects seamlessly with the physical environment, enhancing the perception of mixed reality experiences.

Components and Features

- **Displays:** Utilize high-resolution OLED or LCD screens with a wide field of view (FOV) ranging typically from 90 to 120 degrees.
- **Sensors:** Include accelerometers, gyroscopes, and magnetometers for precise head tracking, adjusting the display based on user movements.
- **Audio:** Employ spatial audio systems to create immersive 3D sound environments that enhance presence and realism.
- **Comfort and Ergonomics:** Designed with adjustable straps, ergonomic padding, and balanced weight distribution for prolonged user comfort and reduced fatigue.

HMDs play a pivotal role in delivering immersive XR experiences by integrating advanced display technologies, precise tracking systems, and high-fidelity audio. These features collectively create a compelling sense of presence, whether fully transporting users to virtual worlds or enhancing their interaction with augmented content in the real world.

Table 2.3: Comparative Table of Main HMDs [42]

Product	Meta Quest 3	PlayStation VR2	HTC Vive Pro 2	Microsoft HoloLens 2	Magic Leap 2
Manufacturer	Meta	Sony	HTC	Microsoft	Magic Leap
Website	meta.com	playstation.com	vive.com	microsoft.com	magicleap.com
Device Type	Standalone VR	Console-powered VR	PC-powered VR	Standalone AR	Standalone AR
Platform	Meta Quest	PlayStation	SteamVR, Viveport	Windows MR	-
Release Date	Oct 2023	Feb 2023	Jun 2021	Nov 2019	Sep 2022
Retail Price	$499 with controllers (128GB) $649 with controllers (512GB)	$550 with controllers	$799 headset only $1399 with controller, base stations	$3500 headset only	$3299 with controllers $4099 Developer Pro $4999 Enterprise
Optics	Pancake lenses	Fresnel lenses	Dual-element Fresnel lenses	Waveguides	Waveguides
Ocularity	Binocular	Binocular	Binocular	Binocular	Binocular
IPD Range	58-71mm	58-73mm	57-70mm	-	-
Adjustable Diopter	Presc. lens inserts sold sep.	-	Assisted via eye tracking	-	-
Passthrough	Dual 18 PPD color passthrough cameras	Grayscale passthrough via tracking cameras	Dual passthrough cameras	Native passthrough, 8MP @ 30fps	Native passthrough, 12.6MP 60fps RGB
Display Type	2 x LCD binocular	2 x OLED binocular	2 x LCD binocular	2 x LBS binocular	2 x LCoS binocular
Resolution	2064x2208 per-eye	2000x2040 per-eye	2448x2448 per-eye	1440x936 per-eye	1440x1760 per-eye
Refresh Rate	120Hz	120Hz	120Hz	60Hz	-

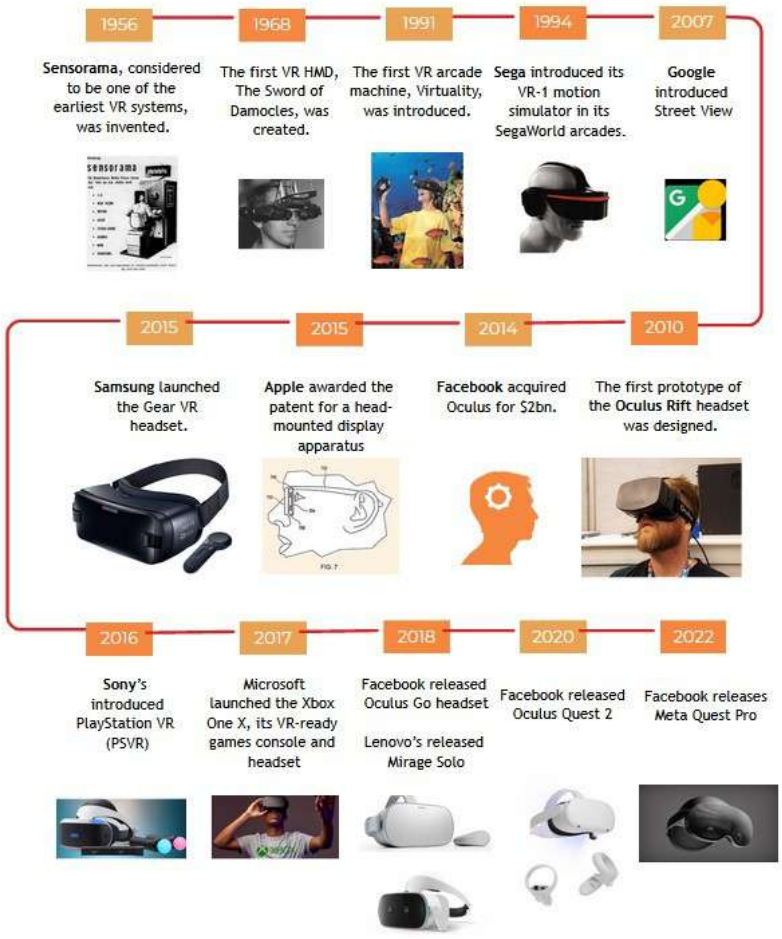


Figure 2.3: VR Evolution from 1956

2.2.2 Tracking Systems

Tracking systems are essential components of Virtual Reality (VR) and Augmented Reality (AR) technologies, as they determine the accuracy and responsiveness of user interactions within the virtual or augmented environment. These systems can be broadly categorized into optical, non-optical, and hybrid tracking methods.

Optical Tracking

Optical tracking utilizes cameras and sensors to monitor the position and orientation of the user and other objects within the environment. This method often involves the use of markers or LEDs that the cameras can track. Examples include the use of external cameras in the Oculus Rift and the PlayStation VR, which track LED patterns on the headsets and controllers. Optical tracking systems offer high accuracy and low latency but can be susceptible to occlusion and require a clear line of sight [43].

Non-Optical Tracking

Non-optical tracking encompasses techniques such as inertial measurement units (IMUs), magnetic tracking, and ultrasonic tracking. IMUs, which include accelerometers and gyroscopes, are commonly integrated into headsets and controllers to track movement and orientation without relying on external sensors. Magnetic tracking, used in systems like the Sixense STEM, employs magnetic fields to determine the position and orientation of objects [44]. Ultrasonic tracking, although less common, utilizes sound waves to measure distances and detect movements [45]. Non-optical methods are generally more robust to occlusion and can function in various environments, but they may suffer from drift and lower accuracy compared to optical systems.

Hybrid Tracking

Hybrid tracking systems combine optical and non-optical methods to leverage the advantages of both. For instance, the HTC Vive Pro employs a hybrid system using both external lighthouse sensors (optical) and IMUs (non-optical) to provide highly accurate and robust tracking [46]. By integrating multiple tracking techniques, hybrid systems can achieve superior performance, reducing the limitations inherent to any single tracking method.

Table 2.4: Comparison of Tracking Systems in XR

Aspect	Optical Tracking	Non-Optical Tracking
Description	Uses cameras to track physical movements and gestures in XR environments.	Relies on gyroscopes and accelerometers to measure orientation and motion.
Technology	Uses visible light spectrum or infrared cameras.	Utilizes MEMS (Micro-Electro-Mechanical Systems) sensors.
Accuracy	High accuracy with precise spatial tracking capabilities.	Accurate for short-term movements but prone to drift over time.
Applications	Suitable for precise hand and body tracking in VR/AR applications.	Ideal for mobile VR/AR experiences and applications requiring quick response times.
Advantages	Provides detailed positional data for realistic interaction in virtual environments.	Compact, lightweight, and power-efficient.
Challenges	Requires line-of-sight between cameras and markers, sensitive to lighting conditions.	Drift issues over prolonged use, limited accuracy for fine-grained movements.

Tracking systems continue to evolve, with advancements aiming to enhance accuracy, reduce latency, and improve robustness in diverse environments. These improvements are critical for delivering more immersive and seamless VR and AR experiences.

2.2.3 Interfaces and Interaction Tools

In virtual reality (VR), augmented reality (AR), and mixed reality (MR), interfaces and interaction tools play a crucial role in facilitating user engagement and control within digital environments. These technologies encompass a variety of input methods and interaction paradigms designed to enhance user immersion and functionality.

Input Methods

The primary input methods in XR (Extended Reality) environments include:

- **Motion Controllers:** Hand-held devices equipped with sensors for tracking hand movements and gestures. Motion controllers allow users to interact with virtual objects

by simulating natural hand movements and gestures.

- **Hand Tracking:** Utilizes cameras or sensors to detect and interpret hand movements directly without the need for physical controllers. Hand tracking enhances user interaction by enabling intuitive gestures and manipulations within virtual environments.
- **Gesture Recognition:** Algorithms that interpret predefined gestures to trigger specific actions or commands. Gesture recognition enhances user interface (UI) interactions by providing a natural and intuitive input method.
- **Voice Commands:** Integration of speech recognition technologies allows users to control applications and interact with virtual environments using voice commands. Voice commands offer hands-free interaction and are particularly useful in scenarios requiring multitasking or accessibility.
- **Eye Tracking:** Monitors the movement and focus of the user's eyes within the virtual space. Eye tracking enables gaze-based interactions, such as selecting objects or triggering actions based on visual attention.

Interaction Paradigms

Interfaces in XR environments employ various interaction paradigms to facilitate user engagement and control:

- **Spatial Interactions:** Users interact with virtual objects and environments in a spatially coherent manner, leveraging motion controllers or gestures to manipulate and navigate within 3D space.
- **Object Manipulation:** Enables users to manipulate virtual objects by grabbing, rotating, scaling, or moving them within the virtual environment. This interaction paradigm supports applications ranging from product design to virtual sculpting and training simulations.
- **Menu Systems and HUDs (Heads-Up Displays):** Overlay digital menus, controls, or information directly within the user's field of view. HUDs provide contextual information and controls without obstructing the virtual environment, enhancing user accessibility and interaction efficiency.
- **Teleportation and Locomotion:** Techniques for navigating virtual environments, such as teleportation (instantaneous movement between locations) and locomotion (simulated movement through physical space). These methods balance user comfort with spatial exploration and are critical for VR applications requiring extensive movement.

Interfaces and interaction tools continue to evolve, driven by advancements in hardware capabilities and user experience design. By optimizing input methods and interaction paradigms, developers aim to create intuitive, immersive, and user-friendly XR experiences across diverse

applications and industries.

2.2.4 Difference Between 3DoF and 6DoF

In the realm of immersive technologies, the terms 3 Degrees of Freedom (3DoF) and 6 Degrees of Freedom (6DoF) refer to the degrees of movement freedom that a user experiences within a virtual environment.

3 Degrees of Freedom (3DoF)

3DoF systems track rotational movements along three axes: pitch, yaw, and roll. These movements correspond to the rotation of the user's head or controller in virtual space. A typical example is a headset that allows users to look around in a virtual environment but does not track positional movement or translational shifts in space. 3DoF devices are simpler and often less expensive, suitable for experiences where full positional tracking is unnecessary, such as 360-degree videos or seated VR experiences.

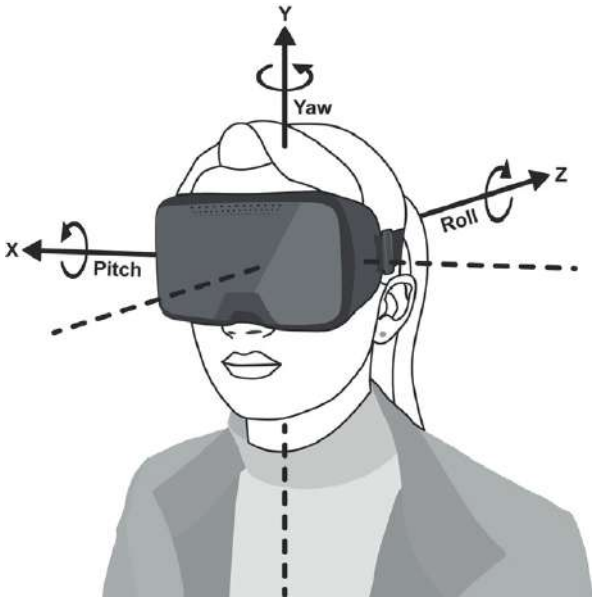


Figure 2.4: Movement in 3DoF [47]

6 Degrees of Freedom (6DoF)

6DoF systems track both rotational movements (pitch, yaw, roll) and translational movements (surge, sway, heave). This means the system can track not only how the user rotates their head but also their physical movement within the physical space. 6DoF provides a more immersive experience as users can lean, crouch, or walk around within the virtual environment, enhancing realism and interaction. High-end VR headsets like the Oculus Rift S or HTC Vive Pro utilize 6DoF tracking for precise spatial positioning, enabling realistic simulations, accurate gameplay, and interactive experiences.

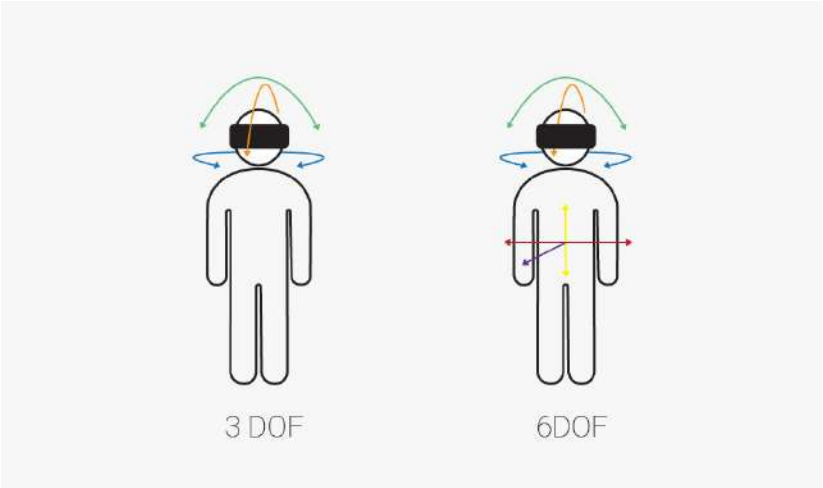


Figure 2.5: Movement in 6DoF

Applications and Considerations

The choice between 3DoF and 6DoF depends on the desired level of immersion and interaction for the application. 3DoF is suitable for experiences where users primarily observe or interact within a fixed viewpoint, such as media consumption or introductory VR experiences. On the other hand, 6DoF is essential for applications requiring full spatial awareness and physical interaction, such as gaming, training simulations, or virtual collaborations.

Understanding the distinction between 3DoF and 6DoF is crucial for designing and selecting appropriate hardware and software solutions in XR applications, ensuring that user experience aligns with the intended level of immersion and interactivity [48].

2.2.5 Field of View (FoV)

Field of View (FoV) is a critical parameter in eXtended Reality (XR) technologies, determining the extent of the observable environment that can be seen at any given moment through a display device. In the context of head-mounted displays (HMDs) and other immersive devices, FoV is typically measured in degrees and encompasses both horizontal and vertical components.

A wider FoV contributes significantly to the sense of immersion and presence within a virtual environment. It allows users to perceive more of their surroundings without needing to move their head or eyes, closely mimicking the natural human vision, which has an approximate horizontal FoV of 200 degrees and vertical FoV of around 135 degrees. Most modern VR headsets, such as the Meta Quest 3, offer a FoV in the range of 90 to 110 degrees, balancing technical feasibility and user comfort.

Device	FOV (Degrees)	Description
Human Eye	200° horizontal, 135° vertical	Natural vision with peripheral awareness
Meta Quest 2	89° horizontal, 93° vertical	Common VR headset, providing immersive experience
Meta Quest 3	110° horizontal, 100° vertical	Enhanced VR headset with wider FOV
HTC Vive Pro 2	120° horizontal, 90° vertical	High-end VR headset with expansive view
PlayStation VR2	110° horizontal, 100° vertical	Latest console VR headset with competitive FOV
Microsoft HoloLens 2	52° horizontal, 45° vertical	AR headset with focused FOV for mixed reality
Magic Leap 2	70° horizontal, 70° vertical	Advanced AR headset with wider FOV

Table 2.5: Comparison of Field of View (FOV) for Various Devices

Achieving a wide FoV presents several technical challenges, including maintaining high resolution and minimizing distortions at the edges of the display. Advances in optics, such as the development of aspherical lenses and fresnel lenses, have been crucial in expanding FoV while ensuring image clarity and reducing aberrations. Additionally, the design of VR content must consider FoV to optimize visual elements and ensure a consistent and immersive experience across different devices.

Understanding and optimizing FoV is essential for the development of effective and immersive XR applications. As technology progresses, increasing the FoV of XR devices will continue to be a focal point, driving enhancements in user experience and expanding the potential of immersive environments [49].

2.2.6 Motion Sickness

Motion sickness, also known as VR sickness or cyber sickness, is a common issue experienced by users of immersive technologies, particularly in virtual reality environments. It arises when there is a disconnect between the visual motion cues perceived by the eyes and the lack of corresponding physical movement sensed by the inner ear. Symptoms can include dizziness, nausea, headaches, and disorientation.

The risk of motion sickness is generally higher in 3DOF systems, where the lack of positional tracking can exacerbate the disconnect between visual and physical motion. Conversely,

6DOF systems, with their ability to track both rotational and translational movements, tend to reduce the incidence of motion sickness by providing a more natural and consistent correspondence between visual and physical sensations [50].

Design considerations to mitigate motion sickness include maintaining a high frame rate, reducing latency, ensuring realistic motion and acceleration within the virtual environment, and providing users with the ability to control their movement. Understanding and addressing the factors that contribute to motion sickness is crucial for developing comfortable and enjoyable XR experiences [51].

The FoV impacts not only immersion but also user comfort and the effectiveness of the VR experience. A limited FoV can result in a ”tunnel vision” effect, reducing the user's peripheral vision and potentially causing discomfort or disorientation. Conversely, a wide FoV enhances spatial awareness and realism, making interactions more intuitive and reducing the likelihood of motion sickness [50].

2.3 Applications of XR

2.3.1 Gaming

Extended Reality (XR) and gaming are pivotal concepts that harness the power of technology to create highly interactive and immersive experiences. These technologies not only enhance user engagement but also offer innovative approaches to entertainment, learning, and problem-solving across various domains.

Gaming, in particular, has evolved significantly with advancements in XR technologies. Virtual Reality (VR) and Augmented Reality (AR) offer new dimensions to gaming experiences by immersing players in richly detailed virtual worlds or by overlaying game elements onto the real world. This results in highly engaging and interactive experiences that transcend traditional gaming formats. Games like Pokémon GO [52] and Beat Saber [53] illustrate how AR and VR can transform the gaming landscape, offering players unique and immersive experiences.

Extended Reality (XR), encompassing VR, AR, and Mixed Reality (MR), enhances gaming experiences by providing immersive environments and interactive elements that make players feel like active participants in the game world. XR technologies enable features such as spatial audio, haptic feedback, and high-resolution visuals, which collectively create a more convincing and engaging experience. These immersive qualities can enhance storytelling, gameplay mechanics, and player immersion, making games more enjoyable and memorable [54].

While serious games are digital games designed with the primary purpose of education, training, or awareness rather than mere entertainment. These games leverage interactive and engaging gameplay mechanics to promote knowledge acquisition, skill development, and

behavioral change. By simulating real-world scenarios, serious games provide users with a safe and effective platform to learn and practice new concepts and techniques in various fields such as healthcare, military, and education [55].

2.3.2 Medicine

In the field of medicine, Extended Reality (XR) technologies are revolutionizing various aspects of healthcare by providing innovative solutions for training, rehabilitation, and therapy [56]. Virtual Reality (VR) and Augmented Reality (AR) are particularly impactful in medical education and surgical training. VR simulations create highly realistic environments where medical professionals can practice surgical procedures, enhancing their skills and reducing the risk of errors in real operations [57].

Additionally, AR is used to overlay critical information during surgeries, such as patient data and anatomical guides, directly onto the surgeon's field of view, thereby improving precision and outcomes [58]. Rehabilitation programs also benefit from XR technologies; VR-based exercises make physical therapy more engaging for patients, leading to higher compliance and better recovery rates [59].

Furthermore, XR is utilized in pain therapy, where VR environments can distract patients from pain and anxiety during medical procedures or chronic pain treatment. This innovative approach not only enhances patient comfort but also reduces the need for pain medications, contributing to better overall health management [60].

2.3.3 Education

Extended Reality (XR) technologies are transforming the landscape of education by offering immersive and interactive learning experiences. Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR) are used to create virtual classrooms and simulations that engage students in hands-on learning activities.

VR environments allow students to explore complex subjects through interactive simulations, such as historical events, scientific experiments, or architectural designs. This experiential learning approach enhances comprehension and retention by providing a more engaging and memorable educational experience [61].

AR enhances traditional learning materials by overlaying digital content onto physical objects or environments, making learning interactive and contextual. For example, AR applications can provide real-time translations of text, interactive 3D models for anatomy lessons, or virtual field trips to historical landmarks [62].

MR combines elements of both VR and AR to create collaborative learning environments where students and teachers interact with virtual objects while maintaining awareness of the physical classroom. This technology supports teamwork, problem-solving skills, and

creativity by allowing students to manipulate virtual elements and collaborate on projects in real-time [63].

Overall, XR technologies in education promote active learning, critical thinking, and creativity by providing students with immersive and interactive learning environments that transcend traditional teaching methods.

2.3.4 Industry

Extended Reality (XR) technologies are increasingly integrated into industrial applications, offering transformative solutions for training, maintenance, design, and collaboration. Virtual Reality (VR) is utilized for immersive training simulations, allowing employees to practice complex procedures and scenarios in a safe and controlled environment. This not only reduces training costs but also minimizes risks associated with hands-on training [64].

Augmented Reality (AR) and Mixed Reality (MR) are employed in predictive maintenance, where technicians can visualize real-time data and diagnostic information overlaid onto machinery. This capability enhances efficiency by enabling proactive maintenance and minimizing downtime [65].

In collaborative design, AR and MR technologies facilitate real-time collaboration among teams by overlaying virtual models onto physical spaces. This enables engineers and designers to visualize, interact with, and modify designs collectively, regardless of geographical location. Such collaborative environments foster innovation, improve decision-making, and accelerate product development cycles in various industries [66].

2.3.5 Training

Extended Reality (XR) technologies are transforming training methodologies across various industries by offering immersive and interactive simulations. Virtual Reality (VR) simulations allow organizations to train employees in realistic scenarios without the constraints and risks associated with traditional training methods. This approach is particularly valuable in high-risk environments such as aviation [67] and industrial settings [64], where XR can simulate emergency situations or complex procedures.

Augmented Reality (AR) and Mixed Reality (MR) are also employed in training contexts to enhance on-the-job learning and skill development [68]. AR overlays real-time data and instructions onto physical equipment or environments, providing immediate guidance and support to workers during tasks like equipment maintenance or assembly processes. MR, with its ability to blend virtual and physical elements seamlessly, enables collaborative training environments where teams can interact with virtual prototypes or visualize complex systems in real-time.

Moreover, XR-based training programs are scalable and cost-effective compared to traditional

methods, offering customized learning experiences that adapt to individual learner needs. By incorporating interactive elements, spatial awareness, and realistic feedback mechanisms, XR enhances retention rates and practical skills acquisition, ultimately improving workforce efficiency and safety standards.

Table 2.6: Applications of Extended Reality (XR)

Sector	VR	AR	MR
Gaming	Enhances gaming experiences through immersive environments	Overlays game elements onto real world	Combines virtual and real-world interactions
Medicine	Surgical training, rehabilitation, pain therapy	Overlay of patient data and anatomical guides during surgeries	Enhanced collaborative learning and simulation environments
Education	Virtual classrooms, interactive simulations	Augmented learning materials, real-time translations	Collaborative learning environments with virtual objects
Industry	Immersive training simulations, maintenance	Predictive maintenance, real-time data visualization	Collaborative design, virtual prototypes
Training	Realistic training simulations, on-the-job learning	Overlays real-time data and instructions on physical environments	Blends virtual and physical elements for complex systems

Chapter 3

State of the Art in Photogrammetry and Innovative Techniques

3.1 Introduction to Photogrammetry

3.1.1 Definition of Photogrammetry

Photogrammetry is a technique used to obtain reliable information about physical objects and the environment by recording, measuring, and interpreting photographic images. It is a method of creating 3D models from 2D images based on the principle of triangulation, which involves capturing multiple overlapping images from different angles and using them to determine the spatial coordinates of points on the object's surface.

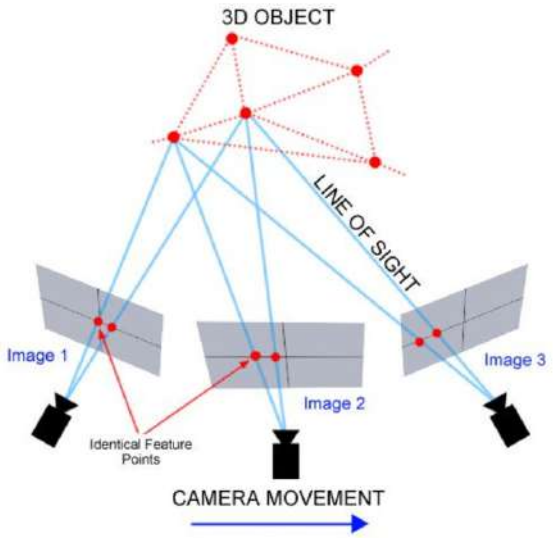


Figure 3.1: Image acquisition in photogrammetry

The fundamental concept behind photogrammetry is the extraction of geometric information from photographic images. This process involves capturing multiple images of the object or scene from various angles, identifying and marking key points or features in these images,

establishing correspondences between the key points in different images, and then calculating the 3D coordinates of these key points by triangulation. Finally, these 3D coordinates are used to create a detailed 3D model of the object or scene. Each of these steps will be elaborated in detail in the subsequent sections.

Photogrammetry can be broadly classified into two categories:

- **Aerial Photogrammetry:** This involves taking photographs from an aerial perspective, typically using drones or aircraft. It is widely used in surveying, mapping, and environmental monitoring.
- **Terrestrial Photogrammetry:** This involves taking photographs from the ground level, often using handheld or tripod-mounted cameras. It is commonly used in applications like architecture, archaeology, and forensics.

One of the key advantages of photogrammetry is its ability to capture detailed and accurate 3D information from images taken with relatively inexpensive equipment, such as consumer-grade digital cameras. Additionally, photogrammetry can be used in combination with other 3D reconstruction techniques, such as LiDAR, to enhance the accuracy and detail of the final models.

In recent years, advancements in computer vision and machine learning have significantly improved the capabilities and automation of photogrammetry processes, making it an increasingly powerful tool in various fields of research and industry.

3.1.2 Traditional Photogrammetry Techniques

Traditional photogrammetry techniques have been used for decades to derive spatial information from photographic images. These techniques are grounded in the principles of geometry, optics, and projective mathematics. Here we outline some of the classical methods and their evolution over time [69].

Analog Photogrammetry (1850s - 1980s)

Analog photogrammetry, also known as film-based photogrammetry, involves the use of photographic film and optical instruments. The primary tools in analog photogrammetry include:

- **Stereoplotters:** Devices that use pairs of photographs (stereo pairs) to create three-dimensional models. Stereoplotters allow for manual measurement and plotting of features from stereo images.
- **Aerial Cameras:** High-resolution cameras mounted on aircraft to capture large-scale images of the Earth's surface. These images are used in mapping and surveying.

The process involves capturing overlapping photographs, typically from aerial surveys, and

using stereoscopic viewing techniques to interpret the 3D structure of the terrain or object. This method requires significant manual intervention and skilled operators to ensure accurate measurements.

Analytical Photogrammetry (1960s - 1990s)

With advancements in computing technology, analytical photogrammetry emerged as an improvement over traditional analog methods. Analytical photogrammetry utilizes mathematical models to process photographic images and extract spatial data. Key components of this approach include:

- **Computers and Software:** Use of digital computers to perform complex calculations and adjustments for image rectification and triangulation.
- **Digital Scanners:** Conversion of analog photographs into digital format for easier manipulation and processing.

Analytical photogrammetry offers greater accuracy and efficiency by automating many of the manual tasks involved in analog methods. It allows for more precise adjustment of images and correction of distortions caused by camera lenses and other factors.

Evolution of Traditional Techniques (1990s - Present)

The evolution from analog to analytical photogrammetry marked a significant shift in the field, paving the way for modern digital photogrammetry. Key developments include:

- **Transition to Digital Imagery:** The shift from film-based to digital cameras improved the resolution and ease of image acquisition.
- **Automation of Processes:** Introduction of software tools to automate tasks such as image matching, triangulation, and model generation.
- **Integration with Geographic Information Systems (GIS):** Enhanced capabilities for storing, analyzing, and visualizing spatial data in conjunction with other geographic information.

Traditional photogrammetry techniques laid the foundation for contemporary practices by establishing core principles and methodologies that continue to be refined and expanded upon with new technologies.

Modern Photogrammetry Approaches

Modern photogrammetry approaches have evolved significantly with the advent of digital technologies and computational advancements. These approaches leverage sophisticated algorithms and software tools to enhance the accuracy, efficiency, and scope of photogrammetric applications. Here we explore some of the key methods:

- **Structure-from-Motion (SfM):** SfM is a photogrammetric technique that estimates 3D structures from 2D image sequences. It reconstructs the 3D geometry of a scene by establishing correspondences between image features and then computing the camera positions and scene structure [70].
- **Multi-view Stereo (MVS):** MVS extends the principles of stereo vision to multiple images captured from different viewpoints. It aims to generate dense 3D reconstructions by matching corresponding points across multiple images and estimating their 3D positions.
- **Digital Surface Model (DSM):** DSMs are representations of the Earth's surface or other objects, including buildings and vegetation, created by interpolating point data obtained from aerial photographs or satellite imagery. They provide detailed information about terrain elevation and object heights.
- **Point Cloud Processing:** Advances in computing power have enabled the processing of large-scale point clouds generated from photogrammetric data. Point cloud processing involves segmentation, classification, and analysis of 3D point data to extract meaningful information.

Table 3.1: Comparison of Photogrammetry Techniques

Parameter	SfM	MVS	DSM	Point Cloud Processing
Computational Demand	Medium	High	Medium	High
Software Availability	High	Medium	High	Medium
Equipment Cost	Low	High	Medium	Medium
Result Production Time	Medium	High	High	High

These modern photogrammetry approaches have revolutionized applications such as urban planning, environmental monitoring, archaeology, and cultural heritage preservation. They offer higher precision and automation compared to traditional methods, making them indispensable tools in contemporary spatial data acquisition and analysis.

3.2 Photogrammetry Workflow and Tools

3.2.1 Standard Photogrammetry Workflow

Step 1: Data Acquisition

1.1 Environment Setup The environment setup phase in photogrammetry is fundamental as it establishes the conditions under which images will be captured. This step involves meticulously preparing the physical surroundings to optimize the quality and consistency of the captured images.

Ensuring proper lighting is paramount during environment setup. Adequate lighting conditions across the scene help minimize shadows and ensure uniform illumination, which is crucial for accurate feature extraction and image matching. Techniques such as diffused lighting or strategically placed light sources may be employed to achieve optimal results.

Stability is another critical consideration. Securing the camera and any supporting equipment to prevent unintended movement is essential. Even slight vibrations or shifts can compromise the alignment of images, leading to inaccuracies during the reconstruction process. Methods to enhance stability include the use of tripods, mounts, or fixed platforms.

Careful consideration of the background is also part of environment setup. A background with suitable contrast and minimal distractions can facilitate easier identification and tracking of key features in the images. This helps in ensuring robust image matching and triangulation, contributing to the overall accuracy of the 3D model generated.

Furthermore, calibration of the camera settings is essential. Adjusting parameters such as focus, aperture, and exposure settings ensures optimal image clarity and consistency across all captured images. Calibration procedures may involve using calibration charts or software tools provided by the camera manufacturer.

Image Acquisition Image acquisition involves the capture of multiple images of the object or scene from diverse angles using cameras. This process requires meticulous planning to ensure comprehensive coverage and adequate overlap between images.

Effective image acquisition begins with strategic placement of cameras to capture the subject matter from various perspectives. This approach facilitates the collection of sufficient visual data, enabling the photogrammetry software to accurately triangulate and reconstruct the 3D model.

During this phase, considerations such as camera settings (for example, aperture, shutter speed, ISO sensitivity) are carefully adjusted to achieve optimal image quality. Additionally, factors like lighting conditions and environmental stability play crucial roles in ensuring uniformity and clarity across the captured images.

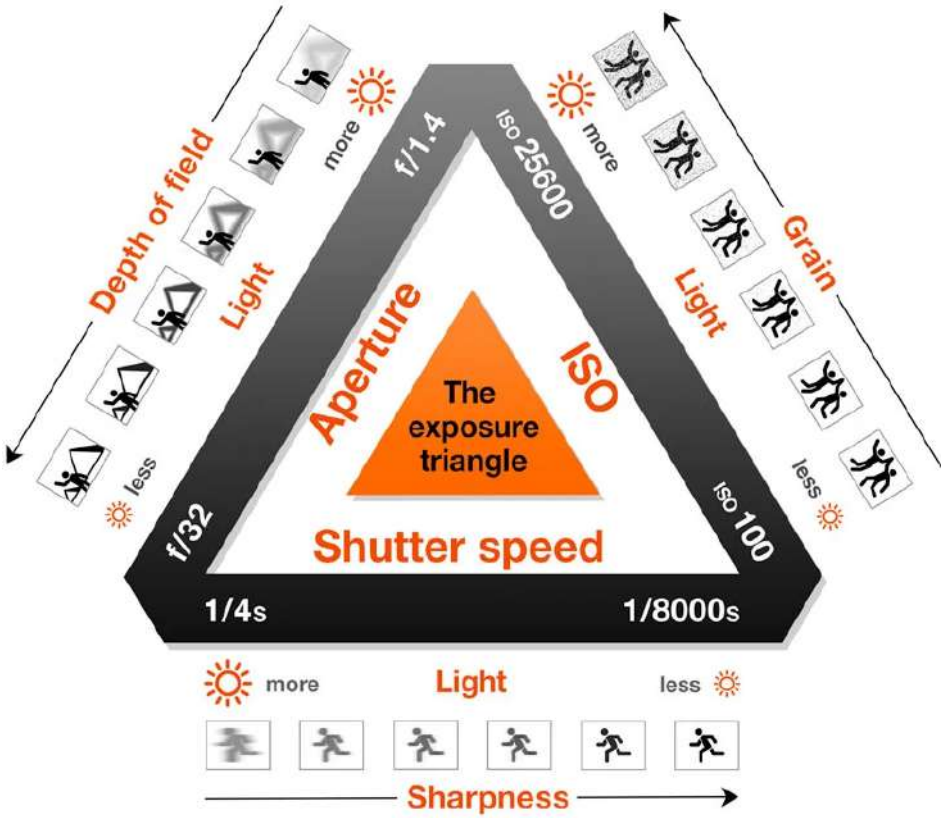


Figure 3.2: Illustration of the Exposure Triangle in Image Acquisition.

Proper planning and execution of image acquisition are essential for overcoming potential challenges such as occlusions, reflections, or varying lighting conditions that could affect the quality and consistency of the image dataset. By capturing a robust set of images with adequate overlap, photogrammetry software can effectively extract key features and generate accurate 3D models of the object or scene.

When scanning an environment, such as a room or outdoor landscape, the approach to image acquisition involves strategically placing cameras to capture the entire area from various viewpoints. This method aims to encompass the entire scene with overlapping images, ensuring that all aspects and details are adequately represented in the dataset.

In contrast, when capturing a single object or subject, the focus shifts to capturing detailed views from multiple angles around the object. This approach emphasizes close-range photography to capture fine details and textures, often requiring precise control over lighting and camera positioning to minimize shadows and reflections.

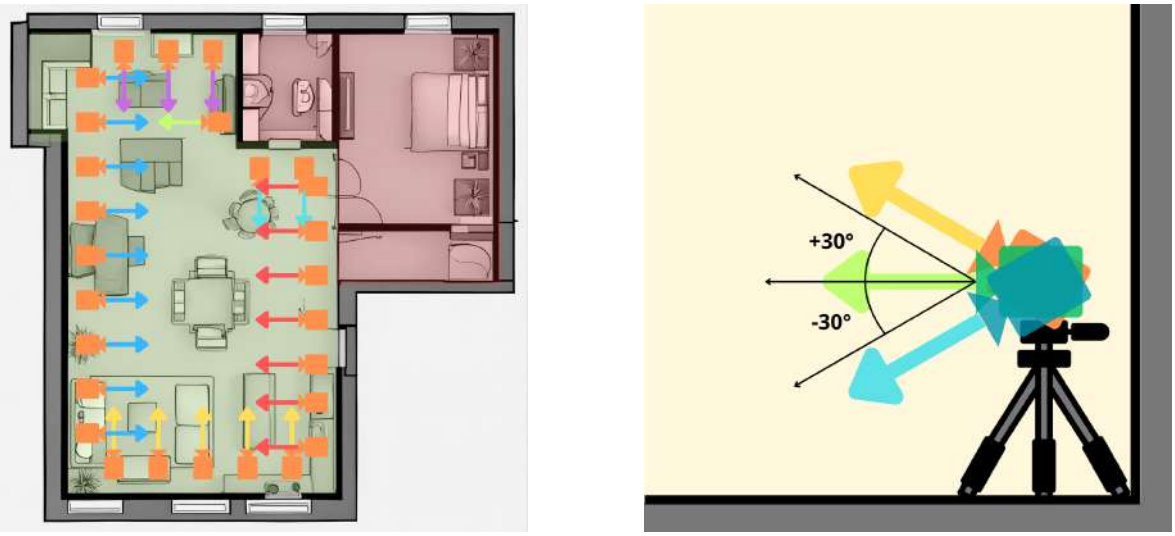


Figure 3.3: Illustration of Image Acquisition Techniques for Indoor Environments.

Figure 3.4: Images acquisition from multiple angles around the object

Image Pre-processing (Optional) Image pre-processing is an optional but beneficial step in the photogrammetry workflow aimed at improving the quality and accuracy of the captured images. This stage involves several techniques and adjustments to enhance the images before they are processed for feature extraction and 3D model generation.

One crucial aspect of image pre-processing is enhancing image quality. This can include adjustments to brightness, contrast, and color balance to ensure uniformity across the image

set. Additionally, techniques such as noise reduction and sharpening may be applied to mitigate artifacts and enhance details.

