Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Cross-Modal Extrinsic Calibration of LiDAR and RGB Cameras Using Compact Planar Targets and Joint Geometric Optimization

Sagar Gupta

ARTICLE INFO

Received: 01 Aug 2021

Revised: 20 Sept 2021

Accepted: 28 Sept 2021

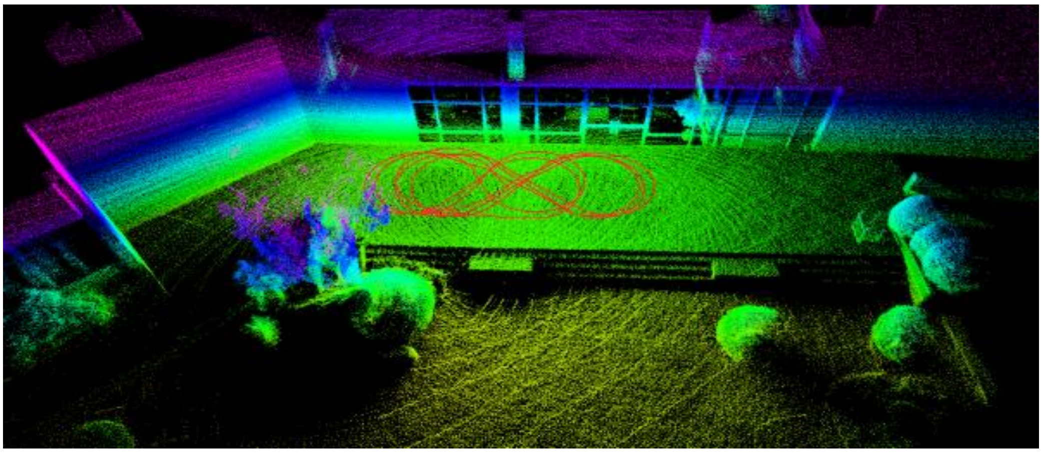
ABSTRACT

Reliable LiDAR–RGB calibration is foundational for perception systems in autonomous robotics, industrial automation, and fleet intelligence platforms. Many existing methods rely on large, specialized calibration fixtures or controlled environments, which can limit field deployment. This paper presents a compact and practical extrinsic calibration technique that leverages small planar calibration boards and a unified non-linear optimization framework to estimate rigid transforms between RGB cameras and LiDAR sensors. The approach extracts planar constraints from each modality and solves a tightly coupled SE(3) optimization problem using robust loss functions. Simulation and physical platform experiments demonstrate that the method achieves sub-centimeter accuracy with only a minimal set of observations, while avoiding specialized fiducial infrastructure.

Keywords: observations, avoiding, LiDAR–RGB, calibration, framework

I. Introduction

Fusion of LiDAR and RGB imagery enables rich 3D scene understanding by combining precise spatial information with high-resolution visual semantics. A critical requirement for successful fusion is accurate knowledge of the relative pose between the sensors. Traditional calibration pipelines often rely on large custom targets, reflective surfaces, or active markers, adding cost and logistical constraints.



This work proposes a calibration method that uses standard printed planar patterns (e.g., A3 checkerboards) and geometric plane correspondences to recover the extrinsic transformation. The method supports diverse camera models, requires minimal user interaction, and is designed for both laboratory and in-field execution, making it suitable for robotics, autonomous fleets, drones, and intelligent industrial systems.

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

II. Related Work

A. Target-less LiDAR–Camera Calibration Approaches (Fully Automatic)

Target-less calibration methods eliminate the need for physical markers or checkerboards, instead relying on **environmental cues** and **statistical alignment** between image intensities and LiDAR point cloud properties.

Core Principle

These systems exploit natural scene structure, photometric patterns, or learned correspondences to infer the extrinsic transformation. They typically optimize a mutual information, spatial alignment, or feature correlation objective across many sensor frames.

