AI –Based 3D Crime Scene Reconstruction from Multi Model inputs

Myadagoni Rishik[1], Sanga Sree Vallabh[2], Jakka Srikanth Reddy[3]*

[1,2,3]B.Sc. (Hons) Digital Forensic Science, Malla Reddy University, Hyderabad.*

Mail id :- rishikmyadagoni35@gmail.com.

ABSTRACT: -

This study presents a comprehensive analysis on the re-construction of crime scene by multiple inputs sources; using Colmap. Which mainly focuses on the basis of LiDAR data, notes, motion sensors, video footage, images from the crime scene. Investigation totally relies on examination, observation and analysis of the physical evidence. In many cases, the crucial evidence remains invisible to eyes, hard to find or vanished to the environmental conditions. Recent developments in AI, computer vision and sensory technology play vital role reconstruction of the crime scene. Methodology was data gathering, preprocessing, feature extraction, Structure-from-Motion (SfM), Multi view stereo (MVS), object identification, mesh creation, post process, and output. It also compares the AI– based construction and traditional construction methods based on authenticity, coherence, latent evidence identification. Phenomena aids in the flawless reconstruction of crime scene. we clearly write the evidence and examine the evidence with high authenticity, then we used cloud fusion and sematic segmentation. Additionally, we used tool like YOLOv3 (2018) [7]. This paper explores the crime scene reconstruction which aids forensic investigation and courtroom presentations, and overall crime scene analysis for effective investigation.

KEYWORDS: - Structure-from-Motion (SfM), Multi view stereo (MVS), mesh creation, cloud fusion, sematic segmentation.

D. R. Reddy et al. (eds.), *Proceedings of the First International Conference on Advances in Forensics and Cyber Technologies (ICFACT 2025)*, Advances in Computer Science Research 127,
https://doi.org/10.2991/978-94-6239-610-4_19

2. INTRODUCTION: -

Crime scene reconstruction and documentation is one of the most crucial and vital roles in forensic investigation, which helps the investigators to identify the crucial evidence. As the traditional technique's, in clue manual sketches, photography and laser measurement devices, often show human error and lack of evidence [14],[15]. Many of the crime scenes contain damaged, hidden, or invisible to the naked eye; which make difficulty in reconstruction and investigation. Then the AI 3d crime scene reconstruction helps with AI Driven images and 3d construction tools [8],[10] to reconstruct the crime scene for the legal proceedings. As pre the recent advancements in artificial intelligence (AI), computer vision, and sensor technologies have changed the dynamics of forensic crime scene reconstruction. We get multi model inputs for reconstruction from forensic photographs, videos, LiDAR scans, drone footage, UR\IR imaging [15],[16], and police notes (sketching and description). By these inputs we can generate and reconstruct the crime scene in detail and realistic 3d models. This process is known as multi model data fusion [8], which helps in understanding the complex crime scene by combining visual and spatial information. In this we use a free open-source tool COLMAP [3], which uses photogrammetry and structure from motion (SFM) which has a ability to reconstruct 3d crime scene from unstructured 2d images and sensors data. It's also uses the multi model inputs such as CCTV, drones, and photographs for the accuracy of reconstruction improves. This paper introduces a complete methodology for AI based 3d crime scene reconstruction by COLMAP. It allows automatic camera position estimation, sparse and dense point generation and reconstruction from image data. It also combines additional sensor inputs and AI based analysis to build realistic 3d crime scene models.

3.LITERATURE REVIEW: -

2.1. Image-based reconstruction: Structure-from-Motion and Multi-View Stereo:

Image-based 3D reconstruction has become a technique in computer vision and photogrammetry, enabling the recovery of 3D scene geometry form 2D images. Most widely used are Structure from Motion (SFM), Multi view Stereo (MVS), are the backbone of reconstruction which include Colmap.

2.1.1. Structure from Motion

A major advancement in SFM was presented by Schonberger and Frahm (2016) [1], in Structure from Motion, which improved accuracy, robustness, Scalability, to enhance matching, reconstruction strategies, and refined bundle adjustment. Their work contributes to image noise, overlap images, and make sfm more reliable for real world examples. This helps to lay a foundation for practical, large scale SFM pipelines. They begin with detection and matching, where key points using SIFT [12] or ORB and match images. These are used to find camera poses and develop a Spare Point cloud [1],[4] for scene geometry. Where it shows accuracy camera position and global structure.