Correcting distortions is another important aspect of image pre-processing. Lens distortions, perspective distortions, and vignetting can affect the accuracy of feature extraction and subsequent 3D reconstruction. Techniques such as lens correction and perspective adjustment are employed to rectify these distortions.

Furthermore, image alignment and registration are essential in cases where multiple images need to be seamlessly integrated. Aligning images ensures that they are correctly positioned relative to each other, which is crucial for accurate feature matching and triangulation in the photogrammetric process.

By conducting thorough image pre-processing, photogrammetry software can effectively extract key features and generate high-fidelity 3D models. This process optimizes the image dataset, resulting in more accurate spatial data and enhanced realism in the final reconstructed models [71].

Upload Images The process of uploading captured images to the photogrammetry software represents a crucial step in the workflow, facilitating further processing and reconstruction of 3D models. After images have been meticulously acquired and, if necessary, pre-processed to enhance quality and correct distortions, they are transferred to the photogrammetry software.

Uploading involves transferring the image dataset from the storage medium, such as a camera memory card or computer drive, to the software interface. This step typically includes organizing the images within the software, ensuring that they are correctly sequenced and labeled for subsequent processing stages.

During the upload process, the photogrammetry software may perform initial checks and analyses on the images. These checks can involve assessing image resolution, clarity, and consistency to determine their suitability for accurate feature extraction and model generation. Ensuring that all images are correctly uploaded and verified sets the foundation for reliable 3D reconstruction.

Moreover, uploading images involves setting parameters and configurations within the software interface. Users may specify image alignment preferences, resolution settings, and other parameters to optimize the reconstruction process based on the specific characteristics of the image dataset.

Step 2: Creating 3D Model through Photogrammetry Software

Image Matching Image matching is a critical process in photogrammetry where the software identifies corresponding points (features) across multiple images. This step is essential for establishing how different images relate spatially to each other, which is crucial for accurate 3D reconstruction.

During image matching, the software employs advanced algorithms to analyze the images and detect similarities in key features such as corners, edges, or texture patterns. These features are often identified based on their distinctiveness and repeatability across multiple views.

Once potential correspondences are identified, the software calculates the similarity between the features in different images. Various techniques such as feature descriptors (for example, SIFT, SURF) and matching algorithms (for example, RANSAC) are used to robustly match the features, even in the presence of noise, occlusions, or changes in viewpoint.

The accuracy and reliability of image matching directly impact the quality of the final 3D model. Properly matched features enable precise triangulation and reconstruction of the object or scene, ensuring that the spatial relationships between points in the real world are accurately represented in the virtual model.

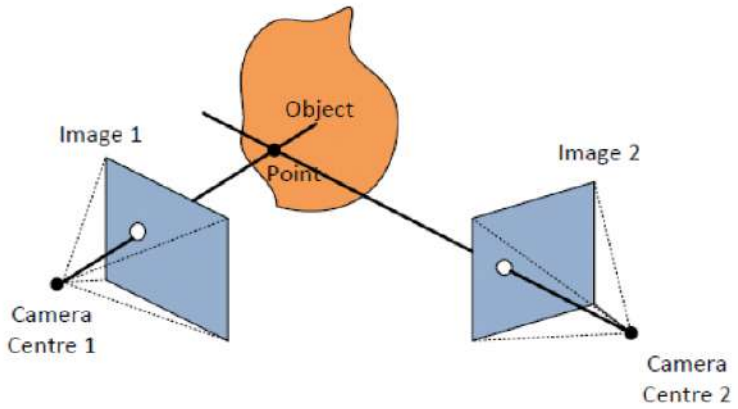


Figure 3.5: Image matching

Feature Extraction After identifying corresponding points through image matching, the next step in the photogrammetry workflow is feature extraction. This process involves extracting detailed information about the identified features from each image in the dataset.

During feature extraction, the photogrammetry software retrieves precise coordinates of the identified features in 2D space. Additionally, the software may extract additional descriptors or attributes associated with each feature, such as scale, orientation, or texture information. These descriptors help in distinguishing and matching features across multiple images with varying viewpoints and lighting conditions.

Feature extraction plays a crucial role in the accuracy and robustness of the photogrammetric reconstruction process. By accurately extracting and characterizing features, the software enhances its ability to triangulate the 3D coordinates of these features and subsequently generate a detailed and accurate 3D model of the object or scene.

Advanced feature extraction algorithms and techniques, often based on machine learning and computer vision principles [72], are employed to handle complex scenes and ensure reliable feature identification across diverse image datasets. This ensures that the final 3D

model maintains fidelity to the original object or scene, capturing intricate details and spatial relationships effectively.

Triangulation Software calculates the precise 3D coordinates of points on the object or scene based on their corresponding 2D positions in multiple images. This process relies on geometric principles and mathematical algorithms to reconstruct the spatial geometry from the collected image data.

During triangulation, the photogrammetry software utilizes the known positions of the camera stations (from where the images were taken) and the corresponding points identified through image matching and feature extraction. By triangulating the rays extending from each camera center through the identified features, the software determines their 3D coordinates in real-world space.

Various triangulation methods exist, including direct linear transformation (DLT) [73], bundle adjustment, and iterative methods. These methods aim to minimize errors and optimize the alignment of 3D points across multiple images, taking into account factors such as lens distortions, camera calibration, and image perspective.

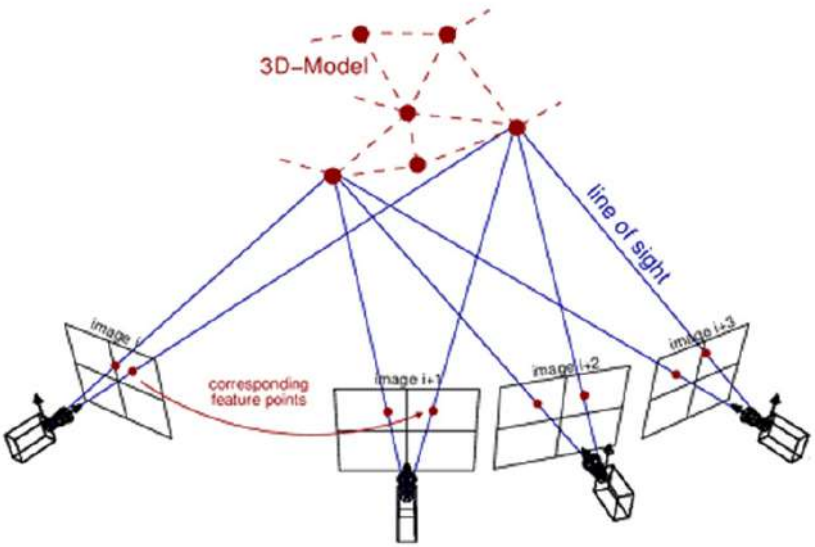


Figure 3.6: Example of bundle adjustment technique [74]

The accuracy of triangulation directly influences the quality and precision of the final 3D model. Properly triangulated points ensure that the spatial relationships and dimensions of the object or scene are faithfully represented in the digital reconstruction. Advanced photogrammetry software employs sophisticated algorithms to enhance triangulation accuracy and handle challenging scenarios such as occlusions, varying lighting conditions, and complex geometries.

Model Generation Model generation is the process where the photogrammetry software constructs a detailed 3D model from the triangulated points obtained in the previous steps.

This stage involves creating a mesh that represents the surface geometry of the object or scene and applying textures to enhance visual realism.

The first step in model generation is mesh construction. The software connects the triangulated 3D points to form a network of polygons, typically triangles or quadrilaterals, that define the surface geometry. This mesh accurately represents the shape and contours of the captured subject, providing a geometric framework for the 3D model.

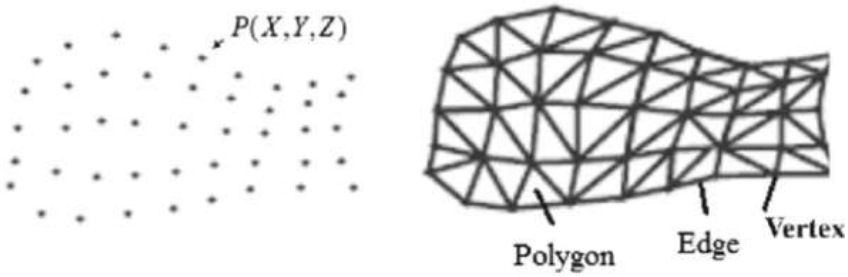


Figure 3.7: From point cloud to mesh costruction

After mesh construction, the software applies textures to the mesh. Texturing involves projecting the original photographic images onto the 3D surface, effectively wrapping the images around the mesh. This process adds color, detail, and realism to the 3D model, making it visually similar to the real-world object or scene [75].

The accuracy and quality of the final 3D model depend on various factors, including the quality of the input images, the effectiveness of the feature extraction and triangulation processes, and the capabilities of the photogrammetry software. A well-generated model can serve a wide range of applications, from virtual reality and augmented reality experiences to engineering analysis and cultural heritage documentation.

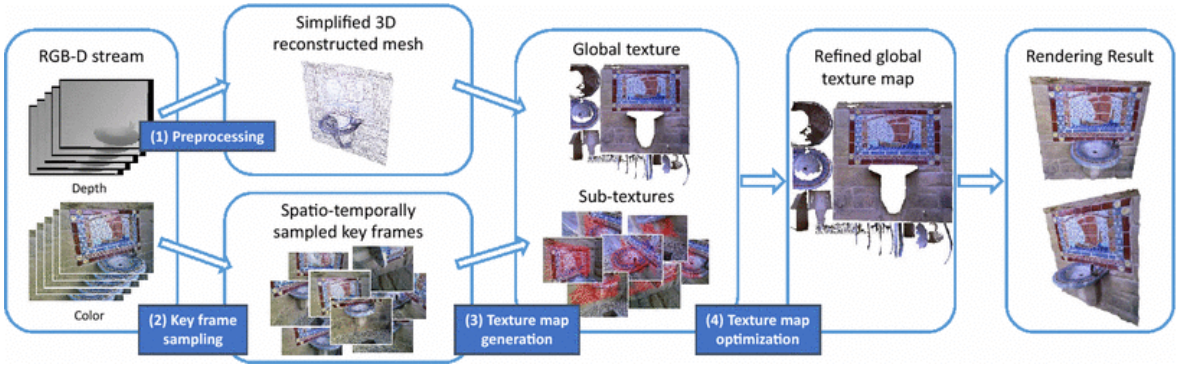


Figure 3.8: A general texture mapping process [76]

Step 3: Mesh Cleanup (Optional Post-Processing)

Repairing Repairing involves refining the initial 3D model by addressing issues such as noise, redundant points, and irregularities that detract from the overall quality. One primary

task during this phase is the elimination of stray points and artifacts that do not contribute to the accuracy or visual quality of the model [77]. These extraneous elements can result from various sources such as image noise, mismatches during feature extraction, or occlusions during image capture. Removing these points helps in reducing the complexity of the mesh and improves the subsequent processing steps. Typical software used for repairing includes MeshLab, Autodesk Meshmixer, and Blender, which offer tools for cleaning up and refining the mesh.

Remeshing Remeshing involves reconstructing the mesh to improve its quality and structure. This process includes optimizing the mesh topology by simplifying the geometry where possible while preserving critical details. Remeshing helps to reduce the number of polygons in less important areas, enhancing computational efficiency and making the model easier to manipulate and render in downstream applications [78]. Typical software used for remeshing includes MeshLab, Autodesk Meshmixer, and Blender, which provide advanced tools for adjusting the mesh topology and improving overall model quality 3.9.

Figure 3.9: The result of a remeshing

Resculpting Resculpting involves enhancing and adding details to the 3D model to achieve better accuracy and realism. This process can include refining surface textures, adding intricate details, and correcting any remaining inaccuracies. Resculpting helps to ensure that the final model is as true to the original object or scene as possible, making it suitable for high-fidelity applications such as digital archiving, animation, and virtual reality. Commonly used software for resculpting includes ZBrush, Autodesk Mudbox, and Blender, which offer a range of powerful tools for detailed 3D sculpting and fine-tuning.

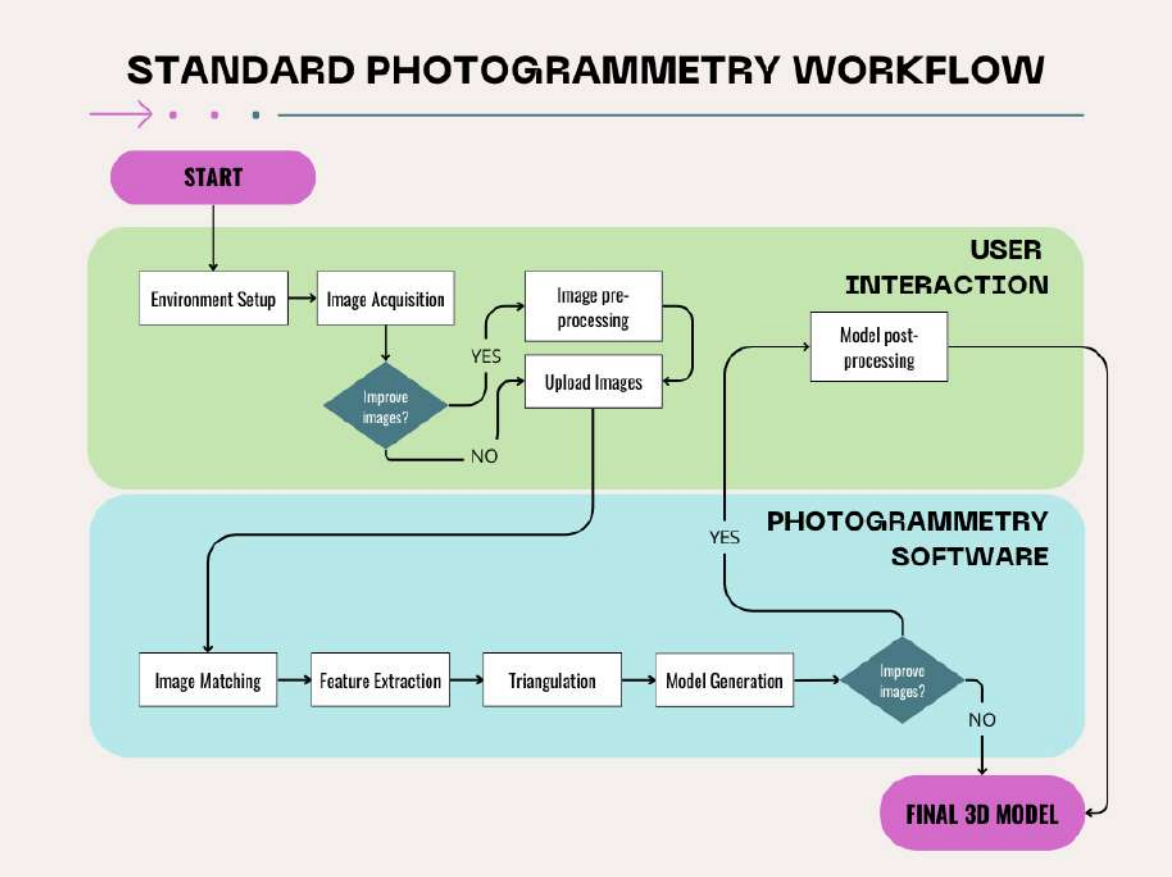


Figure 3.10: Standard Photogrammetry Workflow

3.2.2 Tools and Software for Photogrammetry

Several software tools are available for photogrammetry, each offering unique features and capabilities that cater to various needs and applications. Below are some of the most prominent tools in the field:

Meshroom [79]

Meshroom is an open-source photogrammetry software developed by AliceVision. It automates the process of 3D reconstruction, making it accessible to users with varying levels of expertise. Meshroom employs a node-based workflow, which provides flexibility and allows users to customize the reconstruction process to suit their specific requirements. This modular approach enables the integration of different algorithms and processes, making Meshroom a powerful tool for various applications. Despite being free, Meshroom offers robust features and produces high-quality 3D models, making it a popular choice among hobbyists and professionals alike.

Agisoft Metashape [80]

Agisoft Metashape is a commercial photogrammetry software known for its high accuracy and comprehensive feature set. It supports both aerial and terrestrial photogrammetry, making

it suitable for a wide range of applications, from surveying and mapping to cultural heritage preservation and visual effects production. Metashape offers advanced editing tools for post-processing, including dense point cloud generation, mesh editing, texture mapping, and orthophoto creation. Its user-friendly interface and detailed documentation make it accessible to both novice and experienced users. Additionally, Metashape provides options for batch processing and network processing, enhancing its efficiency for large-scale projects.

RealityCapture [81]

RealityCapture, developed by Capturing Reality, is renowned for its speed and efficiency in processing large datasets. This commercial software excels in handling extensive image collections and producing high-quality 3D models in a relatively short time. RealityCapture leverages advanced algorithms to achieve precise alignment and reconstruction, making it ideal for applications requiring rapid turnaround times, such as forensic analysis, virtual reality content creation, and urban planning. The software also supports a wide range of input formats, including images and laser scans, and offers comprehensive tools for post-processing and export.

Pix4D [82]

Pix4D is another commercial software that provides specialized solutions tailored to various industries, including construction, agriculture, and mapping. It offers both desktop and cloud-based processing options, allowing users to choose the platform that best fits their workflow. Pix4D's cloud processing capabilities enable users to upload images and perform complex photogrammetric analyses remotely, making it convenient for fieldwork and collaborative projects. The software includes advanced analytics tools, such as vegetation indices for agriculture and volumetric measurements for construction, which provide valuable insights for industry-specific applications. Pix4D's modular approach also allows users to select and combine different software components to create customized solutions for their needs.

3D Zephyr [83]

3D Zephyr, developed by 3Dflow, is a versatile photogrammetry software known for its user-friendly interface and powerful reconstruction capabilities. It supports both automatic and manual workflows, allowing users to adjust parameters and refine the reconstruction process according to their needs. 3D Zephyr excels in handling various types of input data, including images and laser scans, and provides advanced tools for mesh generation, texture mapping, and post-processing. Its flexible licensing options make it accessible to both professionals and hobbyists. Additionally, 3D Zephyr offers specialized features for industries such as archaeology, architecture, and industrial inspection, making it a valuable tool for diverse applications.

Table 3.2: Comparison of Photogrammetry Software

Software	**Platform**	**Cost**	**Accuracy**	**Processing Speed**	**Data Source**	**Notable Features**
Meshroom	Windows Linux	Free	High	Moderate	Standard images Equirectangular images	Node-based workflow Flexibility Open-source
Agisoft Metashape	Windows macOS Linux	$179 for Standard $3,499 for Professional	Very High	High	Standard images Equirectangular images	Advanced editing tools Supports aerial and terrestrial photogrammetry
RealityCapture	Windows	Free for individuals, small business, and students $1,250/year for Standard $1,850/year Unreal Subscription	Very High	Very High	Standard images Equirectangular images	High-speed processing Handles large datasets efficiently
Pix4D	Windows macOS Linux	Different plans and prices Free limited version available	High	High	Standard images Equirectangular images	Cloud-based processing Specialized industry solutions
3D Zephyr	Windows	$199 for Lite $250 for 30 days $4,200 for Pro Free limited version available	High	High	Standard images Equirectangular images	User-friendly interface Robust reconstruction algorithms

These tools automate much of the photogrammetry workflow, making it accessible to users with varying levels of expertise. The choice of software often depends on the specific requirements of the project, such as the desired accuracy, processing speed, and available budget. By leveraging these sophisticated tools, users can efficiently create detailed and accurate 3D models for a wide range of applications.

3.3 Applications and Case Studies

Photogrammetry finds applications in diverse fields, offering precise and cost-effective solutions for various industries. Below are some key areas where photogrammetry has made significant impacts.

3.3.1 Archaeology

In archaeology, photogrammetry is invaluable for documenting and reconstructing archaeological sites and artifacts. High-resolution 3D models enable archaeologists to preserve the exact state of a site or artifact at a given time, facilitating detailed analysis and virtual conservation. By capturing intricate details and textures, photogrammetry aids in the study of wear patterns, tool marks, and other fine details that are crucial for understanding historical usage and cultural significance [84]. This technology also supports virtual repatriation projects, allowing digital replicas of artifacts to be shared with originating communities or displayed in virtual museums [85].

Figure 3.11: Digital replica of the Hypogeum of Crispia Salvia (Marsala, Italy) [86]

3.3.2 Architecture and Engineering

In the fields of architecture and engineering, photogrammetry is employed to capture and analyze structures and landscapes. This technology supports various stages of the construction lifecycle, from initial planning and design to ongoing maintenance and preservation. By creating accurate 3D models of buildings and infrastructure, professionals can perform structural analyses, monitor changes over time, and identify potential issues before they become critical [87]. Additionally, photogrammetry facilitates the creation of detailed as-built models, which are essential for renovation and restoration projects. The ability to generate precise

topographical maps and digital elevation models also aids in site planning and environmental impact assessments [88].

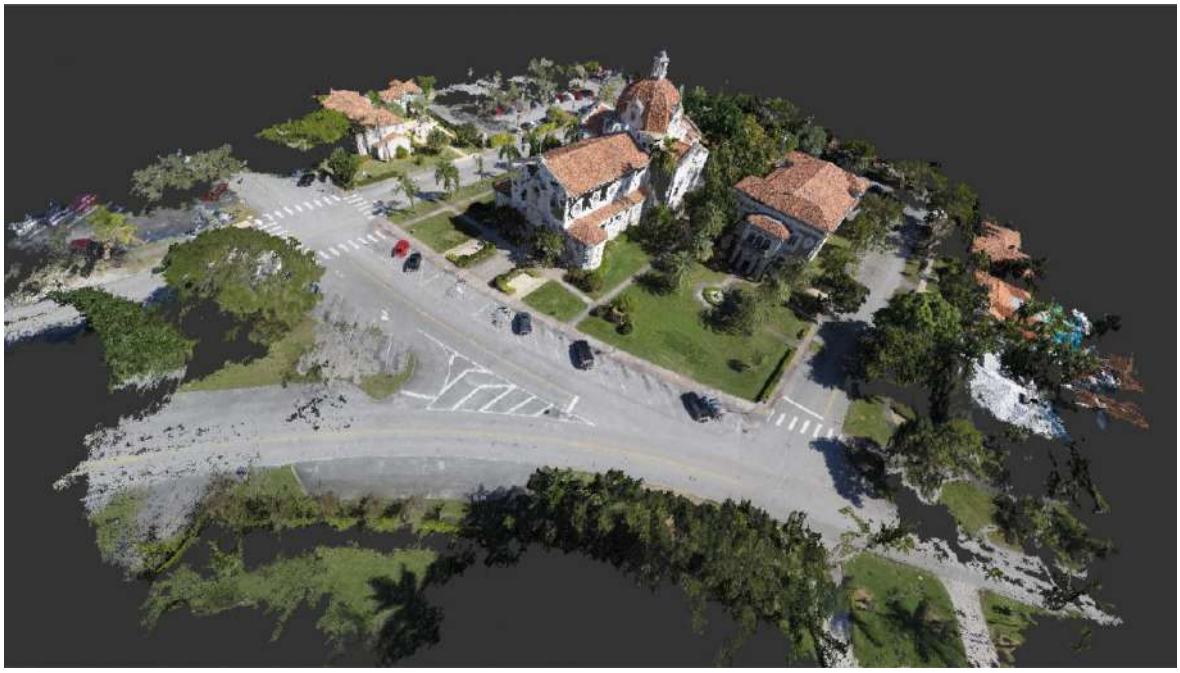


Figure 3.12: Point cloud of the Church of the Little Flower (Coral Gables, Florida) [89]

3.3.3 Extended Reality (XR) and Virtual Reality

Photogrammetry is extensively used in the creation of immersive virtual environments for Extended Reality (XR) and Virtual Reality (VR) applications. By transforming real-world objects and scenes into highly detailed 3D models, photogrammetry enhances the realism and authenticity of virtual experiences (Fig.: 3.13). This technology is pivotal in industries such as gaming, film, and education, where lifelike simulations and interactive environments are paramount. In XR, photogrammetry enables the seamless integration of digital and physical worlds, enhancing user engagement and providing more intuitive interfaces for virtual design and training programs [90].

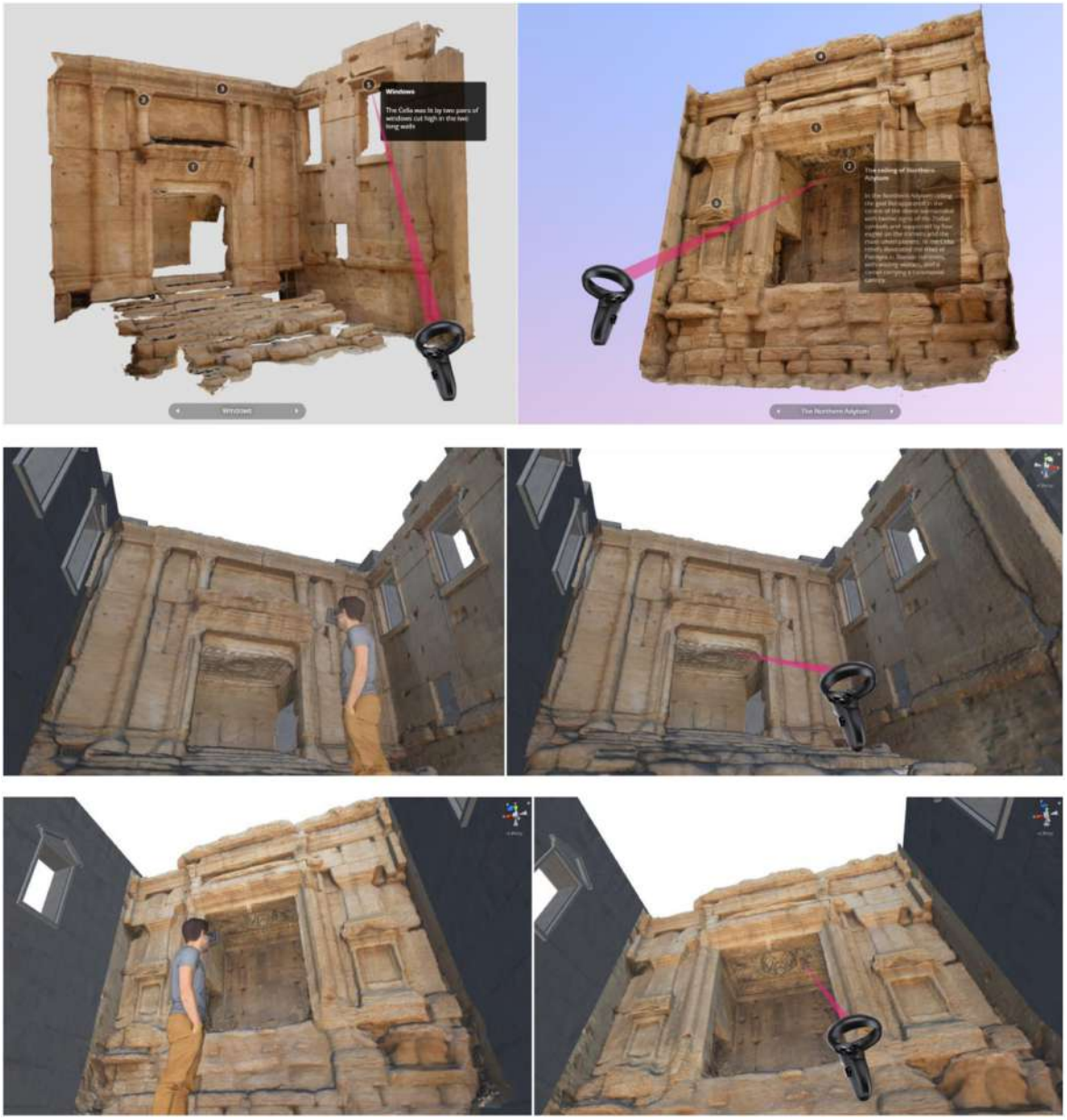


Figure 3.13: Visitor's interaction with the Northern and Southern Adytum inside the Cella viewing visual data through proposed interactive virtual hot points and annotations [90]

3.3.4 Forensic Science

In forensic science, photogrammetry is applied to document and analyze crime scenes, accident sites, and evidence. This non-invasive method ensures that critical details are preserved accurately, enabling investigators to revisit the scene virtually at any time [91]. High-resolution 3D models can be used to perform measurements, reconstruct events, and present findings in court. The precision and reliability of photogrammetric data are crucial for ensuring the integrity of forensic analyses and providing compelling visual evidence that supports expert testimonies.

Figure 3.14: First-person snapshot of a reconstruction shared amongst three users [91]

3.3.5 Medical Applications

Photogrammetry also finds applications in the medical field, particularly in areas such as prosthetics design, surgical planning, and patient monitoring. By creating detailed 3D models of patients' anatomy, healthcare professionals can design custom prosthetic devices that fit precisely and function effectively [92]. In surgical planning, photogrammetry aids in visualizing complex anatomical structures, allowing surgeons to plan and rehearse procedures with greater accuracy [93]. Additionally, this technology is used in dermatology and reconstructive surgery to monitor the progression of conditions and evaluate the outcomes of treatments over time [94].

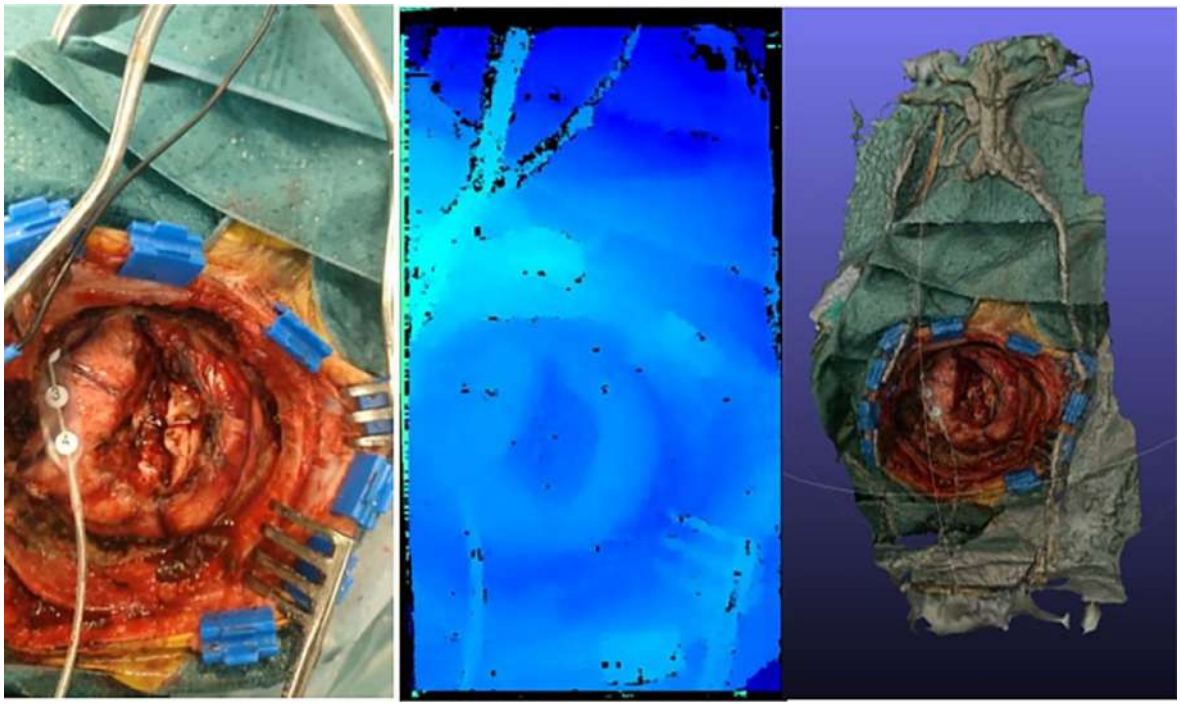


Figure 3.15: Example of Meshroom reconstruction. From the extracted frames of the video (a), the software obtains the depth maps (b), which are used to create the final mesh (c). [92]

3.3.6 Environmental Monitoring

In environmental monitoring, photogrammetry is used to track changes in natural landscapes, assess the health of ecosystems, and monitor the impacts of human activities. High-resolution 3D models and orthophotos provide valuable data for studying erosion, deforestation, glacier retreat, and other environmental phenomena [95]. This technology supports conservation efforts by offering precise, up-to-date information that aids in the management and protection of natural resources. Furthermore, photogrammetry is instrumental in disaster response, helping to assess damage, plan recovery efforts, and improve resilience to future events [96].

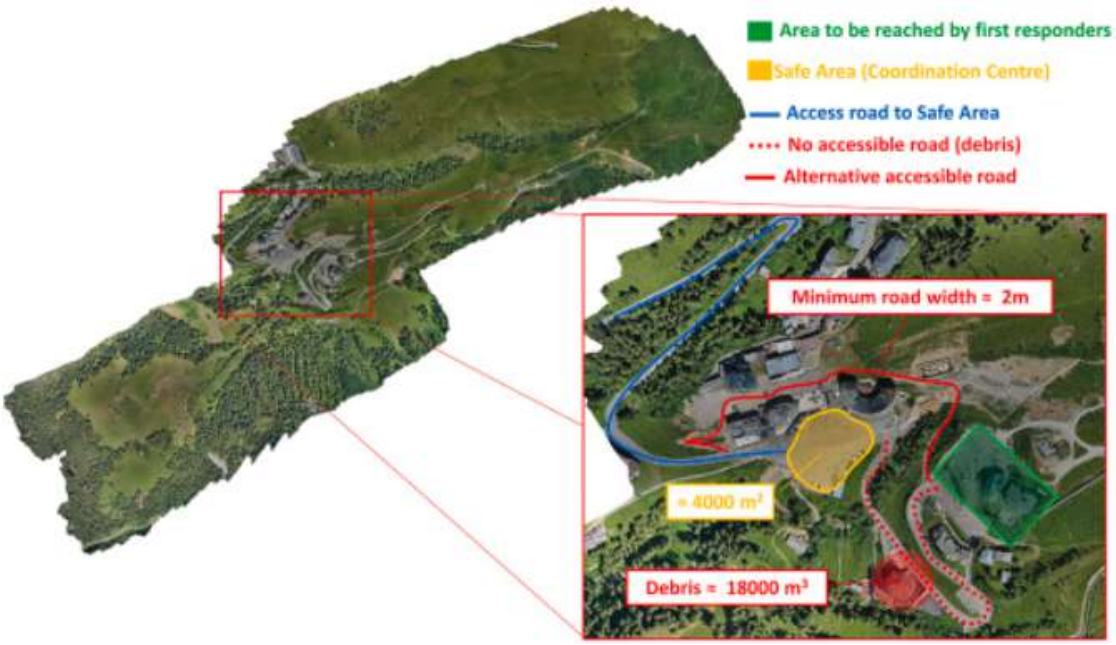


Figure 3.16: Orthophoto: simulation of a fast investigation of the overall situation of a portion of the affected area [96]

3.4 Innovative Techniques in Photogrammetry

Although photogrammetry is a well-established and extensively used methodology, it remains a field of significant interest and ongoing innovation within the scientific community. As advancements in technology and computing continue to evolve, new techniques and enhancements are being developed to push the boundaries of what photogrammetry can achieve. This section explores some of the most recent and innovative techniques that are shaping the future of photogrammetry, reflecting the dynamic nature of the field and its potential for further advancement.

3.4.1 Deep Learning in Photogrammetry

The integration of deep learning and artificial intelligence (AI) into photogrammetry represents a significant leap forward in the automation and enhancement of 3D reconstruction

processes. Deep learning techniques, particularly convolutional neural networks (CNNs), are increasingly used to automate feature extraction, model refinement, and error correction.

In traditional photogrammetry, feature extraction from images often relies on manual or semi-automated processes, which can be time-consuming and error-prone. Deep learning algorithms, however, can learn from large datasets of annotated images, allowing them to automatically detect and classify complex features with high accuracy. For instance, these algorithms can identify keypoints, edges, and textures in images, significantly speeding up the feature extraction process and improving the reliability of the data used for 3D reconstruction [97].

Model refinement benefits from deep learning as well, with AI algorithms capable of predicting and correcting distortions in the 3D models. By training on diverse datasets, these models can learn to recognize common artifacts and anomalies in photogrammetric models and apply corrective measures. This leads to more accurate and visually appealing 3D models by minimizing errors such as noise, inconsistencies, and artifacts [98].

Additionally, the use of deep learning enhances the scalability of photogrammetry applications. AI-driven algorithms can process and analyze vast amounts of image data efficiently, making it feasible to tackle large-scale projects and complex environments. Innovations in deep learning also pave the way for new applications, such as real-time 3D reconstruction and automated mapping, where high-speed processing and accuracy are crucial [99].

3.4.2 Real-time Photogrammetry

Real-time photogrammetry is an advanced technique that enables the instantaneous capture and processing of 3D data, facilitating immediate feedback and interactive experiences. This approach involves using high-speed cameras and sophisticated algorithms to reconstruct 3D models dynamically as images are captured [100].

The core challenge of real-time photogrammetry lies in handling the substantial amount of data generated and ensuring high accuracy while processing data at such speeds. Advanced algorithms must be optimized to perform feature extraction, matching, and triangulation quickly, without sacrificing the quality of the 3D models. Real-time photogrammetry requires powerful computational resources and efficient data management strategies to address these challenges.

Applications of real-time photogrammetry are diverse and include virtual reality (VR) and augmented reality (AR), where interactive and immersive experiences benefit from real-time updates of the 3D environment. In industries such as gaming and simulation, real-time photogrammetry enables the creation of highly detailed and responsive virtual worlds [101]. Additionally, it has potential applications in fields like robotics and autonomous navigation, where immediate environmental feedback is essential for decision-making and control [102].

Despite its advantages, real-time photogrammetry faces limitations, such as the need for high-performance hardware and potential trade-offs between processing speed and model detail. Researchers and developers continue to explore solutions to enhance the efficiency and accuracy of real-time photogrammetry systems, aiming to overcome these challenges and expand their applications.

3.4.3 Neural Radiance Fields (NeRF)

Neural Radiance Fields (NeRF) represents an advanced approach in 3D reconstruction that leverages deep learning to generate photorealistic 3D scenes from 2D images. NeRF models use a neural network to represent a scene as a continuous volumetric field. The core idea is to encode scene geometry and appearance into a neural network that predicts the color and density of each point in space, enabling the synthesis of novel views from any perspective.

NeRF operates by training a neural network to optimize the volumetric representation of a scene, which is achieved by minimizing the discrepancy between the rendered views from the model and the actual input images. The model processes input images taken from various angles, learning to interpolate and generate consistent, high-quality imagery.

