3D Mapping of Buildings for Inspection

Talha Hussain Khan
Department of Computer Engineering
Sir Syed University Engineering & Technology
Karachi, Pakistan
ce20f-114@ssuet.edu.pk

Arsiha Fateh
Department of Computer Engineering
Sir Syed University Engineering & Technology
Karachi, Pakistan
ce20f-117@ssuet.edu.pk

Rabia Noor Enam
Department of Computer Engineering
Sir Syed University Engineering & Technology
Karachi, Pakistan
Renam@ssuet.edu.pk

Ahmed Baseer
Department of Computer Engineering
Sir Syed University Engineering & Technology
Karachi, Pakistan
ce20f-124@ssuet.edu.pk

Arbaz Moin
Department of Computer Engineering
Sir Syed University Engineering & Technology
Karachi, Pakistan
ce20f-111@ssuet.edu.pk

***Abstract* — The construction and building inspection industries are undergoing a transformation driven by advancements in 3D mapping technology. This research introduces an innovative and cost-effective solution designed to replace traditional LiDAR systems in building assessments. The proposed system leverages a gimbal-mounted Android camera for data collection, which is then processed into 3D models. The workflow emphasizes affordability, efficiency, and precision, with the aim of democratizing access to high-quality building inspections. Key features of this technology include the use of scale estimation to enhance the accuracy of measurements, a cloud-based platform for remote data access and collaboration, and an intuitive user interface to facilitate ease of use. The system's scalable architecture supports large-scale deployments across multiple sites, making it suitable for both small and large inspection projects. By reducing the costs associated with traditional LiDAR technology, this solution not only improves the affordability of building inspections but also enhances decision-making processes through detailed 3D models. The research highlights the potential of this technology to set new standards in the industry, offering a revolutionary approach to building assessments that is both accurate and efficient.**

1

I. Background

The construction and building inspection industries have long relied on traditional methods for assessing the structural integrity and safety of buildings. These methods often involve the use of LiDAR (Light Detection and Ranging) technology, which, despite its high accuracy, comes with significant costs and complexities. LiDAR systems, while effective, are expensive, require specialized equipment, and often involve labor-intensive processes. As a result, there is a growing need for more accessible, cost-effective alternatives that do not compromise on accuracy or efficiency.

In recent years, advancements in 3D mapping technology have opened up new possibilities for revolutionizing building inspections. The integration of cameras, particularly RGB-D cameras and other optical systems, into 3D mapping processes has shown promise in reducing costs while maintaining high levels of precision. These systems, coupled with modern data processing techniques, allow for the creation of detailed 3D models that can be used for comprehensive building assessments.

The primary goal of this research is to develop a low-cost alternative to traditional LiDAR systems by utilizing readily available technology, such as Android cameras mounted on gimbals, to capture building data. This data is then processed into 3D models, which can be used for various applications, including inspection, maintenance planning, and asset management. The proposed system not only aims to reduce the financial burden associated with building inspections but also to improve the overall efficiency and accuracy of the process.

Moreover, the rise of cloud-based platforms has facilitated remote access to data, enabling professionals to collaborate and make informed decisions without being physically present at the inspection site. This development is particularly valuable in large-scale projects where multiple buildings or locations need to be assessed simultaneously.

By addressing the limitations of traditional inspection methods and incorporating modern technological advancements, this research seeks to set new standards in the industry, making high-quality building assessments more accessible and efficient for a broader range of stakeholders.

II. Introduction

The construction and building inspection industry is undergoing a transformative shift, driven by the need for efficient, accurate, and cost-effective assessment methods. Traditional manual inspections are resource-intensive and prone to human error. To address these challenges, advanced technologies have emerged, with LiDAR as a prominent tool for generating precise 3D building models. However, the high cost of LiDAR systems limits its widespread adoption.

