International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 10 | Oct - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

AI-Assisted Camera-Based 3D Object Scanning with Intelligent Capture Guidance and Geometric Accuracy Validation

Prof. Vaibhav Ingole

Atharva Wagh, Aditya Kumbhar, Prajwal Ghule, Sanket Jamdade

Department of Artificial Intelligence and Machine Learning, Zeal Polytechnic, Pune, India
Email: atharvawagh1102@gmail.com

Abstract—Camera-based 3D reconstruction offers a relatively low-cost way to create digital models from photographs, without requiring a dedicated industrial scanner. However, the quality of the final model depends strongly on the photographs supplied to the reconstruction software. Blurred images, inconsistent lighting, repeated views, and insufficient coverage can lead to incomplete geometry or unreliable dimensions. This paper presents a proposed AI-Motion and Multi-View Stereo, while Open3D will support point-cloud inspection and measurement. The project will be evaluated by comparing a basic reconstruction workflow with the assisted workflow using the same objects and controlled capture conditions. Dimensional error, surface completeness, reconstruction success, and processing time will be recorded. At this stage, the paper describes the literature review and proposed method; experimental results have not yet been collected, and no performance improvement is claimed.

Keywords—3D scanning, photogrammetry, artificial intelligence, Structure-from-Motion, Multi-View Stereo, image quality, geometric validation.

I. INTRODUCTION

Three-dimensional models are used in product design, inspection, reverse engineering, education, and digital documentation. A conventional 3D scanner can produce useful geometry, but its cost and availability may be a barrier for small laboratories and student projects. Photogrammetry provides another option: several overlapping photographs are processed to estimate camera positions and recover the visible surface of an object. Open-source software has made this workflow accessible to students and researchers [1], [2].

A basic scanner, however, does not automatically guarantee a reliable model. The software can only reconstruct information that is sufficiently visible in the input images. Shiny metal, transparent surfaces, plain surfaces with little texture, shadows, blur, and missing viewpoints can reduce reconstruction quality. These issues are especially important for mechanical parts, where a model may look acceptable while small holes, edges, or dimensions are inaccurate.

This project therefore focuses on improving the capture process and checking the resulting geometry, rather than claiming that photograph-to-model conversion is itself new. The proposed contribution is a practical assistance layer that checks image quality, recommends missing viewpoints, and reports geometric differences against known measurements.

cautions against specular highlights and notes that adding more images does not always improve the result if processing time increases without useful new coverage [2].

assisted workflow that combines image-quality assessment, capture guidance, photogrammetric reconstruction, and geometric validation for small mechanical components. The planned system will use computer-vision checks to flag unsuitable images and guide users toward additional viewpoints before reconstruction. COLMAP is proposed for Structure-from-

II. PROBLEM STATEMENT

In a typical low-cost photogrammetry workflow, users capture photographs and submit them to reconstruction software. If the images are unsuitable, the problem may only become clear after processing has finished. Users may then need to repeat the capture and reconstruction steps. For engineering use, a further problem is that a visually convincing mesh does not necessarily have accurate real-world dimensions. The project addresses these issues through pre-reconstruction image checks, capture guidance, and post-reconstruction validation.

III. OBJECTIVES

- To develop an affordable camera-based scanning workflow for small mechanical components.
- To detect common image-quality problems such as blur, poor brightness, and highly repetitive views.
- To provide simple guidance on additional viewpoints when coverage appears insufficient.
- To reconstruct a point cloud and mesh using an established photogrammetry pipeline.
- To compare reconstructed dimensions with reference measurements and report the observed errors.
- To compare the basic workflow and assisted workflow using consistent objects and capture conditions.

IV. LITERATURE REVIEW

A. Structure-from-Motion and COLMAP

Schönberger and Frahm presented an incremental Structure-from-Motion (SfM) approach that estimates camera poses and sparse 3D structure from unordered images [1]. Their work is important to this project because COLMAP provides an open-source implementation of a general SfM and Multi-View Stereo pipeline. Its documentation recommends adequate visual overlap, varied viewpoints, consistent illumination, and visible surface texture. It also

B. Learning-Based Multi-View Stereo

Yao et al. introduced MVSNet, a learning-based method for estimating depth maps from multiple views [3]. It uses image features and a cost-volume representation to estimate depth. This work shows that learned methods can support dense

reconstruction, but deploying a deep model may require more computation and careful dataset selection. For a student prototype, the recommended first step is to use a proven reconstruction engine and apply AI or computer vision to the capture-quality stage.