One of the main advantages of NeRF is its ability to produce highly detailed and photorealistic renderings of complex scenes, capturing fine-grained textures and intricate lighting effects. Unlike traditional methods that rely on discrete points or meshes, NeRF provides a continuous representation, allowing for smoother and more accurate reconstructions [103].

NeRF is particularly effective in scenarios requiring high-fidelity scene reproduction and can be applied in virtual reality, digital content creation, and other fields where realistic scene representation is crucial. However, NeRF requires significant computational resources and training time, which can be a limitation for real-time applications.

By incorporating NeRF into the workflow, it is possible to achieve unprecedented levels of detail and realism in 3D scene reconstruction, making it a valuable tool for advanced applications in computer vision and graphics.

3.5 Other Techniques for 3D Reconstruction

While photogrammetry remains a prominent technique for 3D reconstruction, other advanced methods offer complementary advantages in various applications. This section provides a concise overview of these alternative techniques, highlighting their unique capabilities and use cases.

3.5.1 Laser Scanning

Laser scanning technologies are widely used for capturing precise 3D data of environments and objects by measuring the distance between the scanner and the surface being scanned.

LiDAR (Light Detection and Ranging)

LiDAR (Light Detection and Ranging) is a remote sensing technology that uses laser pulses to measure distances and generate precise 3D models of environments. By emitting rapid pulses of laser light and measuring the time it takes for each pulse to return, LiDAR systems create detailed point clouds representing the scanned area [104].

The primary advantage of LiDAR is its ability to capture high-resolution data across various terrains and environmental conditions. This technology excels in providing accurate topographical data, which is invaluable in applications such as geospatial mapping, forestry management, and archaeological site documentation. LiDAR's ability to penetrate vegetation and capture ground surfaces makes it particularly useful for surveying large and complex areas [105].

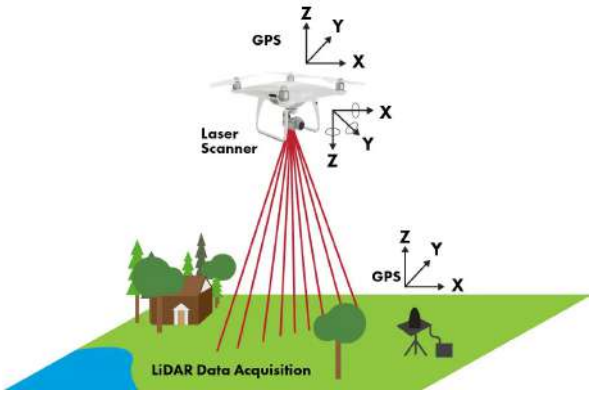


Figure 3.17: How usually works UAV-LiDAR

LiDAR data is often used in combination with other techniques, such as photogrammetry, to enhance the overall quality and completeness of 3D models. For example, LiDAR can provide detailed elevation data, while photogrammetry can contribute to texture and color information. This integration of methods allows for the creation of highly accurate and visually rich 3D representations [106].

Terrestrial Laser Scanning

Terrestrial laser scanning involves the use of ground-based laser scanners to capture high-resolution 3D data of buildings, landscapes, and other structures. These scanners rotate to capture data from multiple angles, producing dense point clouds that represent the scanned area. The process typically involves setting up the laser scanner at various positions around the target area to ensure comprehensive coverage and accurate data collection.

The technology operates by emitting laser pulses and measuring the time it takes for each pulse to return after hitting an object. This time-of-flight measurement is then used to calculate the distance between the scanner and the object, allowing for precise 3D coordinates to be determined for millions of points. The resulting point cloud can be used to create detailed 3D models and maps of the environment.

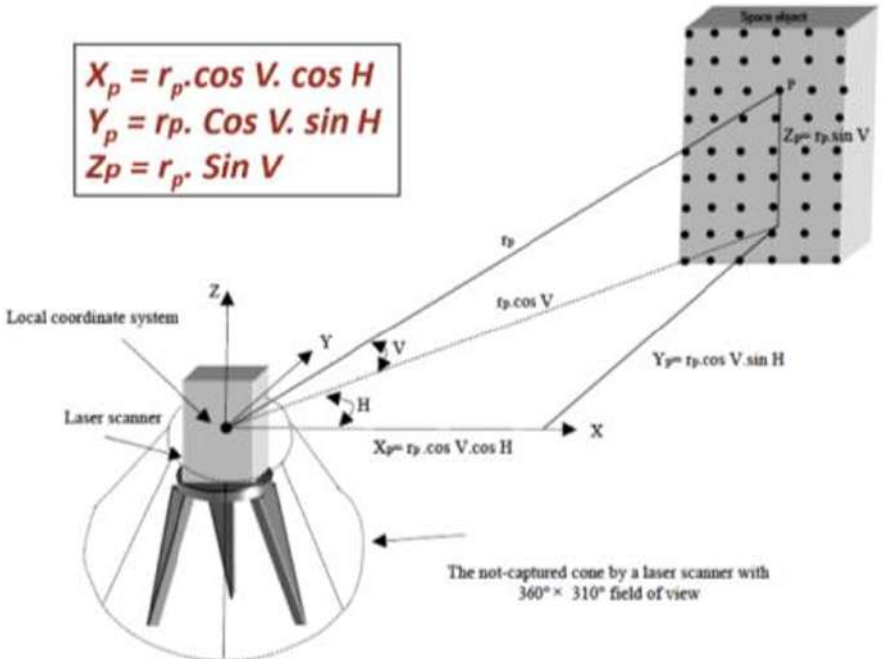


Figure 3.18: A Terrestrial Laser Scanning diagram [107]

Terrestrial laser scanning is widely employed in various fields due to its accuracy and efficiency. In architecture and civil engineering, it is used for creating as-built documentation, performing structural analysis, and aiding in construction planning and monitoring [108]. Heritage preservation benefits from this technology through the accurate documentation of historical sites and artifacts, enabling conservation efforts and virtual reconstruction [109].

Additionally, terrestrial laser scanning is utilized in mining [110] and environmental monitoring [111]. Its ability to capture intricate details and produce high-resolution data makes it an invaluable tool for surveying complex and large-scale areas. Advanced processing software can further enhance the data, allowing for the extraction of meaningful information such as measurements, contours, and surface models.

The integration of terrestrial laser scanning with other technologies, such as photogrammetry, can provide even more comprehensive and versatile datasets. This integration facilitates more robust analysis and visualization, supporting decision-making processes in various industries [112].

3.5.2 Computer Vision-based Reconstruction

Computer vision-based reconstruction involves using algorithms to infer 3D shapes from multiple 2D images or video sequences. This technique relies on analyzing overlapping images taken from different viewpoints to reconstruct the three-dimensional structure of the scene.

Key methods in computer vision-based reconstruction include Structure-from-Motion (SfM) and Multi-View Stereo (MVS). SfM algorithms are crucial as they estimate the camera positions and orientations, and reconstruct the sparse 3D point cloud by matching features across images. This sparse point cloud provides a general structure of the scene but lacks the detailed surface information.

MVS algorithms are used optionally after SfM to refine the initial 3D model. They create a dense point cloud by improving depth estimation and resolving ambiguities. The dense point cloud adds detailed surface information and enhances the resolution of the 3D model, resulting in a more accurate and complete representation of the scene.

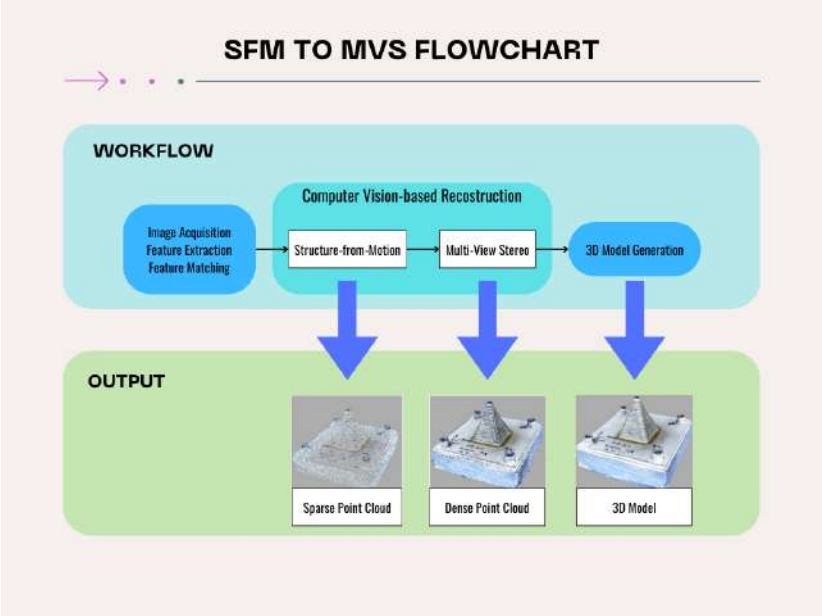


Figure 3.19: A basic 3D Reconstruction Workflow that use Computer Vision algorithm (model images from [113]

These techniques leverage computer vision algorithms to handle complex scenes and varying lighting conditions, providing flexibility and adaptability. SfM is essential for creating the foundational 3D structure, while MVS is optional and enhances the detail and precision of the reconstruction. Computer vision-based reconstruction is applicable to a wide range of scenarios, including mobile and aerial platforms, and is often used in conjunction with other techniques to further enhance model accuracy and completeness [114].

Chapter 4

Enhancing Quality of Experience (QoE) through IoT, XR, and Photogrammetry

4.1 Definition of QoE

With the advent of broadband networks and high-performance user devices, Internet services have evolved to include a diverse array of applications such as video streaming, virtual reality, and cloud gaming. As these offerings are fundamentally service-oriented, service providers prioritize the assurance of high service quality. Concurrently, end-users assess the services based on their perceived quality. Consequently, substantial efforts have been directed towards the rigorous measurement and continuous enhancement of service quality metrics.

The importance of understanding service quality from the end user's perspective has been recognized. To describe the end users' viewpoints about the service quality, the new terminology, Quality of Experience (QoE), was introduced at the beginning of this century [115]. Since QoE deals with subjective matters, it is challenging to establish a generic theory or model, and it does not depend solely on Quality of Service (QoS). To grasp user satisfaction, feelings, complaints, and so on, collaboration with other research fields including usability engineering, psychology, and biometrics is increasingly necessary.

The term QoE is officially recognized by the International Telecommunication Union and is defined as follows: "The degree of delight or annoyance of the user of an application or service." This definition encompasses the type and characteristics of the application or service, the context of use, the user's expectations with respect to the application or service and their fulfilment, the user's cultural background, socio-economic issues, psychological profiles, emotional state, and other factors that will likely expand with further research [116].

The process of measuring or estimating the QoE for a set of users of an application or a service involves a dedicated procedure, considering the influencing factors which may be controlled, measured, or simply collected and reported. The output of this process may be a scalar value, a multi-dimensional representation of the results, and/or verbal descriptors.

All assessments of QoE should be accompanied by the description of the influencing factors included.

A comprehensive assessment of QoE includes many specific factors, such as a majority of the known factors. In contrast, a limited QoE assessment includes only one or a small number of factors.

4.1.1 Techniques for Measuring QoE

Quality of Experience (QoE) can be assessed through subjective and objective methodologies [117]. Subjective assessment involves obtaining direct user feedback, while objective metrics rely on automated measurements and data analytics.

Subjective Approaches

A commonly used subjective approach is the Mean Opinion Score (MOS), as defined by ITU-T [118]. Users rate their experience on a numerical scale, providing insights into perceived quality. However, subjective methods have drawbacks. They require significant resources, including time and costs associated with gathering a sufficient number of user opinions. Furthermore, these tests necessitate controlled laboratory conditions to ensure consistent and reliable results.

Table 4.1: Mapping of Mean Opinion Scores (MOS) to Opinion Labels

MOS	Opinion Label
1.0 - 1.9	Very Bad
2.0 - 2.9	Bad
3.0 - 3.9	Fair
4.0 - 4.9	Good
5.0	Excellent

Table 4.2: Example of Mean Opinion Scores (MOS) with Three Conditions

Subject	Condition A	Condition B	Condition C
Participant 1	4.5	3.8	4.2
Participant 2	4.2	4.0	3.9
Participant 3	3.9	4.3	4.1
Participant 4	4.1	3.7	4.0
Participant 5	4.4	4.5	4.3
Average	**4.2**	**4.0**	**4.1**

Objective Metrics

Objective Quality of Experience (QoE) assessment leverages automated measurements and data analytics to infer quality from usage patterns, errors, and interaction data. Unlike

subjective assessments that rely on user opinions, objective metrics provide quantifiable measures based on observable data points.

Objective parameters often rely on Quality of Service (QoS) metrics [119]. However, it is important to distinguish between the two concepts:

- **Quality of Service (QoS)** refers to the measurement of technical performance in a network, service, or system, encompassing measurable parameters such as latency, jitter, throughput (data transfer capacity), packet loss, and others. QoS aims to ensure that a service meets specific performance requirements, such as data transmission speed, maximum acceptable delay, and reliability;
- **Quality of Experience (QoE)** pertains to the end user's subjective perception of a service or application quality during usage. QoE is influenced by various factors, including QoS parameters, as well as emotional, contextual, and usability aspects that can affect user perception. Measuring QoE often requires a blended approach of objective methods (like QoS parameters) and subjective methods (such as surveys and user evaluations);

Table 4.3: Mapping of Technical Parameters to User Perception

Technical Parameter	Explanation	User Feedback without Technical Expertise
Latency	Delay in data transmission; higher latency can lead to perceived delays in interactions or actions.	Perceived as slow responsiveness or delay in application tasks.
Jitter	Variability in packet arrival times; high jitter may cause irregularities or interruptions in audio/video playback.	Noticed as stuttering or pausing in audio/video streams.
Packet Loss Rate	Percentage of lost data packets during transmission; high loss rates can result in gaps or freezes in communication.	Seen as frozen screens or pixelated images during video calls or streaming.
Throughput	Amount of data transferred per unit of time; affects the smoothness of data streams.	Experienced as buffering or pauses during multimedia streaming.
Bitrate	Rate of data transfer in a multimedia stream; higher bitrate typically correlates with better audio/video quality.	Noted as clearer audio or sharper video quality.

This approach offers several advantages:

- **Scalability**: Objective metrics can be applied efficiently across large user bases and

diverse contexts without the need for extensive human resources.

- **Continuous Monitoring**: Automated systems can continuously monitor QoE indicators in real-time, allowing for immediate identification of issues and performance trends.
- **Granular Insights**: By analyzing detailed interaction data such as session duration, latency, packet loss rates, and playback quality metrics, objective metrics can pinpoint specific aspects affecting user experience.
- **Comparative Analysis**: Objective assessments enable comparative analysis between different services, versions of applications, or network configurations, facilitating informed decision-making for service optimization.
- **Standardization**: Objective metrics can be standardized across industries, enabling benchmarking and compliance with quality standards and regulatory requirements.

However, objective metrics also have limitations. They may not fully capture subjective aspects of user experience, such as emotional responses or contextual factors that influence perception. Moreover, the accuracy and relevance of objective measures depend on the quality and comprehensiveness of the data collected and analyzed.

Objective QoE assessment complements subjective methods by providing a robust foundation for quantitative evaluation, enhancing the overall understanding and management of user experience in telecommunication and digital services.

Hybrid Approaches

Hybrid approaches combine subjective and objective methods to provide a comprehensive assessment of Quality of Experience (QoE). This methodology integrates user-centric feedback obtained through surveys, interviews, or focus groups with automated monitoring of technical metrics. The integration of subjective and objective data sources allows for a more holistic evaluation of user experience across various digital services and applications.

Key examples of hybrid approaches include:

- **Correlation of Survey Responses with System Performance Metrics:** Integrates subjective survey results, such as Mean Opinion Score (MOS), with objective metrics like latency, jitter, and packet loss rate to assess overall QoE.
- **Contextual Analysis with Real-time Monitoring:** Combines real-time monitoring of network conditions with contextual information (for example, user location, device type) to understand how environmental factors influence user experience.
- **Quality of Experience Models Incorporating User Feedback:** Develops predictive models that incorporate both user feedback and technical metrics to forecast QoE across different usage scenarios.

These hybrid approaches enhance the reliability and comprehensiveness of QoE assessments by bridging the gap between user perceptions and technical performance metrics. By leveraging both subjective insights and objective data analytics, organizations can make informed decisions to optimize service quality and enhance user satisfaction. Innovative hybrid approaches are employed to assess the quality of 3D video in immersive experiences [120].

However, hybrid approaches also present challenges. Implementing these methodologies requires integrating diverse data sources and methodologies, which can increase complexity and resource requirements. Ensuring compatibility and consistency between subjective feedback and objective metrics may pose technical challenges, affecting the reliability and comparability of QoE results. Moreover, the interpretation of hybrid QoE assessments may vary due to different weighting of subjective and objective components, potentially leading to variability in decision-making processes.

Table 4.4: Advantages and Disadvantages of QoE Assessment Approaches

Approach	Advantages	Disadvantages
Subjective	Provides direct insights into user perception and experience, capturing emotional and contextual factors.	Resource-intensive, requiring time and costs for gathering user opinions. Results influenced by individual biases and preferences. Requires controlled environments for consistent evaluations.
Objective	Offers scalable evaluation across large user bases and diverse contexts. Provides real-time monitoring capabilities for immediate issue identification. Quantifiable metrics enhance comparative analysis and standardization.	May not capture subjective aspects of user experience such as emotional responses and situational contexts. Accuracy depends on data quality and comprehensiveness.
Hybrid	Integrates subjective insights with objective metrics for a comprehensive assessment. Bridges the gap between user perception and technical performance metrics. Enhances reliability and relevance of QoE assessments.	Requires integration of diverse data sources, increasing complexity and resource demands. Challenges in ensuring compatibility and consistency between subjective and objective data. Interpretation may vary due to different weighting of components.

4.1.2 Distinguishing QoS and QoE

Quality of Service (QoS) encompasses a set of technologies and techniques aimed at managing and optimizing the performance of computer networks. The primary objective of QoS is to allocate network resources in a manner that guarantees a specific level of performance for designated network traffic.

QoS implementations can occur at various layers of the network stack, including the data link, network, and transport layers. Common QoS techniques comprise traffic shaping, packet prioritization, and congestion control.

One of the critical advantages of QoS is its ability to enable network administrators to ensure that mission-critical applications and services receive the necessary resources to function correctly. This capability can lead to enhanced productivity, minimized downtime, and an improved user experience.

Table 4.5: Key Parameters for Evaluating QoS

Parameter	Description	Unit of Measure
Latency	The time it takes for a data packet to travel from the source to the destination across a network. Lower latency indicates faster data transmission and is crucial for real-time applications such as VoIP and online gaming.	Milliseconds (ms)
Jitter	The variation in the time taken for packets to reach their destination. High jitter can cause irregularities in audio and video streams, leading to a poor user experience.	Milliseconds (ms)
Throughput	The amount of data successfully transmitted over a network in a given period, usually measured in bits per second (bps). Higher throughput signifies better network performance and capacity.	Bits per second (bps)
Packet Loss	The percentage of data packets that are lost during transmission. High packet loss can result in incomplete data transfer and degraded performance, particularly for multimedia applications.	Percentage (%)
Bandwidth	The maximum rate at which data can be transmitted over a network path. Adequate bandwidth ensures sufficient capacity to handle traffic loads without congestion.	Bits per second (bps)
Error Rate	The frequency of errors in data transmission, typically measured as a percentage. Lower error rates indicate higher quality and reliability of the network.	Percentage (%)

Quality of Experience (QoE) refers to the overall satisfaction of a user when utilizing a computer network. QoE encompasses factors such as network speed, reliability, and usability.

In the contemporary digital era, where technological advancements are relentless, it is imperative for computer networks and systems to deliver a high QoE. This requirement is particularly crucial for businesses, as a subpar QoE can result in decreased productivity and revenue losses.

Network reliability is another vital component of QoE. A reliable network ensures unin-

terrupted access for users, eliminating downtime. For businesses that depend on internet connectivity for operations, even brief periods of downtime can significantly impact productivity and revenue.

Usability also plays a pivotal role in determining QoE. A user-friendly network facilitates easy access and navigation, contributing to a more efficient and productive work environment.

To better understand the difference between Quality of Service (QoS) and Quality of Experience (QoE), consider the following example. Imagine are on a video call where you struggle to understand the person speaking, and the presentation slides are so pixelated that you cannot decipher the words. Despite these issues, the transmission of the video meets the minimum requirements outlined in the Service Level Agreement (SLA).

From the Quality of Service (QoS) perspective, metrics such as latency, jitter, throughput, and packet loss indicate that the network is performing within acceptable limits. The data packets are transmitted and received successfully, meeting the SLA standards. Thus, from a QoS standpoint, the network performance is deemed satisfactory.

However, from the Quality of Experience (QoE) perspective, the user's overall experience is poor. The difficulty in hearing the speaker clearly and the unreadable presentation slides lead to frustration and dissatisfaction. Despite the satisfactory QoS metrics, the user's experience is negative, highlighting the discrepancy between network performance and user satisfaction.

This example demonstrates the critical difference between QoS and QoE. QoS focuses on technical performance metrics and ensures that network resources are allocated efficiently, while QoE emphasizes the user's subjective experience and satisfaction. Both aspects are essential for delivering high-quality digital services, but they address different dimensions of service quality. Effective network management requires a balanced approach that considers both QoS and QoE to ensure optimal user experiences.

From a technical perspective, it is evident that Quality of Service (QoS) constitutes a necessary yet insufficient condition for achieving optimal Quality of Experience (QoE). QoS parameters, such as latency, jitter, and packet loss, form the foundational metrics that contribute to network performance. However, QoE encompasses a broader spectrum, including user-perceived factors like usability, satisfaction, and contextual influences. Hence, while QoS is a critical subset of QoE, ensuring high QoS alone does not inherently guarantee superior QoE, as the latter demands a holistic approach integrating both technical performance and user-centric factors [121].

Figure 4.1: QoS is necessary for QoE, but non sufficient

In recent studies, researchers have endeavored to establish correlations between Quality of Service (QoS) metrics and Quality of Experience (QoE) assessments [122], aiming to enhance understanding and optimize user satisfaction in various technological domains, including IPTV [123] services and 5G networks [124].

4.2 Quality of Experience (QoE) vs. User Experience (UX)

In the context of modern technological advancements, particularly with the integration of IoT, XR, and photogrammetry, it is essential to distinguish between Quality of Experience (QoE) and User Experience (UX). Although these terms are often used interchangeably, they encompass different aspects of user interaction and perception.

4.2.1 User Experience (UX)

User Experience (UX) refers to a person's overall experience when interacting with a product, system, or service. It encompasses all aspects of the end-user's interaction, including the interface, graphics, design, usability, and emotional response. The primary goal of UX is to enhance customer satisfaction and loyalty by improving the usability, ease of use, and pleasure provided in the interaction with the product.

The key elements of UX include:

Usability

Usability focuses on how easily users can achieve their goals when interacting with a product or service. It involves the efficiency, effectiveness, and satisfaction with which users accomplish tasks. A highly usable product minimizes user effort and reduces the likelihood of errors, contributing to a seamless and intuitive interaction.

Design

Design pertains to the visual and interactive aspects that affect user interaction. This includes the layout, typography, color schemes, and overall aesthetics of the interface. Good design enhances the user's ability to navigate and interact with the system, making it visually appealing and easy to understand.

Accessibility

Accessibility ensures that the product is usable by people with various disabilities. This involves designing interfaces that are navigable and functional for users with visual, auditory, motor, or cognitive impairments. Accessibility is a critical component of UX, as it broadens the user base and ensures inclusivity.

Emotional Response

The emotional response of users during their interaction with a product plays a significant role in UX. This includes the feelings of joy, satisfaction, frustration, or disappointment that users experience. Positive emotional responses are often linked to higher user satisfaction and can lead to increased loyalty and positive word-of-mouth.

Satisfaction

Satisfaction measures the user's overall contentment with the interaction. It is influenced by all other aspects of UX, including usability, design, accessibility, and emotional response. A high level of user satisfaction indicates that the product meets or exceeds user expectations and provides a valuable experience.

4.2.2 Distinguishing QoE from UX

While UX focuses on the design and usability aspects that affect the user's direct interaction, QoE extends to include technical parameters that influence the overall perception of service quality. For instance, in an XR application, UX addresses the intuitiveness of the interface and ease of navigation. In contrast, QoE encompasses system responsiveness, visual fidelity, and the smoothness of the user experience under varying network conditions.

QoE is influenced by factors such as network performance, latency, and reliability, which directly impact how the service is perceived by the user. Thus, while UX aims to optimize the interaction from a design and usability perspective, QoE aims to optimize the experience from a technical performance standpoint.

Understanding the distinction between QoE and UX is crucial for a holistic improvement of user satisfaction. Optimizing technical aspects to enhance QoE while ensuring a user-friendly and engaging interface for a better UX can significantly elevate the overall user experience in applications involving IoT and XR technologies.

4.3 Intermediate Position of QoE between QoS and UX

Quality of Experience (QoE) occupies an intermediary position between Quality of Service (QoS) and User Experience (UX) as it integrates elements from both. QoS provides the technical foundation and performance metrics, focusing on measurable aspects such as latency, jitter, throughput, and packet loss. These metrics are crucial for assessing the technical performance of a system or network.

On the other hand, QoE evaluates how these technical aspects impact the overall perception and experience of the user. It goes beyond mere technical measurements to encompass the subjective perception of quality, which includes emotional, contextual, and usability factors. Thus, QoE bridges the gap between the technical quality assessed by QoS and the comprehensive, qualitative aspects characterized by UX.

In essence, QoE integrates the technical quality measured by QoS with the more subjective and qualitative dimensions that define UX, providing a holistic assessment of user satisfaction and service quality. By doing so, QoE offers a more complete evaluation that reflects both technical performance and user perception, making it a crucial metric in understanding and improving overall user experience.

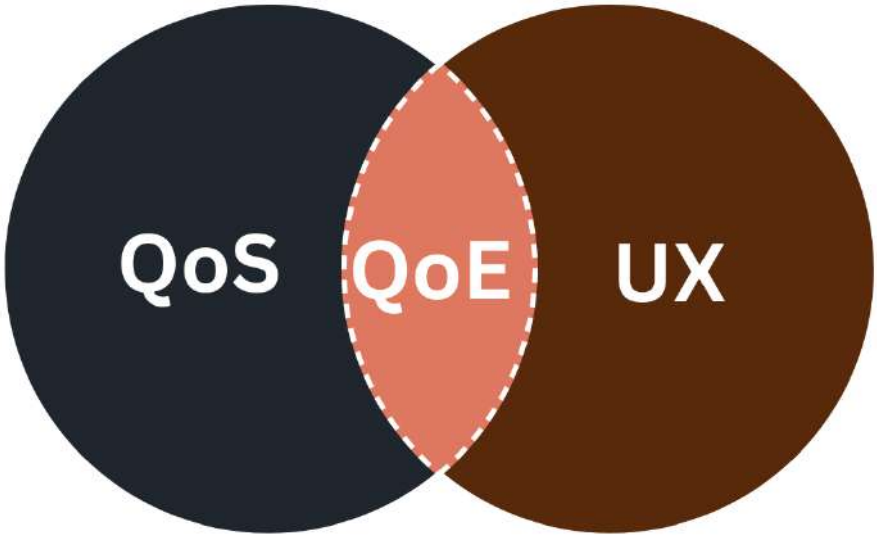


Figure 4.2: QoE as an intersection of QoS and UX: QoE integrates technical performance metrics from QoS (latency, reliability, accuracy) with subjective aspects of UX (usability, design, satisfaction), providing a comprehensive view of overall user experience.

Table 4.6: Comparison between Quality of Experience (QoE) and User Experience (UX)

Aspect	Quality of Experience (QoE)	User Experience (UX)
Definition	Overall perception of the quality of a service or application Influenced by technical performance	Overall experience when interacting with a product, system, or service Focusing on design and usability
Key Elements	Latency Jitter Bandwidth Reliability	Usability Design Accessibility Emotional Response Satisfaction
Evaluation	Measured through objective metrics Subjective user feedback on service quality	Assessed through user testing, surveys Feedback on interaction and satisfaction
Application	Enhancing technical aspects to improve perceived quality Network Services Video Streaming Gaming	Improving interface design, ease of use Overall user engagement in software and hardware products
Objective	Optimize service delivery Ensure high-quality performance under various conditions	Enhance the overall user interaction Enhance overall user satisfaction

Chapter 5

Doctoral Research Progress and Applied Case Study Analysis

5.1 Introduction to the Doctoral Research Path

The doctoral research journey began on January 3, 2022, at the University of Catania, within the PhD program in Systems Engineering, Energy, Computer Science, and Telecommunications. The research is funded by TIM S.p.A., under the topic "Cloud-native and edge-powered verticals in a 5G ecosystem." This broad theme allowed for a multidisciplinary approach that explores applications of IoT, XR, and photogrammetry from both academic and industrial perspectives.

Initially, the focus was on Internet of Things (IoT) applications in underwater environments, integrating the data produced by the underwater system developed by the university's telecommunications research group into a service provided by TIM for smart and sustainable cities. As the research evolved, the idea emerged to leverage the multimedia content generated by the underwater system, particularly video and images, to create models that could be accessed and explored at any time.

The potential of photogrammetry, introduced through a course on its applications in cultural heritage and archaelogy, became apparent as a method for constructing 3D models. A significant amount of time was then dedicated to mastering photogrammetry techniques, which, along with XR technologies, aimed to enhance users' Quality of Experience (QoE) by improving the realism and accessibility of virtual environments.

A substantial portion of the research was conducted at the TIM Innovation Lab, where new methods were developed to enhance Proof of Concept (PoC) presentations through IoT and XR. This effort culminated in the development of the Digital Twin for the Smart Water Loop, a TIM prototype designed for efficient water monitoring and fraud detection.

The final phase of the research took place at the University of Hertfordshire's VR & Robotics

Laboratory, focusing on comparing 3D reconstructions made from equirectangular images with innovative and efficient photogrammetry methods. Additionally, the research involved depth perception studies within these environments to assess the potential of equirectangular images for precise photogrammetric modeling.

5.2 Research Objectives and Methodology

The primary objective of this doctoral research is to investigate efficient 3D reconstruction techniques that utilize 360-degree images to optimize scene exploration and immersive observation. Moving beyond traditional DSLR-based photogrammetry, the research explores how equirectangular images can enhance the realism and usability of 3D models. The research applies novel Standard Operating Procedures (SOPs) for indoor photogrammetry—originally developed using DSLR cameras at the University of Dundee—with equirectangular image technology in the VR & Robotics Laboratory at the University of Hertfordshire.

5.2.1 Motivation for Technology Choices

Photogrammetry was chosen as the primary technique for 3D reconstruction because of its versatility and the accessibility of open-source software and equipment. Photogrammetry supports a wide range of devices, from entry-level to high-end, allowing for flexibility based on user needs. The existing devices in the research labs were sufficient for the experiments, eliminating the need for additional hardware purchases.

XR technologies were also incorporated, driven by recent growth in consumer-grade Head-Mounted Displays (HMDs) and the increasing emphasis on Quality of Experience (QoE) in fields outside telecommunications. Evaluating QoE in this context is a novel approach, with potential to improve user interaction within a variety of applications.

5.2.2 Methodological Approach

The research employed a structured methodology, beginning with a comprehensive literature review to evaluate the state of existing technologies and identify research gaps. This iterative process enabled refinement of the conceptual framework and adaptation of effective methods.

A mixed-methods approach combined both quantitative and qualitative data collection. Quantitative data, such as system performance metrics, provided insights into technical feasibility, while qualitative data from observational studies offered an in-depth understanding of user interactions. This methodology ensured the research remained both technically robust and focused on enhancing user engagement.

5.3 Research Timeline and Milestones

The doctoral journey involved several stages that combined professional development with introspective learning. Key milestones, documented in Table 5.1, reflect essential activities undertaken to address knowledge gaps and meet research goals.

The initial phase required a focus on updating skills and knowledge, particularly following time spent in private and public sectors. A transformative point occurred in the second year with the introduction of photogrammetry for 3D modeling, specifically in conjunction with XR technologies. This led to an exploration of applications across cultural heritage, forensics, and health.

Another pivotal experience was a research visit to the VR & Robotics Lab at the University of Hertfordshire, where innovative problem-solving methods were adopted. Exposure to different approaches provided invaluable perspectives, contributing to both personal growth and technical expertise.

These activities have collectively advanced the aim of this research: to identify methods that enhance QoE by integrating accessible and efficient 3D modeling techniques in XR environments.

Table 5.1: Summary of Activities

Activity	Provider	Instructor	Workload	Month	Year
Seminar about Game Theory	UniversiTIM	PhD Giorgia Cappello	2 h	January	2022
HR Welcome 37° cycle PhD	TIM Academy	N/A	1 h	July	2022
Job Shadowing	UniversiTIM	Ing. Giorgio Castelli	10 h	November-December	2022
European Researcher Night	SHARPER Project	N/A	6 h	September	2022
Seminar in Cognitive Smart & Green Cities	Joint Open Lab WAVE Catania	N/A	3 h	December	2022
Course in TinyML Learning Laboratory	University of Catania	PhD Joannes Sam Mertens	3 h	April	2022
European Green Deal: the role of AI	Joint Open Lab WAVE	N/A	1 h	May	2023
Digital Applications for Cultural Heritage	University of Catania	PhD Stephan Hassam	24 h	April-July	2023
Course in 5G Architectures and Services	TIM Academy	N/A	5 h	July	2023
Photogrammetry and Mixed Reality Project Meeting	Joint Open Lab WAVE Catania	N/A	2 h	July	2023
Course in Digital Game Development	University of Catania	Prof. Filippo Stanco	48 h	March-June	2024
Counter Terror Expo and Forensics Europe Expo 2024	ExCeL London (London, UK)	N/A	8 h	June	2024
International Forensics Summer School IFOSS 2024	University of Catania	Prof. Sebastiano Battiato	26 h	July	2024
Visiting Research Fellow	University of Hertfordshire	Prof. Salvatore Livatino	172 days	June-November	2024
Conference Attende and Staff Member	IEEE MetroXRAINE 2024	Prof. Salvatore Livatino	24 h	October	2024

In conclusion, the doctoral journey has advanced both professional expertise and personal insights, paving the way for future contributions in the technological landscape. These milestones are integral to achieving the overarching goal of this research: to enhance the Quality of Experience (QoE) through innovative, efficient, and accessible 3D reconstruction techniques.

5.4 Overview of Case Studies

The following section provides an overview of the three case studies, each targeting different aspects of the applications of photogrammetry and XR technologies for enhanced immersive experiences:

- **Case Study 1: Advancing Proof of Concept Demonstrations through IoT and XR Integration**
  This study involved creating an effective approach for presenting PoC demonstrations within a corporate environment, utilizing foundational photogrammetry to model and showcase digital scenes. The process helped establish photogrammetry's practical use and laid essential groundwork for more advanced applications.
- **Case Study 2: Efficient 3D Reconstruction of Indoor Environment Using Equirectangular Images in Photogrammetry**
  This case study presented a significant contribution to indoor photogrammetry by using a Standard Operating Procedure (SOP) for 360-degree images. Unlike traditional applications, this approach enables efficient acquisition and processing for quick yet detailed 3D reconstructions, with practical uses in virtual tours, architecture, and forensic investigation.
- **Case Study 3: Comparative Analysis of Distance Estimation in Spherical Images and 3D Photogrammetry Models**
  This study assessed the applicability of equirectangular images for photogrammetric reconstructions, providing a comparative analysis of distance estimation in both spherical images and photogrammetric models. The findings aim to validate the broader applicability of equirectangular imaging for realistic virtual environment creation.

5.5 Case Study 1: Advancing Proof of Concept Demonstrations through IoT and XR Integration

5.5.1 Background

Proof of Concept (PoC) demonstrations serve as pivotal tools for showcasing the feasibility and potential of emerging technologies or innovative solutions. By providing a preliminary, tangible example of how a technology or concept works, PoCs help stakeholders, including potential investors, partners, and clients, to grasp the value and functionality of the proposed solution. They are particularly crucial in validating ideas before committing to full-scale development or deployment.

The Smart Water Loop is a Proof of Concept (PoC) developed by the TIM Innovation Lab in Catania. This prototype originates from an industrial research project aiming to leverage NB-IoT technology for monitoring water usage and possible fraud in the context of smart and sustainable cities. The system employs a hydraulic circuit that, through a network of pipes, pumps, and three-way valves, simulates either a direct (correct) flow or a reverse (fraudulent) flow. The flow direction is detected by the Quadraplus telemeter and displayed on a screen. Data is collected by the NB-IoT board and then transmitted to the open-source IoT platform Thingsboard CE via TIM's proprietary store and forward platform, TIM Icon. To enhance the PoC experience through extended reality, a communication layer has been added to an IoT-XR platform developed by TIM. Specifically, TIM Icon sends MQTT messages to a NodeJS backend, which provides APIs used by the frontend to activate animations and send direct messages to the NB-IoT board, and thus to the actual PoC (see Figure 5.1).