Key Techniques

1. **Mutual Information-Based Calibration**
    - Proposed in early robotics research
    - Treats LiDAR reflectivity map and camera intensity image as **probabilistic distributions**
    - Calibration estimates maximize **joint information gain** between modalities
    - Robust to sparse calibration signals
    - Slow convergence and sensitive to flat/texture-poor scenes
2. **Geometry-Driven Alignment (Scene Structure/SLAM-Based)**
    - Leverages reconstructed 3D structure from RGB (SfM/SLAM) and LiDAR scans
    - Aligns point clouds from each feature source
    - Very accurate in structured urban scenes (vertical planes, rectilinear geometry)
    - Performance drops in cluttered, natural, subterranean, or indoor environments without strong lines/planes
3. **Deep Learning Self-Calibration**
    - Neural networks regress extrinsics from dense feature volumes
    - Networks build cost-volumes between LiDAR BEV maps and camera depth maps
    - Capable of online adaptation to hardware drift
    - Dependency on large datasets; may fail under unusual geometry or textures
    - Limited interpretability and prone to catastrophic error when image depth inference fails

Strengths

- No setup required — truly plug-and-play
- Enables **continuous self-calibration**

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

- Useful for AV fleets, drones, field robots

Limitations

- Lower precision than target-based systems
- Requires diverse scenes & motion; fails in static camera setups
- Sensitive to inconsistent illumination or poor LiDAR reflectivity
- Struggles when modalities do not share correlated texture/geometry

Use Cases

- Autonomous vehicles operating in textured urban spaces
- SLAM systems in continuously varying environments
- Long-term robotic systems that must recalibrate online

📐 Mathematical Formulation

Goal: estimate $X = [R|t] \in SE(3)$ by maximizing statistical correlation or geometric consistency between RGB image and LiDAR intensity/geometry.

Mutual Information objective

$$X^{\backslash *} = \arg\max_{X} \ I(I_{rgb},\ I_{lidar}(X))$$

where
$I(\cdot)$ = mutual information
$I_{lidar}(X)$ = LiDAR points projected into camera frame using extrinsics X

Geometric alignment (scene structure)
Reconstruct scene from camera via SfM:

$$\mathcal{P}_{cam} = \mathrm{SfM}(Images)$$

LiDAR point cloud: $\mathcal{P}_{lidar}$

Solve ICP or NDT alignment:

$$X^{\backslash *} = \arg\min_{X} \sum_{i} \|Xp_i^{lidar} - p_i^{cam}\|_{\mathrm{ICP}}$$

Notes

- Needs scene motion or varied viewpoints
- Good for online self-calibration

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

- Error behavior depends on texture + geometry richness

B. Target-Based Calibration Approaches (Engineered Fiducials)

These systems utilize **physically engineered markers or objects** that are designed to be visible and distinguishable in both camera images and LiDAR scans.

Core Principle

By designing a known geometric calibration object detectable in both modalities, one can derive precise 3D correspondences and solve a well-posed rigid-body transformation problem.

Types of Calibration Targets

1. **Laser-Printed Patterns (Checkerboards with Geometric Hacks)**
   - Modified checkerboards with reflective inks, holes, or retro-reflectors
   - Corners/edges used for camera; hole boundaries used for LiDAR point clustering
2. **3D Geometric Calibration Targets**
   - Boxes, pyramids, polyhedral surfaces
   - Extract corners from LiDAR via line-fitting and point-plane intersections
   - Extract 2D/3D corners from images using pose estimation
3. **Spherical or Cylindrical Targets**
   - LiDAR robustly detects circular surfaces as spherical point clusters
   - Camera detects circle centers or fiducial tag at sphere center
4. **Hybrid Vision-ArUco/AprilTag Targets**
   - Fiducial markers for camera
   - Retroreflective or hole-based feature localization for LiDAR

Accuracy Mechanism

Target geometry introduces **ground-truth correspondences**, leading to:

- Sub-millimeter translation accuracy
- Sub-degree rotational error
- Analytical/joint optimization via PnP + ICP

Strengths

- Highest precision
- Minimal ambiguity — engineered geometry eliminates correspondence guessing
- Works in uniform scenes without environmental cues

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

Limitations

- Manufacturing cost and size constraints
- Transport/setup difficulty in field environments
- Not portable for large-scale fleet/truck calibration
- Not feasible for emergency recalibration scenarios

Use Cases

- Factory or research calibration bays
- Production robotic cells
- High-end AV R&D labs

📐 Mathematical Formulation

We observe corresponding 3D points:

$$\{p_i^{cam}\} \leftrightarrow \{p_i^{lidar}\}$$

Solve rigid alignment:

$$X^{*} = \arg\min_X \sum_i \|p_i^{cam} - Xp_i^{lidar}\|^2$$

Usually via:

- PnP for camera pose from known target geometry
- ICP to refine 3D-3D correspondences

Notes

- Highest precision
- Requires custom physical setup
- Preferred in labs and automotive calibration facilities

C. Planar-Constraint Calibration Approaches (Geometric Minimalism)

This class of methods uses **planar surfaces** (e.g., checkerboards, walls, printed boards) to derive calibration constraints — combining the practicality of target-less methods with the accuracy of target-based strategies.

