Spatiotemporal Automatic Calibration of Infrastructure Lidar, Radar, and Camera with a Global Navigation Satellite System

Preprint

Faizan Mir, Stanley Young, Rimple Sandhu, and Qichao Wang

National Renewable Energy Laboratory

Presented by 27th IEEE International Conference on Intelligent Transportation Systems
Edmonton, Canada
September 24-27, 2024

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5400-8978
August 2024

Spatiotemporal Automatic Calibration of Infrastructure Lidar, Radar, and Camera with a Global Navigation Satellite System

Preprint

Faizan Mir, Stanley Young, Rimple Sandhu, and Qichao Wang

Suggested Citation
Mir, Faizan, Stanley Young, Rimple Sandhu, and Qichao Wang. 2024. *Spatiotemporal Automatic Calibration of Infrastructure Lidar, Radar, and Camera with a Global Navigation Satellite System: Preprint*. Golden, CO: National Renewable Energy Laboratory. NREL/CP-5400-89785. https://www.nrel.gov/docs/fy24osti/89785.pdf.

NREL is a national laboratory of the U.S. Department of Energy
Office of Energy Efficiency & Renewable Energy
Operated by the Alliance for Sustainable Energy, LLC

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

Contract No. DE-AC36-08GO28308

Conference Paper
NREL/CP-5400-8978
August 2024

National Renewable Energy Laboratory
15013 Denver West Parkway
Golden, CO 80401
303-275-3000 • www.nrel.gov

NOTICE

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Transportation. The views expressed herein do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

This report is available at no cost from the National Renewable Energy Laboratory (NREL) at www.nrel.gov/publications.

U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via www.OSTI.gov.

Cover Photos by Dennis Schroeder: (clockwise, left to right) NREL 51934, NREL 45897, NREL 42160, NREL 45891, NREL 48097, NREL 46526.

NREL prints on paper that contains recycled content.

Spatiotemporal Automatic Calibration of Infrastructure Lidar, Radar, and Camera with a Global Navigation Satellite System

Faizan Mir,[1] Stanley Young,[1] Rimple Sandhu,[2] and Qichao Wang[2]

***Abstract*— Robust and accurate perception is important for modern intelligent transportation systems (ITS), which use sensors of various modalities for data fusion to create a digital twin of an intersection. Sensor calibration is an important process that creates a unified coordinate frame for the sensor output data so that it can be used for data fusion. Classical approaches for sensor calibration are time-consuming, require an overlapping field of view for feature matching, and are not feasible for ITS application as they cause disruptions in the flow of traffic. In this paper, we present a spatiotemporal automatic calibration approach to calibrate multiple infrastructure lidar, radar, and cameras installed at a traffic intersection. The approach uses global navigation satellite system (GNSS) positioning information shared by connected vehicles, and when the vehicle is detected by the sensor, we match the sensor detections with the GNSS coordinates. The proposed algorithm is evaluated with a real-world dataset utilizing detections from two radars, cameras, and lidars with a test vehicle instrumented with a post-processing kinematic (PPK)-corrected GNSS driving past the sensors installed at a four-way traffic intersection. The experimental results show that the proposed automatic calibration approach can achieve the transformation with a root mean squared error of less than 0.5 for radar and lidar and less than 2 for camera detections. The ability to rapidly calibrate sensors not only benefits initial installations, but can also be used for system health monitoring, while utilizing available connected vehicle data to test the real-time sensor fidelity and operational status.**

I. INTRODUCTION

Modern intelligent transportation systems (ITS) employ multiple sensors to obtain a robust estimate of the perceived environment at a traffic intersection. This allows the system to balance the shortcomings of one sensor type with the advantages of another by utilizing sensors with different modalities. The National Renewable Energy Laboratory's infrastructure perception and control concept proposes a cooperative perception engine that leverages data available from both infrastructure-based sensors (e.g., lidar, radar, cameras) and cooperatively shared information from connected autonomous vehicles and connected vehicles to support a wide variety of infrastructure applications such as trajectory-based optimized signal control, eco-approach and departure, curbside management, and safety-affirmative signaling [1][2]. In order to create a robust digital twin of the intersection, accurate spatial registration is required from these sensors for data fusion. The procedure of manually calibrating a sensor is costly and time-consuming. Automatic calibration is essential to manage the increasing number of sensors, particularly in multisensor systems. There are about 300,000 signalized intersection in the United States [3], and for a large-scale deployment, manually calibrating each sensor at an intersection is not feasible. Hence, a robust system is required to calibrate these sensors. Classical methods for sensor calibration have been previously developed, such as Zhang et al. [4] for multiple laser scanners and Xing et al. [5] for a limited overlapping multicamera setup that requires a calibration object be moved within the field of view of the sensor. Therefore, during the calibration process, the route that is being observed must be blocked for a few hours. These techniques are ineffective, as frequent road closures might negatively impact traffic flow and the calibration needs to be performed each time a sensor moves significantly due to vibrations or changing weather conditions. Multimodal sensor calibration is a complex problem due to the different physical measuring conditions and the difficulty in obtaining corresponding features from different sensor modalities.

