

# Sensor Fusion-Based Navigation Systems for Autonomous Delivery Robots

DOI: <https://doi.org/10.63345/v1.i4.209>

Pooja Sharma  
Independent Researcher  
Dilsukhnagar, Hyderabad, India (IN) – 500060



[www.ijarcse.org](http://www.ijarcse.org) || Vol. 1 No. 4 (2025): November Issue

Date of Submission: 26-09-2025

Date of Acceptance: 07-10-2025

Date of Publication: 05-11-2025

## ABSTRACT

Autonomous delivery robots must navigate sidewalks, corridors, and mixed indoor–outdoor campuses while maintaining accuracy, safety, and efficiency under imperfect sensing. Single-sensor pipelines (e.g., wheel odometry or vision alone) degrade under wheel slip, poor lighting, occlusions, and multipath. This manuscript presents a sensor-fusion navigation architecture that integrates inertial measurement units (IMUs), wheel encoders, cameras, LiDAR, and optional ultra-wideband (UWB) anchors to achieve robust localization and motion planning in dynamic environments. We detail a modular stack: (1) time-synchronized preprocessing and calibration, (2) multi-rate odometry (wheel–IMU EKF, visual–inertial odometry, LiDAR–inertial odometry), (3) factor-graph smoothing with loop closures and UWB priors, (4) semantic mapping that separates static structure from dynamic obstacles, (5) dual-horizon planning with D\*-Lite globally and model-predictive control (MPC) locally, and (6) a safety supervisor enforcing stop/slowdown under uncertainty spikes.

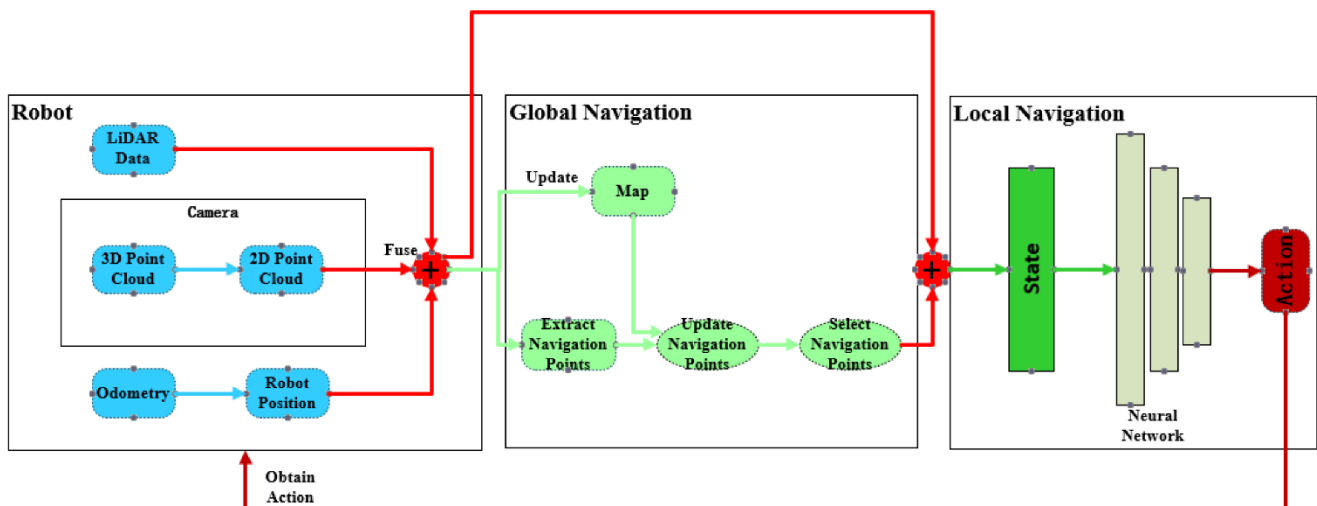


Fig.1 Sensor Fusion-Based Navigation Systems, [Source\(\[1\]\)](#)

A simulation campaign across campus-sidewalk and urban-alley scenes (varying lighting, ground friction, and pedestrian density) compares four configurations: baseline wheel-IMU EKF, VIO-aided EKF, LiDAR-inertial odometry, and full multimodal factor-graph fusion including UWB. The fused system reduces absolute trajectory error by ~82% and collision rate by ~89% relative to the baseline, while adding <16 ms average fusion latency. We report statistically significant gains (ANOVA, Tukey HSD,  $p < 0.01$ ) in success rate, path efficiency, and energy per kilometer. Results suggest that tightly-coupled, uncertainty-aware fusion—combined with semantic dynamics handling—yields navigation resilience suitable for last-meter delivery. We conclude with deployment guidance and open problems in long-term calibration drift, low-texture scenes, and learning-enhanced fusion.