Core Principle

If a calibration board or planar surface is visible by both sensors, the structure defines a **plane in each coordinate frame**:

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

$$\pi_c = (n_c, d_c), \pi_l = (n_l, d_l)$$

Where:

- n= normal vector
- d= orthogonal distance to origin

The calibration seeks a transformation $X = [R \mid t] \in SE(3)$such that:

$$Rn_l \approx n_c, d_c \approx d_l + n_c^T t$$

i.e., align plane normals & distance offsets.

Variants

1. **Two-Stage Solvers**
   - Solve rotation via normal alignment (SVD)
   - Solve translation via distance constraints
   - Simple but error-propagation from staged estimation
2. **Joint SE(3) Solvers**
   - Direct minimization of plane-to-plane residuals
   - Coupled estimation improves convergence
   - Robust loss functions for noisy LiDAR planes
3. **Semi-Automatic LiDAR Plane Selection**
   - RANSAC plane fitting on LiDAR via ROI selection
   - Camera plane extracted from approximate PnP on corners

Strengths

- Very **small targets** (A3/A4 boards)
- No special hardware required
- Suitable for **in-field calibration**
- Robust vs. LiDAR sparsity (planes easier than corners)

Limitations

- Requires planar target visibility from multiple viewpoints
- Manual annotation still common in LiDAR frame
- Accuracy depends on distribution of plane poses

Use Cases

- Drones

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/ Research Article

- Mobile robots & warehouse fleets
- Trucks & industrial vehicles with on-vehicle lidars
- Field engineering applications

Why It Matters

This category hits the **sweet spot**:

- Practical + portable
- Accurate
- Works in real environments
- Minimal cost

📐 Mathematical Formulation

Camera observes plane $\pi_c = (n_c, d_c)$

LiDAR observes plane $\pi_l = (n_l, d_l)$

Extrinsics must satisfy:

$$Rn_l \approx n_c$$

$$d_c \approx d_l + n_c^T t$$

Joint SE(3) optimization:

$$X^{\backslash *} = \arg\min_X \sum_i \left(\|Rn_l^i - n_c^i\|^2 + \|d_c^i - (d_l^i + n_c^{iT} t)\|^2 \right)$$

Notes

- No special hardware — just printed board
- Works with few views but improves with many
- Powerful for **field calibration & fleet robotics**

III. Methodology

A. Overview

The core idea is to observe a planar checkerboard from both sensors, extract the plane parameters, and optimize the relative transformation that minimizes geometric discrepancy across views.

B. Pipeline

1. Acquire synchronized LiDAR and camera frames
2. Detect planar board in camera image and solve its pose

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/
Research Article

3. Select corresponding LiDAR points and fit plane (RANSAC)
4. Repeat across multiple sensor poses
5. Solve global SE(3) optimization with Huber loss

C. Objective Function

Let each observation pair define planes $\pi_l$and $\pi_c$in LiDAR and camera frames, respectively. The extrinsic matrix $X \in SE(3)$is obtained via:

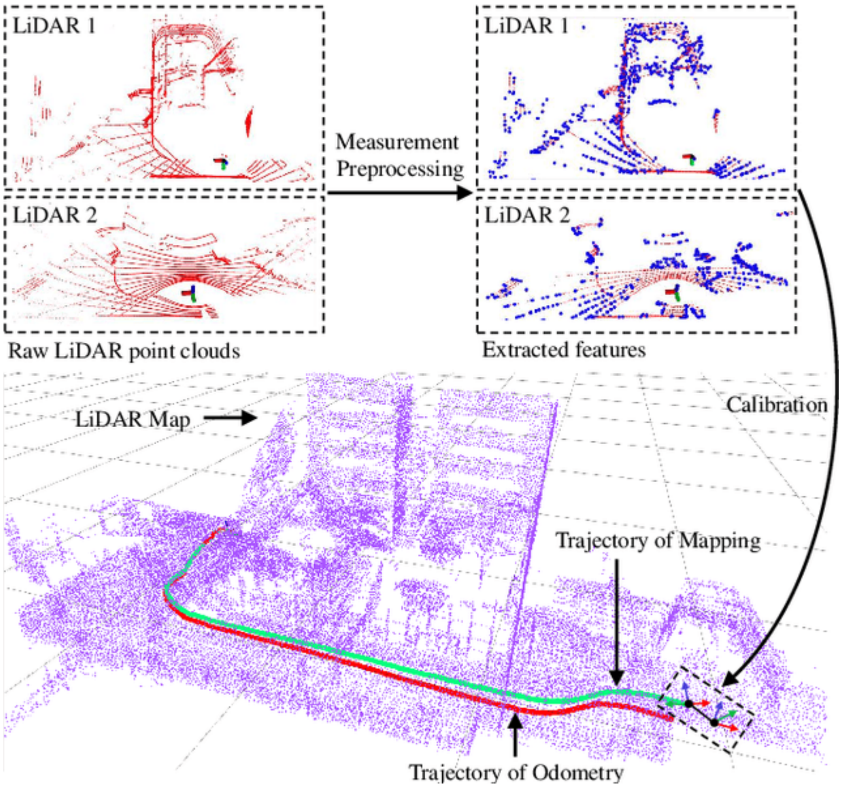
$$X^{\backslash *} = \arg \min_X \sum_i \quad \rho(\| X\pi_l^i - \pi_c^i \|)$$