Without a calibration object, an association problem must be solved on the basis of relative spatial and temporal alignment of cars in the intersection. In this paper we present a technique for multimodal sensor calibration at a traffic intersection without an explicit use of calibration objects. Calibration is achieved by tracking the test vehicle instrumented with a post-processing kinematic (PPK)-corrected global navigation satellite system (GNSS) within the sensor frame. The positional data from the GNSS-equipped test vehicle are shared over to the infrastructure with a vehicle-to-infrastructure (V2I) setup. Most modern vehicles are equipped with a GPS, although these positional data are not commonly available for use in traffic infrastructure systems. However, the approach in this work can be currently used by city traffic departments for calibration and will be convenient for future adoption as vehicle geographic data become available via V2I for traffic infrastructure systems from connected vehicles. To solve the problem of auto-calibration we used the Rauch-Tung-Striebel (RTS) smoother for the temporal alignment of sensors and solved an optimization problem to calculate the translation and rotational alignment of the sensors. We evaluate our approach on real-world data utilizing detection from two radars, lidars, and cameras installed at a four-way traffic intersection in the city of Colorado Springs, Colorado, USA, to show that it can achieve precise sensor calibration while utilizing detection from multiple sensor modes.

[1] Faizan Mir and Stanley Young are with the Center for Integrated Mobility Sciences, National Renewable Energy Laboratory, Golden, CO 80401 USA faizan.mir@nrel.gov; stanley.young@nrel.gov

[2] Rimple Sandhu and Qichao Wang are with the Computational Science Center, National Renewable Energy Laboratory, Golden, CO, USA rimple.sandhu@nrel.gov; qichao.wang@nrel.gov

II. RELATED WORK

Sensor calibration has been extensively researched in both academia and industry. The goal of calibration, more especially extrinsic calibration, is to determine the spatiotemporal transformation—i.e., the relative rotation, translation, and system clock delay—between two sensors. However, the focus has been to calibrate homogeneous sensors (e.g., camera to camera, radar-radar, lidar to lidar) [6][7][8]. Most of the research on sensor calibration from different modalities has been done with a focus on simultaneous localization and mapping (SLAM) [9]; however, these methodologies are not suitable for ITS applications. Generally, the sensor calibration can be classified into target-based and targetless methods. Target-based approaches require a calibration target such as a checkered board or a polygon board that can be accurately tracked within the sensor [10][11][12] and provides a reference for calibration. These methods can provide precise sensor calibration but can be hard to implement in a continuously moving traffic environment. On the contrary, targetless methods do not rely on a calibration target and extract features from the environment and then apply feature matching to find the correspondences between the sensors [13], [14], [15]. However, using these approaches on a traffic intersection presents two common challenges: (1) roadside point clouds are relatively sparse, as sensors are mounted high up on the poles to get a better field of view of the intersection, and (2) these approaches cannot be applied to sensors from different modalities such as radars, which are commonly used for traffic applications but lack descriptive visual features and output data in the form of detected objects.

These calibration ideas have recently been expanded to include different types of sensors and optimized for infrastructure-based perception. Peršić et al. [16] designed a triangular retroreflector calibration target for a radar and a 3D lidar. The calibration method used two-step optimization: reprojection error optimization followed by field of view optimization, leveraging radar cross-section measurements. Ge et al. [14] presented a targetless approach for calibrating an RGBD camera and millimeter-wave radar that involved extracting geometric constraints to provide initial estimates of extrinsic parameters using simulated annealing, and then using object velocity to fine-tune the calibration. Domhof et al. [17] proposed a joint lidar, radar, and camera calibration method by using a specially designed calibration target that uses four circular holes on a planar object to represent a unique geometric 3D shape detected by all three sensors. Ren et al. [6] presented TrajMatch, a spatiotemporal calibration methodology for roadside lidars that used a semantic matching feature and trajectory-level matching to calibrate the lidars. However, the paper only addresses the calibration for lidars and does not include sensors of other modalities. In Schöller et al. [18], an auto-calibration method for roadside radar and camera was presented by using a convolutional neural network to estimate the rotational calibration between sensors. However, the paper only focuses on rotational calibration and cannot be employed on a conventional four-way intersection.