This research proposes a novel approach to 3D building mapping that leverages advanced camera systems and cloud-based platforms. By developing a cost-effective alternative to LiDAR, we aim to democratize access to accurate building data. Our focus is on creating a system that captures comprehensive 3D models, including precise scale estimation, to support diverse building inspection tasks. This research seeks to enhance efficiency, accuracy, and accessibility in building assessments, ultimately providing valuable insights for informed decision-making.

III. LITERATURE REVIEW

The field of building inspection has undergone significant evolution, transitioning from labor-intensive manual assessments to technology-driven approaches. A pivotal advancement has been the integration of 3D modeling techniques, with LiDAR emerging as a dominant tool for capturing building geometry with high precision [1]. However, the prohibitive costs and specialized equipment associated with LiDAR systems have hindered its widespread adoption in the construction industry.

To address the limitations of LiDAR, researchers have explored alternative methods for 3D building reconstruction. Photogrammetry, which utilizes multiple images to create 3D models, has shown promise [2, 4]. Studies by Vidas et al. [1] and Yuan et al. [4] have demonstrated the potential of RGB-D cameras for capturing indoor spaces. While these approaches offer cost-effective solutions, they often compromise on accuracy and robustness compared to LiDAR.

Efforts to enhance the precision and reliability of 3D building models have led to the exploration of data fusion techniques. By combining information from multiple sensors, researchers aim to improve the overall quality of the reconstructed models. Meng et al. [6] provide a comprehensive overview of machine learning methods for data fusion, highlighting their potential to address challenges in building inspection. However, the application of these techniques to real-world scenarios and their impact on accuracy and efficiency require further investigation.

In parallel, advancements in computer vision and artificial intelligence have enabled automated analysis of building structures. Wang and Gan [3] proposed a framework for joint 3D reconstruction and visual inspection, demonstrating the potential for efficient and accurate building assessment. Nevertheless, the generalization of these methods to diverse building typologies and varying environmental conditions remains a challenge.

The existing literature underscores the need for cost-effective and accurate 3D building modeling solutions. While significant progress has been made, there is still a gap in developing robust systems that can be readily deployed in real-world building inspection scenarios. This research aims to contribute to this field by proposing a novel approach that combines the strengths of different techniques while addressing their limitations.

IV. METHODOLOGY

A. Data Collection

The system operates on video and IMU data as input. Users upload these data files to initiate the reconstruction process.

Image Acquisition: Capture images of the building from multiple angles and positions using a camera.