2.1.2. Multi-View Stereo (MVS)

To solve the errors in spare reconstruction, Multi view Stereo (MVS) is applied after SFM to produce dense 3D points clouds [5]. It aids in the depth analysis of each pixel of image through multiple views. This aids in increase the detail and creation of dense point cloud and mesh. Schonberger (2016) has done work on pixelwise view selection for unstructured MVS [2], which improved depth analysis and improved accuracy by selecting optimal image subsets for each pixel. This increased robustness in images with multiple view points and lighting conditions which makes suitable for indoor and outdoor crime scenes. It's also helps in adding details like object boundaries, surface irregularities, spatial evidence relationship. It is sensitive to illumination changes, reflective surfaces (mirrors), and low texture regions [5],[13].

2.1.3. Integration of SFM and MVS in COLMAP

Morden reconstruction pipeline contains SFM and Mvs in workflow. COLMAP is a free open-source tool which in used in reconstruction that combines SFM and Mvs tool algorithm into fully automatic system[3]. Its aid in the estimation of camera position, spare reconstruction, dense point cloud generation and surface reconstruction. Due to its accuracy, open, free, so it is used as backbone in 3D reconstruction. When we combine addition data like LiDAR or AI sematic models, SFM-Mvs pipeline provide a strong geometric for multi model 3D reconstruction of any crime scene.

2.2. Semantic Understanding in Reconstructed Scenes

Algorithms such as Structure-from-Motion (SfM) and Multi-View Stereo (MVS) focus on recovering the spatial structure of a scene, they do not inherently provide semantic meaning [6]. A reconstructed 3D model typically consists of point clouds or meshes without contextual information about what the objects represent. For forensic investigations, however, understanding *what* an object is—such as a weapon, bloodstain, or entry point—is often more critical than simply knowing its location. This limitation has led to increased research interest in semantic understanding within reconstructed scenes.

Semantic understanding refers to the ability of a system to identify, classify, and label objects or regions within a 3D environment. Early approaches relied on manual annotation, which was time-consuming and prone to human error. With advances in artificial intelligence, particularly deep learning, automated semantic analysis has become increasingly feasible. Convolutional Neural Networks (CNNs) have been widely adopted for 2D object detection and segmentation [17] tasks, enabling the identification of forensic-relevant objects in images and video frames.

In forensic applications, semantic understanding plays a crucial role in evidence detection and analysis. Research has shown that AI-assisted semantic labelling [8],[18] can improve the detection of weapons, body outlines, bloodstain patterns, and scene boundaries, even when such evidence is partially occluded or visually subtle. By combining semantic interpretation with geometric reconstruction, investigators gain a more comprehensive and objective understanding of crime scenes. Despite significant progress, challenges remain, including handling occlusions, class imbalance, and ensuring robustness across diverse environments. These limitations highlight the need for continued research into AI-driven semantic integration for reliable forensic reconstruction.

3. METHODOLOGY:

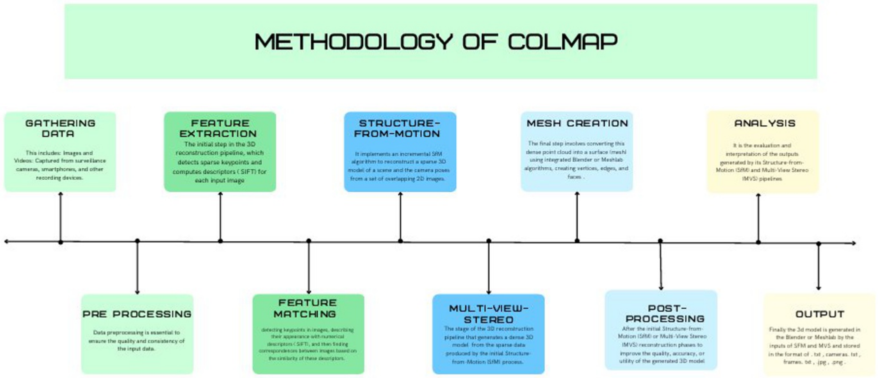

fig 1 :- Method of colmap.