A. Temporal Calibration

Time-based calibration is usually required because of the different sampling frequencies of sensors. Researchers have presented various approaches for temporal alignment of asynchronous signals such as Gaussian process regression [19] and data interpolation [20]. In Tanzmeister and Steyer [21], an alignment algorithm is proposed that takes the sensor with the highest sampling rate as the reference. In Du et al. [22], a vehicle motion-fitting model was proposed for temporal alignment between a camera and a millimeter-wave radar.

B. Spatial Calibration

After establishing the temporal correspondence, the spatial alignment can be calculated analytically by using commonly employed point-set registration techniques such as iterative closest point (ICP) [23]. The ICP algorithm works very well for SLAM applications where the pose between two transformations is small. It does not provide reasonable results for ITS applications, as it is heavily dependent on a good initial guess of the transformation.

III. CALIBRATION METHODOLOGY

In this section, we describe the proposed auto-calibration algorithm. The problem formulation and required presumptions are mentioned first, followed by a detailed explanation of the calibration algorithm.

A. Problem Formulation

Consider a multisensor (S_i) setup at an intersection with N sensors such that $S_i, i \in [1, N]$. We can define a set of 3D data points in world coordinates as $W = \{w_1, w_2, w_3, ...w_i\}$, and we can define a set of N sensors as $P_{S_N} = \{P_{S_1}, P_{S_2}, P_{S_3}, ...P_{S_N}\}$, where $P_{S_1} = \{p_1^1, p_2^1, p_3^1, ...p_i^1\}$ represents a set of 3D data points for the first sensor. The points in the sensor and world coordinates have the same dimension and are given as:

$$p_i^N = [x_i^N, y_i^N, t_i^N]^T, \quad p_i^N \in \mathbb{R}^3$$
$$w_i = [x_i^w, y_i^w, t_i^w]^T, \quad w_i \in \mathbb{R}^3$$

Therefore, each point in the sensor coordinate frame can be mapped to a point in the world coordinate frame:

$$w_i = \underbrace{\begin{bmatrix} R_{S_i}^w & T_{S_i}^w \\ 0 & 1 \end{bmatrix}}_{H_{Si}^w} . p_i^N \tag{1}$$

where $R_{S_i}^w$ is the rotation matrix and $T_{S_i}^w$ is the translation matrix. For a point, we can define the coordinate transform as:

$$\hat{p} = \begin{bmatrix} cos\theta & -sin\theta & 0 \\ sin\theta & cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ t \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ dt \end{bmatrix} \tag{2}$$

where θ is the angle between the two points, t_x and t_y are the translation in x and y, and dt is the time offset.

B. Algorithm Overview

An object is detected as soon as it is in the field of view of the sensor, assigned an object ID, and classified into categories such as cars, trucks, bicycles, pedestrians, etc. with its position and velocity information. The proposed framework takes the object trajectory information from multiple sensors and extracts features from the object list, which are invariant to object rotation and transformation, as the detections are made by each sensor in their localized coordinate frame. As with any multisensor setup, each sensor has a different detection frequency, and it is important to align these to a common time resolution to find the spatial alignment. Next, we use an RTS smoother for the temporal alignment of the selected object trajectory from the sensor and the PPK-corrected GNSS-equipped test vehicle. To increase tracking performance, the motion of the vehicles must be precisely represented with their dynamic behavior. Here we employ a constant velocity motion model to model the vehicle traveling through the intersection.