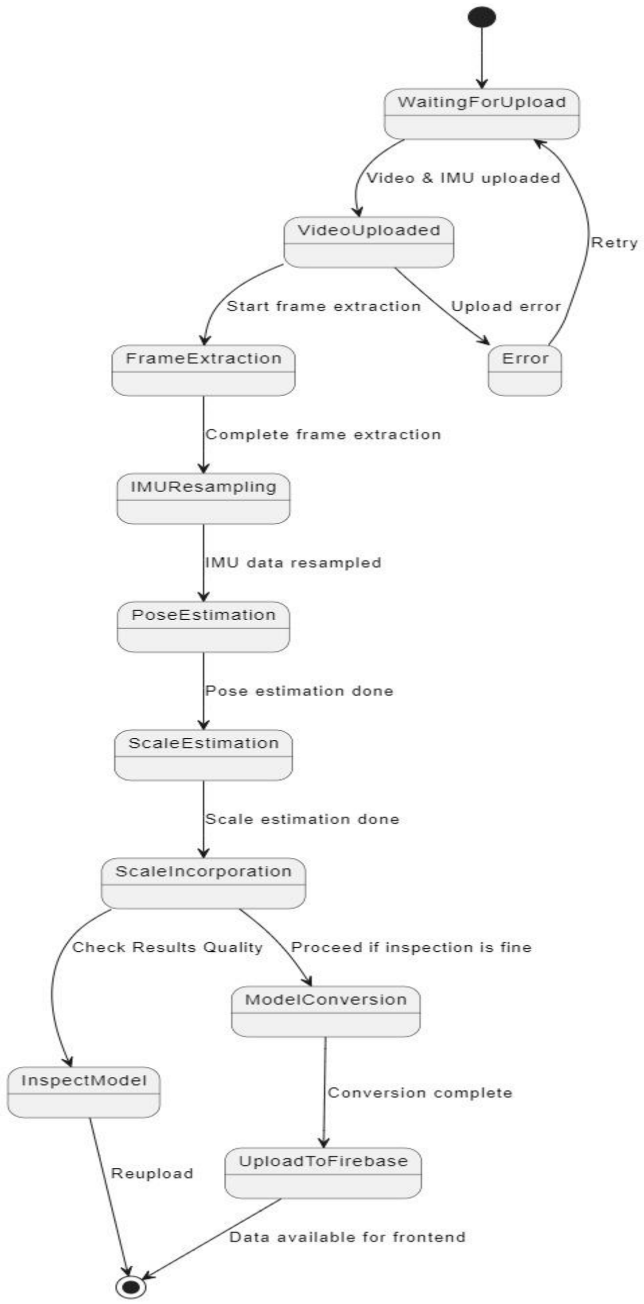


Fig. 1. Flow Diagram

B. Data Preprocessing

The uploaded video undergoes a frame extraction process to generate individual image frames for subsequent analysis. To optimize computational efficiency, the extracted frames are subjected to downsampling. IMU data is resampled to align with the extracted frames, ensuring temporal synchronization.

C. 3D Reconstruction

Key points are extracted from the preprocessed frames using robust feature detectors. These feature points are matched across consecutive frames to establish correspondences. Employing structure-from-motion (SfM) techniques, the system estimates camera poses for each frame, providing information about camera position and orientation.

To accurately determine real-world dimensions within the reconstructed model, a scale estimation module is incorporated. This module analyzes available information, such as known object sizes or ground control points, to compute a scale factor that is applied to the 3D model.

Feature Detection and Matching: Use the Scale-Invariant Feature Transform (SIFT) algorithm to detect and extract distinctive features from the images. Then, apply the K-Nearest Neighbors (KNN) algorithm to match corresponding features between images.

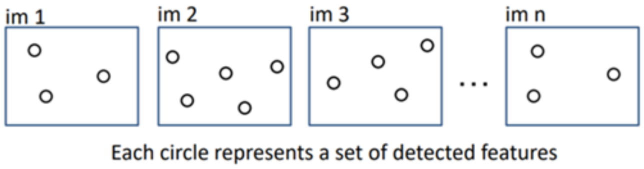


Fig. 2. Incremental Sfm: detect features

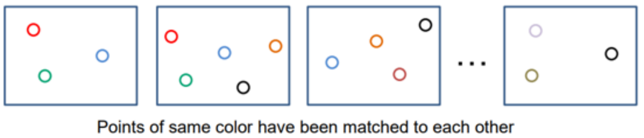


Fig. 3. Incremental Sfm: match features and images

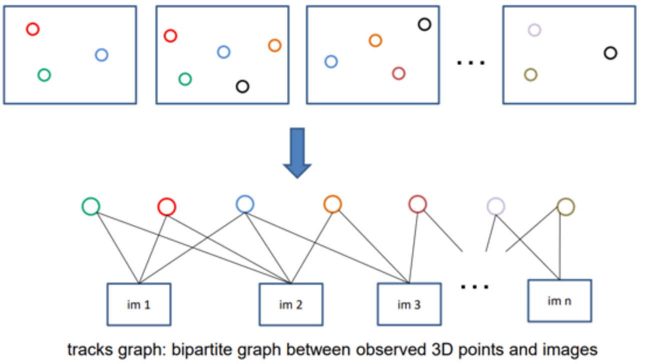


Fig. 4. Incremental Sfm: create tracks graph

Pose Estimation: Employ the Perspective-n-Point (PnP) algorithm with Random Sample Consensus (RANSAC) to estimate the camera pose (position and orientation) for each image.