The AI – Driven forensic analysis includes a series of interconnected steps to enhance crime scene reconstruction through the application called Colmap. The description of each step is mention below in brief explanation highlighting the techniques used in it according through fig 1.

3.1. DATA GATHERING:

First step involves the collection of the data from various sources associated with the crime scene. The includes Images, videos, notes, LiDAR Images, motion sensor data.

Textual data: - witness statements, police notes, sketching etc.

Image data: - photography, device photos.

Video data: - CCTV footage, recordings, smartphone videos also include LiDAR image.

Sensory data: - motion sensor data, IoT devices data [15],[16].

3.2. PREPROCESSING DATA: -

It is the critical step in the 3d reconstruction, in this the enhancement of the image is done. In this the high-resolution images are generated or captured from multiple viewpoints using cameras. The image overlap is between 60% to 80% [13],[14],[15] is maintained to feature matching the image for the 3d generation. And the images are captured from hr different angles and hights to add the details to the image. The lighting, motion blur of the image is corrected and shadow, reflections will be cleared. And noise of the image is corrected by the mathematical equations(matrix). And the images are ensured to be captured with same focus and exposure. Additionally, the meta data of the image is preserved, to analyse the parameters of the camera.

3.3. FEATURE EXTRACTION: -

It is the core step in the Colmap scene reconstruction in which it identifies unique patters with the image. It relies on the Scale-Invariant Feature Transform [12] D.G. Lowe(2004). These helps in identifying unique image locations, include conners, edges, and textured regions, scales, rotation. It helps in the examination of the multiple images captured from different angles. The core point in feature extraction is to support feature matching of image and estimate camera positions. Limitations are found if the all images or not captured from same angle, and have changes in the lighting of the image [12],[5].

3.4. FEATURE MATCHING: -

After the future extraction next crucial step is the feature matching, in this the images and matched between the overlapping images using pairwise matching technique. It mainly aids on the sequential matching, according to the order of images [3],[1]. It is used to verify the robustness and gematric of image by applying Random-Sample-Consensus (RANSAC) [11] to delete the incorrect images. Accurate feature matching aids in the estimation of the camera position and starts triangulation of 3d points in SFM. By removing the borders and preserve high quality matches [11],[1], it improves reconstruction and stability. Its aids the verification of generation during Spare point cloud process. And increase the quality of Sparse and dense 3d models.

3.5. STRUCTURE-FROM-MOTION: - (Sparse reconstruction)

After the data processing in the feature matching the SFM is the core it works on the estimation of the camera parameters and helps in reconstruction of the scene into 3d structure by multiple 2d images. Its aids in the identification of orientation and position of the camera. The recorded image points are triangulated to generate a Sparse 3d point cloud which shows structure of the scene. If any additional images are added, then they are analysed with camera position and aids in accurate reconstruction. To improve accuracy, bundle adjustment is applied to optimize intrinsics, extrinsic, and 3d point locations by minimizing the errors. It captures the overall spatial layout of the scene and provide the foundation for the dense reconstruction using Multi-View-Stereo techniques in various stages[2],[5].

3.6. MULTI-VIEW-STEREO: - (Dense reconstruction)

After the SFM Sparse point cloud then MVS stage will start which aids in the detailed dense 3d reconstructed scene. It uses the accurately estimated camera position and Sparse point cloud from SFM stage, MVS computes depth information of each image by analysing pixels across the multiple viewpoints. Colmap perform dense image matching to generate per-image depth maps, capturing fine-grained surface details that not show Sparse reconstruction. These death map consist of fused dense 3d point cloud of the scene. It helps in filling gaps between the Sparse points, resulting in the realistic 3d models. By analysing the multiple positions, it improves accuracy and noise. It serves as core input for surface reconstruction and texture mapping, which helps in visualization and analysis of real-world scenes [2],[5].