The state vector for the constant velocity model is defined as:

$$\dot{x} = [p_x, p_y, v_x, v_y]^T$$

where p_x and p_y represent the position in the x and y direction, respectively, and v_x and v_y is the velocity in the x and y direction, respectively. The discrete time state-space form can be defined as:

$$x_k = A_k x_{k-1} + q_k$$
$$y_k = H x_k + r_k$$

where $q_k \sim \mathcal{N}(0, Q_k)$; $T = t_k - t_{k-1}$;

$$A_k = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}; Q_k = \begin{bmatrix} \frac{\sigma_1^2 T^3}{3} & 0 & \frac{\sigma_1^2 T^2}{2} & 0 \\ 0 & \frac{\sigma_2^2 T^3}{3} & 0 & \frac{\sigma_2^2 T^2}{2} \\ \frac{\sigma_1^2 T^2}{2} & 0 & \sigma_1^2 T & 0 \\ 0 & \frac{\sigma_2^2 T^2}{2} & 0 & \sigma_2^2 T \end{bmatrix}$$

The discrete-time Kalman smoother, also known as the RTS smoother, is implemented to align the data from the sensors and the PPK-GNSS-equipped test vehicle. The Kalman smoother involves a two-step process: a forward standard Kalman filter and a backward smoothing filter. Equations (3)–(9) describe the Kalman filter and smoother algorithm.

Kalman Filter

Prediction

$$x_{k+1|k} = F_k x_{k|k} \quad (3)$$
$$P_{k+1|k} = F_k P_k F_k^T + Q_k \quad (4)$$

Update

$$K_{k+1} = P_{k+1|k} H^T [H P_{k+1|k} H^T + R]^{-1} \quad (5)$$
$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}[y_{k+1} - H x_{k+1|k}] \quad (6)$$
$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1}[R + H P_{k+1|k} H^T] K_{k+1}^T \quad (7)$$

where K is the Kalman gain and P is the error covariance matrix.

The smoother calculates the state posterior distribution $p(x_k \mid y_{k:k+N})$. The recursive equations for the backward step follow.

Kalman Smoother

$$x_{k|k+N} = x_{k|k} + C_k[x_{k+1|k+N} - x_{k+1|k}] \quad (8)$$
$$C_k = P_{k|k} F^T P_{k+1|k}^{-1} \quad (9)$$

where C_k is the smoother gain. Once we have the sensor points at the same time resolution as the GNSS positional data, we can match the position points from the trajectory at each time stamp to find the transformation that best fits the system until it reaches a threshold. The spatiotemporal problem can be set up as an optimization problem to minimize the matching error between the two sets of data points and is described as:

$$H_{S_i}^w = argmin \sum_i^P \left\| w_i - (R_{S_i}^w \times p_i^N + T_{S_i}^w) \right\| \quad (10)$$

We can use singular value decomposition [24] to calculate the rotational matrix R and translation vector T.

IV. EVALUATION AND RESULTS

A. Experimental Setup

The proposed calibration framework was evaluated at a four-way traffic intersection in the city of Colorado Springs, Colorado, USA. The intersection has a statue in the center that acts as a blind spot, and the sensors lose track of the object when behind the statue and get detected again as they move out of it. The sensors were installed at a corner of the intersection to track the cars in the intersection, as well as the ones approaching from the north and west. Two Econolite EVO radars were set up, with one directed toward the north-south and the second pointed toward the east-west (Figure 1). This ensured a full coverage of the intersection, as well as the approaching vehicles shown in Figure 2. Each EVO radar has a 110° field of view and about 250 m of range and can classify cars, trucks, and pedestrians. The sensor data were sent to an EVO Radar Hub for processing and recorded at a resolution of 10 Hz. Two Axis cameras were also mounted to record the data at the intersection, with each camera pointed in the same direction as the radar, as shown in Figure 1. A deep-learning model for object detection and classification developed by Kapsch was employed for real-time classification of the objects (cars, trucks, and pedestrians) inside and approaching the intersection using an edge device recording at a time resolution of 20 Hz. A set of lidars were also installed at the intersection. The lidars were pointed toward the intersection to detect and classify objects from the 3D point cloud data. This setup containing three different modes of sensors—with radar providing long-range detection, cameras detecting objects inside the intersection, and complementary 3D information from the lidar—provides

Fig. 1: Mounting location of the sensors with a statue in the center at the traffic intersection in the city of Colorado Springs, CO, USA.