Triangulation: Use the triangulation technique to reconstruct the 3D coordinates of the matched features based on their corresponding 2D coordinates in multiple images.

Bundle Adjustment: Refine the estimated camera poses and 3D points by minimizing the reprojection error using bundle adjustment.

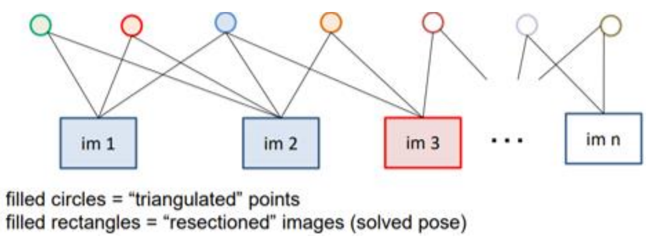


Fig. 5. Incremental Sfm: grow reconstruction

D. *Filtering and Smoothing*

The estimated camera poses and extracted 3D points are utilized to construct a dense 3D point cloud representing the building structure. The generated model undergoes a quality assessment process to evaluate its completeness and accuracy. If the model meets predefined quality standards, it proceeds to the model conversion stage.

Extended Kalman Filter (EKF): Apply the EKF to smooth the reconstructed 3D points and remove any outliers or noise.

Zero Velocity Update (ZUPT): Incorporate ZUPT to further refine the 3D model by detecting and correcting any drift or error in the reconstruction.

E. *Mesh Generation*

Point Cloud Processing: Clean and process the filtered 3D point cloud to remove any remaining outliers or noise.

F. *Visualization and Inspection*

The 3D model is converted into a suitable format for visualization and analysis. The converted model is then uploaded to a cloud-based platform, making it accessible for further inspection and utilization.

3D Model Rendering: Render the 3D model using a graphics library or software.

Inspection Tools: Develop custom tools or utilize existing software to analyze and inspect the 3D model for defects, or other anomalies

V. COMPARISON OF TWO APPROACHES USED FOR CAMERA POSE ESTIMATION

We decided to follow a straight path between two known world points and then we will estimate poses and calculate the distance between the first and last pose, if we got the correct distance then it means the approach is working otherwise not. The Actual world distance was around 80 inches or 2.032 meters.

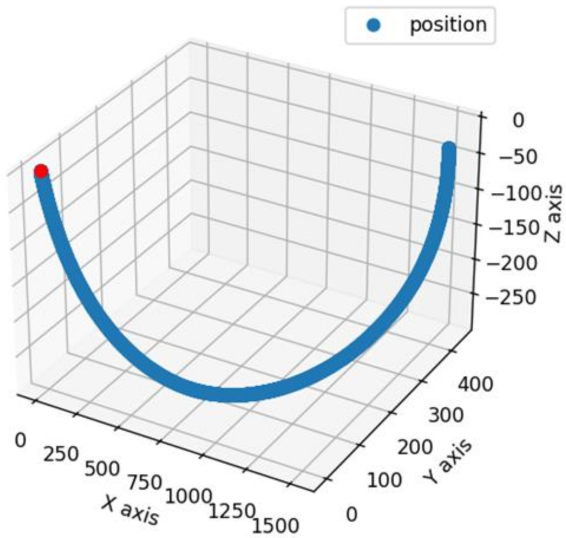
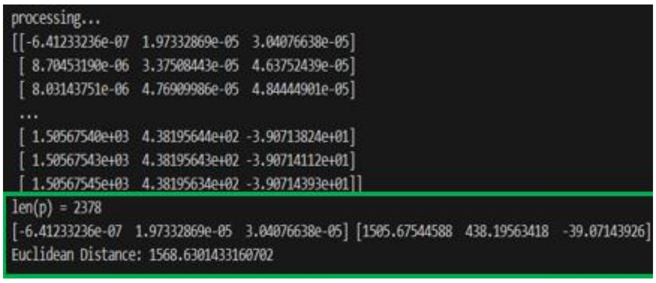