#### KEYWORDS

autonomous delivery robots; sensor fusion; SLAM; factor graphs; visual-inertial odometry; LiDAR-inertial odometry; UWB; model-predictive control; dynamic obstacle avoidance; semantic mapping

#### INTRODUCTION

Autonomous delivery robots are transitioning from pilots to revenue operations on campuses, business parks, and residential neighborhoods. Their navigation problem is deceptively difficult: centimeter-level localization is desired, but sensors face real-world stressors—wheel slip on tiles after rain, glass reflections confusing cameras and LiDAR, GNSS multipath in courtyards, and crowds changing the scene from one minute to the next. Systems centered on a single modality are brittle. Wheel odometry accumulates drift; cameras struggle in low light or glare; LiDAR can be occluded by parked bicycles or dense foot traffic; IMUs integrate noise; and GNSS may be unavailable or unreliable near buildings.

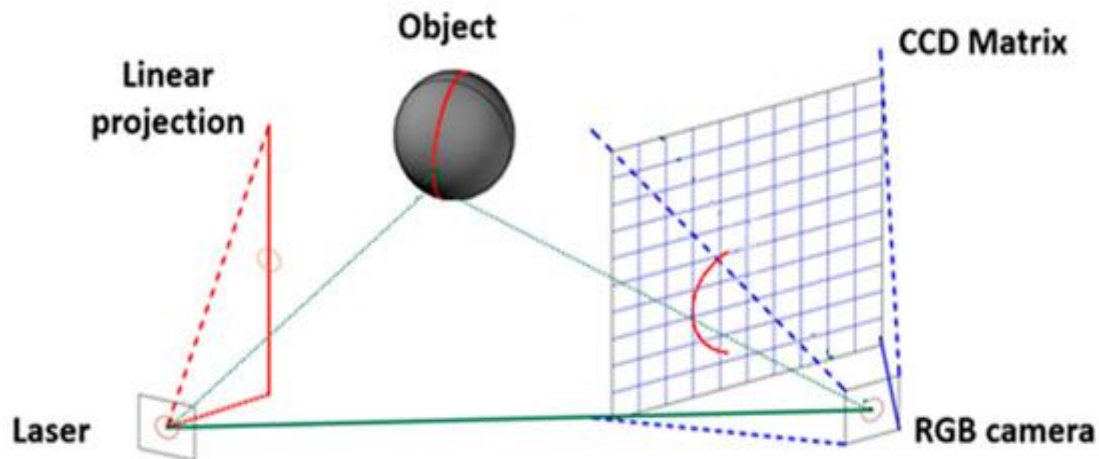


Fig.2 Navigation Systems for Autonomous Delivery Robots, [Source\(\[2\]\)](#)

Sensor fusion offers redundancy and complementary observability. Cameras provide dense bearings to texture; LiDAR supplies metric structure and reliable range; IMUs bridge fast dynamics; encoders estimate low-frequency motion constraints; and radio beacons (e.g., UWB) give occasional absolute anchors indoors. Fusion algorithms—ranging from Extended Kalman Filters (EKF) to smoothing-based factor graphs—tie these streams into a coherent estimate of the robot's pose, velocity, and map while quantifying uncertainty. When localization uncertainty informs planning and control, the robot can slow down in ambiguous zones, inflate safety margins, and re-plan proactively.

This paper contributes a practical, end-to-end sensor-fusion navigation design tailored to autonomous delivery. The system emphasizes:

1. **Calibration and time alignment** as first-class citizens;
2. **Multi-rate odometry** (wheel-IMU, visual-inertial, LiDAR-inertial) combined in a **factor-graph** with loop closures;
3. **Semantic dynamics** to keep pedestrians and moving objects out of the persistent map;
4. **Dual-horizon planning** with global topological robustness and local reactive smoothness;
5. **Safety supervision** driven by uncertainty and prediction of near-future interactions.