3.7. MESH CREATION: - (Surface Reconstruction)

Surface reconstruction is the process which convert the 3d point cloud generated by MVS into a continuous surface representation. It shows the relation between the parallel points and inter surface connectivity and topology. Colmap gives depth and point cloud data which is processed by surface reconstruction algorithms, like passion surface reconstruction to generate Mesh. This mesh shows physical structure of the scene with perfect surfaces and borders. Noise and outline points are cleared to improve smoothness of the Mesh. It aids in the spatial interpretation by converting district point measurements to convert able 3d model. These structure helps for visualization, measurements, and future analysis of the reconstructed scene. The reconstructed surface also aids in analysis of texture mapping, allowing realistic visual presentation of the crime scene in various stages of 3d reconstruction pipeline. Here we use Blender to create the Mesh [5],[12].

3.8. POST PROCESSING OF DATA: -

It is the final stage of the 3d reconstruction of the workflow, which aids in the refining and preparing the model for analysis and presentation. The generated dense point cloud and surface mesh are inspected to remove noise, outlines, borders, and data point, which helps in removing the reconstructed errors. Mesh cleaning options like hole filling, smoothing, and simplification are applied to improve structural details. Scale correction and alignment will work measurements are available [15],[16], to show metric accuracy. Textual maps are used to enhance visual quality. The finalized 3d model is sent for analysis using tools and investigative methods.

3.9. ANALYSIS: -

The reconstructed 3d model is analysed to evaluate accuracy, completeness of crime scene. This includes camera position analysis, point cloud density, error correction, of the generated Mesh. The Sparse reconstruction obtained through Structure-from-motion is examined to verify proper camera alignment, and sufficient scene coverage. And the dense point cloud generated by Multi-View Stereo is analysed for surface continuity and reduced noise levels especially in the low texture. Quantitative measurements, such as distances and angles to assess metric accuracy. Visual inspection of textured surface is performed to identify artifacts, missing regions.

3.10. OUTPUT: -

The output of the proposed Colmap based reconstruction workflow consists of multiple structure and data products. The pipeline generates a Sparse 3d point cloud along with estimate camera intrinsic and extrinsic parameters, representing geometry and camera configuration of the scene. 3D point cloud produced through Multi-view Stereo, capturing fine-grained surface detail. From dense point cloud, a polygonal surface Mesh is reconstruction. High- resolution texture maps derived from the original images are applied to mesh, resulting in photorealistic 3d model. Additional outputs refer depth maps, camera angles. All the outputs are stored in standard formats such as .ply, .bin, .obj. These outputs support for forensic investigation.

4. RESULT: -

This study assists the AI- assisted, Colmap based image reconstruction pipeline for 3d crime scene model. The pipeline generates accurate and detailed 3d model of indoor crime scene. The reconstructed crime scene gives both Sparse and dense point cloud, which aids in analysing the scene. The SFM stage gives the camera position and generate Sparse point cloud. And Multi – view Stereo it gives dense point cloud with surface details, which give detailed view of walls, furniture, and objects. The 3d reconstructed model gives measurements of distance and spatial relationship between objects which helps in forensic analysis. Compared to traditional manual photography, the Colmap reduces the reconstruction time [3],[14]. The automated pipeline minimizes human error and potential measurement error. The reconstructed crime scene supports object level interpretation, enable identification within 3d environment [8],[6],[18]. This aids the detailed reconstruction which do not visible in 2d images. It indicates AI- driven, image-based reconstruction using Colmap provides and efficient and reliable framework for crime scene documentation. It enhances scene understanding, support investigation, and improve visualization for forensic report and courtroom presentation.

5. LIMITATIONS: -

1. The quality of the reconstructed model is dependent on the quality of images. Insufficient image overlap, motion blur, poor lighting conditions, or low-resolution images and feature detection and matching, leading to inaccurate reconstruction.

2. Image based reconstruction method like Structure from motion and Multi view stereo in scene with low texture or reflective surface, includes walls, glass, mirrors, and shiny objects, found in indoor scene, which can result in noise point clouds [5],[13].

3. The methodology depends on the datasets not on forensic case data. So, the result reconstruction accuracy is qualitative than quantitative, and chain of custody are not addressed.

4. Errors in the object detection in 2d images can propagate into 3d model, affect accuracy. The system also do not currently support real time construction and it is time consuming for large datasets.

These limitations highlight area for future improvement, integrating additional sensors, improve accuracy under challenging circumstances and verify the approach using real world forensics cases.