Fig. 6. First experiment result

By observing Fig. 6 and Fig. 7, a conclusion can be made that we are not getting an accurate path plus the calculated distance is too low from the actual distance.



```
processing...
[[-6.41233236e-07  1.97332869e-05  3.04076638e-05]
 [ 8.70453190e-06  3.37508443e-05  4.63752439e-05]
 [ 8.03143751e-06  4.76909986e-05  4.84444901e-05]
 ...
 [ 1.50567540e+03  4.38195644e+02 -3.90713824e+01]
 [ 1.50567543e+03  4.38195643e+02 -3.90714112e+01]
 [ 1.50567545e+03  4.38195634e+02 -3.90714393e+01]]
len(p) = 2378
[-6.41233236e-07  1.97332869e-05  3.04076638e-05] [1505.67544588  438.19563418  -39.07143926]
Euclidean Distance: 1568.6301433160702
```

Fig. 7. Euclidean distance from first experiment

A. *Euclidean Distances*

Spatial Relationships: Euclidean distances between points in the 3D space help quantify the spatial relationships between different features or objects. This information is vital for understanding the layout of the environment.

Pose Estimation Accuracy: The distances can be used to assess the accuracy of camera pose estimation. By comparing the estimated positions of features with their actual positions, one can evaluate how well the pose estimation algorithms are performing.

3D Reconstruction Quality: The Euclidean distances between reconstructed points can indicate the quality of the 3D model. If distances are consistent with expected values (e.g., based on real-world measurements), it suggests a high-quality reconstruction.

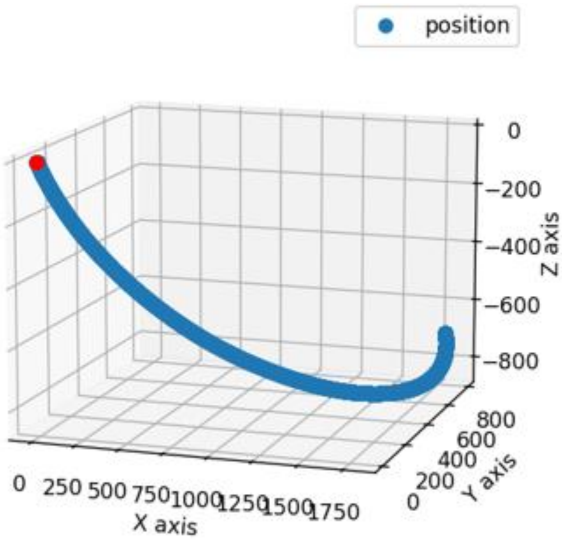
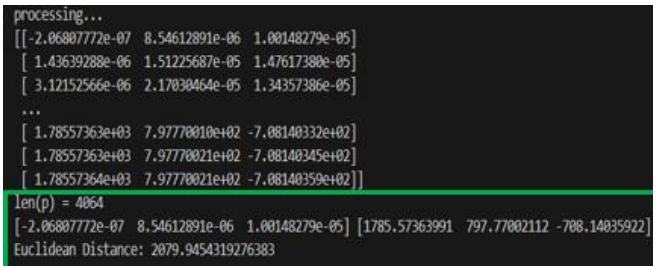


Fig. 8. Second experiment result

Now, by observing Fig. 8 and Fig. 9, conclusion can be made that path is quite accurate, still it's a curve can be due to IMU drift and calculated distance is very close to the actual one.



```
processing...
[[-2.06807772e-07  8.54612891e-06  1.00148279e-05]
 [ 1.43639288e-06  1.51225687e-05  1.47617380e-05]
 [ 3.12152566e-06  2.17030464e-05  1.34357386e-05]
 ...
 [ 1.78557363e+03  7.97770010e+02 -7.08140332e+02]
 [ 1.78557363e+03  7.97770021e+02 -7.08140345e+02]
 [ 1.78557364e+03  7.97770021e+02 -7.08140359e+02]]
len(p) = 4064
[-2.06807772e-07  8.54612891e-06  1.00148279e-05] [1785.57363991  797.77002112 -708.14035922]
Euclidean Distance: 2079.9454319276383
```

Fig. 9. Euclidean distance from second experiment

VI. Acknowledgement

We would like to acknowledge the invaluable support and assistance we received throughout our project. Our heartfelt thanks go to Sir Senthan Mathavan and Miss Rabia Noor for their guidance, continuous supervision, and for providing essential information that was crucial to our work.