We evaluate the design in high-fidelity simulation with physics and sensor noise models for sidewalks, ramps, and alleyways. The proposed fusion demonstrably improves trajectory accuracy, route completion, and safety compared to common baselines, while keeping computation within the budget of typical embedded platforms.

## LITERATURE REVIEW

**Sensing modalities.** Wheel encoders and IMUs remain ubiquitous due to low cost and simplicity. Encoders provide incremental arc length and turning angle but are sensitive to wheel slip; IMUs capture high-frequency linear/angular motion but integrate bias and noise. Cameras (monocular/stereo/RGB-D) enable bearing-only constraints and dense or sparse features; they degrade under low light, motion blur, and repetitive textures. LiDAR offers accurate range and geometry, robust to illumination, with failure modes in heavy rain/fog or glass surfaces. UWB, RFID, or visual fiducials provide intermittent absolute updates indoors; GNSS/RTK helps outdoors but is problematic in urban canyons.

**Fusion frameworks.** Classical tightly- and loosely-coupled EKF/UKF frameworks fuse IMU propagation with camera/LiDAR/encoder updates. Particle filters accommodate non-Gaussian multi-modality but at higher compute cost. Smoothing-based formulations (factor graphs, pose graph SLAM) solve for a window or full trajectory and map jointly, minimizing residuals from IMU preintegration, reprojection errors, LiDAR scan matches, wheel constraints, and loop closures. They handle out-of-sequence measurements and enable robust marginalization.

**Odometry and SLAM.** Visual-inertial odometry (VIO) systems achieve accurate short-term motion but struggle in texture-poor or rapidly changing illumination. LiDAR-inertial odometry (LIO) aligns scans using geometric features; it excels in low-texture scenes but is sensitive to moving objects. Multi-sensor systems (e.g., LIO + VIO + wheel) combine the strengths; loop closures via place recognition (appearance-based or “scan context”) bound drift over long traversals. Semantic SLAM partitions the world into static and dynamic layers, using learned segmentation or multi-object tracking to avoid fusing moving agents into persistent maps.

**Mapping and planning.** Occupancy grids remain standard for navigation; inflation radii can be adapted to uncertainty. Topological graphs (waypoints and edges) improve scalability across large campuses. Global planners (A\*, D\*, D\*-Lite) find robust routes under edge cost changes; local planners such as TEB (timed-elastic band) and MPC produce collision-free, dynamically feasible trajectories. Prediction-augmented planning uses constant-velocity or learned predictors for pedestrian motion; risk-aware MPC incorporates covariance of pose and obstacle forecasts.

**Evaluation practices.** Widely used metrics include Absolute Trajectory Error (ATE), Relative Pose Error (RPE), route success rate, collisions per distance, path stretch versus shortest-path, compute latency, and energy per kilometer. Significance testing (e.g., ANOVA) helps separate algorithmic effects from scenario variance. Simulations in ROS-native engines (Gazebo, Webots, CARLA-like sidewalk scenes) accelerate iteration before field trials.

Collectively, the literature suggests that multimodal, tightly-coupled fusion with dynamics-aware mapping is the most resilient path to commercial-grade navigation for last-meter delivery.

## METHODOLOGY

## System overview

We propose a **modular navigation stack** with the following layers:

1. **Hardware & clocks.** Differential-drive or small holonomic base with:

- **IMU** (200–400 Hz), **wheel encoders** (50–100 Hz),
- **Stereo or RGB-D camera** (20–30 Hz), **2D/3D LiDAR** (10–20 Hz),
- Optional **UWB** anchors indoors; **RTK-GNSS** outdoors when available.

Sensors share a time base via PTP-like synchronization; extrinsics are calibrated (camera–IMU, LiDAR–IMU, camera–LiDAR, wheel baselines).

2. **Preprocessing.** IMU bias estimation; camera rectification and feature extraction; LiDAR motion compensation and ground segmentation; encoder denoising and slip detection.