C. Neural Rendering

Mildenhall et al. proposed Neural Radiance Fields (NeRF), which represent a scene as a continuous function and synthesize novel views from input images with known camera poses [4]. NeRF is relevant to image-based 3D representation, but view synthesis and engineering-grade mesh measurement are not identical goals. Since this project prioritizes exportable geometry and dimensional validation, conventional photogrammetry is a practical baseline, while neural representations can be considered in future work.

D. Existing Student and DIY Scanners

Student and hobbyist projects have already combined cameras, turntables, and photogrammetry. For example, a Fab Academy project used an ESP32-CAM, a motorized camera arc, a rotating platform, and Meshroom to capture objects automatically [5]. A separate open-source project describes a Raspberry Pi camera and stepper-motor turntable workflow that uses COLMAP and surface reconstruction to create meshes [6]. These examples confirm that building a basic camera-based scanner is feasible and that the general idea is not new. The proposed work must therefore be evaluated on the specific value of its image assessment, viewpoint guidance, and accuracy report.

The literature and existing projects support the feasibility of the main pipeline. They do not, by themselves, prove that the proposed combination is novel or improves accuracy. A more focused review of recent papers and a formal comparison with closely related methods will be required before making a novelty claim.

V. PROPOSED METHODOLOGY

A. Image Capture

A smartphone or digital camera will be used to capture a set of overlapping views of each object. The camera settings and lighting should remain as consistent as possible. The object should be photographed from different azimuth and elevation angles, with enough overlap for the same features to appear in multiple images. A simple turntable may be used, but the camera and object movement must be planned so that the images provide useful viewpoint variation.

B. Image Quality Assessment

The first prototype will use interpretable computer-vision checks. Blur can be estimated using a sharpness measure such as the variance of the Laplacian; brightness can be checked from image intensity statistics; and near-duplicate views can be flagged using perceptual similarity or feature matching. Thresholds will be selected through pilot tests rather than treated as universal values. The system will display warnings instead of automatically deleting images, so the user can review the decision.

C. Capture Guidance

The application will summarise which views have been supplied and identify likely coverage gaps using the capture sequence and available reconstruction feedback. The initial version may use a predefined viewpoint checklist (front, rear, sides, top, and bottom where accessible). A later version can use registered camera poses and the distribution of reconstructed points to make more informed recommendations. Any coverage score will be described as an estimate, not a guarantee that every surface has been captured.

D. 3D Reconstruction and Mesh Processing

COLMAP will be used to estimate camera poses and generate sparse and dense reconstructions [1], [2]. The resulting point cloud can be inspected and processed using Open3D. Mesh generation and export will be performed only after checking the point cloud for obvious gaps, isolated points, and alignment problems. Common output formats such as PLY, OBJ, or STL may be supported, depending on the selected reconstruction and meshing tools.

E. Geometric Validation

A reference object will be measured using suitable instruments, such as a vernier caliper or micrometer, depending on the feature size and required precision. The reconstructed model will be scaled using a known reference dimension or another documented scale method. Selected dimensions—such as length, width, height, and hole diameter—will be compared with reference values. If a reference 3D model is available, point-cloud or surface-distance metrics can also be calculated. The report must clearly state the measurement method and its uncertainty.

VI. SYSTEM WORKFLOW

- Capture photographs of a mechanical component from multiple viewpoints.
- Check image sharpness, brightness, similarity, and approximate viewpoint coverage.
- Show warnings and request additional photographs where needed.
- Run COLMAP for camera-pose estimation and dense reconstruction.
- Inspect and process the point cloud or mesh using Open3D or a compatible tool.
- Scale the model using a known reference and measure selected dimensions.
- Generate a report containing the model, warnings, processing time, and dimensional errors.