6. CONCLUSION AND FUTURE WORK: -

This research presented an AI assisted framework for 3d crime scene reconstruction using multi model inputs, with Colmap serving as the primary image-based reconstruction tool. By structure from motion and multi view stereo, the pipeline generates the accurate and detailed 3d model of the crime scene from multi model inputs. It shows the effective identification of the camera position, reconstruction Sparse and dense point cloud, and generate realistic models for forensic analysis. Compared to traditional manual photogrammetry, the proposed approached reduces the time and human error and helps in analysis of the accurate 3d image from 2d image [13],[14].

6.1 FUTURE WORK: -.

1. Integrated Lidar and depth sensors [15],[16] with image-based reconstruction can improve in low-texture.
2. Real time reconstruction can be explored to support rapid on-site analysis.
3. Sematic segmentation models can be trained [6],[18] specifically on forensic data set to improve evidence detection.
4. The proposed framework using real world forensics case data establish metrics which enhance legal investigation and court room proceedings.

Acknowledgement:

We would like to express our sincere gratitude to **Mr. Vinod Kaaparthi**, Department of Digital Forensic Science, Malla Reddy University, for his valuable guidance, constructive suggestions, and continuous support throughout the course of this research. His academic expertise and feedback significantly contributed to the direction, methodology, and overall quality of the study. We also acknowledge the academic environment and resources provided by Malla Reddy University, which facilitated the successful completion of this work.

REFERENCE: -

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

2. J. L. Schönberger, E. Zheng, M. Pollefeys, and J.-M. Frahm, “Pixelwise View Selection for Unstructured Multi-View Stereo,” Proceedings of the European Conference on Computer Vision (ECCV), pp. 501–518, 2016.

3. J. L. Schönberger and M. Pollefeys, “COLMAP: A General-Purpose and Scalable Structure-from-Motion and Multi-View Stereo Pipeline,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

4. N. Snavely, S. M. Seitz, and R. Szeliski, “Photo Tourism: Exploring Photo Collections in 3D,” ACM Transactions on Graphics (SIGGRAPH), vol. 25, no. 3, pp. 835–846, 2006.

5. Y. Furukawa and J. Ponce, “Accurate, Dense, and Robust Multi-View Stereopsis,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1362–1376, 2010.

6. C. Qi, H. Su, K. Mo, and L. J. Guibas, “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation,” Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 652–660, 2017.

7. J. Redmon and A. Farhadi, “YOLOv3: An Incremental Improvement,” arXiv:1804.02767, 2018.

8. C. Villa, M. Olsen, and C. B. Hansen, “A Virtual, 3D Multimodal Approach to Victim and Crime Scene Reconstruction,” Forensic Science International, vol. 336, 2023.

9. R. M. Carew and S. Errickson, “Imaging the Past: Recent Applications of 3D Imaging and 3D Printing in Archaeology and Forensic Science,” Journal of Archaeological Science, vol. 136, 2021.

10. D. Maria Manuel Vianny, R. Mandonna Arieth, S. Manikandana, and Sugumar, “AI Driven Forensic Analysis: Enhancing Crime Scene Reconstruction with Machine Learning,” TIJER, vol. 12, issue 5, 2025.

11. R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed., Cambridge University Press, 2004.

12. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.

13. A. Turner, "Structure from Motion Photogrammetry: A Low-Cost, Effective Tool for Geoscience Applications," Geomorphology, vol. 179, pp. 300–314, 2012.

14. M. Buck, S. Naether, and J. Albertz, "Photogrammetric Documentation of Crime Scenes," The Photogrammetric Record, vol. 23, no. 124, pp. 369–385, 2008.

15. C. Robinson et al., "The Use of 3D Laser Scanning for Crime Scene Reconstruction," Forensic Science International, vol. 202, no. 1–3, pp. 1–7, 2010.

16. J. Shan and C. K. Toth, Topographic Laser Ranging and Scanning: Principles and Processing, CRC Press, 2008.

17. L.-C. Chen et al., "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 4, pp. 834–848, 2018.

18. T. Hackel, J. D. Wegner, and K. Schindler, "Fast Semantic Segmentation of 3D Point Clouds," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. IV-1/W1, pp. 1–8, 2017.