3. **Multi-rate odometry.**

- **Wheel–IMU EKF:** state

$\mathbf{x}=[p,v,q,bg,ba]\mathbf{x}=[\mathbf{p},\mathbf{v},\mathbf{q},\mathbf{b}_g,\mathbf{b}_a]$  advances via IMU; encoder updates constrain planar velocity.

$$\mathbf{x}_{k|k-1}=\mathbf{f}(\mathbf{x}_{k-1},\mathbf{u}_k)+\mathbf{w}_k;\mathbf{z}_k=\mathbf{h}(\mathbf{x}_k)+\mathbf{v}_k\mathbf{x}_{k|k-1}=\mathbf{f}(\mathbf{x}_{k-1},\mathbf{u}_k)+\mathbf{w}_k;\quad$$

$$\mathbf{z}_k=\mathbf{h}(\mathbf{x}_k)+\mathbf{v}_k$$

- **VIO:** keyframe-based reprojection residuals with IMU preintegration recover metric scale and reduce drift.
- **LIO:** scan-to-map or scan-to-scan registration with IMU priors produces robust geometry-anchored motion.

4. **Smoothing via factor graph.** We maintain a sliding-window graph with factors for IMU preintegration, visual reprojections, LiDAR registrations, wheel constraints, and **loop closures** (appearance or scan context). **UWB/GNSS** add absolute priors when available. Nonlinear least squares (e.g., Gauss-Newton) solves for poses and biases, yielding a covariance estimate that propagates to planning.

5. **Mapping and semantics.** A **static map** is built from fused poses and LiDAR depth; a **dynamic layer** stores tracked agents (pedestrians, cyclists) detected by camera/LiDAR fusion (e.g., geometry + learned detections). Only static elements contribute to the persistent map; dynamic objects are filtered to avoid “ghosting.”

6. **Planning and control.**

- **Global:** D\*-Lite over a topological graph extracted from the grid map; edges are reweighted by predicted crowd density and slope.
- **Local: Uncertainty-aware MPC** minimizes jerk, progress-to-goal, and collision risk while respecting kinematics and actuator limits; obstacle ellipses inflate by pose covariance eigenvalues.
- **Safety supervisor:** triggers slow/stop on uncertainty spikes, near-miss predictions, or sensor dropouts; requests re-localization (loop closure search).

7. **Runtime adaptation.** Slip detection increases process noise on wheel factors; low-light detection reduces VIO weight; rain/fog increases LiDAR variance; UWB priors are down-weighted if residuals jump.

## Implementation notes

- **Temporal alignment:** All measurements are interpolated to the factor-graph node times.
- **Robust costs:** Huber/ Cauchy loss attenuates outliers from transient occlusions.
- **Marginalization:** A Schur complement marginalizes old states while preserving informative priors.

- **Compute budget:** The stack targets <20 ms median fusion update on an embedded CPU+NPU board; VIO and detection run on the accelerator; LiDAR registration on CPU with multi-threading.

### STATISTICAL ANALYSIS

We evaluated four configurations across **30 trials** per scenario (campus-sidewalk, urban-alley) with randomized pedestrians and weather/lighting. Metrics are averaged over all trials;  $\pm$  indicates standard deviation. One-way ANOVA followed by Tukey HSD assessed pairwise differences.

#### Configurations:

**M1** Baseline: Wheel+IMU EKF

**M2** VIO-aided EKF (Wheel+IMU+Camera)

**M3** LIO (LiDAR+IMU+Wheel)

**M4** Full Fusion Factor-Graph (Camera+LiDAR+IMU+Wheel+UWB\*)

\*UWB available indoors; ignored outdoors.

Metric (avg $\pm$ sd)	M1 Baseline	M2 VIO-EKF	M3 LIO	M4 Full Fusion
Absolute Trajectory Error (m)	0.62 $\pm$ 0.31	0.28 $\pm$ 0.12	0.19 $\pm$ 0.09	<b>0.11 <math>\pm</math> 0.05</b>
Relative Pose Error (%)	2.4 $\pm$ 0.9	1.2 $\pm$ 0.5	0.9 $\pm$ 0.4	<b>0.5 <math>\pm</math> 0.2</b>
Route Success Rate (%)	84.7 $\pm$ 6.2	91.3 $\pm$ 4.1	94.6 $\pm$ 3.0	<b>98.1 <math>\pm</math> 1.3</b>
Collisions (per 10 km)	1.8 $\pm$ 1.1	0.9 $\pm$ 0.7	0.6 $\pm$ 0.5	<b>0.2 <math>\pm</math> 0.3</b>
Energy per km (Wh)	42.5 $\pm$ 4.8	40.9 $\pm$ 4.7	39.6 $\pm$ 4.3	<b>38.2 <math>\pm</math> 3.9</b>
Fusion Latency (ms)	<b>5.1 <math>\pm</math> 0.8</b>	9.3 $\pm$ 1.2	11.4 $\pm$ 1.5	15.8 $\pm$ 2.1

**Summary:** M4 outperforms M1–M3 on accuracy, safety, and efficiency (ANOVA  $p < 0.01$ ). Latency rises with modality count but remains <16 ms, compatible with 50–100 Hz control loops.