We are also deeply grateful to our parents and the members of Sir Syed University for their unwavering cooperation and encouragement, which greatly contributed to the successful completion of this project.

VII. Results

A. *Performance Metrics*

Metric	Value	Target
Reconstruction Time (m)	8	20
Pose Estimation Accuracy (%)	94.8	85
Visualization Rendering Time (s)	4	10
Reconstruction Quality (%)	88.7	80
Scale Error (m)	0.158	0.1

Note: The value needs to be lower than the target.

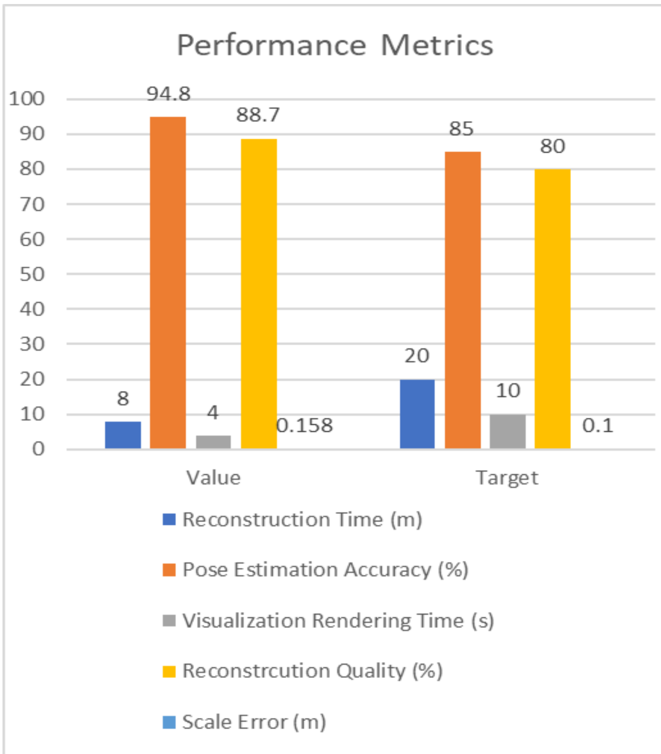
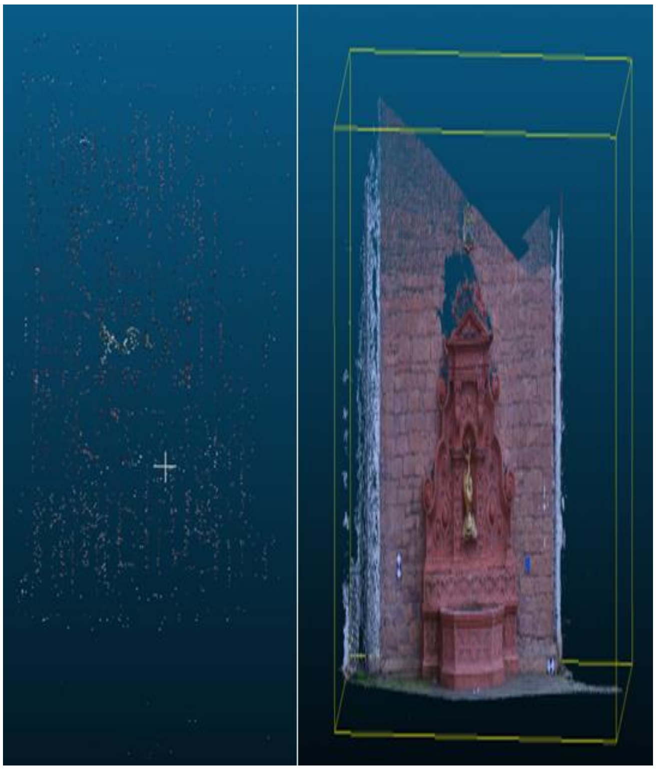