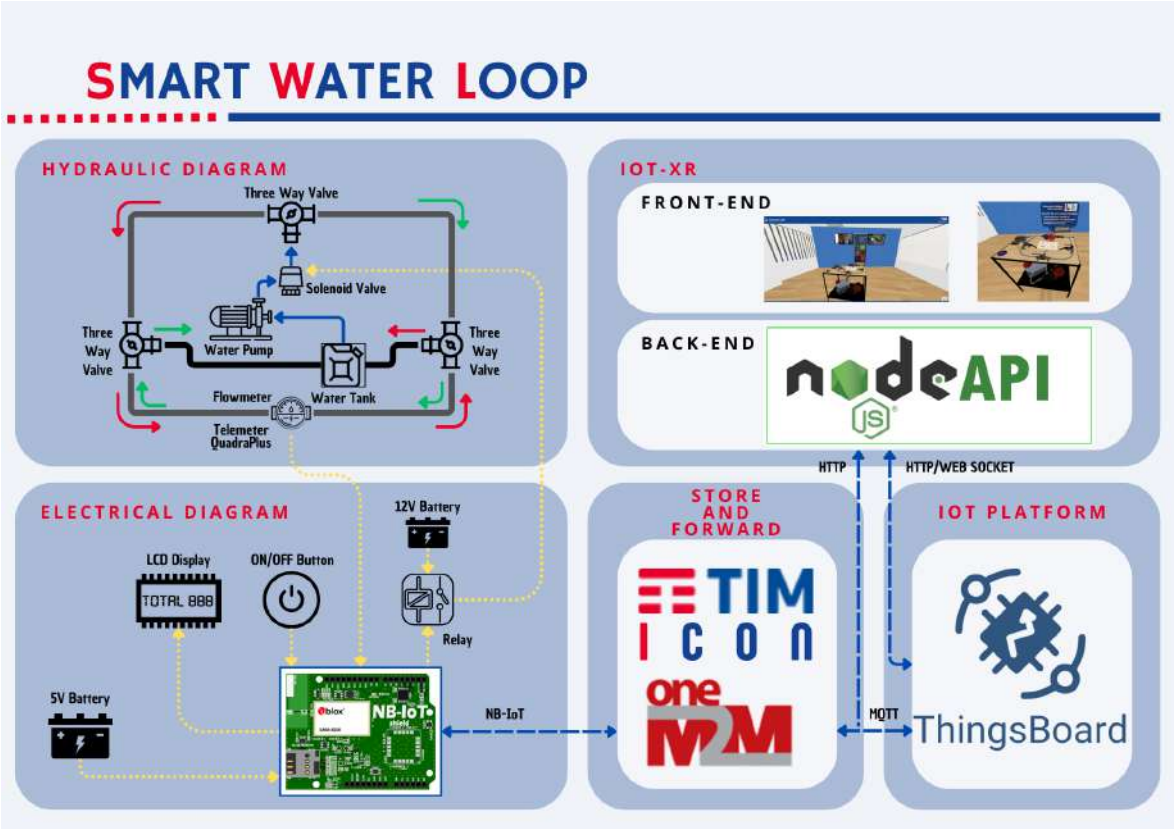


Figure 5.1: Overview of the Smart Water Loop system. Green arrows indicate a direct flow, while red arrows indicate a fraudulent flow. The hydraulic system is designed such that the water is recirculated back into the tank, creating a continuous loop. The direct/reverse flow is controlled by adjusting the three-way valve handles.

To provide a deeper understanding of the core technologies utilized in the Smart Water Loop project, the following sections will focus on the main components that enable its functionality:

NB-IoT Shield by SODAQ

The NB-IoT Shield by SODAQ is Arduino-based and programmed using C code. The programming IDE used is the Arduino IDE. This shield enables low-power, wide-area connectivity for IoT devices, allowing efficient communication over long distances with minimal energy consumption. It supports various features tailored for IoT applications, such as secure data transmission, over-the-air firmware updates, and compatibility with numerous sensors and actuators.

TIM ICON

TIM ICON is a store and forward platform developed by TIM, utilizing the IoT standard oneM2M. This platform facilitates seamless communication between IoT devices and applications by ensuring reliable data transmission and interoperability. It provides robust data management capabilities, including data storage, processing, and forwarding to various IoT services. TIM ICON supports scalable and flexible integration with a wide range of IoT solutions, enhancing the efficiency and effectiveness of IoT deployments.

Thingsboard CE

Thingsboard CE is an open-source IoT platform used for data collection, processing, visualization, and device management. It provides a comprehensive suite of tools for aggregating and analyzing data from various IoT devices in real-time. With its user-friendly dashboards and customizable widgets, Thingsboard CE allows for intuitive data visualization, enabling users to monitor and control their IoT devices effectively. The platform supports various protocols, including MQTT, CoAP, and HTTP, ensuring seamless integration with diverse IoT systems. Additionally, Thingsboard CE offers robust device management features, including remote configuration, firmware updates, and fault detection, making it a versatile solution for managing large-scale IoT deployments.

TIM IoT-XR

The TIM IoT-XR web platform is used to create a navigable and interactive XR environment using NodeJS technology. This platform enables users to explore and interact with IoT data within a virtual reality setting, enhancing the presentation and understanding of IoT systems. By leveraging NodeJS, a highly efficient and scalable JavaScript runtime, TIM IoT-XR ensures real-time communication between IoT devices and the XR environment. NodeJS's event-driven architecture facilitates the handling of multiple connections simultaneously, making it ideal for real-time applications like IoT-XR. The platform supports various XR devices and provides a rich set of tools for creating immersive and engaging user experiences, making it an innovative solution for visualizing and managing IoT data in a virtual space.

5.5.2 State of the Art

With the proliferation of broadband Internet and high-performance devices, a wide range of services has become accessible at any time and from any location. Consequently, the focus has shifted towards service quality, with Quality of Experience (QoE) emerging as a key evaluation metric from the user's perspective. As QoE is inherently subjective and closely tied to user perception and expectations, conducting quantitative and comparative research has proven challenging, given that QoE studies are still in their early stages. However, substantial efforts have been made in this field, leading to a clearer understanding of QoE management architecture [121].

Despite the definitions provided by ITU, QoE still appears to be closely associated with technical aspects. Literature suggests incorporating additional user-centric factors into the analysis, such as physiological conditions, socio-cultural context, demographic data, and psychological state, to gain a more comprehensive understanding of QoE [125].

Quality of Experience (QoE) is explored in both IoT and VR systems, each focusing on different aspects of user interaction and performance. In IoT systems, QoE assessment primarily involves analyzing parameters related to technical performance, such as latency,

reliability, and throughput. These technical metrics provide insights into how well the system meets performance criteria and user expectations from a technical standpoint [126].

In contrast, VR systems consider a range of human factors when evaluating QoE. This includes physiological aspects, such as the user's comfort and physical reactions; psychological aspects, such as cognitive load and emotional responses; and demographic factors, including age and user experience level. These elements contribute to a holistic understanding of how users interact with VR environments and how these interactions affect their overall experience [127].

This approach highlights QoE as an intersection between Quality of Service (QoS) and User Experience (UX), as detailed in Section 4.3. The integration of technical performance metrics from QoS with human-centric factors from UX emphasizes QoE's unique position in bridging the gap between technical performance and subjective user experience.

The combined use of Extended Reality (XR) and Internet of Things (IoT) has been explored in a limited number of pioneering studies. These works investigate integrated IoT-XR systems that enhance the interconnection between real and virtual objects [128]. Notable research in this area includes the utilization of cutting-edge technologies such as Digital Twins [129], which facilitate real-time synchronization and interaction between physical entities and their digital counterparts. These integrated systems aim to bridge the gap between tangible and virtual environments, offering innovative solutions for applications across various domains [130]. The current research highlights the potential of combining XR and IoT to create more immersive and interactive experiences, although this field remains in the early stages of development.

Despite the critical role of Proof of Concept (PoC) in the technology transfer process [131], there is a noticeable gap in the literature concerning innovative or relevant proposals for evaluating Quality of Experience (QoE) in the context of PoCs. Current research often focuses on the technical performance and feasibility aspects of PoCs, but few studies address the integration of IoT and Extended Reality (XR) technologies in assessing QoE [132]. The combination of these technologies presents a unique opportunity to enhance PoC evaluation by providing a more comprehensive understanding of user experience. However, the intersection of IoT and XR for QoE assessment remains underexplored, highlighting a significant area for future research and development.

5.5.3 Motivation and Objectives

When PoC demonstrations are presented to users who lack a technical background, several limitations can arise. These non-technical users may struggle to fully comprehend the technical intricacies and underlying mechanisms of the PoC. Their feedback might be based on superficial impressions rather than a deep understanding of the technology's potential or limitations. This can result in an inaccurate assessment of the PoC's effectiveness and the

value it offers.

For example, consider the PoC of the Smart Water Loop, developed to showcase a new IoT-enabled flowmeter for industrial applications. This PoC involves a physical prototype of the flowmeter that measures fluid flow rates in real-time and integrates with other industrial IoT systems. Users must interact directly with the physical flowmeter to evaluate its performance, such as observing how it displays flow measurements, reacts to changes in flow rate, or integrates with data visualization tools. Without the ability to access and test the device remotely, users may face challenges in fully assessing its capabilities and benefits.

Non-technical users might have difficulty evaluating the PoC's usability, practicality, or integration within existing systems. Their feedback may focus on the overall user experience rather than technical performance metrics such as latency, data accuracy, or system reliability. This can lead to a skewed understanding of the PoC's effectiveness, potentially overlooking critical technical issues that could affect the technology's real-world implementation.

During the studies conducted at the TIM Innovation Lab in Catania, the idea of presenting Proof of Concept (PoC) in an innovative manner emerged, utilizing photogrammetry and Extended Reality (XR) to simplify the usability of the prototype for both technical and non-technical users. Additionally, a web-based IoT-XR platform has been developed by TIM, which can serve as a digital enabler for presenting PoCs in a virtual world. From this initiative, several innovative industrial research objectives were formulated:

1. **Creation of a 3D Model of the Make Room of the TIM Innovation Lab in Catania:** Develop a detailed 3D model of the laboratory to serve as a virtual environment for demonstrations and interactions.
2. **Recreation of the 3D Model of the Smart Water Loop using Photogrammetry:** Utilize photogrammetry techniques to accurately recreate the Smart Water Loop system in a 3D virtual model.
3. **Addition of Animations to Provide Additional Information to Users:** Enhance the virtual model with animations that offer supplementary information to users, such as directional arrows indicating water flow, system manipulations, and informational pop-ups.
4. **Utilization of Digital Twin Data to Trigger the Behavior of the 3D Model:** Integrate real-time data from the digital twin to influence and control the behavior of the 3D model, ensuring it accurately reflects the physical system's state.
5. **Interaction between Virtual and Real Systems:** Enable interactions between the virtual and real prototypes, allowing the virtual prototype's behavior to be altered by the real prototype and vice versa. This includes the use of solenoid valves for operations such as turning the system on and off, and reversing the flow direction.
6. **Survey for Evaluating Quality of Experience (QoE):** Conduct a survey to assess the

Quality of Experience of users interacting with the PoC. This survey aims to evaluate the effectiveness of the new approach by gathering feedback on the usability, engagement, and overall experience of users with the PoC in its virtual and physical forms.

5.5.4 Tools and Software Utilized

The successful execution of this case study relied on a range of specialized equipment to capture, process, and display the 3D models and integrate them with real-time data. The following key pieces of equipment were utilized:

Hardware

Table 5.2: Samsung Galaxy NX Mirrorless Digital Camera Specifications

Digital Camera Specifications	
Feature	**Specification**
Product Name	Samsung Galaxy NX Mirrorless Digital Camera
Image Sensor	20.3 megapixel APS-C CMOS
Image Processor	DRIMe IV
Video Resolution	Full HD 1080p at 30 fps
ISO Sensitivity	ISO 100-25600
Lens System	Samsung NX interchangeable lens system
Operating System	Android 4.2 Jelly Bean
Connectivity	Wi-Fi, Bluetooth 4.0, 3G/4G LTE
Flash	Pop-up flash
GPS	Integrated GPS
Viewfinder	0.46” electronic viewfinder
Screen	4.8” TFT-LCD touchscreen
Autofocus System	Advanced Hybrid Autofocus (phase detection + contrast AF)
Included Lens	Samsung NX-mount 18-55mm f/3.5-5.6 OIS
Lens Equivalent Focal Length	27.7-84.7mm (35mm equivalent)
Lens Stabilization	Optical Image Stabilization (OIS)
Maximum Continuous Shooting Rate	Up to 8.6 fps
Maximum Shutter Speed	1/6000th of a second
Special Features	Smart Mode, Photo Suggest, Story Album, Camera Studio

Table 5.3: ASUS TUF Dash F15 Specifications

Laptop Specifications	
Feature	**Specification**
Model	ASUS TUF Dash F15
Processor	12th Gen Intel(R) Core(TM) i5-12450H 2.00 GHz
Memory	32GB RAM SODIMM 4800 MT/s
Operating System	Windows 11 Home 64bit
Graphics	NVIDIA GeForce RTX 3050 Laptop GPU 4GB
Storage	HD SSD 512GB

1. **Meshroom:** An open-source photogrammetry software developed by AliceVision. Meshroom was used to create 3D models from a series of photographs taken with the DSLR camera. It processes images to generate high-fidelity 3D models through a series of steps including structure from motion, depth map computation, and texturing.
2. **Blender:** A powerful open-source 3D creation suite. Blender was employed for tasks such as remeshing, repairing, and retopology of the 3D models generated by Meshroom. It was also used for creating animations and integrating additional elements to enhance the virtual representation of the Smart Water Loop.
3. **Innoverse Lab:** An XR environment developed by TIM that provides users with immersive experiences. The concept behind utilizing this environment is to serve as a digital enabler for the innovative presentation of Proof of Concept (PoC) demonstrations. By leveraging Innoverse Lab, this research aims to enhance user engagement and understanding, thereby facilitating a more effective evaluation of PoCs as outlined in the objectives of this industrial research work.
4. **Google Forms:** An online survey tool used to design and distribute the Quality of Experience (QoE) survey to users. Google Forms allowed for the collection and analysis of user feedback regarding the usability and engagement of both the physical and virtual PoC presentations.

5.5.5 Methodology

The methodology for this study involves several key steps aimed at developing and evaluating an innovative approach to presenting Proof of Concept (PoC) through Extended Reality (XR) and photogrammetry. This approach leverages advanced technologies to enhance the user experience for both technical and non-technical users. The methodology encompasses the following phases:

Creation of the 3D Model of the Make Room

The creation of the 3D model of the Make Room was accomplished using Blender, selected for its capability to handle the intricate details within the room, including furniture and equipment, which were deemed unnecessary for the final outcome. The goal was to create a spacious and immersive environment, allowing for virtual modifications of the Proof of Concept (PoC) as needed, thereby making the space multipurpose.

Table 5.4: Blender System Requirements [133]

Requirement	Specifications
Minimum Requirements	
Operating System	Windows 8.1 macOS 11.2 (Big Sur) Linux
Processor	4 cores with SSE4.2 support Apple Silicon
Memory	8 GB RAM
Graphics	2 GB VRAM with OpenGL 4.3
Recommended Requirements	
Operating System	Windows 10 or 11 macOS 14 (Sonoma) Linux
Processor	8 cores Apple Silicon
Memory	32 GB RAM
Graphics	8 GB VRAM

The room was recreated with a focus on maintaining the actual dimensions. Initially, the model was created in the .obj format, resulting in a file size of 76.9 MB, which proved to be too large for web application use. Consequently, the model was converted to the .glb format, which is widely adopted in XR web applications due to its lightweight nature [134]. The conversion reduced the file size to 35.8 MB (Fig. 5.2). Despite the reduction in detail, the model remains sufficiently detailed for the purposes of this study (Fig. 5.3).

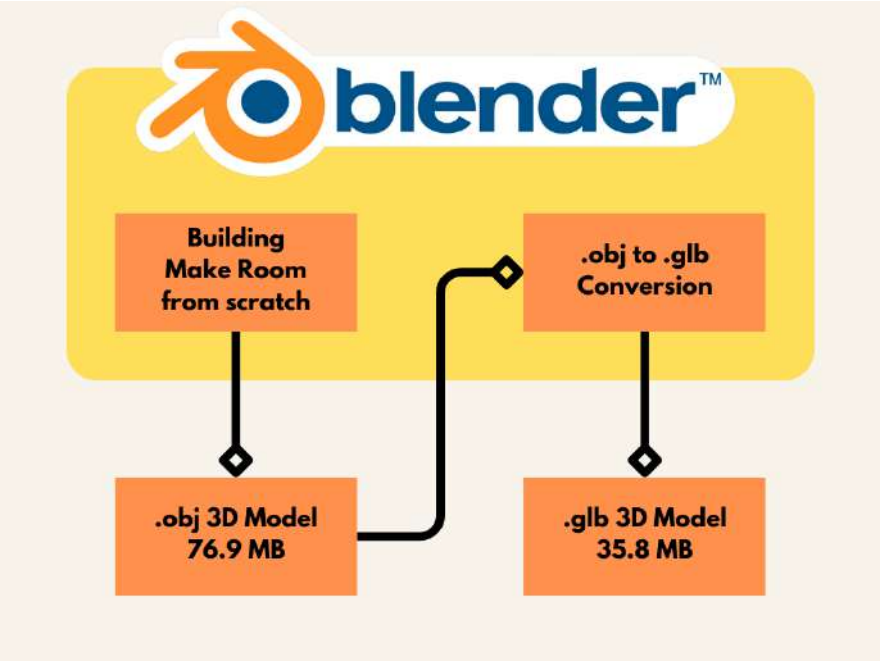


Figure 5.2: Workflow for creating the 3D model of the Make Room using Blender, optimized for web applications by converting from .obj to .glb format.

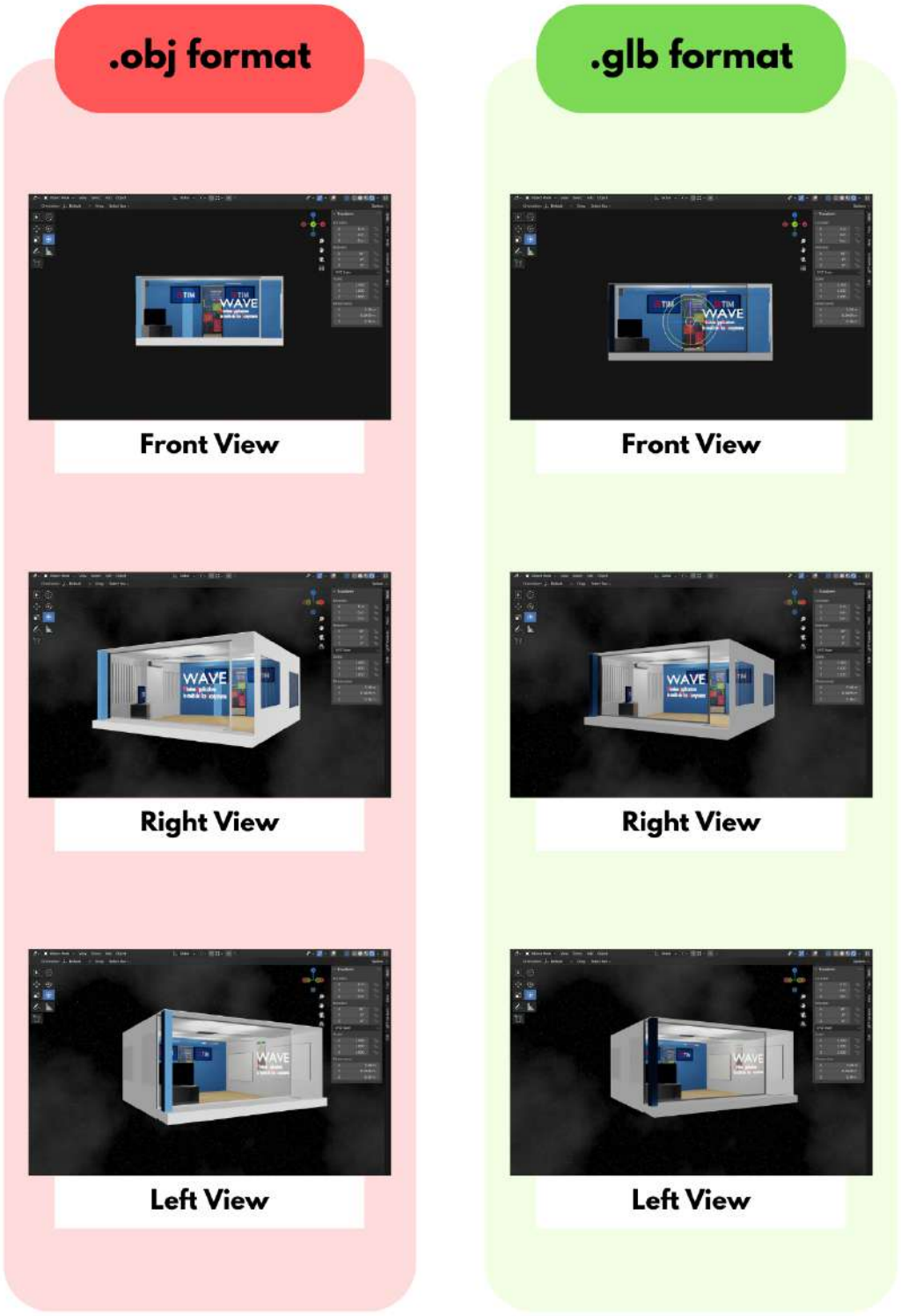


Figure 5.3: Comparison between the .obj and .glb models of the Make Room, highlighting differences in file size and detail.

Recreation of the 3D Model of the Smart Water Loop using Photogrammetry

To reconstruct the Smart Water Loop in 3D, a total of 374 images were captured using the settings specified in Table 5.5. The photogrammetry process faced challenges due to the low quality of the camera sensor and the presence of significant areas of black, white, and reflective surfaces, which disrupted the reconstruction process. As a result, the initial .obj file contained missing parts and various bulges, although the main structure of the Smart Water Loop remained discernible (Fig. 5.5).

Table 5.5: Digital Camera Shooting Settings

Setting	**Value**
F-stop	f/8
Exposure	1.2 seconds
ISO	100
Focal Length	18 mm
Maximum Aperture	3.62

To address these issues, a mesh cleanup phase was conducted using Blender, which involved procedures such as repairing, remeshing, and resculpting the model (Fig. 5.6). This process yielded a high-quality result with a significant degree of realism. Initially, the .obj file had a size of 364 MB, but through the cleanup process and texture reduction, its size was reduced to 76.6 MB (including animations that will be described later), making it optimal for use within a web application (Fig. 5.7).

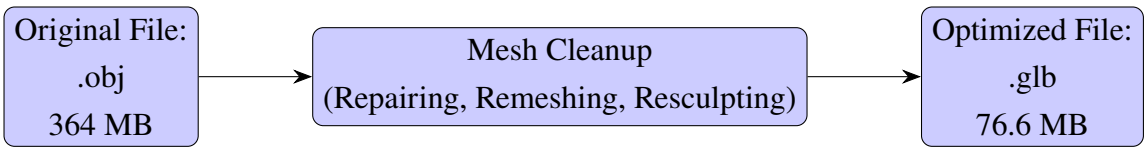


Figure 5.4: Reduction in file size from .obj to .glb through mesh cleanup processes.

Figure 5.5: 3D model of the Smart Water Loop generated through photogrammetry, showcasing the initial reconstruction with identifiable structural elements despite some imperfections.

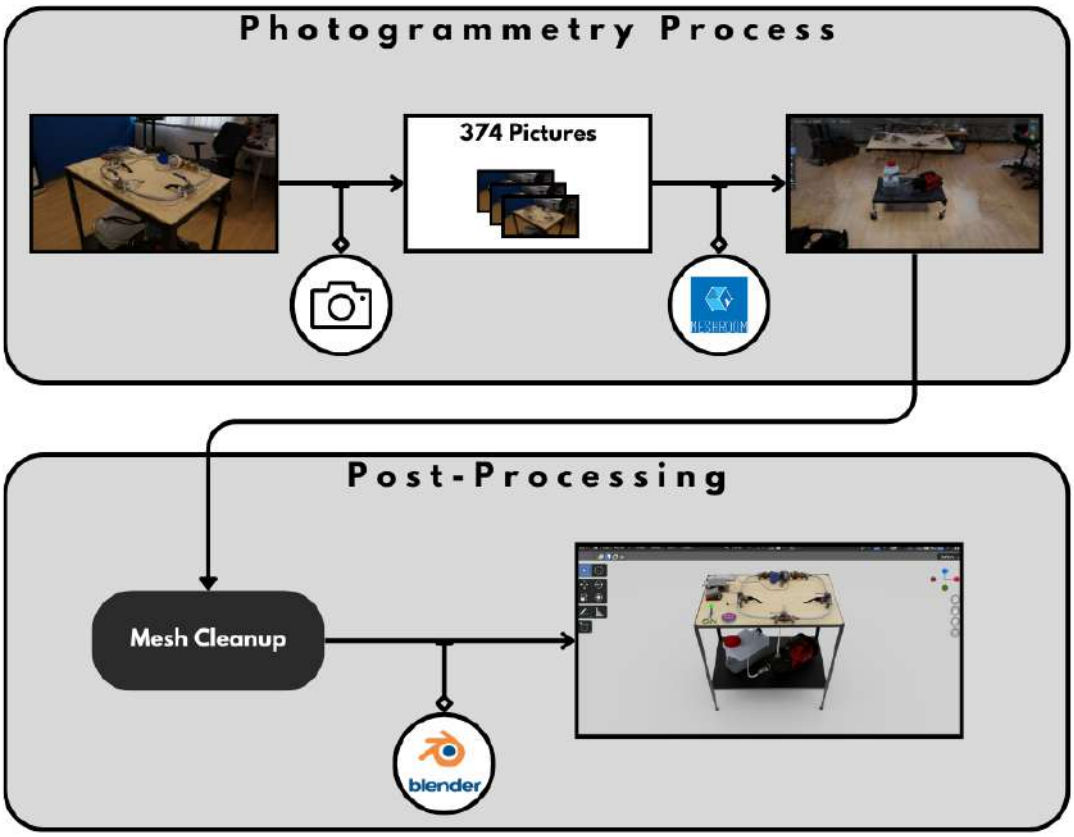


Figure 5.6: Workflow of creating the 3D model of the Smart Water Loop, illustrating the transition from the initial photogrammetry results to the optimized .glb file format, highlighting key steps such as mesh cleanup and texture reduction.

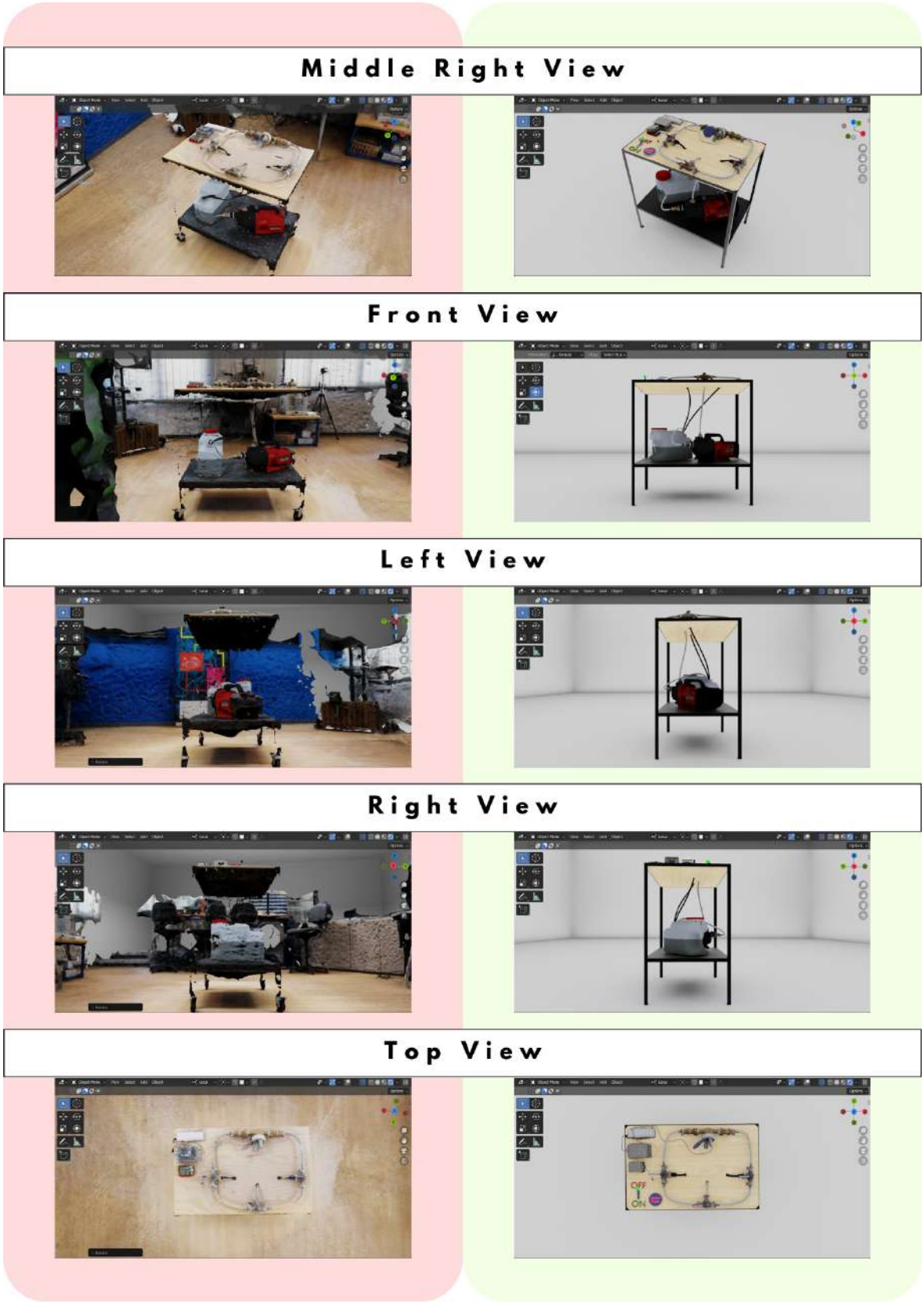
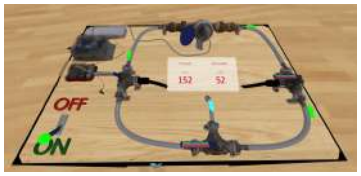


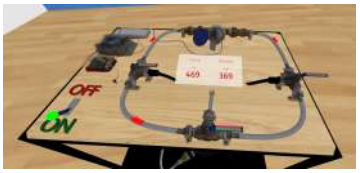
Figure 5.7: Comparison of the Smart Water Loop model from various perspectives: middle right view, front view, left view, right view, and top view. This comparison highlights the reconstructed component details, the elimination of bulges, and the preservation of a high level of realism.

Addition of Animations to Provide Additional Information to Users

To enhance user engagement and understanding, animations are incorporated into the 3D model using advanced 3D modeling software such as Blender. These animations include directional arrows indicating water flow, system manipulations, and informational pop-ups, providing users with supplementary information and interactive elements. An image illustrating these animations is provided for reference (see Figure 5.8).

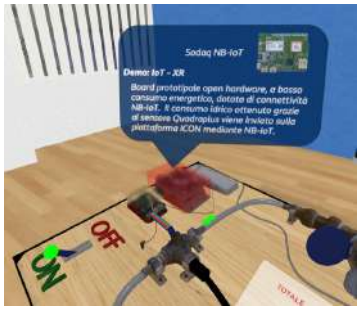


(a) Correct water flow



(b) Wrong water flow

(c) Fraud warning: telemeter removed



(d) NB-IoT information pop-up

(e) Telemeter information pop-up

(f) Switch ON/OFF

Figure 5.8: Overview of the Smart Water Loop animations: (a) Correct flow direction; (b) Reverse flow indicating potential fraud; (c) Fraud warning with the telemeter removed; (d) Informational pop-up for the NB-IoT board; (e) Informational pop-up for the QuadraPlus telemeter; (f) Progression of switching ON/OFF and vice versa.

The animations were created using Blender and are triggered based on user actions within the real Proof of Concept (PoC) environment. This dynamic interaction ensures that users receive real-time feedback and information relevant to their specific actions, thereby enriching their overall experience and comprehension of the smart water loop system.

Exploratory Survey for Quality of Experience (QoE)

To evaluate the Quality of Experience (QoE), a survey (5.6) was developed that considers various influence factors related to the human, system, and context dimensions (5.9) . This approach ensures a comprehensive assessment of user perceptions and experiences, enabling to capture critical insights into how these factors impact overall satisfaction and engagement with the Proof of Concept (PoC). By focusing on elements such as user characteristics, system performance, and contextual variables, the survey provides a holistic view of the

QoE, allowing for targeted improvements in both the physical and virtual interactions with the PoC.

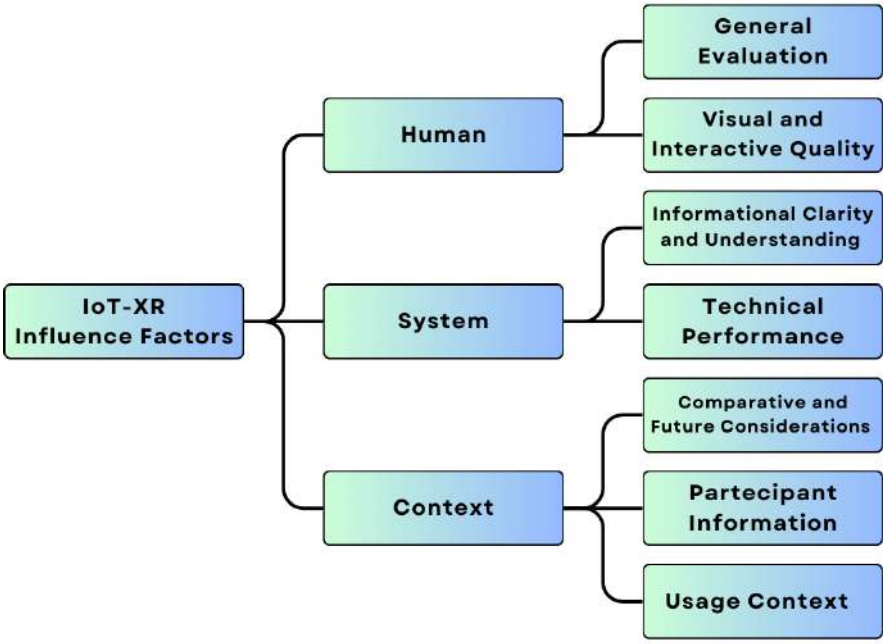


Figure 5.9: Influence factors affecting the Quality of Experience (QoE) as perceived by users [127]. These factors play a significant role in shaping users' immersion and satisfaction levels during their interactions with the PoC.

Table 5.6: Structure of the IoT-XR Proof of Concept Evaluation Survey

Participant Information
Please provide some information about yourself before answering the survey questions.
Age Group
Under 18
18-25
26-35
36-45
46-55
56-65
Over 65
Professional Background
Technical (for example, Engineer, Developer, IT Specialist)
Non-technical (for example, Marketing, Sales, Management)
Student
Other: _____
Experience with IoT Technologies
None

Basic familiarity

Intermediate

Advanced

Experience with XR Technologies

None

Basic familiarity

Intermediate

Advanced

General Evaluation

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Overall Experience

How would you rate your overall experience with the PoC?

1 2 3 4 5

Ease of Use

How easy was it to interact with the PoC?

1 2 3 4 5

Responsiveness

How responsive was the system when interacting with the virtual model and the physical IoT device?

1 2 3 4 5

Visual and Interactive Quality

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Visual Quality

How would you rate the visual quality of the XR interface?

1 2 3 4 5

Immersion

How immersive was the XR experience?

1 2 3 4 5

Functional Clarity

How clear were the specific functionalities like the visualization of flow arrows and pop-up information?

1 2 3 4 5

Informational Clarity and Understanding

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Clarity of Information

How clear and comprehensible was the information provided by the XR interface?

1 2 3 4 5

Enhancement of Understanding

How much did the XR component enhance your understanding of the IoT system?

1 2 3 4 5

Technical Performance

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Latency

How would you rate the technical performance of the PoC in terms of latency?

1 2 3 4 5

Reliability

How would you rate the technical performance of the PoC in terms of reliability?

1 2 3 4 5

Accuracy

How would you rate the technical performance of the PoC in terms of accuracy?

1 2 3 4 5

Comparative and Future Considerations

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Comparison with Traditional Methods

Compared to traditional methods, how effective was the XR experience in demonstrating the IoT system?

1 2 3 4 5

Perceived Value

How valuable do you find the integration of XR in visualizing IoT data and interactions?

1 2 3 4 5

Engagement and Satisfaction

How engaged did you feel while interacting with the PoC?

1 2 3 4 5

Future Adoption

How likely are you to adopt similar IoT and XR technologies in your work or daily life?

1 2 3 4 5

Usage Context

Please rate the following aspects on a scale from 1 to 5, where 1 indicates ”Very Bad” and 5 indicates ”Excellent.”

Location

Where did you use the PoC?

Home

Office

Laboratory

Other: _____

Duration of Use

For how long did you interact with the PoC?

Less than 30 minutes

30 minutes to 1 hour

1 to 2 hours

More than 2 hours

Open-Ended Feedback

Please provide any additional comments or suggestions regarding your experience with the IoT and XR Proof of Concept.

Suggestions for Improvement

Do you have any suggestions for improving the PoC?

Additional Comments

Any other comments or feedback you would like to provide?

5.5.6 Results Discussion

Due to confidentiality constraints imposed by the proprietary nature of the Proof of Concept (PoC), the evaluation survey was restricted to an internal audience within the laboratory environment. Consequently, the survey was administered to a sample of 7 individuals who either utilized or observed the use of the demo. This limited sample size necessitates careful interpretation of the findings, as the insights garnered may reflect the perspectives and experiences of a small, specific group rather than a broader population. The following discussion will delve into the results obtained from this sample, highlighting key observations, patterns, and implications for the development and future application of XR and photogrammetry technologies in the context of the PoC.

Participant Information

From the participant information collected, it is evident that the sample predominantly comprises individuals with a technical background, possessing substantial experience with

both IoT and XR technologies. The variability in the participants' ages further adds diversity to the sample, potentially offering a broader range of perspectives.

This technical expertise and familiarity with IoT and XR could lead to a more critical evaluation of the Proof of Concept (PoC), particularly in the assessment of system performance and interactive quality. The participants' technical proficiency may also result in higher expectations regarding the functionality and realism of the XR interface, which could influence their overall satisfaction and the perceived value of the integration.