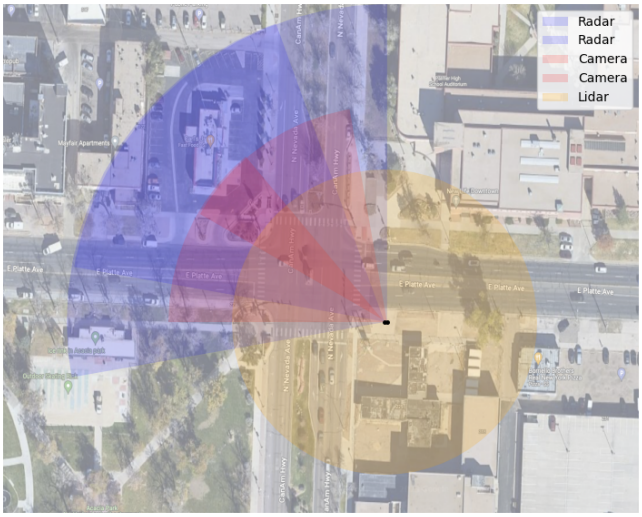


Fig. 2: The field of view of radar (blue) and camera (red) and Lidar (orange) at the intersection, with darker shades representing the overlap between sensors.

a robust image of the intersection that is essential for ITS applications and creating a digital twin.

A high-accuracy Emlid real-time kinematic GNSS was mounted on the top of a test vehicle, while a base unit was set up near the intersection for PPK correction of the geographic coordinates sent by the vehicle as it passed thought the intersection at a time resolution of 30 Hz. Figure 3 shows the GPS tracks of the instrumented vehicle as it passed through the intersection.

B. Results

As the test vehicle moves through the intersection it records its position in WGS-84 geodetic coordinates. First, the geodetic latitude and longitude coordinates are transformed to Earth-centered, Earth-fixed (ECEF) coordinates and then transformed to an east, north, up (ENU) tangent plane in terms of x and y data points. With the GNSS

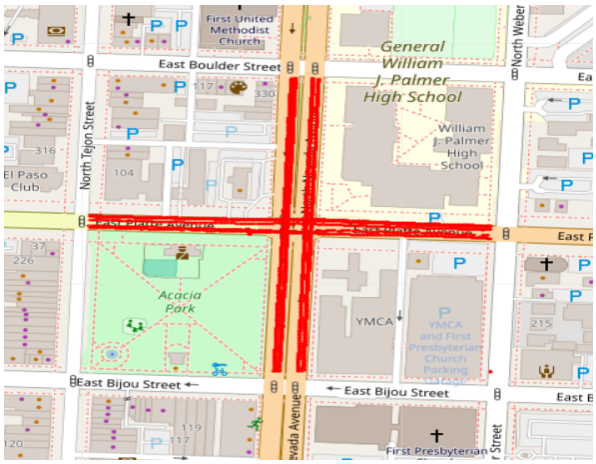
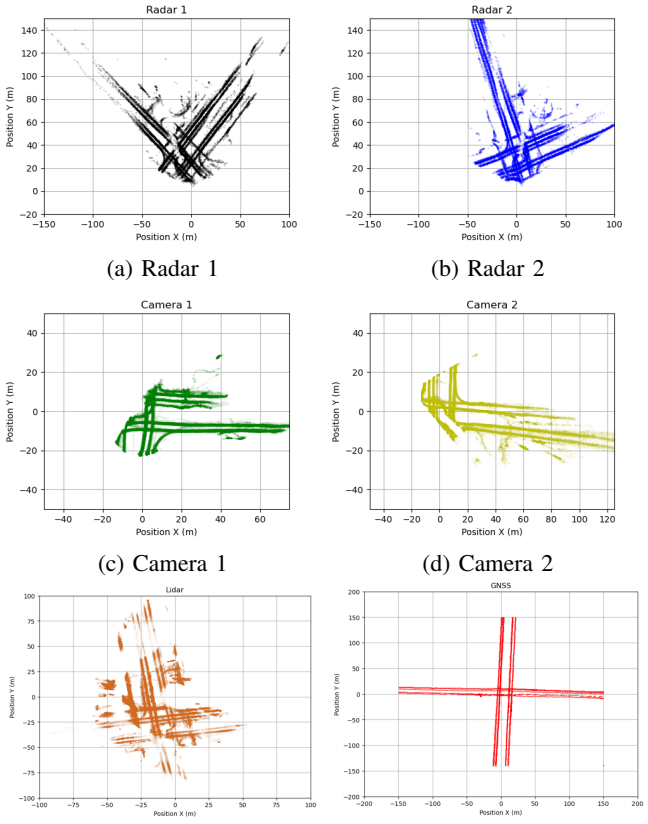


Fig. 3: PPK-GNSS tracks from the instrumented vehicle at the intersection.

positional data as our ground truth, we are able to determine the time-stamped detection from the sensor and apply these transformations to calibrate them. The vehicle and pedestrian tracks recorded for an hour and detected by the multisensor setup at the four-way traffic intersection are shown in Figure 4. The tracks shown are in local sensor coordinate frame.