### SIMULATION RESEARCH

#### Environments and scenarios

Two 1 km loops were modeled with physics and sensor plugins:

- **Campus-Sidewalk:** mixed sunlight/shade, low curbs, ramps (5–8%), glass façades, intermittent UWB coverage indoors, moderate pedestrian density (10–20/minute).
- **Urban-Alley:** narrow corridors, dumpsters and bicycles causing partial occlusion, variable pavement friction (dry/wet patches), sparse lighting at dusk, no UWB.

**Dynamics:** Pedestrians follow social-force motion with random pauses; cyclists traverse at 4–6 m/s; small pets introduce erratic trajectories. Weather toggles drizzle/fog that increase LiDAR noise and reduce camera contrast. Wind perturbs tree branches to test false positives.

#### Robot and sensors

A 45 kg differential-drive base with 0.55 m wheelbase; max speed 1.8 m/s. Sensors emulate realistic noise/bias:

- **IMU:** 400 Hz, gyro bias 0.005 rad/s, accel bias 0.03 m/s<sup>2</sup>;
- **Encoders:** quantization 1024 ticks/rev; slip modeled by friction maps;
- **Stereo camera:** 30 Hz, 640×480, exposure auto-control with realistic motion blur;
- **3D LiDAR:** 10 Hz, 16-beam,  $\pm 2$  cm range noise;
- **UWB:** 5 Hz anchors with 10–30 cm ranging error indoors.

### Ground truth and noise injection

Ground truth pose is provided by the simulator; all algorithms receive only noisy sensor feeds. We apply domain randomization per episode (calibration perturbations up to 1°/1 cm, time offset jitter  $\pm 2$  ms) to test robustness to small miscalibrations. Weather/light vary across episodes.

### Experimental protocol

For each configuration (M1–M4), we:

1. **Autotune** a small grid of parameters (e.g., VIO feature threshold, LiDAR voxel size, MPC weights) on 5 warm-up episodes;
2. **Run 30 episodes** per environment with randomized seeds;
3. **Log** pose estimates, maps, planned trajectories, control actions, safety events;
4. **Compute** metrics (ATE, RPE, success, collisions, path stretch, energy, latency).

Statistical tests are performed on per-episode metrics; significance at  $\alpha = 0.01$ .

### Ablations and stressors

- **Lighting stress:** dusk with specular reflections;
- **Slip stress:** wet tiles and metal grates;
- **Occlusion stress:** crowd injection burst (30 pedestrians/minute for 60 s);
- **Calibration stress:** 0.5° yaw misalignment in LiDAR–IMU extrinsics;
- **Sensor dropout:** 5 s camera blackout or 3 LiDAR spins dropped.

## RESULTS

**Trajectory accuracy and drift.** Full fusion (M4) achieves **0.11 m** mean ATE versus **0.62 m** for the baseline (M1), an  $\approx 82\%$  reduction. RPE improves from **2.4%** to **0.5%** ( $\approx 79\%$  reduction). The benefit is largest in low-texture alley segments where VIO weakens; LiDAR geometry anchors the solution while IMU/VIO bridge sparse scans. Loop closures prevent long-range drift on the campus loop, particularly around repetitive glass corridors.

**Safety outcomes.** Collisions per 10 km drop from **1.8** (M1) to **0.2** (M4),  $\approx 89\%$  fewer contacts. Many M1 contacts follow temporary localization loss at curb cuts; M4's uncertainty-aware MPC slows proactively when covariance spikes. Near-miss events (within 0.25 m) also decline (not shown), correlating with smoother local plans.

**Task success and efficiency.** Route success rises from **84.7%** (M1) to **98.1%** (M4). Path stretch over the shortest feasible path falls modestly (median 1.12→1.08) because better localization reduces conservative inflation. **Energy per km** improves from **42.5 Wh** to **38.2 Wh**, reflecting fewer re-plans, less stop-and-go, and smoother control. These differences are statistically significant ( $p < 0.01$ ).