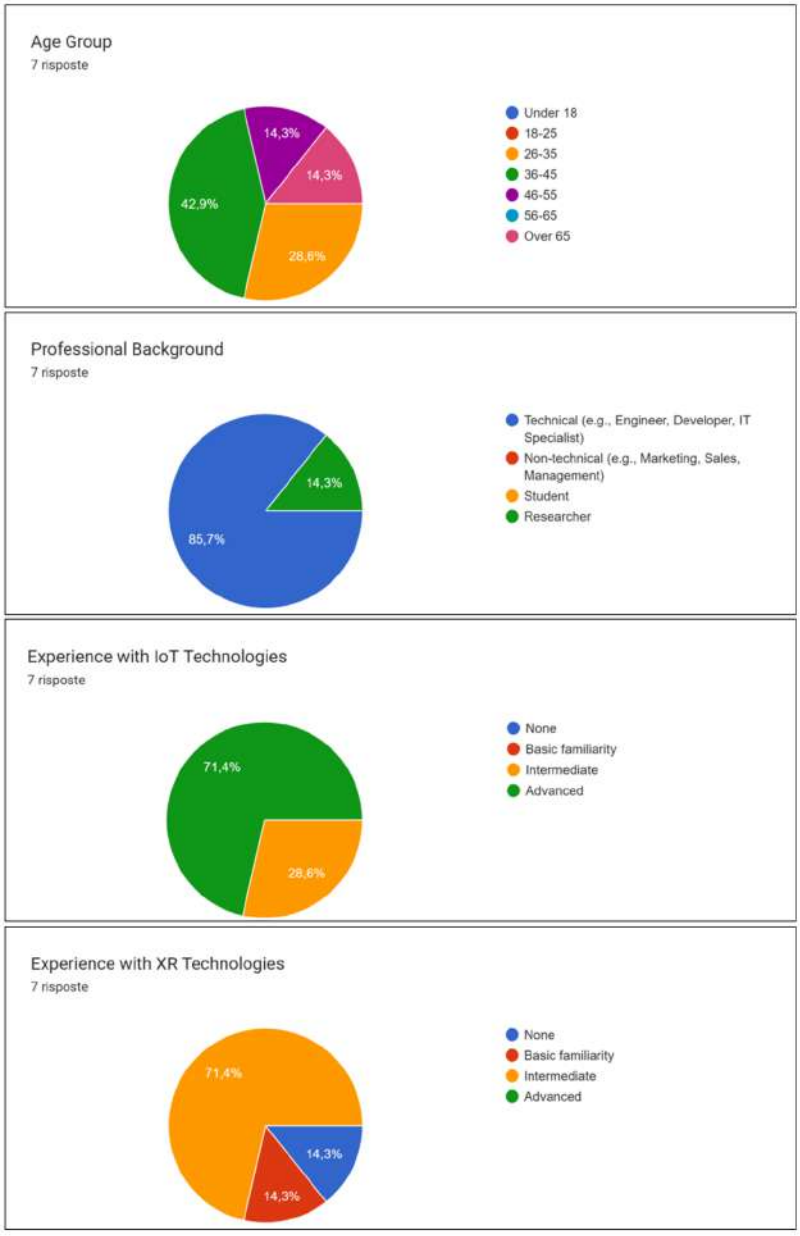


Figure 5.10: Demographic and background information of participants: (a) Age Group, indicating the distribution of participant ages; (b) Professional Background, categorizing participants by their field of expertise; (c) Experience with IoT Technologies, illustrating the participants' familiarity with Internet of Things systems; (d) Experience with XR Technologies, showing the participants' prior exposure to Extended Reality systems.

General Evaluation

The results indicated a general appreciation of the proposed system, with a particularly strong endorsement for its ease of use. This aspect enhances the system's potential for demonstration to a broader audience, including those with varying levels of technical expertise. However, the system's responsiveness was rated less favorably, which is a critical consideration, especially given that the evaluators predominantly have a technical background. This observation suggests that while the system is user-friendly, its performance in terms of responsiveness needs to be addressed to ensure it meets the expectations of technically proficient users. This factor will be important in the evaluation of subsequent sections, particularly those involving real-time interactions and system reliability.

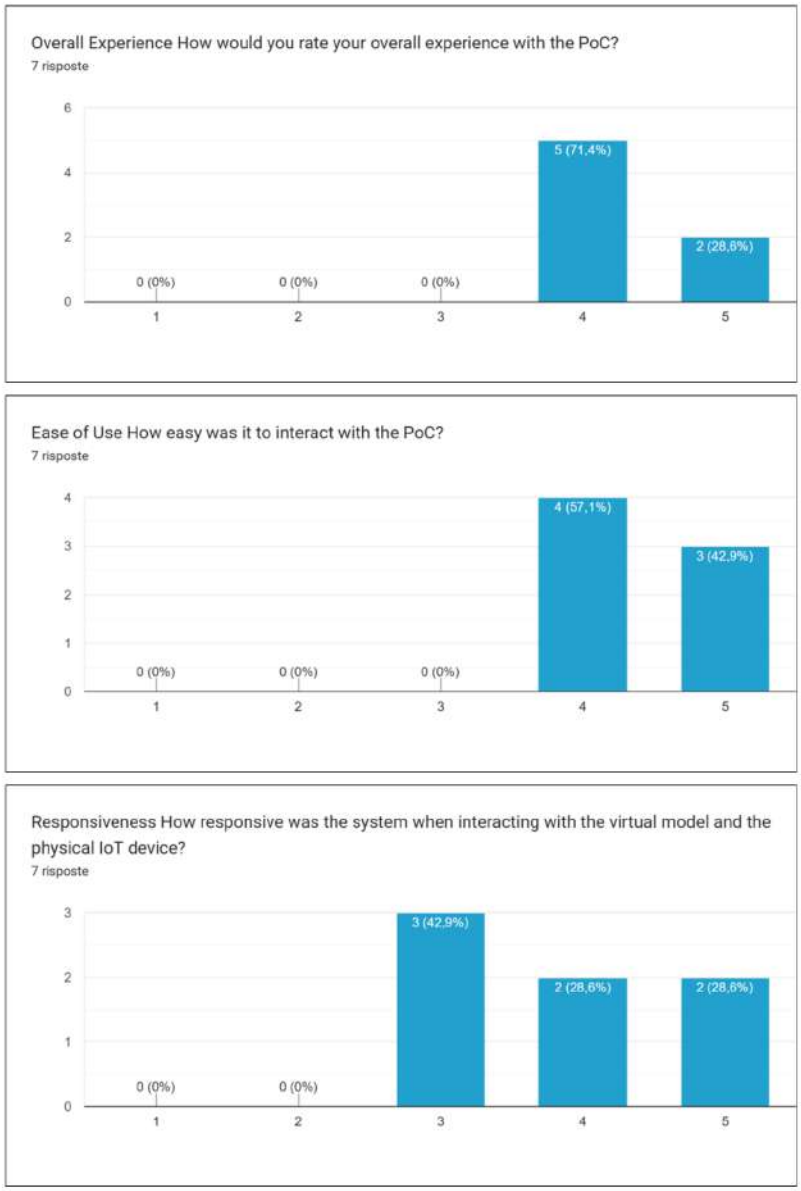


Figure 5.11: Graphs illustrating the general evaluation of the Proof of Concept (PoC), focusing on participants' ratings for Overall Experience, Ease of Use, and Responsiveness. These metrics provide insight into the users' initial impressions and the usability of the system.

Visual and Interactive Quality

The results indicate that the demonstration is graphically qualitative and effectively conveys the functionalities of the PoC. However, some participants reported a lack of immersion in the virtual environment. This shortfall could be addressed by incorporating XR devices, such as VR headsets, to enhance the immersive experience. Given the technical background of the participants, their feedback suggests that improvements in immersion could significantly impact the perceived overall effectiveness of the system. This aspect will be critical to consider in subsequent evaluations, particularly in the context of broader deployment scenarios.

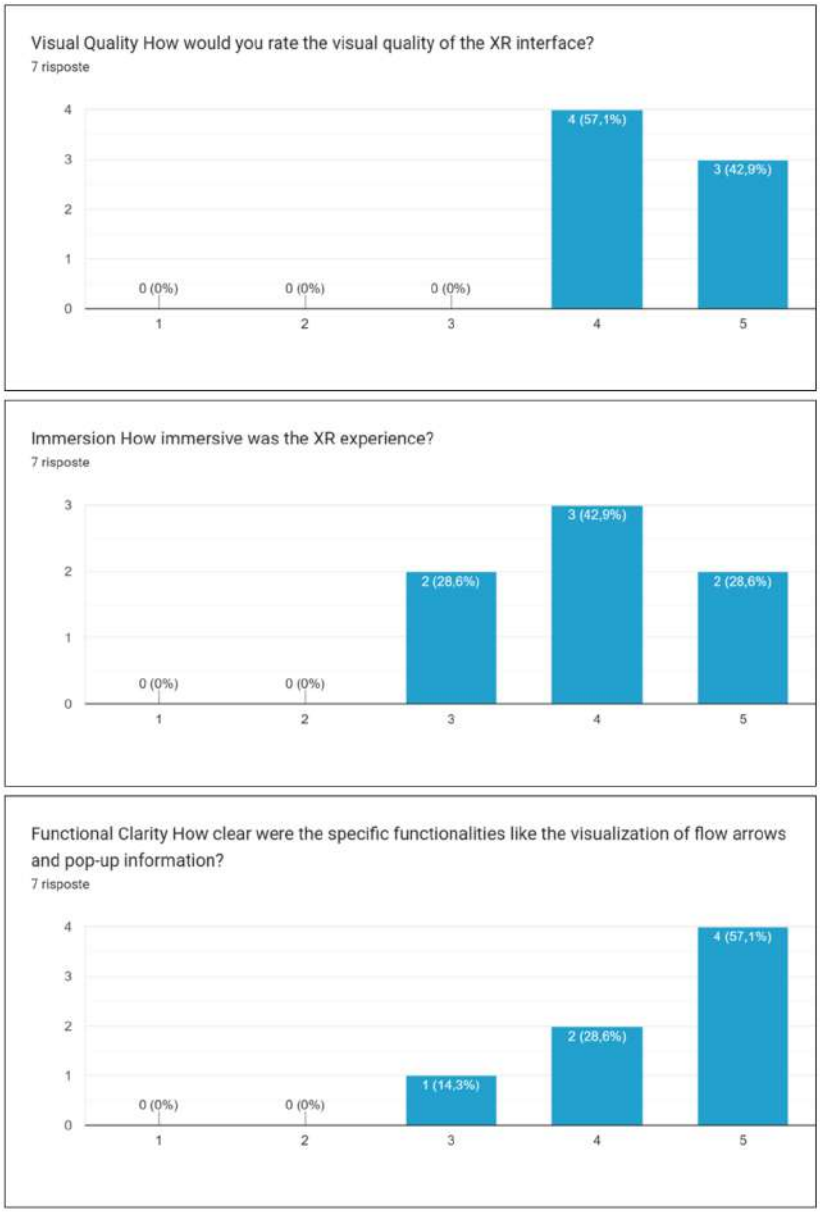


Figure 5.12: Graphical representation of participant evaluations regarding the visual and interactive quality of the system, encompassing: Visual Quality, Immersion, and Functional Clarity. The results reflect the perceived effectiveness of the system's visual presentation, the depth of immersive experience, and the clarity of functional elements during interaction.

Informational Clarity and Understanding

The results indicate that the demonstration effectively communicated the information, providing a tangible enhancement over a standard PoC demonstration. This was achieved through the use of informative pop-ups and animations, specifically related to the water flow and its inversion. These elements contributed to a clearer understanding of the system's functionality, making the demonstration not only visually appealing but also educationally valuable.

This clarity and added value are critical as they suggest that similar interactive elements could be beneficial in other aspects of the system, potentially improving user engagement and comprehension in future iterations or deployments.

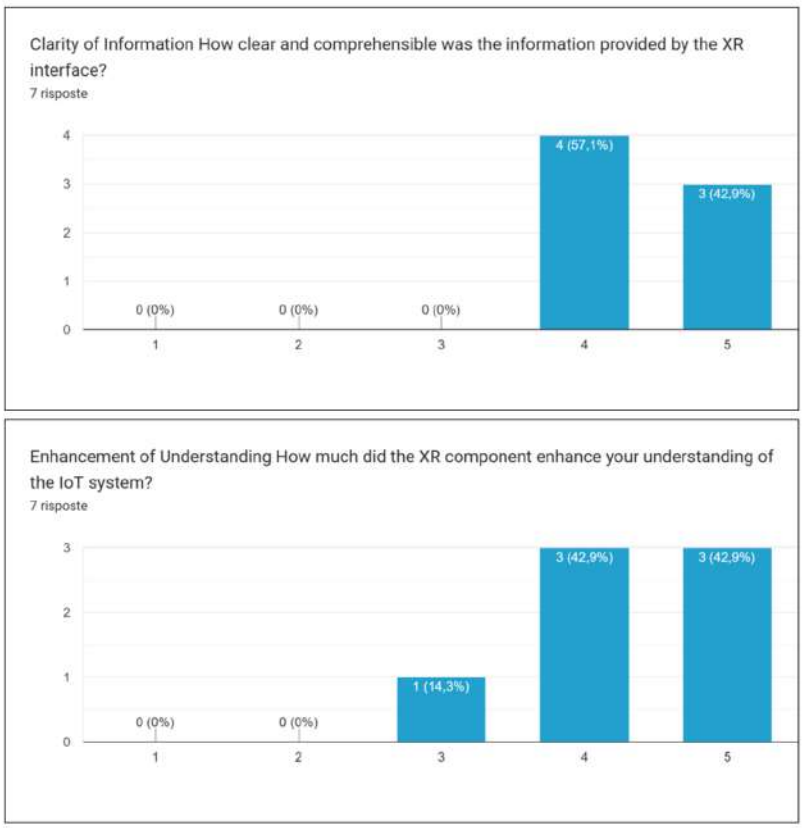


Figure 5.13: Graphs illustrating the evaluation of Informational Clarity and Understanding. The metrics displayed include Clarity of Information, which assesses the participants' perception of how clearly the information was conveyed, and Enhancement of Understanding, which measures the extent to which the demonstration improved their comprehension of the system's functionality.

Technical Performance

The results revealed that the system is perceived as reliable and accurate; however, a significant latency issue was identified. As noted in Section 4.4.1, the system comprises various nodes, which may be designed and managed by third parties. This Proof of Concept (PoC) serves as an effective testbed for evaluating these parameters and refining the product before its production release.

The observed latency could impact user satisfaction and the overall effectiveness of the system

in practical scenarios. Addressing this latency issue is critical as it may affect how users interact with the system and perceive its responsiveness. The technical performance results underscore the need for further optimization, particularly in terms of reducing latency. Such improvements are essential for ensuring that the system meets performance expectations in real-world applications and enhances the user experience in subsequent evaluations.

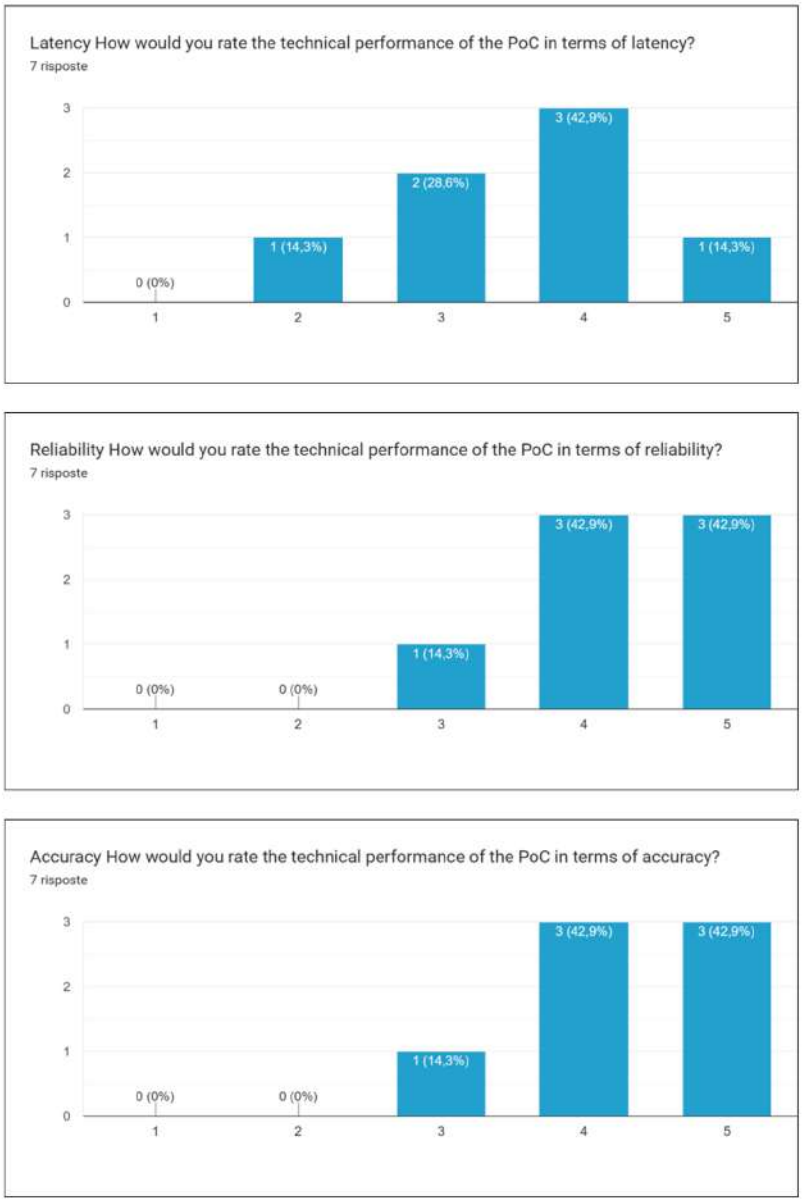


Figure 5.14: Graphical representation of the technical performance metrics, including Latency, Reliability, and Accuracy, as evaluated by the participants. These metrics provide insight into the system's responsiveness, stability, and precision, which are crucial for its practical application and user satisfaction.

Comparative and Future Considerations

The results indicate that the proposal to leverage IoT and XR technologies for enhancing prototype demonstrations is generally well-received and satisfactory. Participants expressed a strong attraction to and satisfaction with the overall experience. Furthermore, the majority

of participants view the integration of IoT and XR technologies as valuable, considering their potential application in their professional or personal lives.

This positive reception and high perceived value suggest that the implementation of IoT and XR in prototype demonstrations holds significant promise. It implies that future enhancements could focus on optimizing the system's responsiveness and immersion, addressing the concerns noted in the earlier evaluations. The favorable response to the integration of these technologies also points to their potential for broader adoption and suggests a beneficial direction for future developments. Enhanced user engagement and satisfaction could lead to increased interest and investment in similar technologies, thereby influencing future design and deployment strategies for XR and IoT solutions.

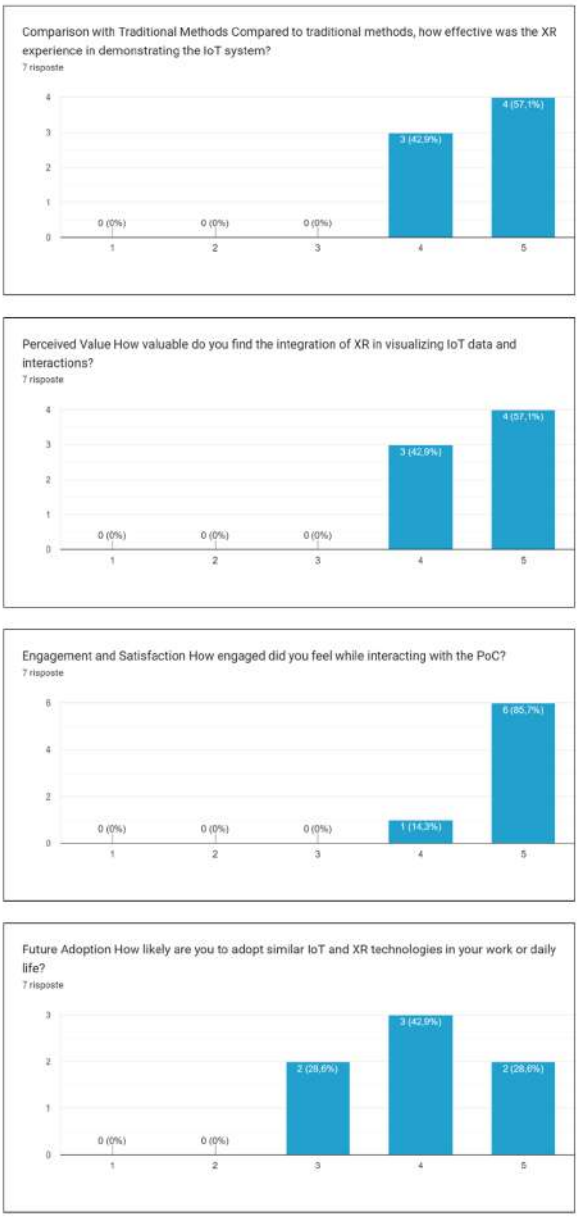


Figure 5.15: Graphs illustrating the results for Comparative and Future Considerations, including: Comparison with Traditional Methods, Perceived Value, Engagement and Satisfaction, and Future Adoption. These metrics assess how the XR Proof of Concept compares to traditional methods, its perceived value, user engagement and satisfaction, and the potential for future adoption.

5.5.7 Conclusions

The exploratory survey conducted to assess the Quality of Experience (QoE) of the IoT-XR Proof of Concept (PoC) has yielded significant insights into user perceptions and system performance. The findings, derived from a sample of seven participants with a predominantly technical background, underscore both the strengths and areas for improvement within the PoC.

Overall, the PoC was positively received, with notable praise for its ease of use and visual quality. Participants appreciated the user-friendly interface and the clarity of the visual elements and interactive features. The informative pop-ups and animations contributed effectively to the understanding of the IoT system's functionalities, enhancing both user engagement and educational value.

However, several areas warrant attention for future enhancements. The system's responsiveness, though generally satisfactory, was rated lower compared to other aspects, indicating that latency issues need to be addressed. This aspect is critical as it affects the system's real-time interaction capabilities and overall user satisfaction. Additionally, the immersive quality of the XR experience was identified as a potential area for improvement. The incorporation of advanced XR devices, such as VR headsets, could enhance immersion and address current deficiencies in user engagement.

The technical performance of the PoC was deemed reliable and accurate; however, latency emerged as a significant challenge. Addressing this issue will be crucial for ensuring that the system meets performance expectations and delivers a seamless user experience. Future developments should focus on optimizing latency, enhancing immersion, and expanding the testing base to include a broader range of users. This expanded testing will provide a more comprehensive understanding of the system's effectiveness across different user groups.

Looking ahead, several challenges need to be addressed. Scaling the system to accommodate a larger user base and more complex scenarios presents technical and logistical challenges. Ensuring seamless integration of XR components with diverse IoT devices is also a key concern, particularly in maintaining compatibility and consistent performance. Additionally, addressing the diverse needs and preferences of a wide user base will be essential to ensure that the system remains intuitive and accessible. Continual improvement based on user feedback and technological advancements will be necessary to sustain and enhance the PoC's effectiveness and impact.

In conclusion, while the PoC demonstrates substantial potential and positive reception, focusing on the identified challenges and future development areas will be essential for optimizing the system and maximizing its applicability in subsequent deployments.

5.6 Case Study 2: Efficient 3D Reconstruction of Indoor Environment Using Equirectangular Images in Photogrammetry

5.6.1 Background

The field of photogrammetry has undergone substantial evolution over the past few decades, driven by advancements in imaging technology and computational methods. Traditional photogrammetric techniques relied heavily on manual processes and a limited number of photographs, often resulting in lengthy data acquisition and processing times. However, recent developments in automation, computer vision, and machine learning have transformed this landscape, enabling faster and more accurate 3D reconstructions from 2D images.

In sensitive contexts, such as crime scene investigations, the need for rapid content acquisition becomes paramount. Timely capture of visual evidence can significantly impact the integrity of investigations and subsequent legal proceedings. Consequently, technological advancements in photogrammetry may pave the way for new techniques in the documentation and crystallization of indoor environments.

A notable innovation in this domain is the adoption of equirectangular imaging techniques [135]. Equirectangular, or spherical, photographs capture a panoramic view of the environment, representing the entire scene in a single image. This capability accelerates data acquisition and enhances the potential for immersive experiences in virtual reality (VR) applications. As VR technology gains traction across diverse fields—including education, training, entertainment, and forensics—the significance of high-quality equirectangular images becomes increasingly apparent.

These advancements facilitate the integration of photogrammetry with VR technologies, allowing users to interact with 3D reconstructions more intuitively. The use of equirectangular images enriches the user experience by providing a comprehensive spatial understanding of the environment. Additionally, as the demand for rapid deployment in sensitive contexts continues to grow, leveraging equirectangular photography as an effective data capture tool is gaining significant attention.

With the evolution of photogrammetric techniques, corresponding processing software has also advanced; many tools, such as Agisoft Metashape Pro, are now capable of processing equirectangular images. This development opens new avenues for the application of equirectangular photogrammetry, allowing for improved data capture and reconstruction in various contexts, including crime scenes [136].

This case study explores the application of contemporary photogrammetric methodologies, with a focus on the efficient reconstruction of indoor scenes. The effectiveness of established protocols, such as the Dundee GT Protocol [137], is examined alongside the capabilities

of advanced imaging devices like the Insta360Pro2. This assessment aims to evaluate the feasibility of integrating these innovative techniques into practical workflows, ensuring both speed and quality in the reconstruction process.

5.6.2 State of the Art

In recent years, advancements in photogrammetry and computer vision have led to significant improvements in the reconstruction of 3D models from 2D images. Traditional methods, such as Structure from Motion (SfM) and Multi-View Stereo (MVS), have paved the way for more sophisticated techniques that leverage equirectangular images. Equirectangular projection, often utilized in spherical and panoramic imaging, offers a comprehensive representation of an environment by encompassing a wide field of view, which is particularly beneficial for indoor scene reconstruction [138].

One notable contribution to this field is the Dundee GT Protocol, which outlines a systematic approach for capturing indoor scenes using DSLR cameras [137]. This protocol emphasizes the importance of achieving high-quality image capture while minimizing the time spent during the acquisition phase. By utilizing this protocol, practitioners can ensure that critical details are preserved, which is essential for applications in fields such as forensic science and architectural documentation.

Recent studies have explored the application of 360-degree cameras in the reconstruction of indoor environments [139]. These cameras offer the advantage of capturing a complete view of the scene in a single shot, significantly reducing the time required for image acquisition compared to traditional DSLR methods [138]. However, challenges remain regarding the quality of the generated 3D models, particularly concerning detail resolution and texture fidelity. While established methods for both subjective and objective assessment of equirectangular image quality exist [140], the literature lacks extensive studies focused on the qualitative evaluation of 3D models generated from equirectangular images, particularly in terms of resolution, accuracy, and usability in practical applications.

Additionally, the extraction of images from video recordings captured by 360-degree cameras has been examined as a potential method for improving workflow efficiency. This approach allows for rapid data collection while maintaining the possibility of high-resolution output, assuming that appropriate post-processing techniques are employed [141].

Despite the progress in these areas, a gap remains in the comparative analysis of traditional and modern methodologies for indoor scene reconstruction. Although some comparative studies have been conducted [138][139], none have specifically tested the Dundee GT Protocol in conjunction with 360-degree imaging technology, such as the Insta360Pro2. Further investigation is necessary to assess the feasibility of adapting the Dundee GT Protocol for use with these advanced imaging systems while ensuring the quality of the resulting 3D models is maintained.

5.6.3 Motivation and Objectives

The motivation behind this case study stems from the critical need for efficient and precise methods of documenting indoor environments, particularly in contexts where time is of the essence, such as crime scene investigations. Traditional photogrammetric techniques, while effective, often require extensive setup and processing time, which may compromise the integrity of the evidence if not managed promptly.

Recent advancements in photogrammetric software have introduced the capability to process equirectangular images, significantly expanding the potential applications of this technology. Software solutions such as Agisoft Metashape Pro and RealityCapture now support equirectangular image processing, facilitating comprehensive environmental capture and more streamlined data acquisition workflows. This development underscores the importance of exploring the integration of equirectangular imaging within established photogrammetric protocols.

The advent of advanced imaging technologies, particularly equirectangular photography, presents a significant opportunity to enhance both the speed and quality of data acquisition. Equirectangular images enable comprehensive environmental capture, allowing for the documentation of an entire scene with a single image, thus eliminating the need for numerous individual photographs. This capability is particularly advantageous in sensitive scenarios where the preservation of evidence and the minimization of disturbance to the scene are paramount.

The primary objectives of this case study are:

- **Evaluation of the Dundee GT Protocol with Insta360Pro2:** Assessing the effectiveness of the Dundee GT Protocol when utilized with the Insta360Pro2.
- **Creation of 3D Models:** Generating 3D models of an indoor environment using photogrammetry with DSLR images, equirectangular images, and images extracted from an equirectangular video.
- **Comparison of Processing Times:** Analyzing the processing times, with a specific focus on equipment preparation time, image acquisition time, and model creation time.
- **Qualitative Comparison of 3D Models:** Evaluating the qualitative aspects of the resulting 3D models to determine any improvements or deteriorations in detail and accuracy when employing the same SOP with different imaging technologies.

By conducting this comparative analysis, the study aims to identify best practices for the efficient 3D reconstruction of indoor scenes, while ensuring that the standards of detail and accuracy are upheld.

5.6.4 Tools and Software Utilized

This case study employs a combination of hardware and software tools to facilitate the efficient 3D reconstruction of indoor scenes. The following sections outline the specific tools and software utilized during the project.

DSLR Nikon D3100

The Nikon D3100 is an entry-level digital single-lens reflex (DSLR) camera characterized by its compact design and robust performance in various imaging conditions. It features a 14.2-megapixel CMOS sensor that operates within the Nikon DX format, facilitating a 1.5x crop factor for effective field-of-view adjustments. The camera's ability to capture high-resolution images in RAW format is particularly advantageous for photogrammetric applications, where precision and detail are critical for accurate 3D reconstructions. Additionally, the D3100 supports Full HD video recording, offering flexibility in documenting dynamic scenes for further analysis. Its user-friendly interface makes it an accessible tool for both novice photographers and professionals engaged in the meticulous documentation of environments for forensic or architectural photogrammetry [142].

Table 5.7: Nikon D3100 DSLR Camera Specifications

Digital Camera Specifications	
Feature	**Specification**
Product Name	Nikon D3100 DSLR Camera
Image Sensor	23.1 x 15.4 mm CMOS sensor, Nikon DX format (1.5x FOV crop), 14.2 million effective pixels
Image Quality	NEF (12-bit compressed RAW), JPEG fine (1:4), JPEG normal (1:8), JPEG basic (1:16)
Video Resolution	1920 x 1080 at 24p, 1280 x 720 at 30p, 1280 x 720 at 25p
ISO Sensitivity	Auto, ISO 100-12800
Lens Mount	Nikon F mount (with AF contacts)
Included Lens	Nikon 18mm-55mm VR Lens
Autofocus System	11 area TTL phase detection, Nikon Multi-CAM1000 autofocus module
Continuous Shooting Rate	Approximately 3.0 frames per second
Maximum Shutter Speed	1/4000th of a second
Viewfinder	Optical fixed eye-level, 95% frame coverage
LCD Monitor	3.0" TFT LCD, 230,000 pixels
Battery Type	Lithium-Ion battery pack EN-EL14
Weight	505g (1.1 lb) including battery
Dimensions	124 x 96 x 75 mm (4.9 x 3.8 x 3.0 in)
Supplied Accessories	Rechargeable Li-ion Battery EN-EL14, Quick Charger MH-24, USB Cable UC-E4

Nikon DX Nikkor 18-55mm f/3.5-5.6G VR II

The Nikon DX Nikkor 18-55mm f/3.5-5.6G VR II is a midrange zoom lens optimized for Nikon DX-format digital SLR cameras. This lens incorporates a silent wave motor (SWM) for efficient autofocus performance and Vibration Reduction (VR) technology to counteract camera shake, thereby enhancing image stability during handheld shooting. With a focal length range equivalent to approximately 27-82.5mm in 35mm terms, this lens is suitable for a diverse array of applications, including landscape, portrait, and architectural photography. In the context of photogrammetry, the lens's ability to produce high-resolution images with minimal optical distortion is crucial for accurate 3D model reconstruction and measurement, making it an invaluable tool for professionals in fields such as surveying and environmental monitoring.

Table 5.8: Nikon DX Nikkor 18-55mm f/3.5-5.6G VR II Specifications

Lens Specifications	
Feature	**Specification**
Product Name	Nikon DX Nikkor 18-55mm f/3.5-5.6G VR II
Autofocus	Yes
Vibration Reduction	Yes
Maximum Aperture	f/3.5-5.6
Minimum Aperture	f/22-36
Weight	265 g
Maximum Diameter	73 mm
Length	79 mm
Filter Diameter	52 mm
Picture Angle	76° - 28°50'
Groups/Elements	8/11
Number of Diaphragm Blades	7
Closest Focusing Distance	0.28 m
Release Date	2014
MSRP	$250

Insta360 Pro2

The Insta360 Pro2 is a state-of-the-art 360-degree camera designed for professional use in high-resolution imaging and video capture. It features six lenses with a maximum resolution of 8K and advanced stitching capabilities for seamless image creation. The camera supports multiple recording modes, including live streaming in 8K, making it suitable for various applications, such as virtual reality content creation, immersive storytelling, and real estate photography. A particular focus has been placed on a new feature compared to its predecessors: the Automatic Exposure Bracket (AEB) mode. This mode allows the camera to continuously capture 3, 5, 7, or 9 groups of photos at different exposure levels, synthesizing them into a single image with a wider dynamic range, commonly referred to as HDR (High Dynamic Range) photography. This functionality is particularly beneficial in static environ-

ments where light differences are significant, ensuring that images retain detail across varying exposure levels, that also provides improvements to photogrammetry [143]. Additionally, the camera boasts impressive stabilization technology, allowing for smooth footage even in dynamic shooting conditions, making it a versatile choice for professionals in diverse fields.

Table 5.9: Insta360 Pro2 Camera Specifications

Digital Camera Specifications	
Feature	**Specification**
Product Name	Insta360 Pro2
Number of Lenses	6 x F2.4 fisheye lenses
Storage	6 MicroSD cards + 1 Full SD card
Power Supply	12V 5A adapter
Material	Aluminum alloy
Exposure Modes	Auto, Manual, Custom per lens, Shutter Priority (photo only), ISO Priority (photo only)
Diameter	143mm
Weight	1550g
ISO Range	100-6400
Battery Capacity	Removable 5100mAh battery
White Balance Modes	Auto, Manual
Bitrate per Lens	Up to 120Mbps
Video File Format	MP4
Video Encoding	H.264 in-camera, H.265 available for post stitching
Video Resolution (Post-Processing)	7680 x 3840 @30 fps HDR (8K 2D), 7680 x 7680 @30 fps HDR (8K 3D)
Streaming Modes	4K 2D or 3D @30fps, 8K 2D @30fps (additional license and PC required)
Photo File Formats	JPEG / DNG
Photo Modes	Single Shot, Burst (10 continuous shots), Timelapse, AEB (HDR)
Video Transmission Range	Up to 300m ground-to-ground, 1000m ground-to-air
Stabilization	9-axis high-precision gyroscope, FlowState stabilization
GPS	Built-in module
Audio Format	AAC, Spatial audio (Ambisonic)

Hahnel Triad 60 Lite Tripod

The Hahnel Triad 60 Lite is a lightweight tripod constructed from aluminum alloy, weighing 1.9 kg with a maximum load capacity of 5 kg. It features a three-way fluid-effect pan and tilt head, dual spirit levels for precise leveling, and quick-release plate for efficient camera mounting. The tripod's legs have quick locks and rubber feet, providing stability and protection on various surfaces. It offers adjustable heights with a maximum of 162 cm and a minimum of 43.5 cm using a reversible center column. These features make it well-suited for use with both the Nikon D3100 DSLR and the Insta360 Pro2 camera.

Table 5.10: Specifications of the Hahnel Triad 60 Lite Tripod

Feature	**Specification**
Head Type	Three-way, fluid-effect pan & tilt
Quick Release Plate	Yes
Spirit Level	Dual
Pan Lock	Yes
Leg Sections	4-section with snap leg locks
Feet Type	Rubber feet
Maximum Height	162 cm
Minimum Height	43.5 cm (low level with reversible column)
Maximum Load Capacity	5 kg
Weight	1.9 kg
Material	Aluminum alloy
Additional Features	Includes case

ASUS TUF Dash F15 Laptop

The ASUS TUF Dash F15 is a high-performance laptop designed for demanding computational tasks, making it ideal for photogrammetric processing. Equipped with an NVIDIA GeForce RTX 3050 dedicated graphics card, it provides exceptional parallel processing power essential for complex photogrammetry algorithms. Its 12th Gen Intel Core i5-12459H CPU and 32GB of RAM ensure efficient handling of large datasets and intensive tasks. The high-speed NVMe SSD storage further enhances data read/write operations, reducing processing time. These hardware capabilities allow the ASUS TUF Dash F15 to efficiently run photogrammetric software, such as Agisoft Metashape and RealityCapture, delivering high accuracy and fast processing speeds.

Table 5.11: ASUS TUF Dash F15 Specifications

Laptop Specifications	
Feature	**Specification**
Model	ASUS TUF Dash F15
Processor	12th Gen Intel(R) Core(TM) i5-12450H 2.00 GHz
Memory	32GB RAM SODIMM 4800 MT/s
Operating System	Windows 11 Home 64bit
Graphics	NVIDIA GeForce RTX 3050 Laptop GPU 4GB
Storage	NVMe HD SSD 512GB

Agisoft Metashape [80]

Agisoft Metashape is a commercial photogrammetry software known for its high accuracy and comprehensive feature set. It supports both aerial and terrestrial photogrammetry, making it suitable for a wide range of applications, from surveying and mapping to cultural heritage preservation and visual effects production. Metashape offers advanced editing tools for post-processing, including dense point cloud generation, mesh editing, texture mapping, and orthophoto creation. Its user-friendly interface and detailed documentation make it accessible to both novice and experienced users. Additionally, Metashape provides options for batch processing and network processing, enhancing its efficiency for large-scale projects.

Table 5.12: Specifications of Agisoft Metashape

Attribute	**Details**
Platform	Windows macOS Linux
Cost	$179 for Standard $3,499 for Professional
Accuracy	Very High
Processing Speed	High
Data Source	Standard images Equirectangular images
Notable Features	Advanced editing tools, Supports aerial and terrestrial photogrammetry

Darktable [144]

Darktable is a free and open-source photography workflow application and raw developer. It is designed to manage digital negatives in a database, allowing photographers to view them through a zoomable lighttable and develop raw images with high-quality adjustments. Darktable supports non-destructive editing, meaning all adjustments are applied on-the-fly without altering the original image. It is highly extensible, supporting plugins and custom scripts, and is cross-platform, running on Linux, Windows, and macOS. As a community-driven project licensed under the GNU GPL, Darktable benefits from continuous contributions and frequent updates. Its powerful features and flexibility make it an excellent choice for professional photographers and enthusiasts alike.