The positional track of the test vehicle as it passes through the intersection is shown in Figure 5. In Figure 5a, the initial positional data of the test vehicle are shown as detected by Radar 1 (black) and Radar 2 (blue). After applying the proposed algorithm, the positional data from the radar sensor are transformed into a common coordinate system. In Figure 5b, the trajectory after applying a Kalman smoother is shown in black and blue for Radar 1 and Radar 2, respectively, whereas the data points represent the actual detection after transformation. The vehicle and pedestrian tracks after calibrating the sensors using the proposed auto-calibration framework are shown in Figure 6. The detections from Radar 1 and Radar 2 (Fig. 6c) have the least calibration error and are able to provide information about the vehicles approaching the traffic intersections as far as 100 m. The detections from the cameras as shown in Figure 6a match with the GPS tracks; however, there are certain detections that do not match with the GPS tracks. This deviation can be explained by the following reasons: (1) the camera suffers with distortion and needs additional post-processing to remove this distortion, or (2) there is an error in the center point of the vehicle in the image plane. The camera uses a fixed vehicle center point based on bounding box dimension to determine the center point of the vehicle, but this method is not consistent if the vehicle approaches the camera at various angles, hence causing the error in detection. Finally, for lidar we have good matching results in the intersections, as shown in Figure 6b. However, as we move away, we start to observe higher errors. This could be because the lidar was pointed down toward the intersection to get more information, and at farther distances it had fewer points reflecting off the objects.

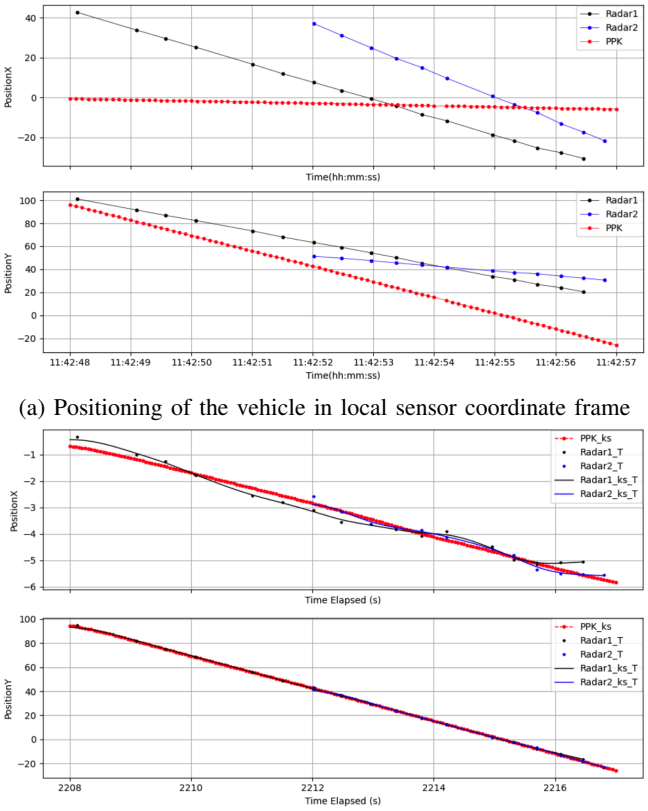


(a) Radar 1 (b) Radar 2

(c) Camera 1 (d) Camera 2

(e) Lidar (f) PPK-GNSS

Fig. 4: Vehicle and pedestrian tracks detected by (a) Radar 1, (b) Radar 2, (c) Camera 1, (d) Camera 2, (e) Lidar, and (f) tracks of test vehicle instrumented with PPK-GNSS passing through a four-way intersection.



(a) Positioning of the vehicle in local sensor coordinate frame

(b) Positioning of the vehicle after transformation

Fig. 5: Positional data of the vehicle: (a) initial output and (b) final solution.