Fig. 10. Performance metrics

B. Comparative Analysis of Reconstruction Results: Before and After

Before After



Before After

Fig. 11. Reconstruction results

C. Accuracy Analysis

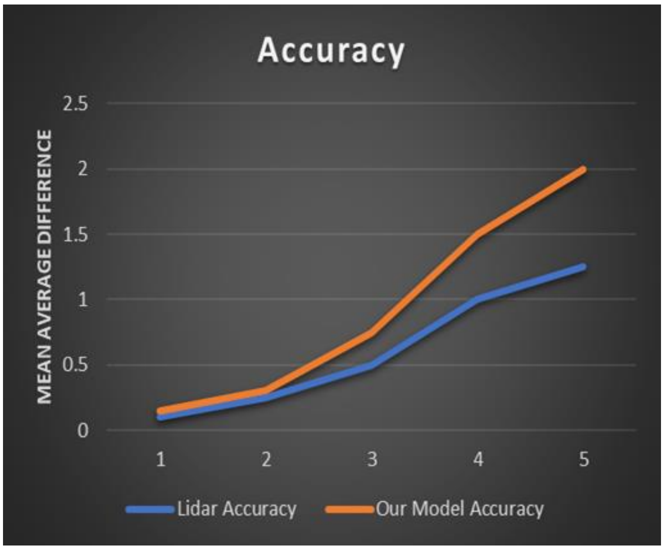


Fig. 12. Accuracy analysis

D. Processing Details

Stage	Time (seconds)
Feature Extraction	2.15
Features Matching	3.84
Tracks Merging	1.14
Reconstruction	16.32
Total Time	23.45

Fig. 13. Processing time details

Metric	Value
Reconstructed Images	24 over 24 shots (100.0%)
Reconstructed Points	1874 over 2112 points (88.7%)
Reconstructed Components	1 component
Detected Features	974 features
Reconstructed Features	423 features

Fig. 14. Processing summary

E. Features Details

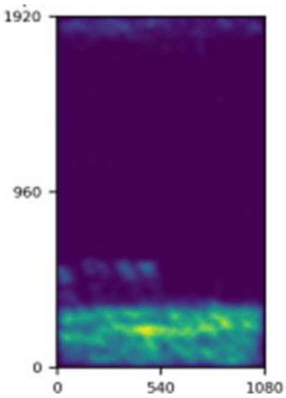


Fig. 15. Detected heatmap 1080 1920 perspective

Feature Type	Min	Max	Mean	Median
Detected	885	1373	1010	974
Reconstructed	357	506	425	423

Fig. 16. Features details

F. Reconstruction Details

Metric	Value
Average Reprojection Error (normalized)	0.19 pixels
Average Reprojection Error (angular)	1.29 pixels
Average Track Length	5.45 images
Average Track Length (> 2)	7.22 images