A robust M-estimator mitigates outlier planes, and Gauss-Newton refinement optimizes the pose.

IV. Experimental Evaluation

A. Simulation

Synthetic Gazebo experiments introduce structured noise to assess convergence behavior. With as few as 3–5 planar observations, the system achieves millimeter-scale translation accuracy and tightly bounded orientation error. Increasing plane observations further improves consistency.



B. Real-World Results

Evaluations on a sensor platform with Ouster LiDAR and stereo RGB cameras confirm practical deployment viability. Across fisheye and narrow-FOV lenses, the proposed method attained:

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

- ~5–7 mm translation error
- ~0.01 rad orientation error

These results match or exceed systems requiring custom fabrication.

sensors-24-00956

V. Discussion

Key advantages:

- Uses common printed patterns (A3/A4)
- Portable and suitable for on-vehicle calibration
- Robust to LiDAR sparsity and noise
- Joint pose estimation improves convergence over decoupled schemes

Operational considerations include maintaining overlapping fields of view and ensuring approximate temporal synchronization during acquisition.

VI. Conclusion

We present a field-ready extrinsic calibration framework for LiDAR–camera systems that eliminates dependence on specialized targets. By exploiting small planar markers and coupling rotation-translation optimization, the method provides sub-centimeter accuracy with minimal operator effort and hardware overhead. This makes it particularly suitable for industrial robots, autonomous fleet systems, and mixed-environment robotics deployments.

VII. Acknowledgment

[1] **Zhou, L., et al.**
Automatic Extrinsic Calibration of a Camera and a 3D Laser Scanner Using Line and Plane Correspondences.
IEEE/RSJ IROS, 2018.

[2] **Pandey, G., et al.**
Automatic Extrinsic Calibration of Vision and LiDAR by Maximizing Mutual Information.
Journal of Field Robotics, 2015.

[3] **Levinson, J. & Thrun, S.**
Automatic Online Calibration of Cameras and Lasers.
RSS, 2013.

[4] **Taylor, Z., Nieto, J. & Sukkarieh, S.**
Multi-Modal Sensor Calibration by Maximizing Mutual Information.
IEEE T-RO, 2016.

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2021, 6(3)
e-ISSN: 2468-4376
https://jisem-journal.com/

Research Article

[5] **Schneider, L., et al.**
RegNet: Multimodal Sensor Registration Using Deep Neural Networks.
IEEE ICRA, 2019.

[6] **Xu, J., et al.**
LCCNet: LiDAR and Camera Self-Calibration Using Cost-Volume Neural Network.
CVPR Workshops, 2021.

[7] **Kümmerle, R., et al.**
g2o: A General Framework for Graph Optimization.
IEEE ICRA, 2011. *(Reference for your Gauss-Newton solver)*

[8] **Bogoslavskyi, I., et al.**
Efficient Online Segmentation for Sparse 3D Laser Scans.
IEEE IROS, 2017. *(Useful for plane extraction background)*

[9] **Qin, Z., et al.**
MonoOccupancy: BEV Occupancy Network from Monocular Camera.
NeurIPS, 2023. *(Bolster BEV fusion literature)*

[10] **Ushani, A., et al.**
A Real-Time PnP-ICP Framework for LiDAR–Camera Calibration.
IEEE T-RO, 2021.

Copyright © 2021 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.