C. Evaluation Metrics

The accuracy for the proposed calibration method is evaluated for each sensor by calculating the error in final transformation with the test vehicle data:

$$Error(S_N) = \|w_i - \hat{p}_1\|_2 \tag{11}$$

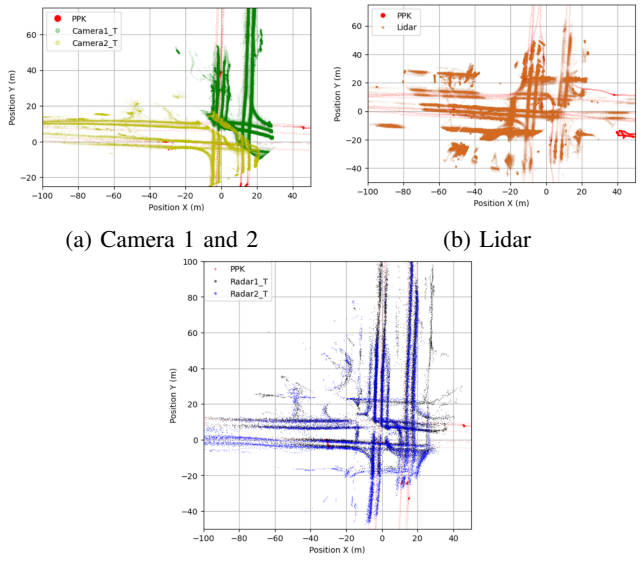
where w_i is the position vector for test vehicle and $\hat{p}_i$ is the position vector of the sensor after transformation.
The above error metric is combined into a single metric per sensor using the root mean square error (RMSE) given as:

$$RMSE(S_N) = \sqrt{\frac{\sum_1^k (Error(S_N))^2}{k}} \tag{12}$$

The RMSE values for each sensor after transformation using the proposed framework are given in the following table:

Sensor	RMSE Value (m)
Radar 1	0.30
Radar 2	0.35
Camera 1	1.7
Camera 2	1.9
Lidar	0.37

The RMSE values are low for radar and lidar, as we were able to get consistent detection for a test vehicle object ID,



(a) Camera 1 and 2 (b) Lidar

(c) Radar 1 and 2

Fig. 6: Vehicle and pedestrian tracks detected by (a) Camera 1 and 2, (b) Lidar, and (c) Radar 1 and 2 transformed after using the proposed framework.

but in the case of the camera, the RMSE values are higher due to the detection errors explained above. While manual sensor calibration using a calibration target outperforms the proposed auto calibration in terms of accuracy, such target-based calibration is not practical for a traffic intersection as already discussed. As our results are within an acceptable range it can be further used for data fusion. The proposed algorithm relies on GNSS for calibration and would suffer if there is a loss of GPS signal and in a dense urban setting as a weak GPS signal would lead to reduced calibration accuracy. Low sensor sampling rate would also affect the calibration accuracy as it would lead to higher motion estimation errors. The sensor time clocks must be accurately synchronized to prevent inaccurate spatial registration due to clock drift, which would also impact calibration performance.

V. CONCLUSIONS

In this work, an automatic calibration method is presented for calibrating infrastructure-based radar, lidar, and cameras that use trajectory information provided by a GNSS-equipped test vehicle to find the extrinsic calibration parameters for the sensors. By calibrating different types of sensors we are able to achieve a robust perception of the intersection. In future work, the algorithm should be tested at different road layouts such as urban, rural and highway setting. To ensure the proposed algorithm is agnostic of sensor brand or model, it should be evaluated with various sensor configurations, brands and further extended to perform system health monitoring and analyze sensor fidelity at the traffic intersection.

ACKNOWLEDGMENT

We are thankful to Dan Sines and Taryn Christoff from the city of Colorado Springs, CO for their help with sensor setup and data collection.

This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Transportation. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

REFERENCES

[1] S. E. Young, E. A. Bensen, L. Zhu, C. Day, J. S. Lott, R. Sandhu, C. Tripp, and P. Graf, *Concept of Operations of Next-Generation Traffic Control Utilizing Infrastructure-Based Cooperative Perception*, 2022, pp. 93–104. [Online]. Available: https://ascelibrary.org/doi/abs/10.1061/9780784484326.010

[2] S. E. Young, "Cutting-edge operations concepts: Intelligent infrastructure, cooperative driving, signal control, and curbside management," 8 2023. [Online]. Available: https://www.osti.gov/biblio/1994290

[3] N. T. O. Coalition, "National traffic signal report card," 2012.