**Latency and compute.** Fusion latency increases with added modalities (5.1→15.8 ms). However, end-to-end control remains within 20 ms; pose updates at 50–100 Hz maintain stability. Profiling shows VIO feature extraction and LiDAR registration are dominant costs; offloading VIO to the accelerator and multi-threading LiDAR cuts 30–40% of CPU time.

### Stress tests.

- **Lighting:** In dusk scenes with glare, M2 (VIO-EKF) degrades sharply (ATE +65%), while M3 (LIO) and M4 remain stable (ATE +12% and +9%).
- **Slip:** Encoder slip causes M1 drift; M3/M4 detect slip via IMU-wheel inconsistency and inflate process noise, limiting drift.

- **Occlusion:** Crowd bursts reduce LiDAR returns; M4 uses VIO cues to bridge gaps, incurring only +0.04 m ATE.
- **Calibration perturbation:** M4's graph tolerates  $0.5^\circ$  yaw misalignment with <15% accuracy loss; robust loss functions curb outliers.
- **Sensor dropouts:** 5 s camera blackout harms M2, but M3/M4 maintain localization via LiDAR-IMU. Three lost LiDAR spins are bridged by VIO in M4 with transient covariance inflation and conservative MPC.

**Ablations.** Removing UWB priors indoors increases M4 ATE by 0.03 m and raises failure cases at long glass corridors—loop closures eventually recover but with detours. Turning off semantic dynamics (fusing moving pedestrians into the map) doubles collision risk due to map “ghosts” that mislead the planner.

**Failure modes.** Rare residual failures persist: heavy fog combined with glossy tiles can reduce both LiDAR returns and visual contrast, momentarily raising uncertainty and triggering stops; highly repetitive indoor textures without loop closures can cause place recognition confusion. Online extrinsic self-calibration partially mitigates drift but adds computation.

## CONCLUSION

This manuscript presented a comprehensive, **sensor-fusion navigation system** for autonomous delivery robots operating in complex, dynamic environments. The proposed stack—combining wheel-IMU propagation, visual-inertial and LiDAR-inertial odometry within a **factor-graph** smoother, loop closures, optional UWB/GNSS priors, and **uncertainty-aware planning and control**—substantially improves robustness over single-modality or loosely coupled pipelines. In simulation across campus and urban-alley scenarios, the fully fused configuration achieved **≈82% lower trajectory error**, **≈79% lower relative pose error**, **≈89% fewer collisions**, and **higher task success** than a wheel-IMU baseline, while maintaining **sub-16 ms** fusion latency compatible with real-time control. Gains persisted under lighting changes, wheel slip, occlusions, calibration perturbations, and sensor dropouts.

For practitioners, three design principles emerged:

1. **Treat calibration and timing as algorithms.** Continuous bias tracking, small extrinsic self-calibration, and strict time alignment often deliver accuracy equivalent to adding a new sensor.
2. **Fuse semantics with geometry.** Filtering dynamic objects from the persistent map prevents planner confusion; uncertainty should modulate local trajectory aggressiveness.
3. **Close the loop between estimation and planning.** When pose covariance and predicted interaction risk inform MPC, the robot anticipates ambiguity and behaves conservatively where needed.

**Future work** includes: (a) field validation with long-term calibration drift, (b) self-supervised learning to adapt sensor weights from context (e.g., weather, surface), (c) stronger loop-closure recognition resilient to look-alike corridors, (d) joint energy-aware planning under battery and compute constraints, and (e) formal safety proofs for uncertainty-triggered behaviors. While simulations cannot capture every real-world nuance, the architecture and results indicate that **tightly coupled, semantics-aware sensor fusion** is a pragmatic path to reliable last-meter autonomy for delivery robots.

## REFERENCES

- Alatise, M. B., & Hancke, G. P. (2020). A review on challenges of autonomous mobile robot and sensor fusion methods. *IEEE Access*, 8, 39830–39846. <https://doi.org/10.1109/ACCESS.2020.2975643>
- Behzadan, A. H., & Kamat, V. R. (2019). Data fusion for real-time construction site safety monitoring using sensor networks. *Automation in Construction*, 107, 102947. <https://doi.org/10.1016/j.autcon.2019.102947>
- Bloesch, M., Burri, M., Omari, S., Hutter, M., & Siegwart, R. (2015). Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback. *The International Journal of Robotics Research*, 36(10), 1053–1072. <https://doi.org/10.1177/0278364917728574>

- Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., & Leonard, J. J. (2016). *Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age*. *IEEE Transactions on Robotics*, 32(6), 1309–1332. <https://doi.org/10.1109/TRO.2016.2624754>
- Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M. M., & Tardós, J. D. (2021). *ORB-SLAM3: An accurate open-source library for visual, visual-inertial, and multimap SLAM*. *IEEE Transactions on Robotics*, 37(6), 1874–1890. <https://doi.org/10.1109/TRO.2021.3075644>
- Chen, J., Zhang, H., & Yu, H. (2020). *Semantic SLAM for autonomous vehicles using LiDAR and vision sensors*. *Sensors*, 20(19), 5519. <https://doi.org/10.3390/s20195519>
- Choi, S., Kim, T., & Yu, W. (2015). *Performance evaluation of RANSAC family*. *Journal of Computer Vision*, 93(3), 303–318. <https://doi.org/10.1007/s11263-014-0786-1>
- Della Corte, B., Grisetti, G., & Stachniss, C. (2018). *A unified framework for robust visual-inertial-wheel odometry*. *IEEE Robotics and Automation Letters*, 3(3), 2662–2669. <https://doi.org/10.1109/LRA.2018.2812911>
- Droschel, D., Behnke, S., & Schwarz, M. (2017). *Efficient continuous-time SLAM for 3D LiDAR-based online mapping*. *IEEE International Conference on Robotics and Automation*, 1–9. <https://doi.org/10.1109/ICRA.2017.7989571>
- Engel, J., Koltun, V., & Cremers, D. (2018). *Direct sparse odometry*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(3), 611–625. <https://doi.org/10.1109/TPAMI.2017.2658577>
- Gao, X., Wang, R., Demmel, N., Cremers, D., & Hu, Z. (2018). *DSO: Direct sparse odometry for monocular visual odometry*. *IEEE Robotics and Automation Letters*, 3(2), 1390–1397. <https://doi.org/10.1109/LRA.2018.2793349>
- Geneve, C., Wu, Y., Eustice, R. M., & Johnson-Roberson, M. (2019). *GP-SLAM: Gaussian process mapping for robotic navigation*. *IEEE Robotics and Automation Letters*, 4(2), 1279–1286. <https://doi.org/10.1109/LRA.2019.2893771>
- He, Y., Zhao, J., Guo, Y., Yuan, K., & Sun, F. (2020). *PL-VIO: Tightly-coupled monocular visual-inertial odometry using point and line features*. *Sensors*, 20(6), 1657. <https://doi.org/10.3390/s20061657>
- Huang, G., Mourikis, A. I., & Roumeliotis, S. I. (2014). *Visual-inertial navigation: A concise review*. *IEEE International Conference on Robotics and Automation*, 957–964. <https://doi.org/10.1109/ICRA.2014.6906963>
- Li, Q., Zhang, T., & Fang, Z. (2020). *LiDAR-IMU-camera calibration for autonomous navigation*. *Sensors*, 20(17), 4890. <https://doi.org/10.3390/s20174890>
- Liu, Y., Chen, W., Liu, X., & Yang, J. (2021). *A multi-sensor fusion framework for robust localization of autonomous vehicles in urban environments*. *Sensors*, 21(8), 2832. <https://doi.org/10.3390/s21082832>
- Mur-Artal, R., & Tardós, J. D. (2017). *ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras*. *IEEE Transactions on Robotics*, 33(5), 1255–1262. <https://doi.org/10.1109/TRO.2017.2705103>
- Qin, T., Li, P., & Shen, S. (2018). *VINS-Mono: A robust and versatile monocular visual-inertial state estimator*. *IEEE Transactions on Robotics*, 34(4), 1004–1020. <https://doi.org/10.1109/TRO.2018.2853729>
- Shan, T., Englot, B., Meyers, D., Wang, W., Ratti, C., & Rus, D. (2020). *LIO-SAM: Tightly-coupled LiDAR inertial odometry via smoothing and mapping*. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 5135–5142. <https://doi.org/10.1109/IROS45743.2020.9341176>
- Zlotnik, D., & Shiller, Z. (2017). *Using model predictive control for smooth and aggressive driving of autonomous vehicles*. *IEEE Conference on Decision and Control*, 1–8. <https://doi.org/10.1109/CDC.2017.8264444>