5.6.5 Methodology

Prerequisites

To successfully carry out this work, several prerequisites are necessary:

- Proficiency in terrestrial photogrammetry and close-range photogrammetry techniques, along with familiarity with the relevant software.
- Access to appropriate equipment capable of capturing high-quality images suitable for photogrammetry in indoor environments.
- Competence in using photography software to enhance the quality of acquired images.
- Hardware with high computational power to efficiently process the photogrammetric data.

Data Acquisition

Equipment Preparation The tripod was configured to position the camera at a height of 140 cm, approximately halfway between the floor and the ceiling of the test environment. Due to the presence of furniture, it was not feasible to place the tripod precisely at the starting vertex of the 60x60 grid. Consequently, it was positioned slightly offset by 20 cm. The setup of the tripod took 5 minutes and 58 seconds.

$$\mathbf{T}_{\text{Tripod Setup}} = 0\,\text{h} + 5\,\text{m} + 58\,\text{s} \tag{5.1}$$

(a) Tripod dimensions

(b) Eye view

(c) Low view

(d) Balance check 1

(e) Balance check 2

Figure 5.16: Various aspects of tripod setup and camera views: (a) Tripod dimensions, showing the height and stability; (b) Eye view, illustrating the camera perspective at eye level; (c) Low view, highlighting the camera perspective from a lower angle; (d) Balance check 1, ensuring the tripod's stability and balance; (e) Balance check 2, confirming the tripod's alignment and leveling.

Site Selection and Setup For the reconstruction, a typical room in an apartment was selected, which can serve as a sample environment for a potential crime scene. The room was left in the state it was found at the start of image capture, with elements out of place that could impede tripod positioning, mirroring the conditions of a real crime scene, where any movement of objects could potentially contaminate the environment. During the photo shoot, care was taken not to move any objects obstructing the tripod's placement. If the tripod could not be positioned exactly as desired, it was shifted by an appropriate number of centimeters, as specified in the Dundee GT Protocol.

The room is fairly spacious, with some areas well-illuminated by natural light from the window on a clear day. The rest of the room was illuminated by two artificial light sources with a yellow hue. Environmental conditions, including lighting, were carefully considered to maintain the authenticity of the scene and ensure consistent image quality.

(a) Perspective View 1

(b) Perspective View 2

(c) Top View

Figure 5.17: Various perspectives of the room chosen for the test reconstruction

Image Capture Protocols To apply the protocol, the grid shown in Figure 5.18 was used, in accordance with the Standard Operating Procedure (SOP). As detailed below, the image acquisition times varied significantly.

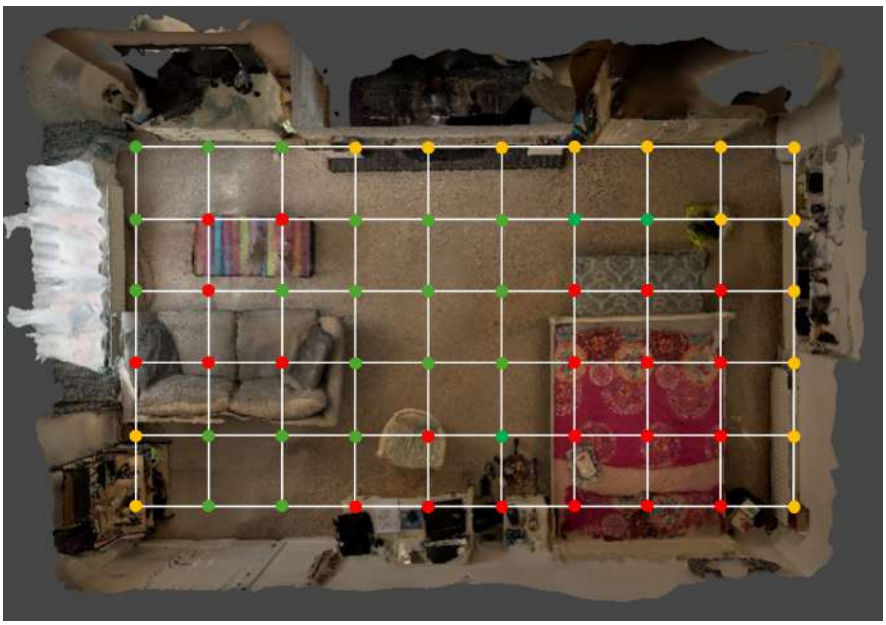


Figure 5.18: Graphical representation of the grid used for capturing images within the indoor test environment. The grid consists of 60x60 cm squares, with the tripod positioned at each vertex. Green points indicate locations where the image capture was feasible without any issues. Orange points represent positions where the image capture was possible, but required repositioning of the tripod due to an obstacle. Red points indicate locations where image capture was not possible, as it would have led to an alteration of the environment.

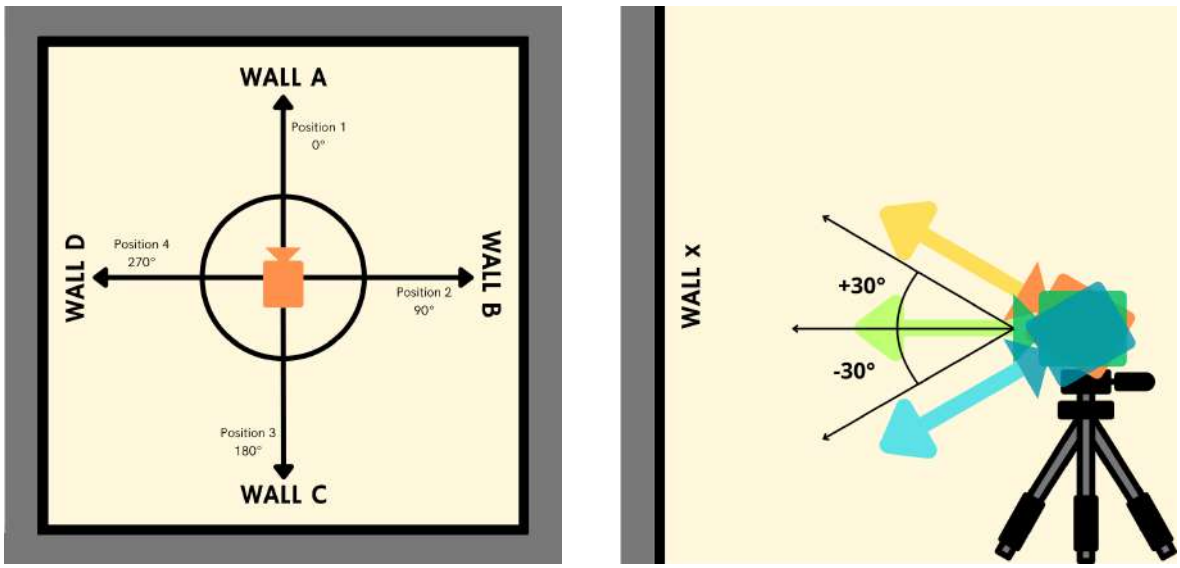
- **DSLR Image Capture:** The DSLR was prepared in accordance with the Dundee GT Protocol (settings are summarized in Table 5.13).

Table 5.13: DSLR Configuration Settings

Setting	**Value**
Image Quality	RAW
White Balance	Preset Manual (Pre-preset Normal)
Release Mode	2s
ISO	100
Metering Option	Spot Metering
Shutter Speed	1.5

The setup of the DSLR took 4 minutes and 22 seconds.

At each acquisition point, 12 photos were taken at eye level: 3 photos per side at angles of +30°, 0°, and -30° (see Fig. 5.19). In some cases, additional shots were taken at a lower height to capture more details.



(a) Top View perspective

(b) Shot procedure per each point

Figure 5.19: Overview of the image capture procedure for the test reconstruction. (a) Depicts the top-view perspective of the shooting process at each capture point. The camera is directed at one wall, three shots are taken as shown in (b), and then the camera is rotated 90° towards the next wall. After completing the shots for all four walls, the tripod is moved to the next capture point. (b) Illustrates the detailed shot procedure for each point, with the camera aimed at a single wall.

A total of 477 photos were taken across 37 acquisition points. The entire process, including taking the photos, measuring the subsequent acquisition point, and repositioning the tripod, took a total of 2 hours, 59 minutes, and 36 seconds.

The total time can be calculated as:

$$\begin{aligned} \mathbf{T}_{\text{Image Capture}} &= \mathbf{T}_{\text{Camera Setup}} + \mathbf{T}_{\text{Shooting}} \\ &= (0\,\text{h} + 4\,\text{m} + 22\,\text{s}) + (2\,\text{h} + 59\,\text{m} + 36\,\text{s}) \\ &= 3\,\text{h} + 3\,\text{m} + 58\,\text{s} \end{aligned} \tag{5.2}$$

- **Equirectangular Image Capture:** The Insta 360 Pro 2 was configured with the settings outlined in Table 5.14.

Table 5.14: Insta 360 Pro 2 Configuration Settings

Setting	Value
Mode	AEB9
Content Type	360 pano
RAW + JPG	ON
Optical Real Time Stitching	OFF
EV Step	±0.3
Delay Time	3 s
Exposure	Auto
Brightness	0
Saturation	64
Contrast	64
GPS	On
Battery	100%

The setup of the Insta 360 Pro 2 took 6 minutes and 27 seconds.

At each acquisition point, one equirectangular image was captured, eliminating the need to manually orient the camera, as was necessary in the DSLR method. This allowed for a more efficient workflow. The images were captured remotely via the dedicated smartphone app, allowing the photographer to exit the scene before shooting.

A total of 37 equirectangular images were taken across 37 acquisition points. The entire process, including image capture, measuring the subsequent acquisition point, and repositioning the tripod, took a total of 1 hour, 20 minutes, and 56 seconds.

The total time can be calculated as:

$$\begin{aligned} \mathbf{T}_{\text{Image Capture}} &= \mathbf{T}_{\text{Camera Setup}} + \mathbf{T}_{\text{Shooting}} \\ &= (0\,\text{h} + 6\,\text{m} + 27\,\text{s}) + (1\,\text{h} + 20\,\text{m} + 56\,\text{s}) \\ &= 1\,\text{h} + 27\,\text{m} + 23\,\text{s} \end{aligned} \tag{5.3}$$

A key aspect to consider is the battery consumption. At the end of the image capture process, the battery charge was at 19%, indicating that 81% of the battery was consumed during the procedure.

- **Equirectangular Video Capture:** The camera was configured with the settings outlined in Table 5.15.

Table 5.15: Insta 360 Pro 2 Video Configuration Settings

Setting	Value
Mode	High FPS
Content Type	360 pano
Project FPS	29.97 FPS
Single Lens Resolution	4k
Color Mode	Normal
Exposure	Auto
Brightness	0
Saturation	64
Contrast	64
Battery	100%

The setup of the camera required a card speed test before the video acquisition, which is included in the total setting time. The total setup time was 8 minutes and 27 seconds.

The video acquisition did not require the use of a tripod or its repositioning, as the camera was held by the operator. The path followed covered all the points in the grid shown in Figure 5.20. The procedure produced a single video with a duration of 54 seconds, which is the total acquisition time.

The battery consumption during the video acquisition was minimal, with only an 8% decrease in battery level.

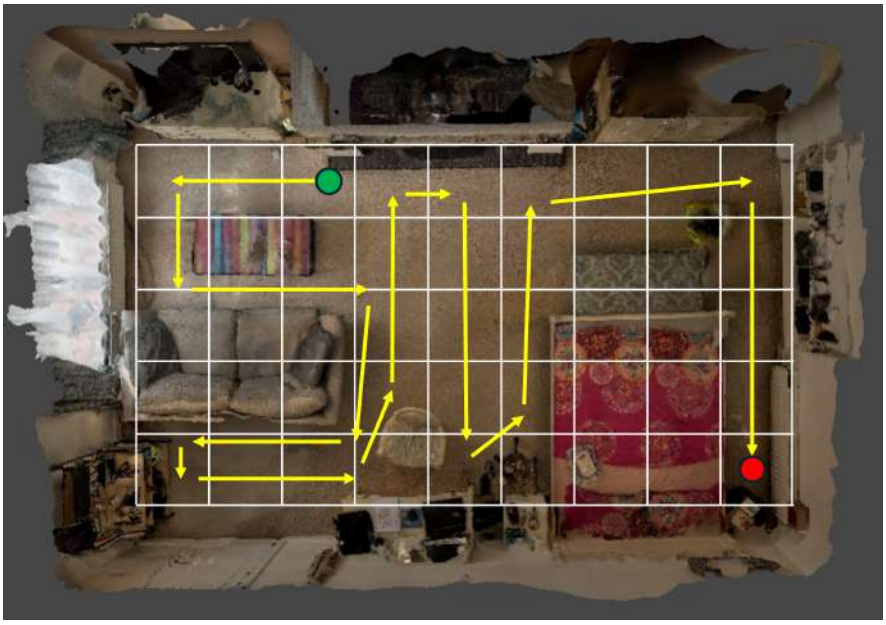


Figure 5.20: The path followed during the equirectangular video capture, covering grid points within the indoor test environment. The trajectory ensured that all relevant areas were recorded according to the grid pattern established for the scene. The green circle indicates the starting point, while the red circle marks the endpoint of the path.

The total time can be calculated as:

$$\begin{aligned}\mathbf{T}_{\text{Video Capture}} &= \mathbf{T}_{\text{Camera Setup}} + \mathbf{T}_{\text{Video}} \\ &= (0\,\text{h} + 8\,\text{m} + 27\,\text{s}) + (0\,\text{h} + 0\,\text{m} + 54\,\text{s}) \\ &= 0\,\text{h} + 9\,\text{m} + 21\,\text{s} \end{aligned} \tag{5.4}$$

Data Processing

It is important to clarify that the aim of this work is neither to describe the photogrammetric process in detail nor to produce a finalized model through extensive post-processing. The primary objective here is to obtain a workable 3D model while minimizing acquisition time by utilizing equirectangular images. The photogrammetric processes and post-processing techniques are assumed to be well-understood and are therefore not covered in depth within this document.

Detailed information on the processing and results can be found in Appendix A, which includes the reports generated by Metashape concerning the reconstructions. These reports provide comprehensive metrics and analysis relevant to the different imaging methods used in this study.

DSLR Image Processing

- **Images Pre-Processing:** The 478 images were pre-processed in Darktable to enhance sharpness and exposure where necessary. This pre-processing step allows the photogrammetric reconstruction software to detect more features within the images, improving the overall quality of the point cloud and subsequent 3D model.
- **Importing and Alignment:** Initially, several photographs were discarded during the alignment process, likely due to the challenging environmental conditions with multiple white surfaces. To address this, 19 markers were employed to assist the software in aligning the images, requiring significant effort and time. Despite these efforts, only 356 of the 478 photos were successfully aligned. Additional markers did not increase the number of photos that could be processed. The alignment process resulted in the generation of a sparse point cloud, with key metrics summarized in Table 5.16.

Table 5.16: Alignment Metrics for DSLR Image Processing

Metric	Value
Number of Photos	356 / 478
RMS Reprojection Error	0.143424 (0.81485 pix)
Max Reprojection Error	0.443242 (35.6196 pix)
Average Tie Point Multiplicity	2.63285
Key point limit	40,000
Tie point limit	4,000

- **Dense Point Cloud Generation:** Following alignment, the sparse point cloud was refined into a dense point cloud, containing 22,320,959 points. This refinement significantly enhances the detail and accuracy of the 3D reconstruction, making it more suitable for subsequent modeling processes.
- **3D Model Creation:** The depth maps was then used to generate a 3D model. After decimation using Metashape's tools, the final mesh contained 999,999 faces and 500,534 vertices, optimizing it for further processing. The 3D model was exported in three formats: OBJ (81.2 MB), GLB (56.6 MB), and FBX (21.2 MB), to ensure compatibility with different applications.

Table 5.17: 3D Model Metrics for DSLR Image Processing

Metric	Value
Faces	999,999
Vertices	500,534
OBJ File Size	81.2 MB
GLB File Size	56.6 MB
FBX File Size	21.2 MB

Equirectangular Image Processing

- **Images Pre-Processing:** The equirectangular images were pre-processed using Darktable to enhance sharpness and detail and to correct exposure where necessary. No distortion correction was applied, as Metashape is capable of processing spherical images directly. This pre-processing step is crucial to optimize the alignment process by improving the detection of features uniformly across the 360-degree view.
- **Importing and Alignment:** The equirectangular images were imported into the photogrammetry software. All 39 images were processed and aligned, and the alignment was further refined by using 7 markers. Although the alignment was already good without markers, their addition improved precision. The process of inserting markers

was much simpler compared to the DSLR case, as all markers were visible in every equirectangular image, and the reduced number of photos also facilitated the process. An optimization method was employed to improve the quality of the sparse point cloud and reduce errors [145], which resulted in significant improvements. The lowest errors achieved were an RMS reprojection error of 0.222384 (0.295767 pix) and a Max reprojection error of 0.828279 (1.56548 pix), with a key point limit set at 80,000 and a tie point limit set to 0, indicating that all detected tie points were used. The alignment metrics are summarized in Table 5.18.

Table 5.18: Alignment Metrics for Equirectangular Image Processing

Metric	**Value**
Number of Photos	39 / 39
RMS Reprojection Error	0.222384 (0.295767 pix)
Max Reprojection Error	0.828279 (1.56548 pix)
Average Tie Point Multiplicity	2.80933
Key Point Limit	80,000
Tie Point Limit	0 (all detected tie points used)

- **Dense Point Cloud Generation:** After alignment, the software generated a dense point cloud from the sparse data. The point cloud was computed using point confidence values, and the confidence filter provided by Metashape was applied to remove noise and reduce the number of points. Initially, the dense point cloud consisted of 1,165,593 points. After applying the filtering process, the number of points was reduced to 315,584, resulting in a cleaner and more refined dataset.
- **3D Model Creation:** The depth maps cloud was used to create a 3D model. The generated mesh consisted of 148,451 faces and 74,338 vertices. For the texturing process, not all equirectangular images were used, as it was empirically observed that areas very close to the capture point (for example, the room perimeter) exhibited significant distortion. A good visual result was achieved by using only the images taken from the central area of the room, where fewer distortions were present. The final 3D model was then exported in OBJ, GLB, and FBX formats, with the file sizes detailed in Table 5.19.

Table 5.19: 3D Model Metrics for Equirectangular Image Processing

Metric	Value
Faces	148,451
Vertices	74,338
OBJ File Size	11.2 MB
GLB File Size	7.35 MB
FBX File Size	3.16 MB

Equirectangular Video Processing

- **Frame Extraction:** Frames were extracted from the equirectangular video at regular intervals. The original video was cropped to a duration of 49 seconds with a frame rate of 60 fps. To extract 300 frames, the video was accelerated using a video-editing software. This approach ensured a consistent representation of the scene across frames, which is critical for accurate alignment and point cloud generation. The interval for frame extraction was chosen to balance detail and computational efficiency. Additionally, exposure adjustments were made to address the dark conditions of the original video. The maximum achievable resolution with the hardware used was 3840x1920 pixels.
- **Images Pre-Processing:** The extracted frames were pre-processed using Darktable to enhance sharpness, detail, and exposure consistency. No distortion correction was applied, as Metashape can directly handle spherical images. Despite these enhancements, the quality of the frames was lower compared to the single shots described in the previous section. This pre-processing step was crucial to improve the quality of the point cloud by ensuring uniform feature detection across the frames.
- **Importing and Alignment:** The pre-processed frames were imported into the photogrammetry software, where they were automatically aligned. The comprehensive 360-degree coverage provided by the equirectangular video facilitated the alignment process. An optimization method, similar to the one used for equirectangular images, was applied to improve alignment quality [145]. However, due to excessive noise caused by poor lighting conditions, it was not possible to use markers, and noise reduction using Darktable was unsuccessful. The lowest errors achieved were an RMS reprojection error of 0.492285 (0.852767 pix) and a Max reprojection error of 1.29742 (3.90654 pix), with a key point limit of 80,000 and a tie point limit of 0 (which uses all detected tie points). The alignment metrics from this process are summarized in Table 5.20.

Table 5.20: Alignment Metrics for Equirectangular Video Processing

Metric	Value
Number of Frames	300 / 300
RMS Reprojection Error	0.492285 (0.852767 pix)
Max Reprojection Error	1.29742 (3.90654 pix)
Average Tie Point Multiplicity	3.29791
Key Point Limit	80,000
Tie Point Limit	0 (all detected tie points used)

- **Dense Point Cloud Generation:** Following alignment, a dense point cloud was generated from the aligned frames. The dense point cloud captured detailed spatial information from the equirectangular video, providing a comprehensive representation of the scene. Similar to the process used for equirectangular images, the confidence-based filtering tool in Metashape was applied to the dense point cloud to remove noise and improve accuracy. The original dense point cloud consisted of 1,984,797 points, which was reduced to 858,609 points after filtering.
- **3D Mesh Creation:** The depth maps was used to create a 3D mesh. The mesh generation process transformed the point cloud into a three-dimensional model, representing the spatial details of the scene. To reduce texture distortions, only central images were used, as the overall texture quality was affected by the poor quality of the images. The generated 3D mesh consisted of 231,659 faces and 116,250 vertices. The final 3D model was exported in OBJ, GLB, and FBX formats with the file sizes detailed in Table 5.21.

Table 5.21: 3D Model Metrics for Equirectangular Video Processing

Metric	Value
Faces	231,659
Vertices	116,250
OBJ File Size	17.9 MB
GLB File Size	8.65 MB
FBX File Size	5.41 MB

5.6.6 Results Discussion

Challenges of the Chosen Site

The selected site is a typical indoor domestic environment, containing a variety of objects of different sizes and materials, illuminated by both natural and artificial light sources. This set-

ting was specifically chosen due to the significant challenges it presents for photogrammetric algorithms. The primary difficulties encountered include:

- **Environmental Conditions:** The lighting conditions posed a notable challenge during the data acquisition process. The natural light source was limited to a south-facing window, which only provided direct sunlight for a brief period each day. The acquisition was conducted under typical overcast weather conditions, further complicating the capture of high-quality images. Artificial lighting was provided by two ceiling-mounted warm light lamps, which created uneven lighting across the scene and added complexity to the image processing.

Figure 5.21: Image of the room highlighting the light sources (indicated in orange). The natural light source is a south-facing window, while artificial light is provided by two ceiling-mounted warm light lamps. The distribution and intensity of these light sources significantly influenced the photogrammetric data acquisition process.

- **Site Accessibility:** The physical arrangement of the room, including the placement of furniture and other obstacles, restricted the flexibility in positioning the tripod and capturing images from optimal angles (see Fig. 5.18). This limitation necessitated compromises in the coverage of certain areas, potentially affecting the completeness and accuracy of the photogrammetric model.
- **Surface Characteristics:** The site featured a high number of surfaces that are notoriously problematic for photogrammetry. The presence of both white and black surfaces, as well as transparent and reflective materials, significantly complicated the feature extraction process. These surfaces often led to inaccuracies in depth estimation, as photogrammetric algorithms struggle to identify reliable keypoints on such surfaces, resulting in potential errors in the final reconstruction.

Comparison of Acquisition Times

This section compares the total acquisition times for each imaging method used. The analysis includes:

- **DSLR Image Capture:** Total time for camera setup, image capture, and repositioning.
- **Equirectangular Image Capture:** Time required for camera setup, image capture, and repositioning.
- **Equirectangular Video Capture:** Total time for camera setup and video capture.

Table 5.22 provides a detailed comparison of these times to evaluate the efficiency of each method.

Table 5.22: Comparison of Acquisition Times for Different Imaging Methods

Time Metric	DSLR	Equirectangular Images	Equirectangular Video
Tripod Setup Time	5 m 58 s	5 m 58 s	5 m 58 s
Device Setup Time	4 m 22 s	6 m 27 s	8 m 27 s
Shooting Time	2 h 59 m 36 s	1 h 20 m 56 s	54 s
Total Time	**3 h 9 m 56 s**	**1 h 33 m 21 s**	**15 m 19 s**

An analysis of the total acquisition times for obtaining the data required for photogrammetric processing using the Dundee GT Protocol and the grid proposed for equirectangular image acquisitions reveals that using the Insta 360 Pro 2 drastically reduces the total time. This reduction is attributed to the fact that spherical image capture requires only a single shot per acquisition point, compared to the 12 shots needed with a DSLR for each point. For video capture, it is sufficient to follow a path that covers the grid points, as illustrated in Fig. 5.20. However, the reduced data acquisition time necessitates a greater effort in post-production to extract frames from the resulting video. This aspect should be considered if the goal is to minimize post-production time, while keeping in mind that the aim of this case study is to identify a method that allows for rapid acquisition and acceptable final results.

Quantitative Comparison of Methods

The quantitative analysis provided in this section aims to offer a comprehensive evaluation of the various photogrammetry methods utilized. By comparing key metrics, this analysis facilitates an objective assessment of the performance and effectiveness of each method. Such an evaluation is essential for understanding the relative strengths and limitations of the different approaches and for determining the most suitable method for specific applications. To understand the significance of the analyzed parameters, refer to Table 5.25.

The comparison focuses on several critical aspects:

- **Alignment Metrics:** This includes RMS Reprojection Error, Max Reprojection Error and Average Tie Point Multiplicity. These metrics are crucial for evaluating the accuracy of the alignment process, which directly affects the quality of the 3D reconstruction.

Table 5.23: Comparative Alignment Metrics for Different Methods

Parameter	DSLR Images	Equirectangular Images	Equirectangular Video
RMS Reprojection Error	0.143424 (0.81485 pix)	0.222384 (0.295767 pix)	0.492285 (0.852767 pix)
Max Reprojection Error	0.443242 (35.6196 pix)	0.828279 (1.56548 pix)	1.29742 (3.90654 pix)
Average Tie Point Multiplicity	2.63285	2.80933	3.29791

The alignment metrics show that DSLR Images generally achieve the lowest RMS and Max Reprojection Errors, indicating better accuracy in the alignment process compared to Equirectangular Images and Equirectangular Video. Specifically, the RMS Reprojection Error for DSLR Images is significantly lower than that for Equirectangular Video, highlighting the superior alignment quality. Similarly, the Max Reprojection Error for DSLR Images is also lower, suggesting fewer deviations in the projected points.

The Average Tie Point Multiplicity is higher for Equirectangular Video compared to the other methods. This suggests that more image observations are used per tie point, which may contribute to the higher accuracy in 3D reconstructions but at the cost of increased computational complexity.

In detailed environments, such as indoor domestic scenes, a higher number of images and better sensor quality are critical to achieving accurate reconstructions. Such environments typically feature intricate details, occlusions, and varying lighting conditions, which require higher resolution and precision. Therefore, devices with superior sensors, like DSLR cameras, are generally more effective in capturing the necessary detail. This also explains why more image acquisitions are often necessary for complex indoor spaces, as fewer images or lower-quality sensors may result in higher reprojection errors and suboptimal reconstructions.

- **3D Model Metrics:** This involves the number of faces and vertices in the generated 3D models, as well as the file sizes of the exported models. These metrics are used to evaluate the complexity of the 3D model and the efficiency of the file formats employed for storage and exchange.

Table 5.24: 3D Model Metrics for Different Methods

Parameter	DSLR Images	Equirectangular Images	Equirectangular Video
Faces	999,999	148,451	231,659
Vertices	500,534	74,338	116,250
OBJ File Size	81.2 MB	11.2 MB	17.9 MB
GLB File Size	56.6 MB	7.35 MB	8.65 MB
FBX File Size	21.2 MB	3.16 MB	5.41 MB

The comparative analysis of 3D model metrics reveals notable differences in complexity and file sizes among DSLR Images, Equirectangular Images, and Equirectangular Video.

Models generated from DSLR Images have a significantly higher number of faces (999,999) and vertices (500,534), reflecting greater detail but resulting in more complex models that can be challenging to process. In contrast, Equirectangular Images and Video produce models with fewer faces and vertices, with Equirectangular Video (231,659 faces, 116,250 vertices) providing more detail than Equirectangular Images (148,451 faces, 74,338 vertices). This reduced complexity makes these models easier to handle during post-processing.

File sizes also differ, with DSLR Images resulting in larger files (OBJ: 81.2 MB; GLB: 56.6 MB; FBX: 21.2 MB), which may be cumbersome in terms of storage and processing. Equirectangular Images produce smaller files (OBJ: 11.2 MB), enhancing efficiency in storage and processing. Equirectangular Video offers a balanced solution with moderate file sizes (OBJ: 17.9 MB), suitable for applications requiring both detail and manageable file sizes.

In summary, while DSLR Images provide the highest detail, their increased complexity can complicate post-processing. Equirectangular methods offer a more manageable workload, making them advantageous where efficiency and ease of handling are priorities.

Table 5.25: Parameter Description for Quantitative Comparison

Parameter Name	**Description**	**Value Meaning**
RMS Reprojection Error	Root Mean Square error of the reprojection, measured in pixels. Indicates the average deviation of projected points from the observed image points.	Optimal: < 0.1 pixel Good: 0.1 - 0.3 pixels Acceptable: 0.3 - 0.6 pixels Problematic: > 0.6 pixels
Max Reprojection Error	Maximum deviation of any projected point from the observed image points, measured in pixels.	Optimal: < 0.2 pixels Good: 0.2 - 0.5 pixels Acceptable: 0.5 - 1.0 pixels Problematic: > 1.0 pixels
Average Tie Point Multiplicity	Average number of image observations per tie point. This metric assesses the redundancy of tie points across images.	Optimal: > 8 Good: 5 - 8 Acceptable: 3 - 5 Problematic: < 3
Faces	Total number of faces in the 3D model. Faces are polygonal elements that define the surface of the model.	Excessive faces can complicate post-processing and may not be necessary. Too few faces can degrade the quality of the final result.
Vertices	Total number of vertices in the 3D model. Vertices are points where faces meet and contribute to the model's detail.	Excessive vertices can complicate post-processing and may not be necessary. Too few vertices can degrade the quality of the final result.
OBJ File Size	Size of the generated 3D model file in OBJ format. The OBJ file format is commonly used for exchanging 3D models between different applications.	Larger file sizes can indicate higher detail but may affect performance during loading and manipulation.
GLB File Size	Size of the generated 3D model file in GLB format. The GLB file format is optimized for web-based applications, providing efficient storage and streaming.	Larger file sizes can indicate higher detail but may affect performance in web-based applications.
FBX File Size	Size of the generated 3D model file in FBX format. The FBX file format is commonly used in Unity and similar game engines for storing 3D models with animations.	Larger file sizes can indicate higher detail and more complex animations but may affect performance in game engines.

Qualitative Comparison of Results

A qualitative evaluation is crucial for understanding the practical implications of each imaging technique's performance. This section provides a comparative analysis of the final outputs from the different methods, focusing on three key aspects: Mesh Quality, Texture Application, and Detailing Quality. Visual aids will be used to highlight the differences and to illustrate the quality of the results obtained from each technique.

- **Mesh Quality:** This section involves a visual inspection of the 3D meshes to evaluate their precision and completeness. The analysis of the meshes from different perspectives corroborates the findings of the quantitative assessment. Specifically, it demonstrates that the meshes obtained using the DSLR method exhibit significantly higher accuracy, while those generated from equirectangular video frames show a noticeable degradation in quality. However, as observed in the DSLR method images, this approach is not free from errors; notably, it produces large holes (blue areas in images d and g, see Fig. 5.22), unlike the equirectangular image-based photogrammetry.

(a) DSLR Images - View 1 **(b)** Equirect. Images - View 1 **(c)** Equirect. Video - View 1

(d) DSLR Images - View 2 **(e)** Equirect. Images - View 2 **(f)** Equirect. Video - View 2

(g) DSLR Images - View 3 **(h)** Equirect. Images - View 3 **(i)** Equirect. Video - View 3

(j) DSLR Images - View 4 **(k)** Equirect. Images - View 4 **(l)** Equirect. Video - View 4

Figure 5.22: Mesh quality comparison across different views and acquisition methods. (a)-(d)-(g)-(j) Meshes from DSLR images, (b)-(e)-(h)-(k) Meshes from equirectangular images, and (c)-(f)-(i)-(l) Meshes from equirectangular video. Each view provides a unique perspective to evaluate the mesh quality and detailing of the 3D reconstructions.

The use of equirectangular images offers notable advantages in mesh generation. These images provide a more consistent coverage of the scene, reducing the likelihood of missing data and resulting in a more uniform mesh. Moreover, the equirectangular images facilitate easier alignment and integration, leading to a smoother and more continuous mesh surface.

It is also important to note that the models have not yet undergone post-processing, which could substantially improve their quality. In terms of subsequent post-processing, the mesh produced from the equirectangular video appears too fragmented, whereas the mesh from equirectangular images is more uniform and amenable to further refinement. This inherent uniformity in the mesh from equirectangular images highlights their

effectiveness for detailed and high-quality 3D reconstruction tasks.

- **Texture Application:** As illustrated in Figure 5.23, the texture application quality varies notably across different methods. DSLR images provide superior texture quality, offering the most realistic results among tested methods. In contrast, textures derived from equirectangular video frames are considerably darker due to poor illumination conditions, affecting their visual fidelity. On the other hand, equirectangular images offer a commendable balance, providing well-lit and consistent textures that stand out as a superior compromise between detail and processing efficiency. Notably, using a subset of images from the Insta 360 Pro 2 significantly reduced processing times while mitigating distortions, positioning the equirectangular images as a highly effective and efficient method for texture application. This highlights their capability to deliver high-quality textures with practical processing demands.

(a) DSLR Images - View 1 **(b)** Equirect. Images - View 1 **(c)** Equirect. Video - View 1

(d) DSLR Images - View 2 **(e)** Equirect. Images - View 2 **(f)** Equirect. Video - View 2

(g) DSLR Images - View 3 **(h)** Equirect. Images - View 3 **(i)** Equirect. Video - View 3

(j) DSLR Images - View 4 **(k)** Equirect. Images - View 4 **(l)** Equirect. Video - View 4

Figure 5.23: Texture quality comparison across different views and acquisition methods. (a)-(d)-(g)-(j) Textures applied to meshes from DSLR images, (b)-(e)-(h)-(k) Textures applied to meshes from equirectangular images, and (c)-(f)-(i)-(l) Textures applied to meshes from equirectangular video. Each view provides a unique perspective to evaluate the texture quality and detailing of the 3D reconstructions.

- **Detailing Quality:** This section examines the quality of intricate details captured in the 3D models, highlighting the precision and accuracy of the final results. Thanks to the larger sensor aperture compared to the Insta 360 Pro 2, the DSLR method was able to capture and process a greater level of detail. For instance, in some images, distinct texts are clearly visible, such as the header of a badge (see Fig. 5.24a) and the titles of several books (see Fig. 5.24d). While the equirectangular images method also provided some level of detail, the equirectangular video method suffered from poor exposure, which limited the visibility of finer details. Therefore, if the objective is to achieve a highly detailed scene, the DSLR approach is preferable. Conversely, for a broader environment where minute details are less critical, the equirectangular images

method offers a good compromise.

(a) DSLR Images - Detail 1 **(b)** Equirect. Images - Detail 1 **(c)** Equirect. Video - Detail 1

(d) DSLR Images - Detail 2 **(e)** Equirect. Images - Detail 2 **(f)** Equirect. Video - Detail 2

(g) DSLR Images - Detail 3 **(h)** Equirect. Images - Detail 3 **(i)** Equirect. Video - Detail 3

(j) DSLR Images - Detail 4 **(k)** Equirect. Images - Detail 4 **(l)** Equirect. Video - Detail 4

Figure 5.24: Detailed view comparison of texture application across different acquisition methods, focusing on specific scene details. (a)-(d)-(g)-(j) Display textures applied to meshes from DSLR images, (b)-(e)-(h)-(k) Show textures applied to meshes from equirectangular images, and (c)-(f)-(i)-(l) Present textures applied to meshes from equirectangular video. Each detail view offers insight into how textures are applied across various imaging techniques.

5.6.7 Conclusions

The primary objective of this work was to test the GT Dundee Protocol approach for creating a 3D model that is not final but workable in post-processing, with a focus on speed. The results demonstrated that equirectangular image acquisition significantly reduces capture times while still producing an acceptable 3D model. This efficiency highlights the strengths of the equirectangular imaging approach, which allows for faster and more practical model generation compared to traditional methods.

The equirectangular imaging approach has shown particular promise due to its ability to

capture a full 360-degree view in a single shot, thus minimizing the time required for image acquisition and simplifying the process. Although the level of detail may not match that of high-resolution DSLR images, the reduced acquisition times and ease of use make it a valuable tool for scenarios where speed and convenience are crucial.

This work opens the door to new challenges and opportunities for further research, including:

- **Testing in Outdoor Environments:** Evaluate the performance of the GT Dundee Protocol in outdoor settings with varying lighting conditions. This will help determine the robustness of the approach under natural light variability and refine the method to handle diverse environmental influences.
- **Applying to Larger Environments:** Extend the approach to larger and more complex scenes to assess its scalability and effectiveness in managing increased data volumes and spatial complexity while maintaining efficiency and accuracy.
- **Exploring Different Equirectangular Imaging Devices:** Compare results from various equirectangular imaging devices to understand their impact on the quality of 3D models and determine if specific devices offer advantages in data acquisition and model fidelity.
- **Capturing Images from Different Levels:** Conduct equirectangular image captures at both tripod level and ground level to obtain diverse perspectives. This will allow for testing photogrammetric approaches with different subsets of photographs, potentially improving the depth and accuracy of 3D reconstruction.

Future research will be essential to address these challenges and enhance the applicability of the equirectangular imaging approach in diverse contexts.

5.7 Case Study 3: Comparative Analysis of Distance Estimation in Spherical Images and 3D Photogrammetry Models

5.7.1 Background

Extended Reality (XR), which encompasses Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), is a rapidly expanding field due to advancements in both hardware and software technologies [36]. These immersive environments allow users to interact with digital content in ways that were previously impossible. The growth of XR is driven by its broad range of applications, from entertainment and gaming to more practical uses in professional industries, education, and scientific research. The increasing availability of consumer-grade head-mounted displays (HMDs) and the growing interest in immersive experiences have further accelerated its adoption.

Exploring a virtual environment offers significant advantages in a variety of fields. In tourism, XR enables users to experience distant or inaccessible locations, providing a sense of presence without physical travel [35]. Similarly, XR has made hazardous environments accessible for training or research purposes, such as in the fields of industrial safety and emergency response [64]. In forensic investigations, virtual reconstructions of crime scenes allow for detailed examination without contaminating evidence, while cultural heritage applications facilitate the preservation and exploration of historical sites in a non-invasive manner [91]. These examples highlight how virtual environments enhance accessibility, safety, and educational opportunities.