[4] J. Zhang, Q. Lyu, G. Peng, Z. Wu, Q. Yan, and D. Wang, "Lb-l2l-calib: Accurate and robust extrinsic calibration for multiple 3d lidars with long baseline and large viewpoint difference," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 926–932.

[5] Z. Xing, J. Yu, and Y. Ma, "A new calibration technique for multi-camera systems of limited overlapping field-of-views," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 5892–5899.

[6] H. Ren, S. Zhang, S. Li, Y. Li, X. Li, J. Ji, Y. Zhang, and Y. Zhang, "Trajmatch: Toward automatic spatio-temporal calibration for roadside lidars through trajectory matching," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 11, pp. 12 549–12 559, 2023.

[7] K. Vuong, R. Tamburo, and S. G. Narasimhan, "Toward planet-wide traffic camera calibration," in *2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024, pp. 8538–8547.

[8] X. Li, Y. Liu, V. Lakshminarasimhan, H. Cao, F. Zhang, and A. Knoll, "Globally optimal robust radar calibration in intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 6, pp. 6082–6095, 2023.

[9] A. R. Vidal, H. Rebecq, T. Horstschaefer, and D. Scaramuzza, "Ultimate slam? combining events, images, and imu for robust visual slam in hdr and high-speed scenarios," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 994–1001, 2018.

[10] Y. Xie, R. Shao, P. Guli, B. Li, and L. Wang, "Infrastructure based calibration of a multi-camera and multi-lidar system using apriltags," in *2018 IEEE Intelligent Vehicles Symposium (IV)*, 2018, pp. 605–610.

[11] J.-K. Huang and J. W. Grizzle, "Improvements to target-based 3d lidar to camera calibration," *IEEE Access*, vol. 8, pp. 134 101–134 110, 2020.

[12] L. Cheng, A. Sengupta, and S. Cao, "3d radar and camera co-calibration: A flexible and accurate method for target-based extrinsic calibration," in *2023 IEEE Radar Conference (RadarConf23)*, 2023, pp. 1–6.

[13] D. Chang, R. Zhang, S. Huang, M. Hu, R. Ding, and X. Qin, "Versatile multi-lidar accurate self-calibration system based on pose graph optimization," *IEEE Robotics and Automation Letters*, vol. 8, no. 8, pp. 4839–4846, 2023.

[14] J. Ge, Y. Zhou, B. Lou, and C. Lv, "Automatic spatial radar camera calibration via geometric constraints with doppler-optical flow fusion," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 1818–1824.

[15] A. Sen, G. Pan, A. Mitrokhin, and A. Islam, "Scenecalib: Automatic targetless calibration of cameras and lidars in autonomous driving," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 7771–7777.

[16] J. Peršić, I. Marković, and I. Petrović, "Extrinsic 6dof calibration of 3d lidar and radar," in *2017 European Conference on Mobile Robots (ECMR)*, 2017, pp. 1–6.

[17] J. Domhof, J. F. P. Kooij, and D. M. Gavrila, "A joint extrinsic calibration tool for radar, camera and lidar," *IEEE Transactions on Intelligent Vehicles*, vol. 6, no. 3, pp. 571–582, 2021.

[18] C. Schöller, M. Schnettler, A. Krämmer, G. Hinz, M. Bakovic, M. Güzet, and A. Knoll, "Targetless rotational auto-calibration of radar and camera for intelligent transportation systems," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 3934–3941.

[19] J. Peršić, L. Petrović, I. Marković, and I. Petrović, "Spatiotemporal multisensor calibration via gaussian processes moving target tracking," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1401–1415, 2021.

[20] J. Rehder, R. Siegwart, and P. Furgale, "A general approach to spatiotemporal calibration in multisensor systems," *IEEE Transactions on Robotics*, vol. 32, no. 2, pp. 383–398, 2016.

[21] G. Tanzmeister and S. Steyer, "Spatiotemporal alignment for low-level asynchronous data fusion with radar sensors in grid-based tracking and mapping," in *2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 2016, pp. 231–237.

[22] Y. Du, B. Qin, C. Zhao, Y. Zhu, J. Cao, and Y. Ji, "A novel spatio-temporal synchronization method of roadside asynchronous mmw radar-camera for sensor fusion," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22 278–22 289, 2022.

[23] P. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[24] V. Klema and A. Laub, "The singular value decomposition: Its computation and some applications," *IEEE Transactions on Automatic Control*, vol. 25, no. 2, pp. 164–176, 1980.