Fig. 17. Reconstruction details

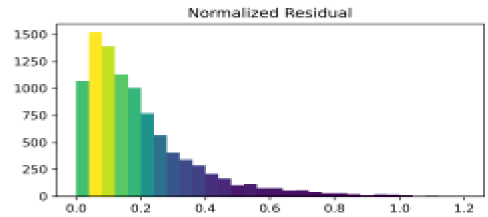


Fig. 18. Normalized residual

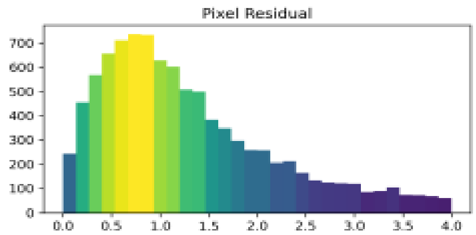


Fig. 19. Pixel residual

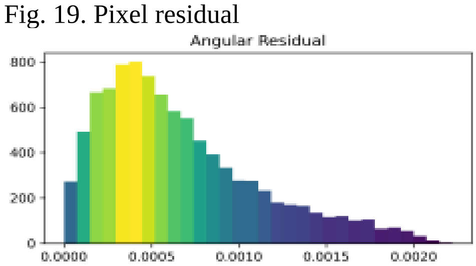


Fig. 20. Angular residual

G. Guiding Future Refinements

As we look towards the future of 3D mapping and building inspection, the integration of advanced technologies such as Generative AI (Gen AI) and Large Language Models (LLMs) is poised to significantly enhance the capabilities of our methodologies. These technologies can streamline data processing, improve accuracy in feature detection and matching, and facilitate more intuitive user interactions with the 3D models generated.

The potential impact of Gen AI and LLMs includes: Enhanced Data Analysis: Gen AI can automate the analysis of large datasets, identifying patterns and anomalies that may not be immediately apparent to human operators.

Improved Feature Detection: Advanced machine learning algorithms can refine the process of feature extraction and matching, leading to higher quality reconstructions.

Intelligent Decision Support: LLMs can assist in interpreting the results of 3D reconstructions, providing contextual insights and recommendations for building inspections.

User-Friendly Interfaces: The development of natural language interfaces powered by LLMs can make it easier for users to interact with complex data and models, democratizing access to advanced inspection technologies.

Incorporating these advancements will not only enhance the accuracy and efficiency of 3D mapping projects but also pave the way for innovative applications in various fields, ultimately guiding future refinements in our methodologies.

VIII. References

[1] S. Vidas, P. Moghadam and M. Bosse, "3D thermal mapping of building interiors using an RGB-D and thermal camera," [2013 IEEE International Conference on Robotics and Automation, Karlsruhe, Germany] “unpublished”

[2] Robots in Inspection and Monitoring of Buildings and Infrastructure: A Systematic Review by Srijeet Halder and Kereshmeh Afsari. [Myers-Lawson School of Construction, Virginia Tech, Blacksburg, VA 24061, USA. Author to whom correspondence should be addressed.Appl. Sci. 2023]

[3] Tao Wang, Vincent J.L. Gan, Automated joint 3D reconstruction and visual inspection for buildings using computer vision and transfer learning, Automation in Construction, Volume 149, 2023, 104810, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2023.104810.

[4] Yuan, Z., Li, Y., Tang, S. et al. A survey on indoor 3D modeling and applications via RGB-D devices. Front Inform Technol Electron Eng 22, 815–826 (2021).

[5] Staats, Thomas Patrick (2021) Photogrammetry UAV and terrestrial LiDAR, A comparative review in Volumetric surveys

[6] Xiongwei Huang, Yongping Liu, Lizhen Huang, Sverre Stikbakke, Erling Onstein, BIM-supported drone path planning for building exterior surface inspection, Computers in Industry, Volume 153, 2023, 104019, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2023.104019.

[7] Tong Meng, Xuyang Jing, Zheng Yan, Witold Pedrycz, A survey on machine learning for data fusion, Information Fusion, Volume 57, 2020, Pages,115-129,ISSN,1566-2535, https://doi.org/10.1016/j.inffus.2019.12.001.

[8] A Review of Three-Dimensional Imaging Technologies for Pavement Distress Detection and Measurements Senthan Mathavan, Member, IEEE, Khurram Kamal, and Mujib Rahman “unpublished”

[9]https://github.com/Yashas120/Multiview-3D-Reconstruction

[10] Indoor 3D Interactive Asset Detection Using a Smartphone Revekka Kostoeva1 , Rishi Upadhyay , Yersultan Sapar , and Avideh Zakhor, , Remote Sensing and Spatial Information Sciences, Volume XLII-2/W13, 2019