A key element in creating realistic virtual environments is the accurate perception of depth and distance. Distance estimation plays a fundamental role in how users navigate and interact within a digital space [140]. Improving depth perception in XR systems significantly enhances the user's sense of immersion, providing a more intuitive and lifelike experience. This is particularly important in applications such as remote collaboration, architecture, and urban planning, where spatial awareness is crucial for effective decision-making.

Despite the widespread use of spherical images and photogrammetric models in XR, the efficacy of these representations for accurate distance estimation remains underexplored. This research seeks to address this gap by comparing four distinct environments: a spherical image, a spherical image with a novel distortion correction algorithm [49], a photogrammetric model obtained using a DSLR, and a photogrammetric model generated from equirectangular images. Each environment will feature standardized objects placed at varying distances, enabling a comparative analysis of distance perception.

5.7.2 State of the Art

Monocular and binocular cues play a crucial role in how individuals estimate distances in their environments. Monocular cues, such as size, texture gradient, and occlusion, allow for distance estimation using information from a single eye, while binocular cues rely on the disparity between the images perceived by both eyes [146]. Numerous studies have investigated these cues to understand their effectiveness in various contexts, including real-world [147] and virtual environments [148].

In virtual environments, the simulation of these cues is essential for enabling accurate distance estimation. Effective representation of monocular and binocular cues significantly contributes to the user's ability to navigate and interact within XR applications. Understanding how these cues are perceived in immersive settings is critical for designing experiences that feel intuitive and realistic.

Distance estimation techniques in XR environments have been explored [140] [49] [149], focusing on various methods for improving user interaction and spatial awareness [148]. Traditional approaches often rely on visual and auditory cues to assist users in accurately estimating distances. However, within the field of extended reality, there is a notable absence of studies comparing distance estimation between spherical environments and photogrammetric models.

Several methodologies for conducting distance estimation experiments have been developed to ensure reliable and valid results. These methodologies typically involve controlled environments where factors such as lighting, perspective, and motion are standardized [150]. Among the most commonly used techniques is the blind walking method, where participants are asked to view a target, close their eyes, and walk to the estimated location of the target. This method is considered highly effective for evaluating perceived distances because it minimizes visual feedback during the walking phase, allowing researchers to isolate the participant's internal distance estimation without the influence of continuous visual cues. The blind walking technique has been widely adopted due to its simplicity, ease of implementation, and ability to yield reliable data regarding spatial perception in both real-world and virtual environments [151].

Several studies have examined how users perceive visual distortions using various subjective methods. The Maximum Likelihood Difference Scaling (MLDS) ranks perceived differences between stimuli, but faces limitations when scales differ across content types [152]. Other methods, such as ACR-HR, DSIS, and SAMVIQ, are widely used to assess 2D image and video quality, and recent studies have applied them to 3D graphics in virtual environments, with DSIS and SAMVIQ proving more accurate [153]. Most research uses reference models to evaluate distortions, while this work explores whether distortions from the employed technology affect user experience without a reference model.

Moreover, guidelines for evaluating XR systems are essential for ensuring robust and accurate

results. Recent work has highlighted the need for standardized evaluation frameworks, especially for non-experts conducting usability studies in virtual reality. One study proposes comprehensive evaluation guidelines that assist researchers and non-experts in performing usability assessments in a structured and effective manner, ensuring that even novices in the field can carry out valid evaluations [154].

Evaluation Through Questionnaires

The literature offers a wide range of questionnaires designed to evaluate the overall user experience when interacting with XR technologies. These questionnaires serve as essential tools for assessing various aspects of the user's perception, immersion, and comfort within virtual environments. A key factor in the choice of a particular questionnaire is its level of standardization. Standardized questionnaires are often preferred due to their reliability, ease of comparison across studies, and established validity. However, the use of innovative and non-standardized questionnaires should not be dismissed, as they may offer valuable insights into novel aspects of user experience that standardized tools might overlook.

Examples of widely adopted, standardized questionnaires include the Igroup Presence Questionnaire (IPQ) [155], which assesses the user's sense of presence within a virtual environment, and the Simulator Sickness Questionnaire (SSQ) [156] [157], which is used to evaluate symptoms of discomfort, such as nausea or dizziness, that may arise during immersive experiences. On the other hand, more innovative questionnaires, such as the Depth Perception Questionnaire (DPQ) [158] and the Image-Lighting Visual Realism Questionnaire (ILVRQ) [140], offer insights into specific aspects of XR systems that are not always addressed by traditional tools. These innovative questionnaires, although not yet fully standardized, contribute to a more nuanced understanding of user interactions with immersive technologies and may prove beneficial in evaluating emerging XR applications.

Moreover, the Visual Distortion Questionnaire (VDQ) has been developed specifically to assess whether distortions introduced by acquisition tools alter the user's experience. Unlike traditional questionnaires that compare user distortion perception against a reference model, this tool focuses on identifying how such distortions may impact user interaction and immersion within immersive environments, providing a more targeted evaluation of these effects.

5.7.3 Motivation and Objectives

The rapid evolution and widespread adoption of 360° multimedia acquisition technologies raise several key questions regarding their effectiveness in delivering immersive experiences. Specifically, this study seeks to answer the following research questions:

- Can the navigation experience within a spherical environment be improved by applying distortion correction algorithms?

- Is the overall user experience in a spherical environment comparable to the experience within a fully reconstructed 3D environment?
- Is the experience in a photogrammetric model created using DSLR-acquired images comparable to that of a model created using equirectangular images?

To address these questions, a series of experiments focused on depth perception and distance estimation will be conducted in a virtual indoor environment. The virtual scene is created using four distinct conditions: a spherical image, a spherical image with the distortion correction algorithm, a 3D model reconstructed from DSLR-acquired images, and a 3D model created from equirectangular images. It is noted that spherical images exhibit significant distortions when users navigate within the scene; however, efforts will be made to enhance the results by applying correction algorithms [49]. Conversely, a 3D model constructed from equirectangular images is generally less accurate compared to a traditional photogrammetric model but allows for faster realization, as demonstrated in Case Study 2. By comparing user experiences across these environments, the study seeks to provide insights into the relative effectiveness of each method in delivering a realistic and immersive XR experience.

The results of this investigation are expected to contribute to the ongoing development of XR technologies, which are increasingly being applied across various sectors. Should the study demonstrate that 360° acquisition technologies provide an experience comparable to more resource-intensive methods, it could support the broader use of these tools due to their cost-effectiveness, ease of use, and rapid data acquisition capabilities. Furthermore, the outcomes of this experiment may pave the way for photogrammetry using spherical acquisition devices in hostile environments or situations where extensive time for image capture is impractical, such as underwater exploration and archaeology, traffic accident documentation, and locations with restrictions.

5.7.4 Tools and Software Utilized

This section outlines the essential tools and software utilized in the experiments conducted for this study, highlighting their roles in facilitating the evaluation of user experience within immersive virtual environments. Each tool has been selected or developed based on its capabilities to support the experimental design and enhance the overall quality of data collected.

Distance Estimation XR (developed in Unity)

The primary software employed in this research is a custom application developed in Unity. This versatile platform offers essential functionalities to conduct experiments effectively, facilitating the integration of various media formats and interactive elements crucial for evaluating user experience in immersive environments.

ASUS TUF Dash F15 Laptop

The ASUS TUF Dash F15 is a high-performance laptop designed for demanding computational tasks, making it ideal for photogrammetric processing. Equipped with an NVIDIA GeForce RTX 3050 dedicated graphics card, it provides exceptional parallel processing power essential for complex photogrammetry algorithms. Its 12th Gen Intel Core i5-12459H CPU and 32GB of RAM ensure efficient handling of large datasets and intensive tasks. The high-speed NVMe SSD storage further enhances data read/write operations, reducing processing time. These hardware capabilities allow the ASUS TUF Dash F15 to efficiently run photogrammetric software, such as Agisoft Metashape and RealityCapture, delivering high accuracy and fast processing speeds.

Table 5.26: ASUS TUF Dash F15 Specifications

Laptop Specifications	
Feature	**Specification**
Model	ASUS TUF Dash F15
Processor	12th Gen Intel(R) Core(TM) i5-12450H 2.00 GHz
Memory	32GB RAM SODIMM 4800 MT/s
Operating System	Windows 11 Home 64bit
Graphics	NVIDIA GeForce RTX 3050 Laptop GPU 4GB
Storage	NVMe HD SSD 512GB

Meta Quest 2

The Meta Quest 2 is a high-performance head-mounted display (HMD) designed for immersive virtual reality (VR) experiences. Featuring a fast-switch LCD panel with a resolution of 1832 x 1932 per eye, it delivers sharp visuals. Its inside-out tracking capability provides 6 degrees of freedom (6DOF), allowing natural movement within virtual environments. The integrated audio system enhances immersion, while the Qualcomm Snapdragon XR2 system on chip (SoC) ensures robust performance across various applications. The device is available in two storage options, catering to different user needs.

Table 5.27: Meta Quest 2 Specifications

Meta Quest 2 Specifications	
Feature	**Specification**
Panel Type	Fast-switch LCD
Display Resolution (per eye)	1832 x 1932
Refresh Rate	72 Hz, 90 Hz, 120 Hz (experimental)
Connectivity	USB-C, Bluetooth 5, Wi-Fi 6, 3.5 mm headphone jack
Tracking	Inside out, 6DOF
Audio	Integrated, in-strap
SoC	Qualcomm Snapdragon XR2
Memory	6 GB
Storage Options	128 GB, 256 GB
Lens Distance	58 mm, 63 mm, 68 mm
Weight	503 g
Operating System	Proprietary, Android-based

Questionnaires

To assess user experiences, multiple questionnaires will be administered. Specifically, the Igroup Presence Questionnaire (IPQ), the Simulator Sickness Questionnaire (SSQ), a Depth Perception Questionnaire (DPQ) obtained by combining the proposals of [158] and [140], and the Image-Lighting Visual Realism Questionnaire (ILVRQ) [140] will be used to evaluate different aspects of immersive experiences. A detailed description of each questionnaire, can be found in Appendix B.

Table 5.28: Questionnaires Overview

Questionnaire	Purpose	Key Focus Areas
Igroup Presence Questionnaire (IPQ)	Measures the sense of presence in virtual environments	Spatial presence, involvement, and realism
Simulator Sickness Questionnaire (SSQ)	Assesses symptoms of motion sickness in simulation	Nausea, oculomotor issues, and disorientation
Depth Perception Questionnaire (DPQ)	Evaluates users' ability to perceive depth in VR	Spatial awareness, distance judgment, and depth cues
Image-Lighting Visual Realism Questionnaire (ILVRQ)	Assesses the perception of visual realism based on image-lighting factors	Pixel density, lightness, color accuracy, contrast, and photorealism
Visual Distortion Questionnaire (VDQ)	Evaluates the impact of distortions introduced by acquisition tools on user experience	Interaction quality, immersion, and perceptual alterations

5.7.5 Methodology

The methodology for this usability evaluation adopts an approach in alignment with the Evaluation Guidelines proposed by Livatino and Koeffel [154]. This structured framework ensures a comprehensive assessment of user interactions within a virtual reality (VR) environment. The evaluation consists of several phases designed to prepare participants effectively and gather detailed data regarding their experience.

The evaluation process is segmented into distinct stages to facilitate a thorough understanding of participants' perceptions and interactions with the VR technology.

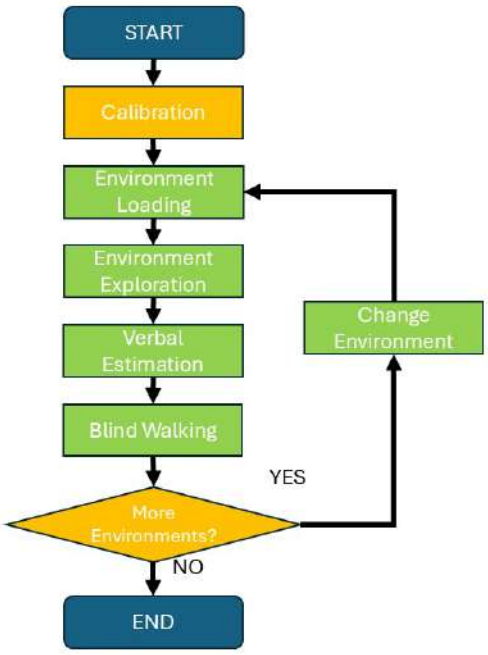


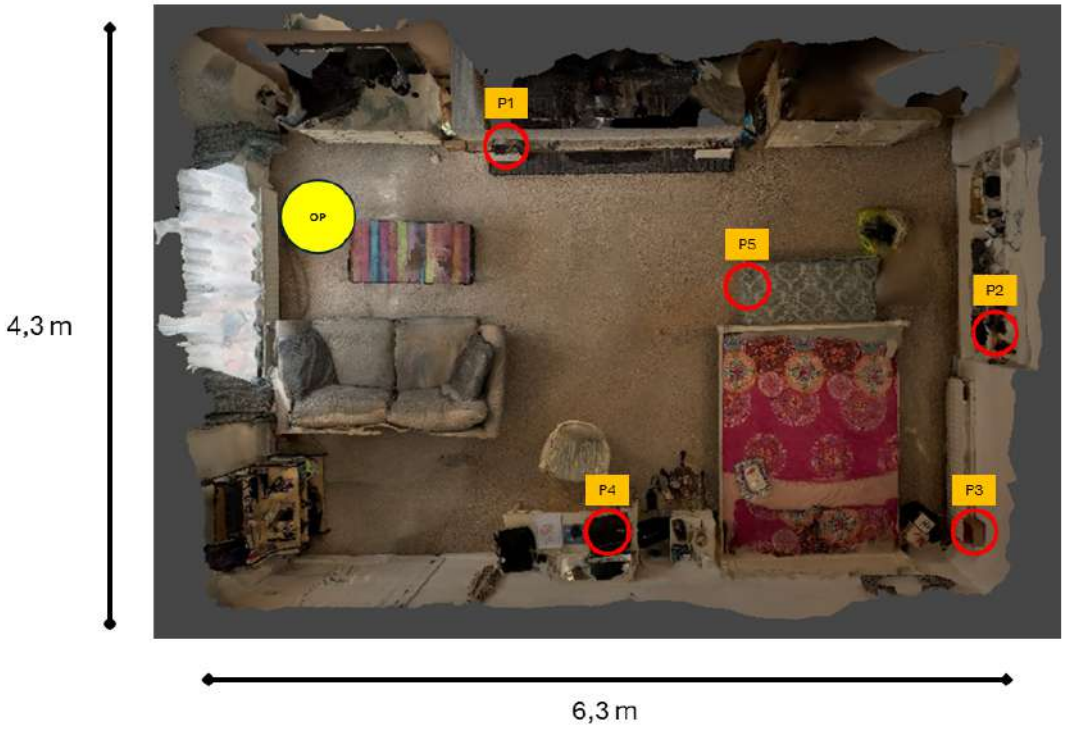
Figure 5.25: Workflow of the usability evaluation experiment, detailing the phases from participant onboarding to environment exploration and distance estimation tasks. Each phase is structured to ensure participant understanding and engagement, facilitating accurate assessment of distance perception in virtual environments.

Real Environment and Observation Points

The environment that will be used for this experiment is the same bedroom utilized in Case Study 2, approximated as a rectangular prism with dimensions of 630 cm in length, 430 cm in width, and 245 cm in height. The room includes a large window on one wall, an entrance door, and standard bedroom furniture.

To ensure that the two observations do not influence each other, two distinct observation points have been selected within the room. In the verbal distance estimation experiment, participants will view the observation points shown in Fig 5.26a, attempting to estimate both egocentric and relative distances, as detailed in Tab. 5.29.

(a) Photo taken from the primary observation point.



(b) Plan view with measurements and distances from the primary observation point.

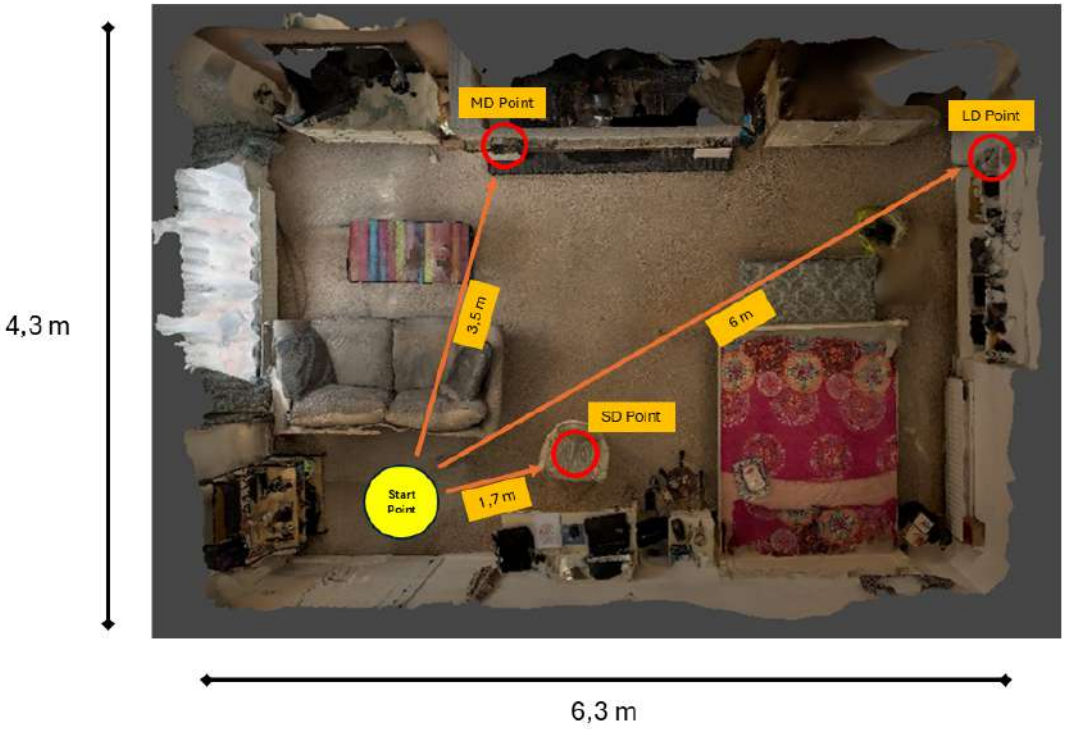
Figure 5.26: Images from the primary observation point.

Table 5.29: Distances Calculated for the Verbal Distance Estimation

Measure	Description	Distance	Type
d(OP, P1)	Distance between OP and ukulele (P1)	1.8 meters	Egocentric
d(OP, P2)	Distance between OP and yellow book (P2)	5.9 meters	Egocentric
d(OP, P3)	Distance between OP and orange lamp (P3)	6.2 meters	Egocentric
d(OP, P4)	Distance between OP and laptop (P4)	1.8 meters	Egocentric
d(OP, P5)	Distance between OP and bedside cabinet (P5)	1.8 meters	Egocentric
d(P2, P3)	Distance between yellow book (P2) and orange lamp (P3)	1.5 meters	Relative
d(P1, P4)	Distance between ukulele (P1) and laptop (P4)	3.3 meters	Relative
d(P3, P4)	Distance between orange lamp (P3) and laptop (P4)	1.8 meters	Relative

For the blind walking distance estimation experiment, the observation and starting point was chosen at the Start Point near the entrance door (see Fig. 5.27a), in order to fully utilize the diagonal length of the room. The selected points cover three distance ranges: short, medium, and long, as detailed in Fig. 5.27b and Tab. 5.30.

(a) Photo taken from the Start Point used for blind walking.



(b) Plan view with measurements and distances from the Start Point for blind walking.

Figure 5.27: Images from the Start Point used for blind walking.

Table 5.30: Observation Points and Distances

Point	Description	Distance
Start Point	Entrance of the room and the starting point for all experiments	0.0 meters
SD Point	A large chair positioned in front of the desk	1.7 meters
MD Point	A small ukulele placed above the fireplace	3.5 meters
LD Point	A white book located on a shelf	6.0 meters

Introduction of the Experiment

Initially, participants are presented with a description of the usability evaluation and are required to sign a consent form. Following this, they complete a demographic questionnaire to collect baseline information. This introductory phase is designed to ensure that participants are informed and agree to the terms of their involvement.

Practice

In the practice phase, participants are given time to acclimatize to the virtual indoor environment. They are encouraged to explore the space and adjust to their new surroundings. The functionality of the VR controllers is explained in detail, ensuring participants understand how to utilize the joypads for interaction within the virtual environment.

For participants new to VR, a brief overview of the technology is provided to facilitate comfort and familiarity. Participants will also have the opportunity to test their distance perception skills both with and without the head-mounted display (HMD). They will estimate distances to specific points in the lab, utilizing both verbal estimations and the blind walking method under supervision in an open and unobstructed testing space.

Once participants feel confident navigating the virtual setting and utilizing the controllers, the researcher will guide them into the operational phases of the experiment.

Operation

At the beginning of the operational phase, participants are informed they will be immersed in a virtual indoor environment that they have not previously encountered. For the verbal distance estimation experiments, participants will be positioned at the Observation Point (see Fig. 5.26b). For the distance estimation experiments using the blind walking technique, participants will be positioned at the Start Point, as shown in Fig. 5.27b. This phase is structured to ensure participants are fully aware of each step before progressing.

(a) VR Lab setting 1 **(b)** VR Lab setting 2

Figure 5.28: Experimental setup area for XR environment testing, with designated boundaries marked within the laboratory to ensure controlled spatial parameters for immersive simulations.

- **Phase 0: VR Headset Calibration**
  The VR headset is calibrated to guarantee accurate visual rendering and alignment with the participant's field of view, ensuring optimal comfort and performance.
- **Phase 1: Scene Loading**
  The initial scene is loaded, with options including a spherical image, a photogrammetric model captured using a DSLR camera, and a photogrammetric model derived from equirectangular images. Participants will be immersed in one of these virtual environments during the experiment.
- **Phase 2: Environment Exploration**
  Participants are guided to walk within the virtual environment while keeping their eyes open. This phase allows them to experience the immersive qualities of the virtual space, encouraging natural movement and exploration to enhance their sense of presence and comfort.
- **Phase 3: Verbal Distance Estimation**
  Participants are asked to verbally estimate the distances indicated by the researcher, which correspond to the pre-measured distances listed in Tab. 5.29. The researcher records the participants' verbal estimates, and these are later compared to the actual distances to assess the accuracy of their distance perception within the virtual environment.
- **Phase 4: Blind Walking Test**
  This phase evaluates the participants' ability to estimate distances and navigate towards a point without visual aid through the blind walking method.
  - **Phase 4.1: Short Distance (1,7 meters)**
    Participants observe the SD Point for 5 seconds. The HMD screen is then turned off, and participants are instructed to walk towards the remembered point. The researcher supervises the movement, measuring the actual distance traveled.

- **Phase 4.2: Medium Distance (3,5 meters)**
  The same procedure is followed for the MD Point, with participants observing the point for 5 seconds before attempting to walk towards it. Actual distances traveled are recorded.
- **Phase 4.3: Long Distance (6 meters)**
  This process is repeated for the LD Point, following the same observational and navigational procedure.

- **Phase 5: End of Trial**
  Upon completing the tasks for the current virtual environment, the first experiment trial concludes, followed by a short break before changing environments or proceeding with questionnaires.

The operation phase will be repeated for fourth different environments, maintaining consistency in the steps outlined. Upon concluding Phase 5 for the first environment and allowing a break, participants will proceed to the second trial through phases 1 to 5. This process will be repeated for the third environment, concluding the operational stage.

5.7.6 Results Discussions

The expected results from this study will provide several avenues for comparison and evaluation. First, the experiment is designed to assess whether the distortion correction algorithm proposed in [49] effectively reduces the perceived distortion experienced by users. This improvement is anticipated to enhance navigation within spherical environments by minimizing visual artifacts, thereby contributing to a more immersive experience.

Additionally, the study will compare the overall user experience of spherical images with that of traditional photogrammetry models. This comparative analysis aims to highlight the strengths and weaknesses of both methods in terms of realism, user comfort, and perceived quality.

A subjective assessment of the phenomena described in Case Study 2 will also be conducted. This evaluation is expected to yield insights into participants' perceptions and how distortions impact the overall experience.

While these experiments are currently ongoing during the writing of this manuscript, final results cannot be presented at this time. However, positive outcomes from this experiment are expected to open new research opportunities regarding the use of equirectangular images and their application in creating 3D scenarios. These findings could pave the way for further studies on the effectiveness and applicability of these techniques across various XR applications.

5.7.7 Conclusions

The anticipated findings from this experiment indicate that the distortion correction algorithm proposed in [49] is likely to substantially enhance user experience within spherical environments by mitigating perceived visual distortions. The comparative analysis between spherical imagery and traditional photogrammetry is expected to yield valuable insights into the optimization of these techniques across various immersive environments.

The subjective evaluation conducted in Case Study 2 is anticipated to provide a comprehensive understanding of the impact of distortions on user perceptions and interactions. Given that the experiments are still in progress, conclusive results cannot be presented at this juncture. Should the findings corroborate the anticipated enhancements, future research may focus on further refining the utilization of equirectangular images and investigating novel methodologies for improving 3D reconstructions via spherical images.

In summary, this study establishes a foundation for subsequent investigations in XR environments, wherein the enhancement of user experience through effective distortion management and the exploitation of alternative imaging techniques could facilitate more immersive and interactive virtual experiences.

5.8 Conclusions

The research presented in this chapter has demonstrated several key insights into the use of XR and photogrammetry techniques in immersive environments. In Case Study 1, the Proof of Concept (PoC) using IoT and XR technologies to enhance user experience showed that integrating real-time IoT data with XR systems is feasible and beneficial. The system demonstrated solid technical performance, though latency and interaction responsiveness emerged as critical challenges. Future efforts should focus on reducing latency and increasing the XR immersion quality, particularly for larger user bases and more complex scenarios.

In Case Study 2, the efficient 3D reconstruction of indoor environments using equirectangular images showcased the potential for reducing acquisition times while maintaining high-quality outputs. The innovative standard operating procedures applied in this case resulted in a significant reduction of processing time, positioning equirectangular images as a valuable tool in environments where speed and quality are paramount, such as in crime scene documentation or archaeological surveys.

For Case Study 3, which focuses on a comparative analysis of distance estimation in spherical images and photogrammetric models, the experiments are still ongoing at the time of this writing. However, it is anticipated that the distortion correction algorithm proposed will reduce perceived visual distortions, improving user navigation in spherical environments. Comparing user experiences across spherical imagery and traditional photogrammetry will provide valuable insights into the strengths and limitations of these methods. This research may pave the way for further exploration of equirectangular image applications in immersive 3D scenarios, potentially extending their use to fields like underwater exploration and traffic accident documentation.

Overall, the findings from these studies contribute to the growing body of research in XR and photogrammetry by introducing innovative methodologies for enhancing user experience, optimizing 3D reconstruction from 360-degree images, and addressing practical challenges in data acquisition. The results suggest that future work should focus on refining these methods to enhance immersion, reduce visual distortion effects, and improve overall interaction within immersive environments. This research highlights the potential of equirectangular imaging as a viable and efficient approach for realistic virtual environment creation.

Conclusion

The research presented in this thesis explores the integration of IoT, XR, and Photogrammetry to enhance Quality of Experience (QoE) in immersive environments, proposing innovative methodologies across three case studies that address both technical and experiential aspects of user interaction within XR platforms.

In **Case Study 1**, an innovative method was developed to present proof-of-concept (PoC) projects within a company setting, serving as a foundational step in learning the basics of photogrammetry techniques. This approach demonstrated how IoT and XR integration can facilitate real-time interaction between physical and virtual components to improve user immersion. Nonetheless, challenges related to latency and system responsiveness, particularly in larger-scale applications, were identified as areas for improvement.

Case Study 2 contributed significantly to the scientific understanding of indoor environment reconstruction through photogrammetry. This study employed equirectangular images within a standardized operating procedure, an approach that, to the best of current knowledge, is unprecedented in the literature. The results showcased that spherical imaging methods can markedly reduce both acquisition and processing time while maintaining a high level of detail in the 3D model, proving advantageous for applications needing rapid yet precise documentation, such as crime scene reconstruction or architectural documentation.

Case Study 3 builds on the findings of Case Study 2, evaluating the practical applicability of spherical images in photogrammetric reconstructions. This study focuses on user perception in distance estimation when comparing spherical images to traditional photogrammetric models. Preliminary results suggest that the application of a distortion correction algorithm improves navigability and realism in spherical environments, further supporting the potential of spherical images in immersive settings.

In conclusion, this thesis establishes a foundational framework for further research on alternative and innovative methods for reality virtualization. The findings underscore the potential for continued development of IoT, Photogrammetry, and XR technologies to enable more interactive, realistic, and efficient immersive experiences, particularly in domains such as education, forensic science, and virtual tourism.

Appendix A

Case Study 2 Metashape Reports

A.1 DSLR Image Processing Report

General

Cameras	478
Aligned Cameras	356
Markers	19
Coordinate System	Local Coordinates (m)
Rotation Angles	Yaw, Pitch, Roll

Point Cloud

Points	130,942 of 237,134
RMS Reprojection Error	0.143424 (0.81485 pix)
Max Reprojection Error	0.443242 (35.6196 pix)
Mean Key Point Size	4.46045 pix
Point Colors	3 bands, uint8
Key Points	No
Average Tie Point Multiplicity	2.63285

Alignment Parameters

Accuracy	High
Generic Preselection	Yes
Reference Preselection	Sequential
Key Point Limit	40,000
Key Point Limit per Mpx	1,000
Tie Point Limit	4,000
Exclude Stationary Tie Points	Yes
Guided Image Matching	No
Adaptive Camera Model Fitting	No
Matching Time	40 seconds
Matching Memory Usage	414.19 MB
Alignment Time	1 minute 3 seconds
Alignment Memory Usage	96.35 MB
Date Created	2024:08:29 17:25:24
Software Version	1.8.5.14752
File Size	13.26 MB

Depth Maps

Count	341
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	4 minutes 22 seconds
Memory Usage	1.24 GB
Date Created	2024:08:29 17:33:32
Software Version	1.8.5.14752
File Size	212.77 MB

Dense Point Cloud

Points	22,320,959
Point Colors	3 bands, uint8
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	4 minutes 22 seconds
Memory Usage	1.24 GB
Processing Time	5 minutes 32 seconds
Memory Usage	5.41 GB
Date Created	2024:08:29 17:39:04
Software Version	1.8.5.14752
File Size	294.13 MB

Model

Faces	999,999
Vertices	500,534
Vertex Colors	3 bands, uint8
Texture	4,092 x 4,092 x 24, 4 bands, uint8
Surface Type	Arbitrary
Source Data	Depth Maps
Interpolation	Enabled
Strict Volumetric Masks	No
Processing Time	2 minutes 10 seconds
Memory Usage	2.46 GB
Mapping Mode	Generic
Blending Mode	Mosaic
Texture Size	4,092
Enable Hole Filling	Yes
Enable Ghosting Filter	Yes
UV Mapping Time	2 minutes 11 seconds
UV Mapping Memory Usage	865.91 MB
Blending Time	8 minutes 9 seconds
Blending Memory Usage	23.38 GB
Date Created	2024:08:29 17:45:37
Software Version	1.8.5.14752
File Size	451.48 MB

System

Software Name	Agisoft Metashape Professional
Software Version	1.8.5 build 14752
OS	Windows 64 bit
RAM	31.63 GB
CPU	12th Gen Intel(R) Core(TM) i5-12450H
GPU(s)	NVIDIA GeForce RTX 3050 Laptop GPU

A.2 Equirectangular Images Processing Report

General

Cameras	39
Aligned Cameras	39
Markers	7
Coordinate System	Local Coordinates (m)
Rotation Angles	Yaw, Pitch, Roll

Point Cloud

Points	10,337 of 88,551
RMS Reprojection Error	0.222384 (0.295767 pix)
Max Reprojection Error	0.828279 (1.56548 pix)
Mean Key Point Size	1.28943 pix
Point Colors	3 bands, uint8
Key Points	No
Average Tie Point Multiplicity	2.80933

Alignment Parameters

Accuracy	Highest
Generic Preselection	Yes
Reference Preselection	Sequential
Key Point Limit	80,000
Key Point Limit per Mpx	1,000
Tie Point Limit	0
Exclude Stationary Tie Points	Yes
Guided Image Matching	No
Adaptive Camera Model Fitting	No
Matching Time	2 minutes 20 seconds
Matching Memory Usage	628.55 MB
Alignment Time	4 minutes 5 seconds
Alignment Memory Usage	151.61 MB
Date Created	2024:09:06 08:21:14
Software Version	1.8.5.14752
File Size	4.29 MB

Depth Maps

Count	37
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	19 seconds
Memory Usage	342.51 MB
Date Created	2024:09:09 10:10:20
Software Version	1.8.5.14752
File Size	21.94 MB

Dense Point Cloud

Points	315,584
Point Colors	3 bands, uint8
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	20 seconds
Memory Usage	428.19 MB
Dense Cloud Generation Processing Time	38 seconds
Memory Usage	1.19 GB
Date Created	2024:09:09 09:51:28
Software Version	1.8.5.14752
File Size	16.96 MB

Model

Faces	148,451
Vertices	74,338
Vertex Colors	3 bands, uint8
Texture	4,096 x 4,096 x 4, 4 bands, uint8
Surface Type	Arbitrary
Source Data	Depth Maps
Interpolation	Enabled
Processing Time	10 seconds
Memory Usage	436.00 MB
Texturing Parameters Mapping Mode	Generic
Blending Mode	Mosaic
Texture Size	4,096
Enable Hole Filling	Yes
Enable Ghosting Filter	Yes
UV Mapping Time	28 seconds
UV Mapping Memory Usage	161.58 MB
Blending Time	20 seconds
Blending Memory Usage	1.25 GB
Blending GPU Memory Usage	1.54 GB
Date Created	2024:09:09 10:05:19
Software Version	1.8.5.14752
File Size	44.89 MB

System

Software Name	Agisoft Metashape Professional
Software Version	1.8.5 build 14752
OS	Windows 64 bit
RAM	31.63 GB
CPU	12th Gen Intel(R) Core(TM) i5-12450H
GPU(s)	NVIDIA GeForce RTX 3050 Laptop GPU

A.3 Equirectangular Video Processing Report

General

Cameras	300
Aligned Cameras	300
Coordinate System	Local Coordinates (m)
Rotation Angles	Yaw, Pitch, Roll

Point Cloud

Points	28,326 of 472,342
RMS Reprojection Error	0.492285 (0.852767 pix)
Max Reprojection Error	1.29742 (3.90654 pix)
Mean Key Point Size	1.67439 pix
Point Colors	3 bands, uint8
Key Points	No
Average Tie Point Multiplicity	3.29791

Alignment Parameters

Accuracy	Highest
Generic Preselection	Yes
Reference Preselection	Sequential
Key Point Limit	80,000
Key Point Limit per Mpx	1,000
Tie Point Limit	0
Exclude Stationary Tie Points	Yes
Guided Image Matching	No
Adaptive Camera Model Fitting	No
Matching Time	10 minutes 8 seconds
Matching Memory Usage	802.11 MB
Alignment Time	4 minutes 3 seconds
Alignment Memory Usage	243.88 MB

Optimization Parameters

Adaptive Camera Model Fitting	No
Optimization Time	0 seconds
Date Created	2024:09:06 16:00:19
Software Version	1.8.5.14752
File Size	25.81 MB

Depth Maps

Count	210
Depth Maps Generation Parameters	
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	2 minutes 18 seconds
Memory Usage	512.88 MB
Date Created	2024:09:06 17:31:11
Software Version	1.8.5.14752
File Size	76.66 MB

Dense Point Cloud

Points	858,609
Point Colors	3 bands, uint8
Depth Maps Generation Parameters	
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	2 minutes 18 seconds
Memory Usage	512.88 MB
Dense Cloud Generation Parameters	
Processing Time	35 minutes 0 seconds
Memory Usage	2.39 GB
Date Created	2024:09:10 08:47:09
Software Version	1.8.5.14752
File Size	28.69 MB

Model

Faces	231,659
Vertices	116,250
Vertex Colors	3 bands, uint8
Texture	4,096 x 4,096 x 4, 4 bands, uint8
Depth Maps Generation Parameters	
Quality	High
Filtering Mode	Mild
Max Neighbors	16
Processing Time	2 minutes 18 seconds
Memory Usage	512.88 MB
Reconstruction Parameters	
Surface Type	Arbitrary
Source Data	Depth Maps
Interpolation	Enabled
Strict Volumetric Masks	No
Processing Time	21 seconds
Memory Usage	674.59 MB
Texturing Parameters	
Mapping Mode	Generic
Blending Mode	Mosaic
Texture Size	4,096
Enable Hole Filling	Yes
Enable Ghosting Filter	Yes
UV Mapping Time	48 seconds
UV Mapping Memory Usage	210.57 MB
Blending Time	1 minute 15 seconds
Blending Memory Usage	2.35 GB
Blending GPU Memory Usage	1.54 GB
Date Created	2024:09:10 09:07:41
Software Version	1.8.5.14752
File Size	50.86 MB

System

Software Name	Agisoft Metashape Professional
Software Version	1.8.5 build 14752
OS	Windows 64 bit
RAM	31.63 GB
CPU	12th Gen Intel(R) Core(TM) i5-12450H
GPU(s)	NVIDIA GeForce RTX 3050 Laptop GPU

Appendix B

Case Study 3 Questionnaires

B.1 Igroup Presence Questionnaire (IPQ) [159]

Table B.1: Igroup Presence Questionnaire (IPQ) Items

No.	IPQ Item Name	Questions	Values Range (1-7)
1	G1	In the computer generated world I had a sense of ”being there”	not at all–very much
2	SP1	Somehow I felt that the virtual world surrounded me.	fully disagree–fully agree
3	SP2	I felt like I was just perceiving pictures.	fully disagree–fully agree
4	SP3	I did not feel present in the virtual space.	did not feel–felt present
5	SP4	I had a sense of acting in the virtual space, rather than operating something from outside.	fully disagree–fully agree
6	SP5	I felt present in the virtual space.	fully disagree–fully agree
7	INV1	How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?	extremely aware-moderately aware-not aware at all
8	INV2	I was not aware of my real environment.	fully disagree–fully agree
9	INV3	I still paid attention to the real environment.	fully disagree–fully agree
10	INV4	I was completely captivated by the virtual world.	fully disagree–fully agree
11	REAL1	How real did the virtual world seem to you?	completely real–not real at all
12	REAL2	How much did your experience in the virtual environment seem consistent with your real world experience?	not consistent-moderately consistent-very consistent
13	REAL3	How real did the virtual world seem to you?	about as real as an imagined world–indistinguishable from the real world
14	REAL4	The virtual world seemed more realistic than the real world.	fully disagree–fully agree

B.2 Simulator Sickness Questionnaire (SSQ) [160]

Table B.2: Simulator Sickness Questionnaire (SSQ)

No.	Symptoms	Values
1	General Discomfort	None = 0, Slight = 1, Moderate = 2, Severe = 3
2	Fatigue	None = 0, Slight = 1, Moderate = 2, Severe = 3
3	Headache	None = 0, Slight = 1, Moderate = 2, Severe = 3
4	Eye strain	None = 0, Slight = 1, Moderate = 2, Severe = 3
5	Difficulty focusing	None = 0, Slight = 1, Moderate = 2, Severe = 3
6	Increased salivation	None = 0, Slight = 1, Moderate = 2, Severe = 3
7	Sweating	None = 0, Slight = 1, Moderate = 2, Severe = 3
8	Nausea	None = 0, Slight = 1, Moderate = 2, Severe = 3
9	Difficulty concentrating	None = 0, Slight = 1, Moderate = 2, Severe = 3
10	Fullness of head	None = 0, Slight = 1, Moderate = 2, Severe = 3
11	Blurred vision	None = 0, Slight = 1, Moderate = 2, Severe = 3
12	Dizzy (eyes open)	None = 0, Slight = 1, Moderate = 2, Severe = 3
13	Dizzy (eyes closed)	None = 0, Slight = 1, Moderate = 2, Severe = 3
14	Vertigo	None = 0, Slight = 1, Moderate = 2, Severe = 3
15	Stomach awareness	None = 0, Slight = 1, Moderate = 2, Severe = 3
16	Burping	None = 0, Slight = 1, Moderate = 2, Severe = 3

B.3 Depth Perception Questionnaire (DPQ) [158] [140]

Table B.3: Depth Perception Questionnaire (DPQ)

No.	Questions	Values (1-7)
1	How much would you score the overall sense of depth impression (i.e. delivered sense of three dimensional appearance)?	1 = Very Poor, 7 = Excellent
2	How much does the tested system help with positions and distance judgements?	1 = Not Helpful, 7 = Extremely Helpful
3	How much does the tested system help with spatial understanding?	1 = Very Poor, 7 = Excellent
4	How much does the tested system help with finding/identifying/classifying symbols or objects?	1 = Very Difficult, 7 = Very Easy
5	How much does the tested system help with memory & recall?	1 = Not Helpful, 7 = Extremely Helpful
6	How do you believe the tested system helps with learning/training/planning or prediction?	1 = Not Helpful, 7 = Extremely Helpful
7	How quickly did you acquire a 3D depth impression?	1 = Very Slow, 7 = Very Fast
8	How much did the depth perception contribute to the sense of realism?	1 = Very Little, 7 = Very Much
9	How much did the lights, shadows, and colors contribute to the perception of depth?	1 = Very Little, 7 = Very Much

B.4 Image-Lighting Visual Realism Questionnaire (ILVRQ)

Table B.4: Visual Realism Questionnaire on Image-Lighting Elements

No.	Questions	Values (1-7)
Q1	How relevant do you find pixel density to the perception of visual realism?	irrelevant–fully relevant
Q2	How much does the lightness of the scene affect the realism of the visual representation?	not at all–very much
Q3	How important is the accuracy of color in influencing visual realism?	not important–very important
Q4	To what extent does contrast impact the perception of realism in the image?	no impact–strong impact
Q5	How does the vividness of the image contribute to its perceived realism?	not at all–very much
Q6	How significant is sharpness for the overall visual realism?	insignificant–very significant
Q7	How relevant is the definition of objects in the scene for realistic perception?	irrelevant–fully relevant
Q8	How would you rate the general level of photorealism in the image?	very low–very high
Q9	How similar is your experience to viewing a real photograph or video?	not similar–very similar

B.5 Visual Distortion Questionnaire (VDQ)

Table B.5: Visual Distortion Questionnaire

No.	Questions	Values
Q1	How noticeable were the visual distortions in the environment?	not noticeable at all–very noticeable
Q2	Which visual elements or features in the environment did you perceive as distorted?	❑ lights ❑ colours ❑ shapes ❑ shadows ❑ none
Q3	How would you rate the level of realism in terms of object or natural element distortion in the environment you observed?	not realistic at all–very realistic
Q4	How would you rate the overall quality of the images seen through the VR display, considering image definition?	very low quality–very high quality
Q5	How sharp (not blurred, focused) did the images appear on the VR display you used to observe the environment?	not sharp at all–very sharp
Q6	Did you notice pixel edges (i.e., areas of the image that looked like a mosaic) while using the VR display to observe the environment?	not at all–very noticeable
Q7	How much do you think the pixel edges affected the realism of the environment you viewed?	no effect at all–strong effect
Q8	To what extent did the distortions interfere with your ability to complete tasks?	no interference at all–strong interference
Q9	To what extent do you think visual distortions affected your overall experience?	no effect at all–strong effect

Bibliography

[1] V. Gupta and G. S. David, "Building the web of things with sun spots," 2010.