VII. EXPERIMENTAL PLAN AND EVALUATION

The evaluation will compare two workflows: (1) a baseline in which the user captures images and runs reconstruction without the proposed assistance, and (2) an assisted workflow that checks image quality and gives capture guidance. Both workflows should use the same objects, camera, lighting setup, and reconstruction settings wherever possible. At least five objects with different shapes and surface characteristics are recommended for an initial student study. Each condition should be repeated when time permits.

A. Measures

Dimensional percentage error will be calculated for each measured feature as:

$$\text{Percentage error} = |\text{reference} - \text{reconstructed}| / \text{reference} \times 100\%$$

Other measures will include reconstruction success rate, number of registered images, estimated missing-surface regions, processing time, and the number of recapture attempts. If a reliable reference mesh is available, geometric distance and completeness metrics can be added. Measurements should be reported for each object as well as in summary form; results should not be averaged in a way that hides failures on difficult objects.

B. Experimental Record

For each trial, record the object type, surface finish, number of photographs, image warnings, number of registered images, reconstruction time, selected reference dimensions, reconstructed dimensions, and any visible defects. No results should be filled in until the experiment has actually been run. The final paper should include sample images, screenshots of the reconstructed models, and a results table based on the collected data.

VIII. EXPECTED OUTCOMES

The expected deliverable is a working prototype that combines image-quality checks with an existing 3D reconstruction engine and produces a basic accuracy report. The system is intended to help users identify poor input images earlier and make the scanning workflow easier to repeat. Whether it reduces dimensional error, improves completeness, or increases reconstruction success remains an experimental question. These outcomes must be demonstrated with measurements rather than stated as completed achievements.

IX. LIMITATIONS

Camera-based reconstruction can struggle with shiny, transparent, reflective, or textureless surfaces. Hidden surfaces cannot be recovered from photographs in which they are never visible. Scale must be established independently, and measurement error may come from both the reference instrument and the reconstruction. Image-quality thresholds may behave differently under different cameras and lighting conditions. A guidance score should therefore be treated as decision support, not as a substitute for inspecting the final model. This project is intended for student-level evaluation and should not be presented as a certified metrology system.

X. CONCLUSION

This paper proposes an affordable camera-based 3D scanning workflow for small mechanical components. Existing research and open-source tools make photogrammetric reconstruction feasible, but successful reconstruction depends on the quality and coverage of the input photographs. The proposed system adds image-quality assessment, capture guidance, and geometric validation around an established reconstruction pipeline. The key research task is to test whether these additions produce measurable benefits compared with a baseline workflow. The next stage is to build the prototype, collect repeatable experimental data, and revise this draft with actual results before journal submission.

ACKNOWLEDGMENT

The authors may acknowledge their project guide, department, laboratory, and any institution or individual that provided equipment or technical support. This section should be completed only after the relevant contributors have been confirmed.

REFERENCES

[1] J. L. Schönberger and J.-M. Frahm, “Structure-from-Motion Revisited,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4104–4113. doi: 10.1109/CVPR.2016.445.

[2] COLMAP, “Structure-from-Motion and Multi-View Stereo: Tutorial and Documentation.” Available: https://colmap.github.io/tutorial.html. Accessed: 28-Oct-2026.

[3] Y. Yao, Z. Luo, S. Li, T. Fang, and L. Quan, “MVSNet: Depth Inference for Unstructured Multi-view Stereo,” in Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 767–783. doi: 10.1007/978-3-030-01237-3_47.

[4] 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,” in Proceedings of the European Conference on Computer Vision (ECCV), 2020. doi: 10.1007/978-3-030-58452-8_24.

[5] Q. Bolsee, “Photogrammetry 3D Scanner Based on ESP32-CAM Module,” Fab Academy 2020 student project. Available: https://fabacademy.org/2020/labs/ulb/students/quentin-bolsee/projects/final-project/4-presentation/. Accessed: 28-Oct-2026.

[6] L. Parkin, “Raspberry Pi Turntable 3D Scanner,” open-source project repository. Available: https://github.com/LouisParkin/3DScanner. Accessed: 28-Oct-2026.

[7] Open3D, “Open3D: A Modern Library for 3D Data Processing.” Available: https://www.open3d.org/. Accessed: 28-Oct-2026.