[2] J. Romkey, "Toast of the iot: The 1990 interop internet toaster," *IEEE Consumer Electronics Magazine*, vol. 6, no. 1, pp. 116–119, 2016.

[3] S. Mann. "Wearcam." (1994), [Online]. Available: `http://wearcam.org/myview.html`.

[4] P. Saffo, "Sensors: The next wave of innovation," *Communications of the ACM*, vol. 40, no. 2, pp. 92–97, 1997.

[5] K. Ashton *et al.*, "That 'internet of things' thing," *RFID journal*, vol. 22, no. 7, pp. 97–114, 2009.

[6] W. Gordon. "2020's smart fridges still aren't worth buying." (2020), [Online]. Available: `https://www.howtogeek.com/33643/2020s-smart-fridges-still-arent-worth-buying/`.

[7] M. E. E. Alahi, A. Sukkuea, F. W. Tina, *et al.*, "Integration of iot-enabled technologies and artificial intelligence (ai) for smart city scenario: Recent advancements and future trends," *Sensors*, vol. 23, no. 11, p. 5206, 2023.

[8] A. A. Khan, A. A. Laghari, Z. A. Shaikh, Z. Dacko-Pikiewicz, and S. Kot, "Internet of things (iot) security with blockchain technology: A state-of-the-art review," *IEEE Access*, vol. 10, pp. 122 679–122 695, 2022.

[9] T. Shenkoya, "Social change: A comparative analysis of the impact of the iot in japan, germany and australia," *Internet of Things*, vol. 11, p. 100 250, 2020.

[10] D. Lupton, "The internet of things: Social dimensions," *Sociology Compass*, vol. 14, no. 4, e12770, 2020.

[11] O. Arshi and S. Mondal, "Advancements in sensors and actuators technologies for smart cities: A comprehensive review," *Smart Construction and Sustainable Cities*, vol. 1, no. 1, p. 18, 2023.

[12] B. Stiller, E. Schiller, C. Schmitt, S. Ziegler, and M. James, "An overview of network communication technologies for iot," *Handbook of Internet-of-Things*, vol. 12, 2020.

[13] S. Sasirekha, A. Priya, T. Anita, and P. Sherubha, "Data processing and management in iot and wireless sensor network," in *journal of physics: conference series*, IOP Publishing, vol. 1712, 2020, p. 012 002.

[14] B. Diène, J. J. Rodrigues, O. Diallo, E. H. M. Ndoye, and V. V. Korotaev, "Data management techniques for internet of things," *Mechanical Systems and Signal Processing*, vol. 138, p. 106 564, 2020.

[15] A. A. Sadri, A. M. Rahmani, M. Saberikamarposhti, and M. Hosseinzadeh, "Fog data management: A vision, challenges, and future directions," *Journal of Network and Computer Applications*, vol. 174, p. 102 882, 2021.

[16] Y. Hajjaji, W. Boulila, I. R. Farah, I. Romdhani, and A. Hussain, "Big data and iot-based applications in smart environments: A systematic review," *Computer Science Review*, vol. 39, p. 100 318, 2021.

[17] M. Merenda, C. Porcaro, and D. Iero, "Edge machine learning for ai-enabled iot devices: A review," *Sensors*, vol. 20, no. 9, p. 2533, 2020.

[18] L. U. Khan, W. Saad, Z. Han, E. Hossain, and C. S. Hong, "Federated learning for internet of things: Recent advances, taxonomy, and open challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1759–1799, 2021.

[19] M. A. Uddin, A. Stranieri, I. Gondal, and V. Balasubramanian, "A survey on the adoption of blockchain in iot: Challenges and solutions," *Blockchain: Research and Applications*, vol. 2, no. 2, p. 100 006, 2021.

[20] H. F. Atlam and G. B. Wills, "Iot security, privacy, safety and ethics," *Digital twin technologies and smart cities*, pp. 123–149, 2020.

[21] L. P. Lavanya P and I. Subba Reddy, "A comprehensive review on lora implementation in iot application domains," in *AIP Conference Proceedings*, AIP Publishing, vol. 2516, 2022.

[22] S. Corporation, "Lora® and lorawan®: A technical overview," Semtech Corporation, Tech. Rep., 2020.

[23] F. Yao, Y. Ding, S. Hong, and S.-H. Yang, "A survey on evolved lora-based communication technologies for emerging internet of things applications," *International Journal of Network Dynamics and Intelligence*, pp. 4–19, 2022.

[24] S. Devalal and A. Karthikeyan, "Lora technology-an overview," in *2018 second international conference on electronics, communication and aerospace technology (ICECA)*, IEEE, 2018, pp. 284–290.

[25] M. Ballerini, T. Polonelli, D. Brunelli, M. Magno, and L. Benini, "Experimental evaluation on nb-iot and lorawan for industrial and iot applications," in *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, IEEE, vol. 1, 2019, pp. 1729–1732.

[26] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band internet of things," *IEEE access*, vol. 5, pp. 20 557–20 577, 2017.

[27] V. Kumar, R. K. Jha, and S. Jain, "Nb-iot security: A survey," *Wireless Personal Communications*, vol. 113, pp. 2661–2708, 2020.

[28] R. S. Sinha, Y. Wei, and S.-H. Hwang, "A survey on lpwa technology: Lora and nb-iot," *Ict Express*, vol. 3, no. 1, pp. 14–21, 2017.

[29] L. Stacchio, A. Angeli, and G. Marfia, "Empowering digital twins with extended reality collaborations," *Virtual Reality & Intelligent Hardware*, vol. 4, no. 6, pp. 487–505, 2022.

[30] I. Wohlgenannt, A. Simons, and S. Stieglitz, "Virtual reality," *Business & Information Systems Engineering*, vol. 62, pp. 455–461, 2020.

[31] H. of Information, *Morton heilig's sensorama, 1962*, https://www.historyofinformation.com/detail.php?id=2785, Accesso: 8 luglio 2024.

[32] S. S. Oyelere, N. Bouali, R. Kaliisa, G. Obaido, A. A. Yunusa, and E. R. Jimoh, "Exploring the trends of educational virtual reality games: A systematic review of empirical studies," *Smart Learning Environments*, vol. 7, pp. 1–22, 2020.

[33] M. Cha, S. Han, J. Lee, and B. Choi, "A virtual reality based fire training simulator integrated with fire dynamics data," *Fire safety journal*, vol. 50, pp. 12–24, 2012.

[34] M. Javaid and A. Haleem, "Virtual reality applications toward medical field," *Clinical Epidemiology and Global Health*, vol. 8, no. 2, pp. 600–605, 2020.

[35] O. S. Itani and L. D. Hollebeek, "Light at the end of the tunnel: Visitors' virtual reality (versus in-person) attraction site tour-related behavioral intentions during and post-covid-19," *Tourism Management*, vol. 84, p. 104 290, 2021.

[36] Statista, *Global augmented & virtual reality market size 2024*, https://www.statista.com/statistics/591181/global-augmented-virtual-reality-market-size/, Accessed: July 8, 2024.

[37] J. Garzón, "An overview of twenty-five years of augmented reality in education," *Multimodal Technologies and Interaction*, vol. 5, no. 7, p. 37, 2021.

[38] J. Gerup, C. B. Soerensen, and P. Dieckmann, "Augmented reality and mixed reality for healthcare education beyond surgery: An integrative review," *International journal of medical education*, vol. 11, p. 1, 2020.

[39] P. Buń, D. Grajewski, and F. Górski, "Using augmented reality devices for remote support in manufacturing: A case study and analysis.," *Advances in Production Engineering & Management*, vol. 16, no. 4, 2021.

[40] X. Zhao, M. Li, Z. Sun, *et al.*, "Intelligent construction and management of landscapes through building information modeling and mixed reality," *Applied Sciences*, vol. 12, no. 14, p. 7118, 2022.

[41] R. P. Kumar, E. Pelanis, R. Bugge, *et al.*, "Use of mixed reality for surgery planning: Assessment and development workflow," *Journal of Biomedical Informatics*, vol. 112, p. 100 077, 2020.

[42] *Vr compare*, https://vr-compare.com/, Accessed: 2024-07-08.

[43] R. Monica and J. Aleotti, "Evaluation of the oculus rift s tracking system in room scale virtual reality," *Virtual Reality*, vol. 26, no. 4, pp. 1335–1345, 2022.

[44] M. Singh, R. A. Shankar, and B. Jung, "Inside-out magnetic tracking for virtual/augmented reality applications," *IEEE Sensors Journal*, vol. 21, no. 24, pp. 28 097–28 106, 2021.

[45] E. Lukosiunas, V. Bulbenkiene, A. Andziulis, L. Pasviestis, and E. Artemciukas, "An ultrasonic tracking method for augmented reality," in *Proceedings of the 17th International Conference on Information and Software Technologies*, 2011, pp. 27–29.

[46] L. A. Groves, P. Carnahan, D. R. Allen, R. Adam, T. M. Peters, and E. C. Chen, "Accuracy assessment for the co-registration between optical and vive head-mounted display tracking," *International journal of computer assisted radiology and surgery*, vol. 14, pp. 1207–1215, 2019.

[47] P. Jones, T. Osborne, C. Sullivan-Drage, N. Keen, and E. Gadsby, *Virtual Reality Methods*. Jun. 2022, ISBN: 9781447360773. DOI: 10.47674/9781447360773.

[48] S. Subramanyam, J. Li, I. Viola, and P. Cesar, "Comparing the quality of highly realistic digital humans in 3dof and 6dof: A volumetric video case study," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, IEEE, 2020, pp. 127–136.

[49] G. Morana, "Impact of imaging and distance perception in vr immersive visual experience," 2024. [Online]. Available: http://hdl.handle.net/2299/27468.

[50] C. H. Lim and S. C. Lee, "The effects of degrees of freedom and field of view on motion sickness in a virtual reality context," *International Journal of Human–Computer Interaction*, pp. 1–13, 2023.

[51] U. A. Chattha, U. I. Janjua, F. Anwar, T. M. Madni, M. F. Cheema, and S. I. Janjua, "Motion sickness in virtual reality: An empirical evaluation," *IEEE Access*, vol. 8, pp. 130 486–130 499, 2020.

[52] Niantic, *Pokémon go live*, Accessed: 2024-07-09. [Online]. Available: https://pokemongolive.com/.

[53] O. D. Team, *Beat saber official website*, Accessed: 2024-07-09. [Online]. Available: https://beatsaber.com/.

[54] G. W. Young, N. O'Dwyer, M. F. Vargas, R. M. Donnell, and A. Smolic, "Feel the music!—audience experiences of audio–tactile feedback in a novel virtual reality volumetric music video," in *Arts*, MDPI, vol. 12, 2023, p. 156.

[55] M. Ullah, S. U. Amin, M. Munsif, *et al.*, "Serious games in science education: A systematic literature," *Virtual Reality & Intelligent Hardware*, vol. 4, no. 3, pp. 189–209, 2022.

[56] B. M. Spiegel, A. Rizzo, S. Persky, *et al.*, "What is medical extended reality? a taxonomy defining the current breadth and depth of an evolving field," *Journal of Medical Extended Reality*, vol. 1, no. 1, pp. 4–12, 2024.

[57] S. Barteit, L. Lanfermann, T. Bärnighausen, F. Neuhann, C. Beiersmann, *et al.*, "Augmented, mixed, and virtual reality-based head-mounted devices for medical education: Systematic review," *JMIR serious games*, vol. 9, no. 3, e29080, 2021.

[58] T. Khan, E. G. Andrews, P. A. Gardner, *et al.*, "Ar in the or: Exploring use of augmented reality to support endoscopic surgery," in *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences*, 2022, pp. 267–270.

[59] A. Demeco, L. Zola, A. Frizziero, *et al.*, "Immersive virtual reality in post-stroke rehabilitation: A systematic review," *Sensors*, vol. 23, no. 3, p. 1712, 2023.

[60] Q. Huang, J. Lin, R. Han, C. Peng, and A. Huang, "Using virtual reality exposure therapy in pain management: A systematic review and meta-analysis of randomized controlled trials," *Value in Health*, vol. 25, no. 2, pp. 288–301, 2022.

[61] R. Liu, L. Wang, J. Lei, Q. Wang, and Y. Ren, "Effects of an immersive virtual reality-based classroom on students' learning performance in science lessons," *British Journal of Educational Technology*, vol. 51, no. 6, pp. 2034–2049, 2020.

[62] V. Geroimenko, "Augmented reality in education," *Springer Series on Cultural Computing. Cham: Springer International Publishing*, 2020.

[63] N. Pellas, I. Kazanidis, and G. Palaigeorgiou, "A systematic literature review of mixed reality environments in k-12 education," *Education and Information Technologies*, vol. 25, no. 4, pp. 2481–2520, 2020.

[64] U. Radhakrishnan, K. Koumaditis, and F. Chinello, "A systematic review of immersive virtual reality for industrial skills training," *Behaviour & Information Technology*, vol. 40, no. 12, pp. 1310–1339, 2021.

[65] D. Mourtzis, J. Angelopoulos, and N. Panopoulos, "Intelligent predictive maintenance and remote monitoring framework for industrial equipment based on mixed reality," *Frontiers in Mechanical Engineering*, vol. 6, p. 578 379, 2020.

[66] P. Wang, X. Bai, M. Billinghurst, *et al.*, "Ar/mr remote collaboration on physical tasks: A review," *Robotics and Computer-Integrated Manufacturing*, vol. 72, p. 102 071, 2021.

[67] P. Dymora, B. Kowal, M. Mazurek, and S. Romana, "The effects of virtual reality technology application in the aircraft pilot training process," in *IOP conference series: materials science and engineering*, IOP Publishing, vol. 1024, 2021, p. 012 099.

[68] A. D. Kaplan, J. Cruit, M. Endsley, S. M. Beers, B. D. Sawyer, and P. A. Hancock, "The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis," *Human factors*, vol. 63, no. 4, pp. 706–726, 2021.

[69] P. R. Wolf, B. A. Dewitt, and B. E. Wilkinson, *Elements of Photogrammetry with Applications in GIS*, 4th. McGraw-Hill Education, 2014, ISBN: 9780071761123.

[70] O. Özyeşil, V. Voroninski, R. Basri, and A. Singer, "A survey of structure from motion*.," *Acta Numerica*, vol. 26, pp. 305–364, 2017.

[71] M. Gaiani, F. Remondino, F. I. Apollonio, and A. Ballabeni, "An advanced preprocessing pipeline to improve automated photogrammetric reconstructions of architectural scenes," *Remote sensing*, vol. 8, no. 3, p. 178, 2016.

[72] R. Qin and A. Gruen, "The role of machine intelligence in photogrammetric 3d modeling–an overview and perspectives," *International Journal of Digital Earth*, vol. 14, no. 1, pp. 15–31, 2021.

[73] Z. Zhao, D. Ye, X. Zhang, G. Chen, and B. Zhang, "Improved direct linear transformation for parameter decoupling in camera calibration," *Algorithms*, vol. 9, no. 2, p. 31, 2016.

[74] J. A. Doumit, "From drones to geospatial data," *KubGU, Krasnodar. Courtesy of assistant professor Steve Scott*, 2018.

[75] U. of British Columbia, *Texturing slides*, Accessed: 2024-07-12, 2013. [Online]. Available: https://www.cs.ubc.ca/~tmm/courses/314-13/slides/textures-4x4.pdf.

[76] J. Jeon, Y. Jung, H. Kim, and S. Lee, "Texture map generation for 3d reconstructed scenes," *The Visual Computer*, vol. 32, pp. 955–965, 2016.

[77] M. Attene, M. Campen, and L. Kobbelt, "Polygon mesh repairing: An application perspective," *ACM Computing Surveys (CSUR)*, vol. 45, no. 2, pp. 1–33, 2013.

[78] D. Khan, A. Plopski, Y. Fujimoto, *et al.*, "Surface remeshing: A systematic literature review of methods and research directions," *IEEE transactions on visualization and computer graphics*, vol. 28, no. 3, pp. 1680–1713, 2020.

[79] AliceVision, *Alicevision: Photogrammetric computer vision framework*, https://alicevision.org/, Accessed: 2024-07-12, 2024.

[80] Agisoft LLC, *Agisoft metashape*, https://www.agisoft.com/, Accessed: 2024-07-12, 2024.

[81] Capturing Reality, *Capturing reality: 3d scanning and photogrammetry software*, https://www.capturingreality.com/, Accessed: 2024-07-22.

[82] Pix4D, *Pix4d: Photogrammetry software for 3d mapping and modeling*, Accessed: 2024-07-22, 2024. [Online]. Available: https://www.pix4d.com/.

[83] 3Dflow. "3df zephyr photogrammetry software." Accessed: 2024-07-22. (2024), [Online]. Available: https://www.3dflow.net/3df-zephyr-photogrammetry-software/.

[84] J. I. Fiz, P. M. Martín, R. Cuesta, E. Subías, D. Codina, and A. Cartes, "Examples and results of aerial photogrammetry in archeology with uav: Geometric documentation, high resolution multispectral analysis, models and 3d printing," *Drones*, vol. 6, no. 3, p. 59, 2022.

[85] Y. Zhou, "Museum digital repatriation and case studies: Exploring guidelines for the future practice and digital bridge to cultural continuity," Ph.D. dissertation, Georgetown University, 2024.

[86] D. Tanasi, S. Hassam, and K. Kingsland, "Underground archaeology: Photogrammetry and terrestrial laser scanning of the hypogeum of crispia salvia (marsala, italy)," in *International Conference on Pattern Recognition*, Springer, 2021, pp. 353–367.

[87] E. Tsoraeva, R. Gadzhiev, S. Kuchiev, A. Pekh, and S. Mezhyan, "Application of photogrammetric methods in architecture, construction and land management," in *IOP Conference Series: Materials Science and Engineering*, IOP Publishing, vol. 1083, 2021, p. 012 052.

[88] S. I. Deliry and U. Avdan, "Accuracy of unmanned aerial systems photogrammetry and structure from motion in surveying and mapping: A review," *Journal of the Indian Society of Remote Sensing*, vol. 49, no. 8, pp. 1997–2017, 2021.

[89] K. R. Mathews, A. Sarafraz, and C. Mader, "3d modeling in the urban classroom: Using photogrammetry for the study of historic architecture in coral gables, florida," *The Journal of Interactive Technology and Pedagogy*, Feb. 2018.

[90] N. Rihani, "Interactive immersive experience: Digital technologies for reconstruction and experiencing temple of bel using crowdsourced images and 3d photogrammetric processes," *International Journal of Architectural Computing*, vol. 21, no. 4, pp. 730–756, 2023.

[91] V. Rinaldi, L. Hackman, and N. NicDaeid, "Virtual reality as a collaborative tool for digitalised crime scene examination," in *International Conference on Extended Reality*, Springer, 2022, pp. 154–161.

[92] R. Ismail, R. B. Taqriban, M. Ariyanto, *et al.*, "Affordable and faster transradial prosthetic socket production using photogrammetry and 3d printing," *Electronics*, vol. 9, no. 9, p. 1456, 2020.

[93] S. Lozares-Cordero, C. Bermejo-Barbanoj, A. Badías-Herbera, *et al.*, "An open-source development based on photogrammetry for a real-time iort treatment planning system," *Physica Medica*, vol. 112, p. 102 622, 2023.

[94] J. Kantor, "Passive stereophotogrammetry and structured light scanning for 3-dimensional imaging in dermatologic surgery," *Dermatologic Surgery*, vol. 46, no. 6, pp. 855–856, 2020.

[95] R. Ferrari, L. Lachs, D. R. Pygas, *et al.*, "Photogrammetry as a tool to improve ecosystem restoration," *Trends in Ecology & Evolution*, vol. 36, no. 12, pp. 1093–1101, 2021.

[96] M. Mandirola, C. Casarotti, S. Peloso, *et al.*, "Guidelines for the use of unmanned aerial systems for fast photogrammetry-oriented mapping in emergency response scenarios," *International Journal of Disaster Risk Reduction*, vol. 58, p. 102 207, 2021.

[97] W. Boonpook, Y. Tan, and B. Xu, "Deep learning-based multi-feature semantic segmentation in building extraction from images of uav photogrammetry," *International Journal of Remote Sensing*, vol. 42, no. 1, pp. 1–19, 2021.

[98] A. Murtiyoso, E. Pellis, P. Grussenmeyer, T. Landes, and A. Masiero, "Towards semantic photogrammetry: Generating semantically rich point clouds from architectural close-range photogrammetry," *Sensors*, vol. 22, no. 3, p. 966, 2022.

[99] M. Chen, A. Feng, K. McCullough, P. B. Prasad, R. McAlinden, and L. Soibelman, "Generating synthetic photogrammetric data for training deep learning based 3d point cloud segmentation models," *arXiv preprint arXiv:2008.09647*, 2020.

[100] L. Chen *et al.*, "Real-time photogrammetry based on parallel architecture for 3d applications," Ph.D. dissertation, The Hong Kong Polytechnic University, 2024.

[101] H. Xia, Z.-H. Lin, W.-C. Ma, and S. Wang, "Video2game: Real-time interactive realistic and browser-compatible environment from a single video," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 4578–4588.

[102] E. Zakeri and W.-F. Xie, "Robust photogrammetry sensor-based real-time pose control of industrial robots," *IEEE/ASME Transactions on Mechatronics*, 2024.

[103] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," *Communications of the ACM*, vol. 65, no. 1, pp. 99–106, 2021.

[104] S. J. Purkis and J. C. Brock, "Lidar overview," in *Coral Reef Remote Sensing: A Guide for Mapping, Monitoring and Management*, Springer, 2013, pp. 115–143.

[105] M. Khanal, M. Hasan, N. Sterbentz, R. Johnson, and J. Weatherly, "Accuracy comparison of aerial lidar, mobile-terrestrial lidar, and uav photogrammetric capture data elevations over different terrain types," *Infrastructures*, vol. 5, no. 8, p. 65, 2020.

[106] P. Zhang, H. He, Y. Wang, *et al.*, "3d urban buildings extraction based on airborne lidar and photogrammetric point cloud fusion according to u-net deep learning model segmentation," *IEEE Access*, vol. 10, pp. 20 889–20 897, 2022.

[107] P. Mat Zam, N. Fuad, A. Yusoff, and Z. Majid, "Evaluating the performance of terrestrial laser scanning for landslide monitoring," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp. 35–55, 2018.

[108] C. Wu, Y. Yuan, Y. Tang, and B. Tian, "Application of terrestrial laser scanning (tls) in the architecture, engineering and construction (aec) industry," *Sensors*, vol. 22, no. 1, p. 265, 2021.

[109] S. Kushwaha, K. R. Dayal, Sachchidanand, *et al.*, "3d digital documentation of a cultural heritage site using terrestrial laser scanner—a case study," in *Applications of Geomatics in Civil Engineering: Select Proceedings of ICGCE 2018*, Springer, 2020, pp. 49–58.

[110] B. Kekeç, N. Bilim, E. Karakaya, and D. Ghiloufi, "Applications of terrestrial laser scanning (tls) in mining: A review," *Türkiye Lidar Dergisi*, vol. 3, no. 1, pp. 31–38, 2021.

[111] K. Calders, J. Adams, J. Armston, *et al.*, "Terrestrial laser scanning in forest ecology: Expanding the horizon," *Remote Sensing of Environment*, vol. 251, p. 112 102, 2020.

[112] J. Jansa, N. Studnicka, G. Forkert, A. Haring, and H. Kager, "Terrestrial laser-scanning and photogrammetry–acquisition techniques complementing one another,"

International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 35, no. 7, pp. 948–953, 2004.

[113] H. Nagamatsu, H. Sasahara, Y. Mitsutake, and T. Hamamoto, "Development of a cooperative system for wire and arc additive manufacturing and machining," *Additive Manufacturing*, vol. 31, p. 100 896, 2020.

[114] M. Pepe, V. S. Alfio, and D. Costantino, "Uav platforms and the sfm-mvs approach in the 3d surveys and modelling: A review in the cultural heritage field," *Applied Sciences*, vol. 12, no. 24, p. 12 886, 2022.

[115] K. Brunnström, K. De Moor, A. Dooms, *et al.*, *Qualinet White Paper on Definitions of Quality of Experience*. Mar. 2013.

[116] I. T. Union, "Recommendation itu-t p.10/g.100," ITU-T, Tech. Rep., 2017. [Online]. Available: https://www.itu.int/rec/T-REC-P.10-201711-I/en.

[117] K. Piamrat, C. Viho, J.-M. Bonnin, and A. Ksentini, "Quality of experience measurements for video streaming over wireless networks," in *2009 Sixth International Conference on Information Technology: New Generations*, IEEE, 2009, pp. 1184–1189.

[118] ITU-T, "Subjective quality evaluation of telephone services," International Telecommunication Union, ITU-T Recommendation ITU-T Recommendation P.800, 1996.

[119] Y. Chen, K. Wu, and Q. Zhang, "From qos to qoe: A tutorial on video quality assessment," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 2, pp. 1126–1165, 2014.

[120] S. Coskun, G. Nur Yilmaz, F. Battisti, M. Alhussein, and S. Islam, "Measuring 3d video quality of experience (qoe) using a hybrid metric based on spatial resolution and depth cues," *Journal of Imaging*, vol. 9, no. 12, p. 281, 2023.

[121] T. Yamazaki, "Quality of experience (qoe) studies: Present state and future prospect," *IEICE Transactions on Communications*, vol. 104, no. 7, pp. 716–724, 2021.

[122] Y. Chen, K. Wu, and Q. Zhang, "From qos to qoe: A tutorial on video quality assessment," *IEEE Communications Surveys 'I&' Tutorials*, vol. 17, no. 2, pp. 1126–1165, 2015. DOI: 10.1109/COMST.2014.2363139.

[123] H. J. Kim and S. G. Choi, "A study on a qos/qoe correlation model for qoe evaluation on iptv service," in *2010 The 12th International Conference on Advanced Communication Technology (ICACT)*, vol. 2, 2010, pp. 1377–1382.

[124] N. Banović-Ćurguz and D. Ilišević, "Mapping of qos/qoe in 5g networks," in *2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 2019, pp. 404–408. DOI: 10.23919/MIPRO.2019.8757034.

[125] Y. Zhu, I. Heynderickx, and J. A. Redi, "Understanding the role of social context and user factors in video quality of experience," *Computers in Human Behavior*, vol. 49, pp. 412–426, 2015.

[126] K. Fizza, A. Banerjee, K. Mitra, *et al.*, "Qoe in iot: A vision, survey and future directions," *Discover Internet of Things*, vol. 1, pp. 1–14, 2021.

[127] S. Vlahovic, M. Suznjevic, and L. Skorin-Kapov, "A survey of challenges and methods for quality of experience assessment of interactive vr applications," *Journal on Multimodal User Interfaces*, vol. 16, no. 3, pp. 257–291, 2022.

[128] A. Morris, J. Guan, and A. Azhar, "An xri mixed-reality internet-of-things architectural framework toward immersive and adaptive smart environments," in *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, IEEE, 2021, pp. 68–74.

[129] K. Kastner, J. Lehmann, F. Wühler, *et al.*, "Leveraging transdisciplinary engineering through the coalescence of digital twins and xr-technologies.," in *TE*, 2023, pp. 343–352.

[130] A. Atanassov, "Vireo: Internet of things integration in extended reality for network platforms," Master's Thesis, Politecnico di Milano, Milan, Italy, 2022.

[131] A. Alunni, *Innovation finance and technology transfer: Funding proof-of-concept*. Routledge, 2019.

[132] P. Truong *et al.*, "Proof-of-concept demo for collaborative machine room planning in virtual reality," M.S. thesis, Aalto University, 2021.

[133] Blender Foundation. "Blender system requirements." Accessed: 2024-07-25. (2024), [Online]. Available: `https://www.blender.org/download/requirements/`.

[134] Aryel, *Gltf/glb is the webar 3d file format best friend and this is why*, Accessed: 2024-07-26, 2020. [Online]. Available: `https://aryelhq.medium.com/gltf-glb-is-the-webar-3d-file-format-best-friend-and-this-is-why-e8a9f11bea1`.

[135] K. SHEPPARD, J. CASSELLA, and S. FIELDHOUSE, "How technology is revolutionising crime scene capture and presentation visualising a crime scene using novel crime scene documentation technology," *CSeye*, vol. 2016, no. Jan, pp. 16–24, 2016.

[136] D. ABATE, I. Toschi, C. STURDY COLLS, and F. Remondino, "A low-cost panoramic camera for the 3d documentation of contaminated crime scenes," in *International Society for Photogrammetry and Remote Sensing*, vol. 42, 2017, pp. 1–8.

[137] S.-h. Yu, G. Thomson, V. Rinaldi, C. Rowland, and N. N. Daeid, "Development of a dundee ground truth imaging protocol for recording indoor crime scenes to facilitate virtual reality reconstruction," *Science & Justice*, vol. 63, no. 2, pp. 238–250, 2023.

[138] M. Janiszewski, M. Torkan, L. Uotinen, and M. Rinne, "Rapid photogrammetry with a 360-degree camera for tunnel mapping," *Remote Sensing*, vol. 14, no. 21, p. 5494, 2022.

[139] S. Herban, D. Costantino, V. S. Alfio, and M. Pepe, "Use of low-cost spherical cameras for the digitisation of cultural heritage structures into 3d point clouds," *Journal of Imaging*, vol. 8, no. 1, p. 13, 2022.

[140] S. Livatino, A. Regalbuto, G. Morana, *et al.*, "Photorealistic true-dimensional visualization of remote panoramic views for vr headsets," *IEEE Access*, vol. 11, pp. 60 305–60 323, 2023.

[141] M. Pepe, V. S. Alfio, D. Costantino, and S. Herban, "Rapid and accurate production of 3d point cloud via latest-generation sensors in the field of cultural heritage: A comparison between slam and spherical videogrammetry," *Heritage*, vol. 5, no. 3, pp. 1910–1928, 2022.

[142] C. Nicolae, E. Nocerino, F. Menna, and F. Remondino, "Photogrammetry applied to problematic artefacts," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, pp. 451–456, 2014.

[143] A. Ntregka, A. Georgopoulos, and M. Santana Quintero, "Photogrammetric exploitation of hdr images for cultural heritage documentation," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 5, W1, 2013.

[144] https://pixls.us/, *Darktable - open source photography workflow application and raw developer*, `https://www.darktable.org/`, Accessed: 2024-08-30.

[145] J. Haines, *Photogrammetry workflow using a dslr camera*, Accessed: 2024-09-05, 2019. [Online]. Available: `https://scholarslab.lib.virginia.edu/blog/documentation-photogrammetry/`.

[146] M. Teittinen, "Depth cues in the human visual system," *The encyclopedia of virtual environments*, vol. 1, 1993.

[147] C. Hendrix and W. Barfield, "Relationship between monocular and binocular depth cues for judgements of spatial information and spatial instrument design," *Displays*, vol. 16, no. 3, pp. 103–113, 1995.

[148] S.-A. Yoo, S. Lee, and S. J. Joo, "Monocular cues are superior to binocular cues for size perception when they are in conflict in virtual reality," *Cortex*, vol. 166, pp. 80–90, 2023.

[149] P. B. Hibbard, A. E. Haines, and R. L. Hornsey, "Magnitude, precision, and realism of depth perception in stereoscopic vision," *Cognitive Research: Principles and Implications*, vol. 2, pp. 1–11, 2017.

[150] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert, "The perception of egocentric distances in virtual environments-a review," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, pp. 1–40, 2013.

[151] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, and L. Anderson, "Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 2, pp. 176–198, 2008.

[152] A. Pastor, L. Krasula, X. Zhu, Z. Li, and P. Le Callet, "Comparison of subjective methodologies for local perception of distortion in videos and impact on objective metrics resolving power," 2024.

[153] Y. Nehmé, J.-P. Farrugia, F. Dupont, P. L. Callet, and G. Lavoué, "Comparison of subjective methods for quality assessment of 3d graphics in virtual reality," *ACM Transactions on Applied Perception (TAP)*, vol. 18, no. 1, pp. 1–23, 2020.

[154] L. Salvatore and K. Christina, "Simple guidelines for testing vr applications," *Advances in human computer interaction*, p. 289, 2008.

[155] T. W. Schubert, "The sense of presence in virtual environments: A three-component scale measuring spatial presence, involvement, and realness.," *Z. für Medienpsychologie*, vol. 15, no. 2, pp. 69–71, 2003.

[156] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness," *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[157] S. A. Balk, D. B. Bertola, and V. W. Inman, "Simulator sickness questionnaire: Twenty years later," in *Driving Assessment Conference*, University of Iowa, vol. 7, 2013.

[158] A. Zocco, S. Livatino, P. Gainley, Y. Iqbal, and G. Morana, "The immersion advantage in command & control: From desktop monitors to vr headsets," in *2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)*, IEEE, 2022, pp. 449–453.

[159] Igroup, *Igroup presence questionnaire (ipq)*, Accessed: 2024-09-26, 1995. [Online]. Available: `https://www.igroup.org/pq/ipq/index.php`.

[160] U. of Minnesota, *Simulator sickness questionnaire*, 2016. [Online]. Available: `https://conservancy.umn.edu/items/646a8ef3-4ce0-4d2e-9426-7ae534b5fe57`.